



VIDYABHARATI MAHAVIDYALAYA AMRAVATI

NAAC Re-accredited with Grade "A" (CGPA 3.23-Third Cycle) | CPE Status (Thrice) by UGC

Mentor College under Paramarsh Scheme by UGC

'Lead College' by S.G.B. Amravati University, Amravati.

**3.3.2. Number of research papers per teachers in
the Journals notified on UGC website during the
year**



Sr. No	Title of paper	Name of the author/s	Name of journal	ISSN number	Link to the recognition in UGC enlistment of the Journal
1.	A combine stress of salinity and high temperature in vigna mungo (l.) Hepper leads to expression of more small heat shock proteins during in-vitro conditions	Dr. P. G. Bansod	The journal of oriental research madras [Vol. XCII-XLIII)	0022-3301	https://www.worldcat.org/formats-editions/1754710
2.	Transition from IPv4 to IPv6 Network in IoT Security Based upon Transition Methods	Prof. S. B. Sarvaiya	International Journal on Orange Technology (IJOT)	2615-8140	https://journals.researchparks.org/index.php/IJOT
3.	The Theme of Human suffering in " The Guide' by R. K. Narayan	Dr. V. P. Shekokar	Royal- An International Multidisciplinary Half Yearly Research Journal	2278-8158	http://www.sjifactor.com/
4.	Bulk Viscous LRS Bianchi Type I Cosmological Model of the Inflationary Universe in General Relativity	A. P. Wasnik, P. P. Khade	International Journal of Physics and Mathematics	2664-8636	https://www.physicsjournal.net/article/view/34/3-3-12
5.	Use of IT in Education	Dr. D. S. Wankhade	International E-research Journal	2348-7143	www.researchjourney.net
6.	Role of Nutrition in Sports Performance and Fitness	Dr. D. S. Wankhade	Multidisciplinary International Research Journal	2278-9308	www.aadharsocial.com
7.	Gender Inequality	Dr. D. S. Wankhade	Sanskriti International Multidisciplinary Research Journal	2455-1511	www.simrj.org.in

8.	A Role of Diet and Nutrition in Sports Performance and Fitness	Dr. D. S. Wankhade	Role of Physical activities, health and fitness in today's crisis	2277-8071	www.ycjournal.net
9.	The Phenotypic Effect of Cardamom oil on the Developmental Stages of <i>Drosophila melanogaster</i>	Dr. Y.D. Akhare	International Journal of Scientific Research in Science and Technology	2395-602X	https://ijsrst.com/paper/8460.pdf
10.	Variability of Nocturnal Insects In Sant Gadge Baba Amravati University Campus, Amravati	Dr. Y.D. Akhare	International Journal of Scientific Research in Science and Technology	2395-602X	https://ijsrst.com/paper/8445.pdf
11.	Effect of <i>Ocimum sanctum</i> Extract on Developmental Stages of <i>Drosophila melanogaster</i>	Dr. Y.D. Akhare	International Journal of All research Education and Scientific Methods	2455-6211	http://www.ijaresm.com/uploaded_files/document_file/Y. D . Akhare_bTkm.pdf
12.	Fourier transform infrared spectroscopy analysis of <i>Sauromatum venosum</i> (Ait.) schott. Tuber extract for alkaloids	Dr. M. U. Ghurde	International Journal of Botany Studies	2455-541X	www.botanyjournals.com
13.	The Core Principles of Management for Business Management	Dr.S.K.Rodde	IDEAL	2319 - 359X	http://www1.ajantaprasashan.com/?subid1=5d434b00-7c44-11ed-83bf-e40eae037a71
14.	Tourism Marketing and Sales Promotion Strategies	Dr. P. B. Upase	IDEAL	2319 - 359X	http://www1.ajantaprasashan.com/?subid1=5d434b00-7c44-11ed-83bf-e40eae037a71
15.	Analysis of IoT Data Transfer Messaging Protocols on Application Layer	Prof. S. B. Sarvaiya	International Journal for Research in Applied Science and Engineering Technology(IJRASET)	2321-9653	www.ijraset.com

16.	Trustworthy IoT Traditional Network Security to Authentication and Access Control Model for Heterogeneous Devices	Prof. S. B. Sarvaiya	International Journal of Innovative Research in Technology	2349-6002	www.ijirt.org
17.	Higher Education in Commerce aligned with Employability Skills: A myth or a reality	Dr. D. S. Rangacharya	Oriental Research	0022-3301	https://portal.issn.org/advancedsearch
18.	Cashless Transactions: Challenges and Remedies	Dr. D. S. Rangacharya	Royal	2278-8158	https://portal.issn.org/resource/ISSN/2278-8158
19.	Financial Inclusion: Solution for Banking Shadow	Dr. D. S. Rangacharya	Ajanta	2277-5730	https://portal.issn.org/resource/ISSN/2277-5730
20.	Empirical Study of Prediction Model in Education Domain	K. P. Raghuvanshi	Journal of Emerging Technologies and Innovative Research (JETIR)	2349-5162	http://www.jetir.org/view?paper=JETIR2211292
21.	An Empirical Study on Virtual Private Network	S. K. Totade	Journal of Emerging Technologies and Innovative Research (JETIR)	2349-5162	http://www.jetir.org/view?paper=JETIR2211292
22.	Cyber Security Awareness	Rana Afreen Sheikh	International Research Journal of Innovations in Engineering and Technology (IRJIET)	2581-3048	https://irjiet.com/
23.	Role of Web Technology in Primary Education	S. B. Bele	International Research Journal of Innovations in Engineering and Technology (IRJIET)	2581-3048	https://irjiet.com/

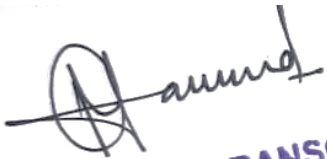
24.	Investigation of Nonlinear I-V characteristics of nanocrystalline tin oxide	Balkhande V.M., Abdul Tantray, Raulkar K.B., Lamdhade G. T	International Journal of Scientific Research in Science and Technology (IJSRST)	2395 - 602X	https://ijsrst.com/
25.	I-V characteristics of Cupric Oxide (CuO) thin film	Balkhande V.M., Abdul Tantray, Raulkar K.B., Lamdhade G.T.	International Journal of Scientific Research in Science and Technology (IJSRST)	2395 - 602X	https://ijsrst.com/
26.	Synthesis and Characterization of LiBaB9O15 : Gd ³⁺ Phosphor by Recrystallization Method	N. D. Kherde, A. O. Chauhan, P. A. Nagpure, S. K. Omanwar	International Journal of Scientific Research in Science and Technology (IJSRST)	2395 - 602X	https://ijsrst.com/IJSRST22291325
27.	Synthesis and Photoluminescence study of LaF ₃ :Gd ³⁺ phosphors for Phototherapy Application	A. O. Chauhan, C. B. Palan, N. S. Sawala, S. K. Omanwar	Journal of Emerging Technologies and Innovative Research (JETIR)	2349 - 5162	www.jetir.org
28.	Synthesis and luminescence properties of MSO ₄ :Eu (M=Ca, Ba) phosphors for radiation dosimetry	C. B. Palan, A. O. Chauhan, N.S. Sawala, S. K. Omanwar	Journal of Emerging Technologies and Innovative Research (JETIR)	2349 - 5162	www.jetir.org
29.	Synthesis and Photoluminescence properties of Eu ³⁺ ions doped LiBaPO ₄ phosphor	N. S. Sawala, S. R. Jaiswal, A. O. Chauhan, C. B. Palan, S. K. Omanwar	Journal of Emerging Technologies and Innovative Research (JETIR)	2349 - 5162	www.jetir.org
30.	Sthanik Swrajya Sanstha Ani Mahila Netrutva	Mr. Amit O. Ingole	Srujan Prabhat	2249-1171	https://portal.issn.org/resource/ISSN/2249-1171
31.	Misconceptions of Covid-19 Vaccine Among College Students	Dr. S. D. Wakode	Printing Area	2394 5303	https://www.vidyawarta.com/03/?p=5036

32.	Misconceptions of Covid-19: The Role of Social Media in it's Spread	Dr. S. D. Wakode	Research Journey	2348-7143	https://www.researchjournal.net/special-issues
33.	Effect of Endosulfanon Intestine and Pancreas of frog Ranatigrina	Rathod S. H, N. R, Thorat	International Journal of Scientific Research in Science and Technology	2395-6011	https://ijsrst.com/paper/9239.pdf
34.	Effects of water extract of <i>Parthenium hysterophorus</i> on Haematological contents of freshwater fish Labeo rohita	(Vikhar) Khedkar, A.P.	Journal of Emerging Technologies and Innovative Research	2349-5162	https://www.jetir.org
35.	Effect of <i>Terminalia Arjuna</i> Bark Extract on Streptozotocin-Induced Diabetic Rats	Akhare Y. D., Shraddha Sharma	International Journal of Current Science Research and Review	2581-8341	https://ijcsrr.org/single-view/?id=5743&pid=5593
36.	Effect of Extract of <i>Emblica officinalis</i> on Developmental Stages of Drosophila melanogaster	Dr. Y. D. Akhare, Shrutika S. Kurhekar	International Journal of Science and Research	2319-7064	https://www.ijsr.net/view.php?id=SR22425144801
37.	Effect of Neem Extract on the Developmental Stages of Drosophila Melanogaster	Y. D. Akhare, H. A. Patharikar, P. R. Kumare	International Journal of All Research Education and Scientific Methods (IJARESM),	2455-6211	www.ijaresm.com
38.	Identification of concealed structural alerts using QSTR modeling for Pseudokirchneriella subcapitata,	Vijay H. Masand , Magdi E. A. Zaki, Sami A. Al-Hussain, Anis Ben Ghorbal, Siddhartha Akasapu, Israa Lewaa, Arabinda Ghosh, Rahul D. Jawarkar,	Aquatic toxicology	0166-445X	https://www.scopus.com/authid/detail.uri?authorId=25626713900
39.	QSAR based virtual screening derived identification of a novel hit as a SARS CoV-229E 3CLpro Inhibitor: GA-MLR QSAR modeling supported by molecular Docking,	R. D. Jawarkar, Ravindra kumar L. Bakal, Magdi E. A. Zaki, Sami Al-Hussain, Arabinda Ghosh, Ajay kumar Gandhi, Nobendu	Arabian Journal of Chemistry	1878-5352	https://www.scopus.com/authid/detail.uri?authorId=25626713900

	molecular dynamics simulation and MMGBSA calculation approaches,	Mukerjee, Abdul Samad, Vijay H. Masand , Israa Lewaa,			
40.	Identification of potential edible mushroom as SARS-CoV-2 main protease inhibitor using rational drug designing approach,	Debanjan Sen, Bimal Debnath, Pradip Debnath, Sudhan Debnath, Magdi E. A. Zaki, Vijay H. Masand ,	Scientific Reports	2045-2322	https://www.scopus.com/authid/detail.uri?authorId=25626713900
41.	Exploring the Prominent and Concealed Inhibitory Features for Cytoplasmic Isoforms of Hsp90 Using QSAR Analysis,	Magdi E. A. Zaki, Sami A. Al-Hussain, Syed Nasir Abbas Bukhari, Vijay H. Masand , Mithilesh M. Rathore, Sumer D. Thakur, Vaishali M. Patil,	Pharmaceuticals	1424-8247	https://www.scopus.com/authid/detail.uri?authorId=25626713900
42.	Identification of potent aldose reductase inhibitors as antidiabetic (Anti-hyperglycemic) agents using QSAR based virtual Screening, molecular Docking, MD simulation and MMGBSA approaches,	Ravindra L. Bakal, Rahul D. Jawarkar, J. V. Manwar, Minal S. Jaiswal, Arabinda Ghosh, Ajaykumar Gandhi, Magdi E. A. Zaki, Sami Al-Hussain, Abdul Samad, Vijay H. Masand , Nobendu Mukerjee, Syed Nasir Abbas Bukhari, Praveen Sharma, Israa Lewaa,	Saudi Pharmaceutical Journal	1319-0164	https://www.scopus.com/authid/detail.uri?authorId=25626713900
43.	Mechanistic Analysis of Chemically Diverse Bromodomain-4 Inhibitors Using Balanced QSAR Analysis and Supported by X-ray Resolved Crystal Structures	Magdi E. A. Zaki, Sami A. Al-Hussain, Aamal A. Al-Mutairi, Vijay H. Masand , Abdul Samad, and Rahul D. Jawarkar	Pharmaceuticals	1424-8247	https://www.scopus.com/authid/detail.uri?authorId=25626713900
44.	Perceiving the Concealed and Unreported Pharmacophoric Features of the 5-Hydroxytryptamine Receptor Using Balanced QSAR Analysis	Syed Nasir Abbas Bukhari, Mervat Abdelaziz Elsherif, Kashaf Junaid, Hasan Ejaz, Pravej Alam, Abdul Samad, Rahul D. Jawarkar, Vijay H. Masand	Pharmaceuticals	1424-8247	https://www.scopus.com/authid/detail.uri?authorId=25626713900

45.	Synthesis and Evaluation of Some New 4H-Pyran Derivatives as Antioxidant, Antibacterial and Anti-HCT-116 Cells of CRC, with Molecular Docking, Antiproliferative, Apoptotic and ADME Investigations	Nahed N. E. El-Sayed, Magdi E. A. Zaki, Sami A. Al-Hussain, Abir Ben Bacha, Malika Berredjem, Vijay H. Masand , Zainab M. Almarhoon, Hanaa S. Omar	Molecules	1420-3049	https://www.scopus.com/authid/detail.uri?authorId=25626713900
46.	QSAR, Molecular Docking, MD Simulation and MMGBSA Calculations Approaches to Recognize Concealed Pharmacophoric Features Requisite for the Optimization of ALK Tyrosine Kinase Inhibitors as Anticancer Leads	Rahul D. Jawarkar, Praveen Sharma, Neetesh Jain, Ajaykumar Gandhi, Nobendu Mukerjee, Aamal A. Al-Mutairi, Magdi E. A. Zaki, Sami A. Al-Hussain, Abdul Samad, Vijay H. Masand , Arabinda Ghosh, Ravindra L. Bakal	Molecules	1420-3049	https://www.scopus.com/authid/detail.uri?authorId=25626713900
47.	QSAR Evaluations to Unravel the Structural Features in Lysine-Specific Histone Demethylase 1A Inhibitors for Novel Anticancer Lead Development Supported by Molecular Docking, MD Simulation and MMGBSA	Rahul D. Jawarkar, Ravindra L. Bakal, Nobendu Mukherjee, Arabinda Ghosh, Magdi E. A. Zaki, Sami A. Al-Hussain, Aamal A. Al-Mutairi, Abdul Samad, Ajaykumar Gandhi, Vijay H. Masand	Molecules	1420-3049	https://www.scopus.com/authid/detail.uri?authorId=25626713900
48.	Repurposing food molecules as a potential BACE1 inhibitor for Alzheimer's disease	Nobendu Mukerjee, Anubhab Das, Rahul D. Jawarkar, Swastika Maitra, Padmashree Das, Melvin A. Castrosanto, Soumyadip Paul, Abdul Samad, Magdi E. A. Zaki, Sami A. Al-Hussain, Vijay H. Masand , Mohammad Mehedi Hasan, Syed Nasir Abbas Bukhari, Asma Perveen, Badrah S. Alghamdi, Athanasios	Frontiers in Aging Neuroscience	1663-4365	https://www.scopus.com/authid/detail.uri?authorId=25626713900

		Alexiou, Mohammad Amjad Kamal, Abhijit Dey, Sumira Malik, Ravindra L. Bakal, Adel Mohammad Abuzenadah, Arabinda Ghosh, Ghulam Md Ashraf			
49.	Quinoline Derivatives with Different Functional Groups: Evaluation of Their Catecholase Activity	Mohamed Moutaouakil, Said Tighadouini, Zainab M. Almarhoon, Maha I. Al- Zaben, Abir Ben Bacha, Vijay H. Masand, Jamal Jamaledine, Rafik Saddik	Catalysts	2073-4344	https://www.scopus.com/authid/detail.uri?authorId=25626713900
50.	Pharmacophore Synergism in Diverse Scaffold Clinches in Aurora Kinase B,	Vijay H. Masand , Sami A. Al-Hussain, Mithilesh M. Rathore, Sumer D. Thakur, Siddhartha Akasapu, Abdul Samad, Aamal A. Al- Mutairi, Magdi E. A. Zaki	International journal of molecular sciences	1422-0067	https://www.scopus.com/authid/detail.uri?authorId=25626713900
51.	Artificial photosynthesis and current status of its application in generating fuels: A patent analytical study	Dr. Pradnya Nalawade	Journal of Emerging Technologies and Innovative Research (JETIR)	2349-5162	http://jetir.org/jetir%20ugc%20approval.pdf
52.	Synthesis and Antimicrobial Studies of newly synthesized 1-substituted-3-substituted propane - 1, 3-Dioness	Dr. P. S. Bodkhe, Sushil Pagariya, Rajendra Pathade, Ram Isankar	Journal of Advanced Scientific Research	0976-9595	http://www.sciensage.info


Prof. P. G. BANSOD
 Co-ordinator
 IQAC COORDINATOR
 Internal Quality Assurance Cell
 Vidya Bharati Mahavidyalaya
 Camp, Amravati-427602 (M.S.)




 SIGNATURE
 PRINCIPAL
 Vidya Bharati Mahavidyalaya
 Amravati.

A COMBINE STRESS OF SALINITY AND HIGH TEMPERATURE IN VIGNA MUNGO (L.) HEPPER LEADS TO EXPRESSION OF MORE SMALL HEAT SHOCK PROTEINS DURING IN-VITRO CONDITIONS*

BY

P. G. Bansod*

Vidya Bharati mahavidyalaya, Amravati

S. N. Malode*

Government Vidarbha Institute of Science and Humanities

Abstract

The analysis was carried out to investigate the induction of heat shock proteins during the seedlings stage of *Vigna mungo* L. Hepper. In this study three day old seedlings of *Vigna mungo* L. Hepper are exposed to high temperature stress along with the integrated high temperature and high salinity stress independently under the control laboratory conditions. Two types of treatments such as dry treatment and pre-soaked water treatments were used for seeds, in dry set seeds were directly treated to various concentrations of NaCl (0.2%, 0.4%, 0.6 % and 0.8%.) for 18 hrs, whereas, pre-soaked water treatments (PSW) seeds were soaked in distilled water for 12 hrs and then subjected the NaCl (0.4%, 0.8%, 1.2% and 1.6%) treatment for 6hrs. The enormous fluctuation in expression of total proteins as well as in SDS-profile has been observed. Many small HSPs of low molecular weight are reported in this analysis such as 7.5 KDa, 5.7 KDa and 3.6 KDa along with a protein with molecular weight 43 KDa which was constitutively expressed in all seedlings irrespective of the high temperature and high salinity treatments.

Keywords – High temperature, salinity, *Vigna mungo*, SDS-PAGE, HSPs.

Received 22 July 2021, Accepted 06 August 2021, Published 12 August 2021

* Correspondence Author: P. G. Bansod

Introduction

Abiotic stresses are serious threats to agriculture and the environment which have been exacerbated in the current century by global warming and industrialization.

Some of the most serious effects of abiotic stresses occur in the arid and semiarid regions where rainfall, high evaporation low, native rocks, saline irrigation water, and poor water management all contribute in agricultural areas. Abiotic stresses lead to series of morphological, physiological, biochemical changes that adversely affect the plant growth and productivity (Wang *et al.*, 2001b). Drought, salinity, extreme temperatures and oxidative stress are often inter connected and may induce similar cellular damage, for example drought and salinity are manifested primarily as osmotic stress,

resulting in the disruption of homeostasis and ion distribution in the cell (Serrano *et al.*, 1999; Zhu, 2001b).

Gisela Jansen (2008) evaluated effects of temperature on yield and protein content of *Lupinus angustifolius* cultivars, showed that higher temperature led to a decrease in yield, but rise in protein content. Don *et al.*, (2005) reported that high temperature affects the high molecular weight fraction of gluten protein in wheat. They reported significant effects of prolonged exposure to high temperatures (up to 40°C) on gluten in macro polymer (GMP) and its constituting gluten in particles showed that changes in dough mixing requirements were directly related to changes in gluten in macro polymer. Similar reductions were observed in starch, protein and oil contents of the maize kernel (Wilhelm *et al.*, 1999). Nayer and Heidari (2008), generated the drought stress in maize plant by PEG 6000, they observed that after 24 hours treatment the total soluble protein content of two maize cultivars (C-704 and C-301) was first increased and then decreased in the roots and leaves. The decrease in total soluble protein content was equal in both varieties, but in the leaves of CV-301 was greater than in CV-704. In drought the accumulation of dehydrin like 38, 50, 57 and 65 KDa root protein and 22 KDa leaf proteins was induced in both varieties. A glass house study was conducted by Bako (2006) to evaluate the effects of high temperature growth conditions on crude protein content in maize (*Zea mays* L.) plants. The crude proteins declined significantly with plant age. Duca and Barsan, (2001) recorded an increase of protein content in the vegetative tissues especially under the action of high concentration of salts. High temperature stress causes a complete loss of low temperature proteins and it is replaced by lower rate of protein synthesis of heat induced proteins. Thus the metabolic inhibition by high temperature is due to a multifaceted impact on enzyme degradation and enzyme specific activity (Nargas and Swarnkar, 2000)

In most of the studies of heat shock induced stress in plants, plant cells, tissues, organs or whole plants have been exposed to only one environmental stress factor i.e. high temperature. However under field conditions plants are often simultaneously exposed to soil drying and high temperature stress. Furthermore in the arid area like Vidarbha many of the soil are saline having considerable concentration of salts. *Vigna mungo* already facing dual stresses, has to be withstand the additional saline stress in arid areas. These multiple stresses could create a novel effect in physiology of crop plants. As the Heat Shock Protein are prime molecules to be expressed during such stresses, it become essential to study the effect of such multiple stresses on expression of Heat Shock Proteins (HSPs). In this investigation we emphasized on high temperature stress independently as well as high temperature in conjunction with salinity.

Materials and Methods

Parental material of *Vigna mungo* (L) Hepper six cultivars were obtained from various Institutes. Four cultivars namely AKU-15, AKU-9904, BDU-1 and TAU-1 were obtained from PKV, Akola and two remaining cultivars i.e. Azad-1 and Shekhar were obtained from Department of Pulses, Pusa, New Delhi. Lethal dose 50 (LD₅₀) is used as a criteria to define the optimum chemical dose. For

determination of (LD_{50}) treatments of various concentrations were given to the dry as well as pre-soaked seeds. Treated seeds then tested for germination. After 48 hours, germinated seeds were counted and recorded the germination percentage.

Methods of treatment:-

Pure and homogenous seeds of *Vigna mungo* (L) Hepper were treated with different concentrations of NaCl. For treatments dry as well as presoaked in water (PSW) seeds were utilized. All the treatments were carried out in triplicates at Cytogenetic and Molecular Biology Laboratory, Department of Botany, Government Vidharbha Institute of Science and Humanities, Amravati. In case of dry treatment seeds were directly treated in various concentrations of chemicals solution for 18hrs, whereas, pre-soaked water treatments (PSW) seeds were soaked in distilled water for 12 hrs and then subjected the chemical treatment for 6hrs. All the germination trays were kept in germination chambers and maintained at 25°C temperature with 70% relative humidity. For dry seeds treatment uniform seeds were directly soaked in the mutagenic solution, for about 18 hrs.

In pre-soaked water (PSW) treatments dry seeds were soaked in distilled water for 12hrs and then exposed to chemical solution for 6 hrs. All treatments were carried out in triplicates at 24±0.5°C in Remi orbital shaking incubator. For presoaked water treatments with NaCl concentrations used were 0.4%, 0.8%, 1.2% and 1.6% and for dry treatments it was 0.2%, 0.4%, 0.6 % and 0.8%.

Quantitative and qualitative analysis of proteins: The quantitative analysis of proteins is carried out by Bradford method for which three day old seedlings of *Vigna mungo* (L) Hepper grown in the germination tray after providing relevant treatments and maintained at 25°C were used. Qualitative analysis of proteins is done by SDS-PAGE (Lammaeli, 1970).

Result and Discussion

1. Temperature stress

When the comparison was made between the quantity of protein in three day old seedling of six cultivars of *Vigna mungo* (L) Hepper treated with temperature alone, it has been observed that, in dry treatment at 35°C, the amount of protein was highest (126.74mg/g) in AKU-15 cv. and lowest (37.08mg/g) in Shekhar cv. which shows 77.36% increase. Whereas at 40°C, it was highest (129.1mg/g) in AKU-15 cv. and lowest (34.72mg/g) in Azad-1 cv. which shows 78.80% increase. At 45°C it was highest (110.22 mg/g) in AKU-15 cv. and lowest (37.08 mg/g) in Shekhar cv. which shows 74.82% increase (Table-1, 3 and 5).

In PSW treatment, at 35°C the quantity of protein among six cultivars was highest (130.46 mg/g) in AKU-15 cv. and lowest (18.2 mg/g) in Shekhar cv. which shows 87.75% increase; whereas at 40°C, the protein quantity was highest (152.68 mg/g) in AKU-15 cv. and lowest (32.36 mg/g) in Shekhar cv. which shows 82.51% increase. At 45°C temperature the protein amount was highest (136.18 mg/g) in AKU-15 cv. which shows 83.12% increase and lowest in (27.64 mg/g) Shekhar cv. (Table-1 and 5).

2. Combine temperature and salinity stress

The comparative study of protein quantity among the three day old seedling of six cultivars of *Vigna*

THE JOURNAL
 The seedling
 salinity str
 and the
 given

mungo (L) Hepper treated with salinity (NaCl) revealed that, in dry treatment at 35⁰C, the amount of protein was highest (136.18 mg/g) in AKU-15 cv. which shows 51.79% increase and lowest (4.04mg/g) in Shekhar cv. Whereas at 40⁰C, it was highest (166.8 mg/g) in AKU-15 cv. which shows 56.36% increase and lowest (11.12 mg/g) in Azad-1 cv. At 45⁰C it was highest (140.9mg/g) in AKU-15 cv. which shows 56.10% increase and lowest (13.48 mg/g) in Shekhar cv. (Table-1 and 5).

In PSW salinity treatment, at 35⁰C the quantity of protein among six cultivars was highest (130.46 mg/g) in AKU-15 cv. which shows 52.15% increase and lowest 4.04mg/g in Shekhar cv.; whereas at 40⁰C, the protein quantity was highest (152.68mg/g) in AKU-15 cv. which shows 54.18% increase and lowest (4.04mg/g) in Azad-1 cv. At 45⁰C temperature the protein amount was highest (136.18mg/g) in AKU-15 cv. which shows 53.22% increase and lowest (8.76mg/g) in Shekhar cv. (Table-1 and 5).

In most of the cases initially during the rise of temperature i.e. up to 40⁰C protein amount also increased. Above 40⁰C temperature this protein amount in seedling gradually or suddenly decreased. Our observations in this regards are same, as at 40⁰C most of the seedlings shows highest amount of proteins. The Azad-1 cultivar was more responsive to temperature and salinity stress in terms of protein production. The increase in protein amount was observed in AKU-15> TAU-1> BDU-1> AKU-9904> Azad-1> Shekhar this trend (Fig. 1)

Fig.1 - Protein quantity of six cvs of *Vigna mungo* (L) Hepper in response to temperature (mg/g)



The literature available on this aspect with reference to this plant is insufficient to fully understand the role of effects of temperature stress but in most of the cases initially during the rise of temperature i.e. up to 40⁰C protein amount also increased. Above 40⁰C temperature this protein amount in seedling gradually or suddenly decreased. Our observations in this regards are same, as at 40⁰C most of the seedlings shows highest amount of proteins.

SDS-PAGE Analysis

The seedlings in which the temperature stress imposed independently as well as in conjunction with salinity stress, many new polypeptides have been induced, which are not found in control seedlings, and these are mostly of low molecular weight. The treatment wise induction of new polypeptides is given below.

Cultivar	Stress applied	New polypeptides obtained
AKU-9904	1.2% NaCl /PSW/ 40°C	21.8 and 5.9 KDa (Plate- 5)
	0.4,0.8,1.2 and 1.6 % NaCl/PSW/ 40°C	169.4, 23.9, 7.5 and 3.6 KDa (Plate- 5)
Azad-1	1.2 % NaCl/PSW/ 45°C	5.7 KDa (Plate- 6)

Table 19 - New polypeptides obtained on SDS-PAGE gel of the three day old seedlings of *Vigna mungo* L. (Hepper) treated with various concentration of NaCl and temperature stress.

The bands obtained in the SDS-PAGE analysis reveal that the heavy bands like 169.4 KDa obtained in AKU-9904 at 40°C in 0.4, 0.8, 1.2 and 1.6 % NaCl - PSW (Plate- 5); are very few as compared to the other bands, which mostly belong to low molecular weight proteins below 30 KDa.

5.9 KDa obtained in AKU-9904 at 40°C in 1.2% NaCl-PSW (Plate-5); 7.5 and 3.6 KDa obtained in AKU-9904 at 40°C in 0.4, 0.8, 1.2 and 1.6 % NaCl-PSW (Plate-5) and 5.7 KDa obtained in Azad-1 at 45°C in 1.2% NaCl-PSW (Plate-6) are the characteristic feature of this investigation.

In some of the treatments it has been observed that the temperature alone i.e. the control seedlings expressed more number of polypeptides than treatment, which indicates the suppression of the synthesis of many proteins in combined stress. These polypeptides are as follows.

Cultivar	Stress applied	New polypeptides obtained in control
Azad -1	Dry / 45°C	17.9 KDa (Plate-2)
BDU-1	PSW/ 45°C	18.9 KDa and 14.6 KDa (Plate-6)
Shekhar	PSW/ 45°C	5.1 KDa (Plate-7)
TAU-1	PSW/ 40°C	94.3, 80.2 and 68.0 KDa (Plate-7)

Table 20 - New polypeptides obtained on SDS-PAGE gel of the three day old control (temperature alone) seedlings of *Vigna mungo* L. (Hepper).

The bands obtained in the SDS-PAGE analysis also reveal that in some cases the seedling without treatment i.e. control shows the bands that did not appeared in the treatment which suggest that some of the proteins are inhibited from being synthesized. This inhibition of protein synthesis may be due to negative regulation of gene expression.

Among five conserved families of HSPs (HSPs 100, HSPs 90, HSPs 70, HSPs 60 and small HSPs), the small Heat Shock Proteins found to be most prevalent in plants. Small HSPs vary in size from 12 to 40 KDa (Vierling, 1991). Various studies have shown that plant small HSPs are not only expressed in response to Heat shock, but also under water, salt and oxidative stress (Almoguera *et al.*, 1993; Alamillo *et al.*, 1995; Sabehat *et al.*, 1998; Harndahl *et al.*, 1999; Hamilton and Heckathorn, 2001).

The first leaf tissues proteins of cultivated tetraploid (*Triticum durum*) and hexaploid (*T. aestivum*) wheat species exposed to 100 mM NaCl stress were separated by two dimensional gel

electrophoresis by Terzi and Yildiz (2008). There was no newly synthesized protein in the NaCl treatment compared to the control treatment. As a result of salt stress the most remarkable change was increase in total proteins.

Meratan (2008) evaluated effect of salinity on growth, proteins and antioxidant enzymes in three *Acanthophyllum* species of different ploidy levels, NaCl generally increased protein content in seedling of all the species. Similar result also obtained in, seedlings of *Nicotiana tabacum* (Niknam *et al.*, 2004) and *Trigonella foenum-graecum* increased under NaCl stress (Niknam *et al.*, 2006).

Gloria Irma (2010) evaluated the effects of saline stress on soluble proteins on *Paulownia imperialis* (Siebold & Zuccarini) and *Paulownia fortunei* (Seemann & Hemsley) grown in vitro. Protein content was higher in plants grown in the lowest concentration of sodium chloride (20 mM) at 15 days. Protein content was statistically similar in plants grown at concentrations of 40, 60, and 80 mM of sodium chloride, and significantly decreased at the highest sodium chloride concentration (160 mM).

Low molecular weight heat shock proteins in both plants and animals are together aggregated into a protein of higher molecular weight (Neumann *et al.*, 1989). The increase in total protein content after the combined stress can be attributed to these aggregations of higher molecular weight protein. This implies that combined stress can be resulted into the increased gene expression.

In the present investigation control refers to the seedlings which have been treated with temperature only, as the main aim of the study was to find out the effect of combined stress. The results revealed that temperature independently is more efficient in inducing the higher protein amount as compared to temperature in conjunction with salinity. Among the plants with salinity induced stress in conjunction with rise in temperature, it has been revealed that at the lower concentrations, these stresses improve the amount of proteins in the seedlings. In the seedling stage of the plants, active localization of metabolites such as DNA and proteins are required which moves from one cellular compartment to other. This movement in turn is regulated by signal transduction mechanism which is the output of signal received from outside the cell, either in the form of chemical or environmental stress. The increase in seedling protein amount may be attributed to increase in signal transduction rate.

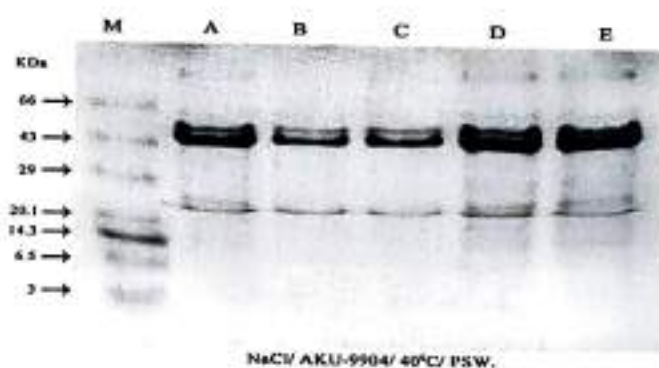
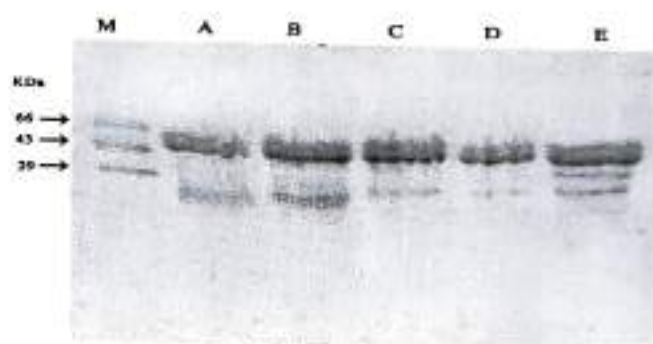
Temperature stress is one of the important environmental factors that may affect morphology, anatomy and plant biochemistry at all levels of organization. Direct injuries due to high temperatures in plants include protein denaturation. Recognition and signaling pathways regulating the responses to abiotic stresses (e.g. drought, salinity, cold and heat) are similar to those used for responding to biotic stresses. The adaptation to one stress condition can affect tolerance to other non-related stresses, a phenomenon referred to as cross-tolerance.

In response to temperature stress various approaches are being used by the plants, which can mitigate the effect of stress and lead to the adjustment of the cellular environment and plant tolerance. In nature stress does not generally come in isolation and many stresses act hand in hand with each

other. In response to these stress signals that cross talk with each other, plants naturally have developed diverse mechanisms for combating and tolerating them.

The developmental stage of the plant is important in order to withstand or sustain the adverse fluctuation in the environment. In fact it is the early development stages of living organism which ultimately decides their fate in future. In the SDS-PAGE profile of the three day old seedlings, in most of the cases, the lane on gel representing control plant proteins showed or expressed less number of bands as compared to treated, where the number ranges from 3 to 9 bands. The high intensity band of molecular weight ~ 40 KDa was common in all the gel irrespective of control and treatment. This protein may be a product of housekeeping gene which performs the normal function of plants. The expression of this protein is not affected by temperature, salinity and mutagens. This band can be utilized to trace phylogenetic relationship between various species of *Vigna*. This band could be a unique feature of our analysis.

PLATE-V



M-Marker, A-Control, B-0.4%, C-0.8%, D-1.2%, E-1.6%.

PLATE-VI

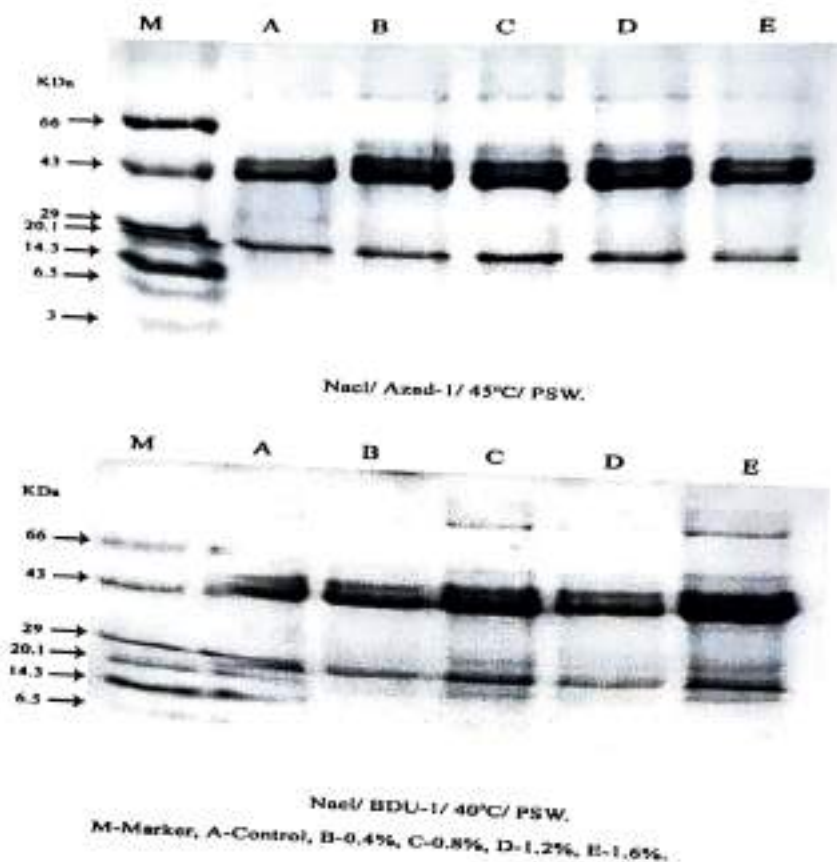
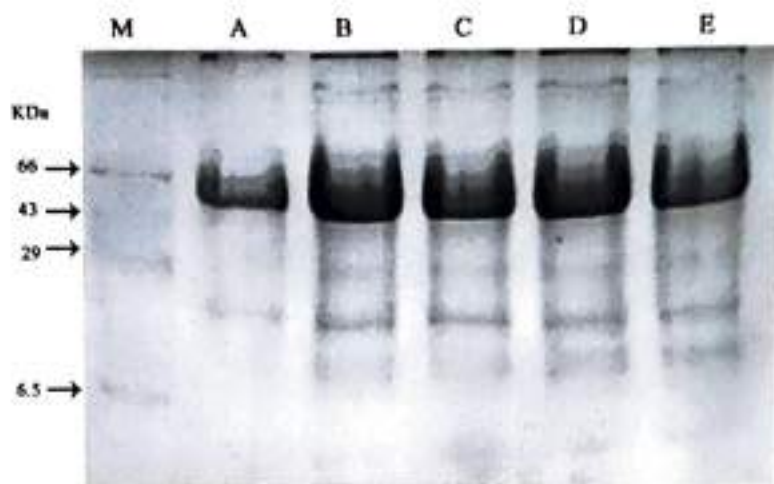
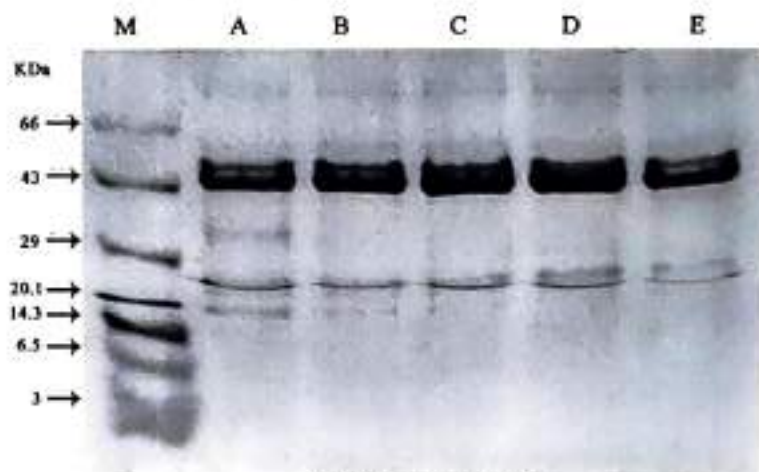


PLATE-II



NaCl/ AKU-9904/ 45°C/ Dry.

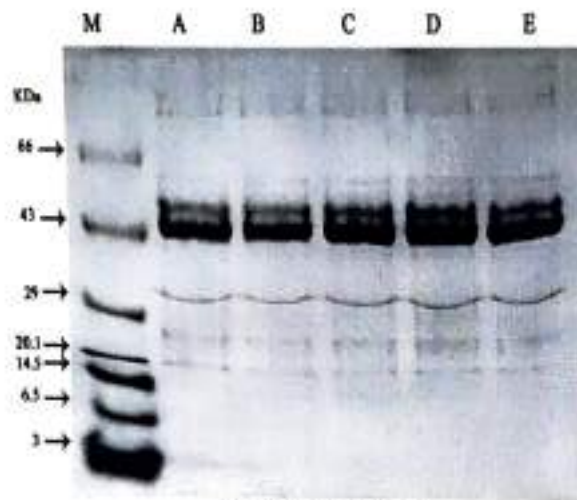
M-Marker, A-Control, B-0.4%, C-0.8%, D-1.2%, E-1.6%.



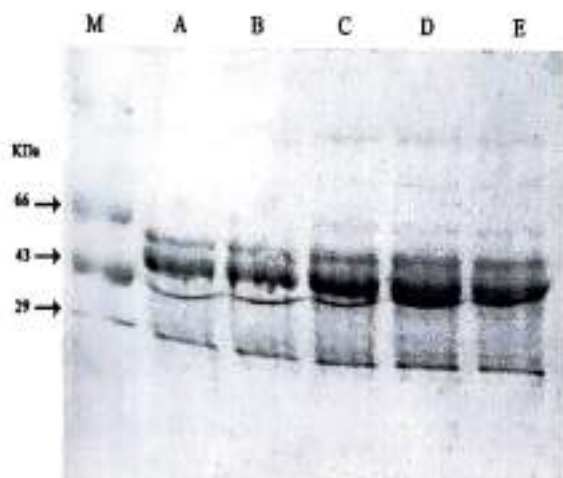
NaCl/ Azad-1/ 45°C/ Dry.

M-Marker, A-Control, B-0.2%, C-0.4%, D-0.6%, E-0.8%.

PLATE-I



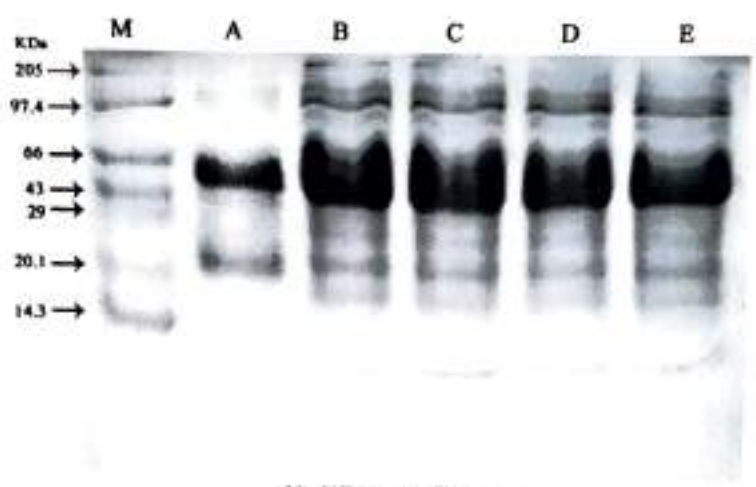
NaCl/ AKU-15/ 45°C/ Dry.



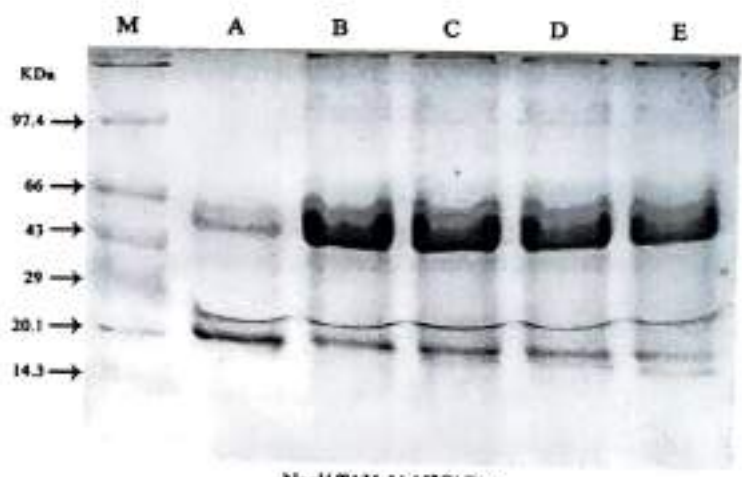
NaCl/ AKU-9904/ 45°C/ Dry.

M-Marker, A-Control, B-0.2%, C-0.4%, D-0.6%, E-0.8%

PLATE-IV



NaCl/TAU-I/40°C/Dry.



NaCl/TAU-I/45°C/Dry.

M-Marker, A-Control, B-0.2%, C-0.4%, D-0.6%, E-0.8%.

References-

1. **Alamillo J., Almoguera C., Bartels D. and Jordano J. (1995).** Constitutive expression of small heat shock proteins in vegetative tissues of the resurrection plant *Craterostigma plantagineum*. *Plant Mol. Biol.* 29: pp. 1093- 1099.
2. **Almoguera E., Coca M.A. and Jordano J. (1993).** Tissue specific expression of sunflower heat shock proteins in response to water stress. *Plant J.* 4: pp. 947-958.
3. **Bako S.P. (2006).** Effects of plant age, ascorbate and Kinetin applications on nitrate reductase activity and leaf protein content of Maize (*Zea mays* L) plant grown under heat stress. *Asian Journal of plant sciences.* 5 (2): pp. 363 – 367.
4. **Duca Maria and Barsan Ana (2001).** The modification of protein metabolism of sunflower plants under saline stress. *Romanian Agricultural Research*, 16, pp.5-16.
5. **Gloria Irma (2010).** Salinity effects on protein content, lipid peroxidation, pigments, and proline in *Paulownia imperialis* (Siebold & Zuccarini) and *Paulownia fortunei* (Seemann & Hemsley) grown *in vitro*. Vol. 13 No. 5.
6. **Hamilton III E.W. and Heckathorn S.A. (2001).** Mitochondrial adaptations to NaCl complex I is protected by antioxidant and small heat shock proteins, whereas complex II is protected by proline and betaine. *Plant Physiology.* 126: pp. 1266-1274.
7. **Harndahl U., Hall R.B., Osteryoung K.W., Vierling E. Bornmann J.F. and Sundby C. (1999).** The chloroplast small Heat Shock protein undergoes oxidation dependent conformational changes may protect plant from oxidative stress. *Cell Stress chaperones.* 4: pp.129-138.

8. **Jansen Gisela (2008)**. Effects of temperature on field and protein content of *Lupinus angustifolius* cultivars. *Proceedings of 12th International Lupin conference 14-18 Sept.* ISBN 0-86476 – pp.183-8.
9. **Laemmli U.K. (1970)**. Cleavage of structure Protein during the assembly of Head of bacteriophage T4. *Nature*. 227: pp. 680-685.
10. **Meratan A.A., Ghaffari S.M. and Niknam V. (2008)**. Effects of salinity on growth, proteins and antioxidant enzymes in three *Acanthophyllum* species of different ploidy levels. *JSUT* 33(4), pp. 1-8.
11. **Nargas, J. and Swarnkar, P.L. (2000)**. Plant responses under heat and chill stresses. In: *Plant Biotechnology: Recent advances* (Ed. Trivedi, P.C.), Panima Publishers.
12. **Nayer Mohammdkhani and Reza Heidari (2008)**. Effects of drought stress on soluble protein in two maize varieties. *Turk J. Biol.* 32: pp. 23-30.
13. **Neumann D., Nover L., Parthier B., Rieger R., Scharf K. D., Wollgiehn R. and Nieden U. (1989)**. Heat shock and other stress response systems of plants, *Biol. Zentralbl.* 108:1-155.
14. **Niknam V., Bagherzadeh M., Ebrahimzadeh H. and Sokhansanj A. (2004)**. Effect of NaCl on biomass and contents of sugars, proline and proteins in seedlings and leaf explants of *Nicotiana tabacum* grown *in vitro*. *Biol. Plant.* 48: pp. 613-615.
15. **Niknam V., Razavi N., Ebrahimzadeh H. and Sharifizadeh B. (2006)**. Effect of NaCl on biomass, protein and proline contents, and antioxidant enzymes in seedlings and calli of two *Trigonella* species. *Biol. Plant.* 50:591-596.
16. **Sabehat A, Lurie S. and Weiss D. (1998)**. Expression of small Heat shock proteins at low temperatures. *Plant Physiol.* 117: pp. 651-658.

17. **Serrano R., Mulet J.M., Rios G., Marquez J.A., de Larinova I.F., Leube M.P, Mendizabal I., Pascual-Ahuir A., Proft M., Ros R. and Moutesinos C. (1999).** A glimpse of the mechanisms of ion homeostasis during salt stress. *J. Exp. Bot.* 50: pp. 1023-1036.
18. **Terzi H. and Yildiz M. (2008).** Effects of NaCl on protein profiles of tetraploid and hexaploid wheat species and their diploid wild progenitors. *Plant Soil Environ.* 54 (6): pp. 227: 233.
19. **Vierling E. (1991).** The roles of heat shock proteins in plants. *Annu. Rev. Plant Biol.* 42: pp. 579-620.
20. **Wang W.X., Vinocur B., Shoseyov O. and Altman A. (2001b).** Biotechnology of plant osmotic stress tolerance: physiological and molecular considerations. *Acta Hort.* 560: pp. 285-292.
21. **Zhu J.K. (2001b).** Cell signaling under salt water and cold stress. *Curr. Opin. Plant Biol.* 4: pp. 401- 406.

Transition from IPv4 to IPv6 Network in IoT Security Based Upon Transition Methods

Ms. Shilpa B. Sarvaiya¹, Dr. D. N. Satange²

¹Department of Computer Science, Vidyabharati Mahavidyalaya, Amravati, Maharashtra, India

²Director, Student Development, S. G. B. A. University, Amravati, Maharashtra, India

Abstract: While deployments of IPv6 networks have increased over recent years, especially in IoT Paradigm. Today there are two types of internet protocol versions that are currently working in the global internet to transfer data from one electronic device to another. IPv4 which consists of 32 bits long addresses and IPv6 which consists of 128 bits long addresses which is more effective as it can handle billions of devices and can assign each device different IP address. This paper will present an overview of the main migration technologies that can be used to transition from an IPv4 network to an IPv6 network, this paper will also research on finding and comparing the effects of IPv6 transition methods such as Dual Stack, Tunneling and Network Address Translation-Protocol Translation will be compared on variant parameters to find the best performing transition method in IoT Network in terms of security.

Keywords: IoT (Internet of Things), IPv4 (Internet Protocol Version 4), IPv6 (Internet Protocol Version 6), Transition Methods, Dual Stack, Tunneling, NAT-PT.

1. Introduction:

In the modern area there are two types of internet protocol versions that are currently working in the global internet to transfer data from one electronic device to another. IP address are assigned to every device and every device has its unique address generated through binary values consists of 0 and 1. Today these two versions of Internet protocol are widely used to connect different networks to each other. IPv4 is the earlier version of IPv6. IPv4 consists of 32 bits long addresses and each unique address is assigned to each device so data can be transmitted to that specific address [1].

The new version of Internet Protocol was published in 1996 called IPv6 which consists of 128 long bits addresses. Due to large number of growths in electronic

devices, an IPv4 address was not enough to cover all the devices. To resolve the issue IPv6 introduced which can handle billions of devices and more than that and assign each device different unique address that is IP address. IPv6 found out much better and efficient in addressing of devices, routing of networks, security of information and data, translation of network address also in support of configuration of protocol. Assign a unique IP address IoT devices establish a secure communication channel, their connection should be bootstrapped through the so-called device binding process and visualize sensor data, users can easily understand the physical environment and operate the devices [2].

A variety of transition methods are available to facilitate the migration to IPv6. These methods have been observed and compared with each other and the effects of these transition methods on IPv6 in IoT security. These transition strategies are observed and compared that are Dual-Stack, Tunneling and NAT-PT. Each method or technique has its own pros and cons and each method performs its own strategy [3].

1.1. Importance of IPV6 Network in IoT Security:

The internet communications have evolved rapidly after the creation of IPv6 Internet Protocol version 6. The major difference of IPv6 is that it allows more unique addresses to create. There are five major reasons why IPv6 is more important and better option for the IoT network paradigm than IPv4 Internet Protocol version 4.

First and most important one is the security:

Security is the most important feature used to secure the communication between the IoT devices from threats, virus, attacks, etc. IPv6 uses end-to-end encryption technology which can encrypt the data so it can be secured and cannot be hacked. IPv6 also supports more secure and safe name resolution than IPv4.

Second is the scalability: Ipv6 protocol provides the connections of devices in more scalable form. It provides large area of devices so they can connect together on a large scale communicating over a long distance as well [4].

Third is the connect ability: which means connecting billions of devices to each other and allow a networking protocol so they can transfer data to each other. IPv6 allows much more addresses than IPv4 so billions of devices can connect to each other.

Fourth is Internet Protocol version 6 uses multicasting: To transmit data packets from one

destination to another means IPv6 supports multicasting of packets at one time in different destinations.

Fifth is IP Protocol version 6 providing Authentication: IP version 4 does not provide authentication whereas IP version 6 provides Authentication as well as Confidentiality, Integration, and Access control of each data packet [5].

The overall graph of the adoption of IPv6 by Internet users since late 2015 to 2025 is shown in figure 1.

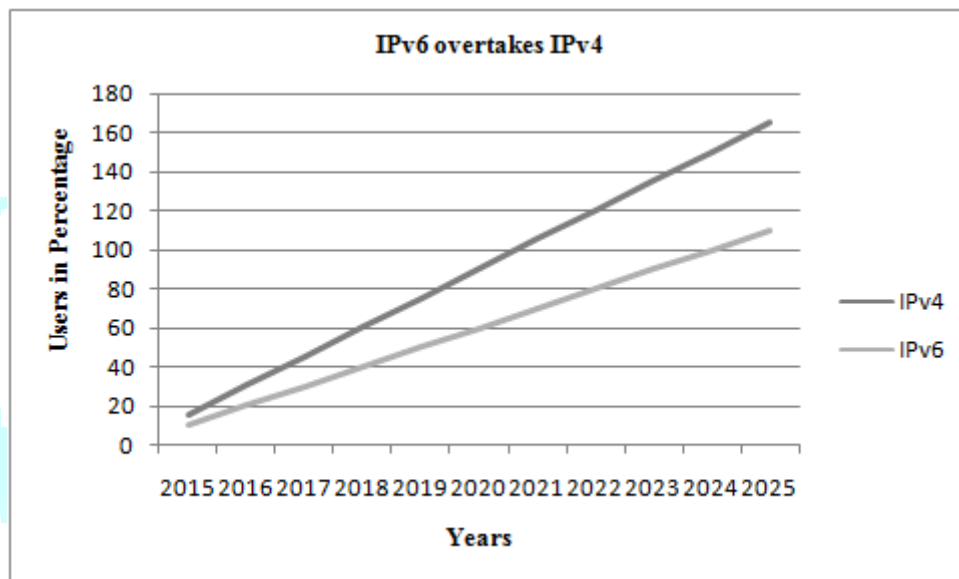


Figure 1: Growth from IPv6 to IPv4 Network

2. History:

In late 1960's there was a great need and demand of research centres and universities to develop a protocol or networking system to exchange data. To overcome this need ARPA (Advanced Research Project Agency) developed a net called ARPANET from 1972 it renamed DARPA from ARPANET [6].

In 1981 the ARPANET developed a transfer control protocol called IPv4 which was a huge success [7].

After 1990's IPv4 address space was getting full and at that time there was not enough addresses left to assign the new devices [8].

The new version of Internet Protocol was published in 1996 called IPv6 which consists of 128 long bits addresses. IPv6 found out much valuable and impressive as compared to IPv4 and it found out much better and efficient in addressing of devices, routing of networks, security of information and data, translation of network address and in support of configuration of protocol [9].

Still now these two versions are using. Both internet protocols have different configurations and are used in different environment. A census of the Internet's connected devices would readily number in the tens of billions of devices. If they all needed a globally unique permanent IP address, IPv6 would have been an imperative over a decade ago [10].

3. Transition from IPV4 to IPV6 Network in IoT Security:

There is not complete transition from IPv4 to IPv6 because IPv6 is not backward compatible. However; there are three methods, which can convert IPv4 to IPv6. The methods that can convert IPv4 to IPv6 are described as below.

3.1. Transition Methods:

One of the most important parts of implementing IPv6 is being able to gracefully transition from IPv4. The methods discussed in this paper can each be used as option when beginning an IPv6 deployment and should each be looked over for applicability depending on the specific requirements of an organization. There are three main methods that can be used when transitioning a network from IPv4 to IPv6 in IoT environment. These methods are 1) Dual-Stack, 2) Tunneling and 3) NAT-PT is explained in this section.

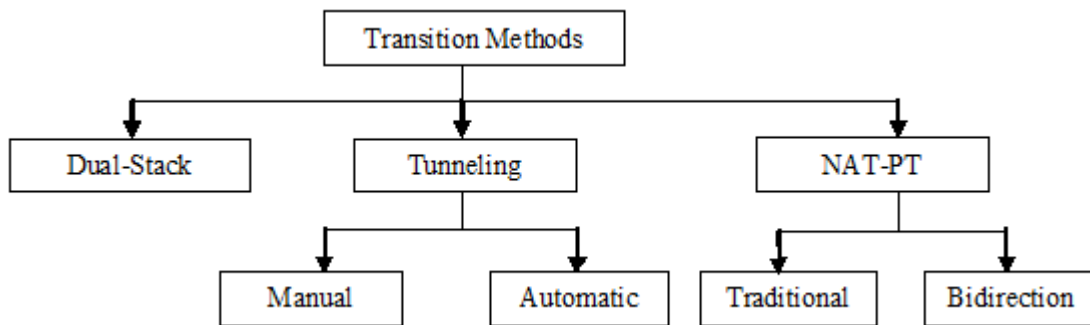


Figure 2: Types of Transition Methods

3.1.1. Dual-Stack Method

Dual Stack can process both IPv4 and IPv6 traffic simultaneously. The increase of devices day by day, it seems we are running out of IP address in IPv4 for each format which seems a big issue. IPv6 is the solution which is a new IP address format. The ISPs (Internet Service Provider) task is to provide net connections to their customers which are IPv4-to-IPv4 or IPv6-to-IPv6 but because of Dual Stack, every network is configured on both IPv4 and IPv6 and data can follow or both protocols simultaneously. Dual Stack equipped with both of the stacks, it can disable any of the stack when required either IPv4 or IPv6 and also can run both at same time [11].

Dual Stack is a simple transition method or solution that supports both internet protocols. Dual Stack devices like PC, a router or a server and other IoT (internet of things) can support both IPv4 and IPv6. This transition method is effective because IPv4 is not compatible sometimes on IPv6 devices and vice versa. Dual Stack includes both protocols working parallel which can be applied on both end system to establish connection and flow [12].

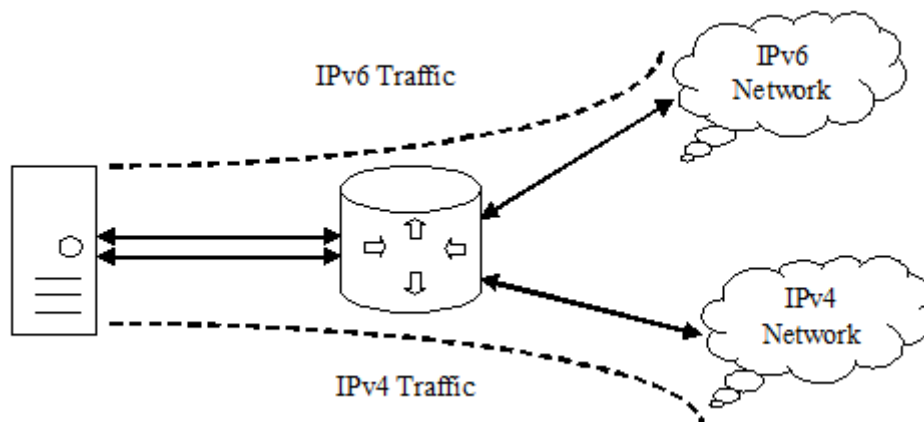


Figure 3: Dual Stack Router Connectivity

In the above figure 3, a server having IPv4 as well as IPv6 address configured for it can now speak with all the hosts on both the IPv4 as well as the IPv6 networks with the help of a Dual Stack Router. The dual Stack Router can communicate with both the networks. It provides a medium for the hosts to access a server without changing their respective IP versions.

3.1.2. Tunneling Method

For minimizing the transitions, all the routers on the way between the two IPv6 nodes do need to support IPv6. This method of transition is called Tunneling. Primarily IPv6 packets are placed inside IPv4 packets then the packets are routed through the IPv4 routers.

Tunneling is another transition method that provides a way or a tunnel to use IPv4 infrastructure to carry traffic of IPv6. This method uses routing infrastructure of one internet protocol to carry internet protocol traffic via channel also called tunnelling. Tunneling can be used as Router-to-Router or Host-to-Router or Host-Host or Router-to-Host. Most of the internet traffic is carried from one router or host to another via tunnels to migrate from IPv4 to IPv6 as the different devices uses different versions. IPv4 which is a 32-bit address can support around 4.3 billion devices where as IPv6 uses 128-bit address and support much more devices i.e. 2 times to 128 power [12].

The Tunneling method is also divided into two types of methods one is Manual Tunneling and another one is Automatic Tunneling are listed below.

Manual Tunneling: Tunnels which uses peer to peer topology and need manual configuration called manual IPv6 tunnel.

Automatic Tunneling: Tunneling uses the embedded address information of IPv4 in IPv6 packet then this type of tunnelling known as Automatic Tunneling.

In a scenario where different IP versions exist on intermediate path or transit networks, tunneling provides a better solution where user's data can pass through a non-supported IP version.

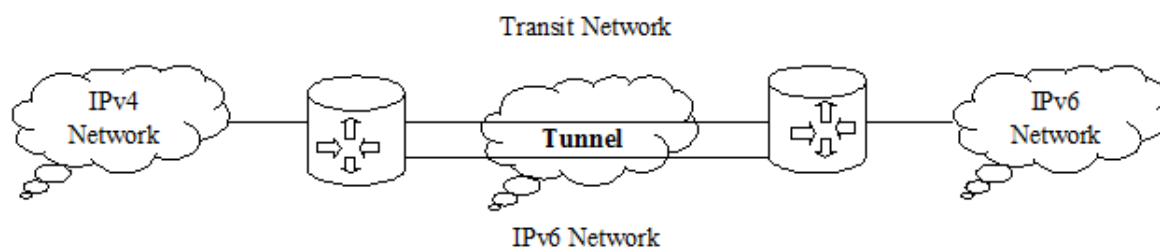


Figure 4: Tunneling Between IPv6 over IPv4 Network

The above figure 4 depicts how remote IPv4 networks can communicate via a Tunnel, where the transit network was on IPv6. Vice versa is also possible where the transit network is on IPv6 and the remote sites that intend to communicate are on IPv4.

3.1.3. Network Address Translation-Protocol Translation (NAT-PT)

This is another important method of transition to IPv6 by means of a Network Address Translation-Protocol Translation (NAT-PT) enabled device. With the help of a NAT-PT device, actual can take happens between IPv4 and IPv6 packets and vice versa.

Network Address Translation (NAT) method facilitates communication between IPv4-only and IPv6-only network by translating two different IP address families. This method translates IPv6 from IPv4 and gives consistent Internet experience to the users by accessing contents over the Internet, which have Ipv4 services. NAT-PT is similar to the NAT system utilized in IPv4 that is frequently used for converting private (RFC 1918) IPv4 address to public IPv4 address and vice versa. It is used to convert IPv4 address to IPv6 address and vice versa. This method should be used only when there are no other techniques to allow IPv6-only devices to communicate with IPv4-only devices [13].

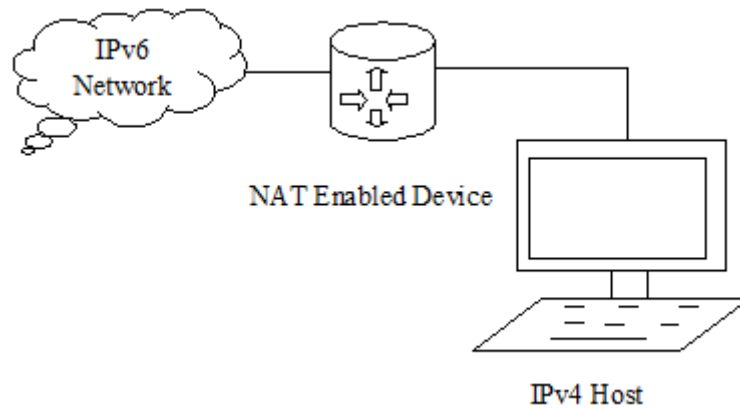


Figure 5: NAT-Protocol Translation Basic Operation

In the above figure 5, a host with IPv4 address sends a request to an IPv6 enabled server on Internet that does not understand IPv4 address. In this scenario, the NAT-PT device can help them communicate. When the IPv4 host sends a request packet to the IPv6 server, the NAT-PT device/router strips down the IPv4 packet, removes IPv4 header, and adds IPv6 header and passes it through the Internet. When a response from the IPv6 server comes for the IPv4 host, the router does vice versa.

In IPv6, there are two types of NAT-PT

Traditional: Traditional NAT-PT where sessions from IPv6 network are unidirectional. In this type, it allows hosts that are in IPv6 network to access the hosts that are in IPv4 network.

Bidirectional: In Bidirectional NAT-PT, sessions from both networks can be initiated i.e. from hosts in the IPv6 network as well as in the IPv4 network.

NAT-PT transition method main task is to migrate from IPv4 to IPv6 and also to provide connection bidirectional between IPv6 and IPv4.

4. Comparison between Transition Methods of IPv6 over IPv4:

All three transition methods are observed, compared and the effect of transition methods in IoT security. All transition methods are useful in some way and all of them has pros and cons according to the system. Comparison has been observed and studied that is presented below [14, 15].

Table1: Comparative Analysis of Three Transition Methods

Parameters/Transition Methods	Dual-Stack	Tunneling	NAT-PT
Latency	Medium	Low	High
Throughput	Moderate	Highest	Lowest
Packet Loss	High	Low	High
Traffic	High	Low	Medium
Packet Delivery	Fast	Fast	Slow
Delay	Minimum	Minimum	Maximum
Security	Higher	Lowest	Average
Connectivity	Bidirectional	Bidirectional	Bidirectional
Transition Approach	Simplest	Complex	More Complex
Flexibility	Greatest	Moderate	Lowest
Cost	Low	High	Medium
Advantages	Easy to implement. Already supported in all Operating System and Devices.	Simple Deployment. No Additional Management.	Solve Network Issues. The Router is used as a Translation Communicator.

Drawbacks	Required additional Memory and CPU Power. Two Routing Tables.	Harder to troubleshooting and Network Management Vulnerable to security attacks	Complexity increases in IP addresses. Reduction in the overall value and utility of the network.
Limitations	Two firewalls sets of policies.	Have single points of failure.	Harder to control on a larger scale
Performance Analysis	Dual-Stack transition method shows better performance in the network compared to Tunneling in terms of Latency and Delay.	Better deliverables are produced in terms of Packet Delivery and Delay.	Rapid Deployment mechanism is convenient and easy to manage. However it has low flexibility.

5. Problems and Discussion:

It is most important that the transition from IPv4 to IPv6 is stable and Non-Interruptive to exiting services. The effect of IPv6 over IPv4 transition methods in IoT security includes the problems in the following majors' areas: Address Architecture, Connectivity, 3) High Availability, 4) Applications and 5) Network Management [16, 17].

- 1. Address Architecture Problems:** IPv6 has much larger address space in comparison with IPv4. Due to the large IPv6 address space, special attention is needed when designing the IPv6 network since it differs from the fragmented and smaller IPv4 address design.
- 2. Connectivity Problems:** While shifting the transition from IPv6 to IPv4 network to provide continuity of services to the users. The Dual-Stack is the natural approach but due to the depletion of IPv4 address, cost and up-gradation of the network to IPv6-only.
- 3. High Availability:** High Availability is the major requirement for every service and network service. An application running on IPv6 may need to failover to IPv4 network due to network failure during transitioning.
- 4. Applications Problems:** During the transition process, IPv4 and IPv6 applications will co-exist in the network. Regardless of what technology providers choose to use, services should be provided to the customers. Users should find out the best for the transition without affecting the services they provide.
- 5. Network Management:** New technologies and methods may be introduced during the transition process. These new technologies and techniques require new operation models.

6. Conclusion:

In this paper, the three transition methods of the IPv4 to IPv6 transition have been discussed, deployed and compare. It has been found that these three methods have distinct advantages, drawback and features. The appropriate transition mechanism will be chosen for the network based on various parameters like the size of the network, the availability of the latest devices, the cost, and the security concern. If Latency, Throughput and Packet Loss are considered then Tunneling method is the best choice as compare to the Dual-Stack and NAT-PT. But the Tunneling method has vulnerable to security attacks, solved these security issues by IPSec (IP Security). So, our recommendation is to use Tunneling mode with IPSec for the transition purpose. The Dual-Stack remains more popular and practical with low cost in implementation and supported by wide range of devices. Transition Methods, like Tunneling and NAT-PT, are not optimally supported for the networks during a transition from IPv4 to IPv6. Thus; Dual-Stack seems the preferable method to begin adopting IPv6 with upgradable devices in order to securely manage the exiting IPv4 infrastructure. The deployment of IPv6 over IPv4 network is the best way for the growth of IoT's devices as well as also improvement in terms of their security.

7. References:

1. Farhan Anwar Ghumman, "Effects of IPv4/IPv6 Transition Methods in IoT (Internet of Things): A survey" SSRN-ID: 3402664, PP. 01-06, 19 June 2019.
2. Abubakar Isa and Idris Abdulmumine, "Design and Comparison Migration between IPv4 and IPv6 Transition Techniques", Proceedings of the

- 3rd Yusuf Maitama Sule University, Kano, PP. 179-189, November 2017.
3. Steffen Hermann and Benjamin Fabian, "A Comparison of Internet Protocol (IPv6) Security Guidelines", Future Internet 2015, ISSN: 1999-5903, PP. 01-61, 24 May 2015.
4. C. V. Ravi Kumar et al, "Performance Analysis of IPv4 to IPv6 Transition Methods", Indian Journal of Science and Technology, ISSN: 0974-6846, Vol. 9, PP. 01-08, May 2016.
5. Aparna Sivaprakash and S. Kayalvizhi, "A Survey on Optimal IPv4 to IPv6 Transition Techniques" International Journal of Research Granthaalayah, ISSN: 2350-0530, Vol. 4, Issue 4; PP. 90-96 April 2016.
6. Iman Akour, "Between Transition from IPv4 and IPv6 Adaption: The Case of Jordanian Government", International Journal of Advanced Computer Science and Applications (IJACSA), Vol. 7, No. 9, PP. 248-252, 2016.
7. Md. Asif Hossain et al, "Performance Analysis of Three Transition Mechanisms between IPv6 Network and IPv4 Network: Dual Stack, Tunneling and Translation", International Journal of Computer (IJC), ISSN: 2307-4523, Vol. 20, No. 1, PP. 217-228, 2015.
8. Daniel Enache et al, " A Study of the Technology Transition From IPv4 to IPv6 For An ISP", Review of the Air Force Academy, Vol. 1, PP. 117-121, 2016.
9. I Kullayamma et al, "Scenarios in Migration of Networks from IPv4 to IPv6", International Journal of Lates Trends in Engineering and Technology-ISSN: 2278-621XVol. 8, Issue 1, PP. 097-105, 2015.
10. Edwin S. Cordeiro, "Comparison between IPv4 to IPv6 Transition Techniques", arXiv, PP. 01-11, 1 Dec 2016.
11. Peng Wu et al, "Transition from IPv4 to IPv6 a State-of-the-Art Survey", In IEEE Communications Surveys and Tutorials, Accepted for Publication, 2015.
12. Pyung Soo Kim, "Analysis and Comparison of Tunneling based IPv6 Transition Mechanisms", International Journal of Applied Engineering Research, ISSN: 0973-4562, Vol. 12, No. 6, PP. 894-897, 2017.
13. Nellore Karthikeyan and K. Chandra Mouli, "Corporate Migration from IPv4 to IPv6 using Different Transition Mechanisms", International Journal of Engineering Science and Research Technology, ISSN: 2277-9655, Vol. 5, PP. 802-808 October 2016.
14. B. I. D. Kumar et al, "A Hybrid Transition for IPv4-IPv6 Co-existence in Small Size Organization" International Journal of Engineering and Advanced Technology(IJEAT), ISSN: 2249-8958, Vol. 9 Issue 2, PP. 3416-3421, December 2019.
15. Fuiang Li et al, "A Case Study of IPv6 Network Performance: Packet Delay, Loss, and Reordering", Hindawi Mathematical Problems in Engineering, Article ID: 3056475, Vol. 2017 PP. 01-11, 2017.
16. Dr. Jitendranath Mungara et al, "Survey on IPv4 and IPv6 Using Dual Stack, Tunneling and Translation", International Journal of Advanced Research in Computer and Communication Engineering, ISSN: 2278-1021, vol. 7, Issue 2PP. 187-189, February 2018.
17. Julianne S. et al, "IPv4 to IPv6 Transition Strategies for Enterprise Networks in Developing Countries", Institute for Computer Sciences, Social Informatics and Telecommunications Engineering, PP. 94-104, 2015

ISSN 2278-8158
AN INTERNATIONAL MULTIDISCIPLINARY
HALF YEARLY RESEARCH JOURNAL

ROYAL

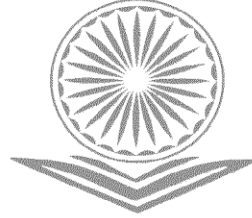
Volume - X

Issue - II

December - May - 2021-22

English / Marathi Part - I

**Peer Reviewed Refereed
and UGC Listed Journal
Journal No. 47037**



ज्ञान-विज्ञान विमुक्तये

**IMPACT FACTOR / INDEXING
2019 - 5.756
www.sjifactor.com**

❖ EDITOR ❖

Assit. Prof. Vinay Shankarrao Hatole
M.Sc (Math's), M.B.A. (Mkt), M.B.A (H.R),
M.Drama (Acting), M.Drama (Prod & Dirt), M.Ed.

❖ PUBLISHED BY ❖



Ajanta Prakashan
Aurangabad. (M.S.)

The information and views expressed and the research content published in this journal, the sole responsibility lies entirely with the author(s) and does not reflect the official opinion of the Editorial Board, Advisory Committee and the Editor in Chief of the Journal “**ROYAL**”. Owner, printer & publisher Vinay S. Hatole has printed this journal at Ajanta Computer and Printers, Jaisingpura, University Gate, Aurangabad, also Published the same at Aurangabad.

Printed by

Ajanta Computer, Near University Gate, Jaisingpura, Aurangabad. (M.S.)

Printed by

Ajanta Computer, Near University Gate, Jaisingpura, Aurangabad. (M.S.)

Cell No. : 9579260877, 9822620877 Ph. No. : (0240) 2400877

E-mail : ajanta6060@gmail.com, www.ajantaprakashan.com

ROYAL - ISSN 2278 - 8158 - Impact Factor - 5.756 (www.sjifactor.com)



EDITORIAL BOARD

**Prof. P. T. Srinivasan**

Professor and Head Dept. of Management Studies,
University of Madras, Chennai.

Dr. Rana Pratap Singh

Professor & Dean, School for Environmental Sciences,
Dr. Babasaheb Bhimrao Ambedkar University
Raebareilly Road, Lucknow.

Dr. P. A. Koli

Professor and Head (Retdf.),
Dept. of Economics, Shivaji University, Kolhapur.

Dr. Kishore Kumar C. K.

Coordinator Dept. of P. G. Studies and Research in
Physical Education and Deputy Director of
Physical Education, Mangalore University, Mangalore.

Dr. Sadique Razaque

University Dept. of Psychology,
Vinoba Bhave University,
Hazaribagh, Jharkhand.

Dr. S. Karunanidhi

Professor & Head, Dept. of Psychology,
University of Madras.

Dr. Uttam Panchal

Vice Principal, Dept. of Commerce and
Management Science,
Deogiri College, Aurangabad.

Dr. Ganesh S. Chandanshive

Asst. Prof. & Head of Dept. in Lokkala Academy,
University of Mumbai, Mumbai.

Dr. Kailas Thombre

Research Guide and Asst. Prof.
Deogiri College Aurangabad.

Dr. Rushikesh B. Kamble

H.O.D. Marathi S. B. College of Arts and
Commerce, Aurangpura, Aurangabad. (M.S.)

Dr. Shekhar Gungurwar

Hindi Dept. Vasant Rao Naik
Mahavidyalaya Vasarni, Nanded.

Dr. Jagdish R. Baheti

H.O.D. S. N. J. B. College of Pharmacy,
Meminagar, A/P. Tal Chandwad, Dist. Nashik.

Dr. Manerao Dnyaneshwar Abhimanji

Asst. Prof. Marathwada College of Education,
Dr. Rafique Zakaria Campus, Aurangabad.

Memon Sohel Md Yusuf

Dept. of Commerce, Nizwa College
of Technology, Nizwa Oman.

PUBLISHED BY



Ajanta Prakashan

Aurangabad. (M.S.)



 **CONTENTS OF ENGLISH PART - I** 

Sr. No.	Name & Author Name	Page No.
1	Biodiversity Vital to Human Welfare P. B. Sirsat	1-6
2	Women Empowerment and Entrepreneurship Development in India Prof. Smt. J. S. Kothiwale	7-13
3	Influence of Mathematics Education in Daily Lives Dr. Varsha D. Chapke	14-18
4	Theme of Women Empowerment in Chetan Bhagat's Novels Vikas R. Selokar Dr. Hitendra B. Dhote	19-22
5	The Third Gender Problems and Soutions Jaiveer Malik Manesh Kumari	23-28
6	Mahatma Gandhi and the Value of Non-Violence in Contemporary Challenges of India Dr. Vijaya Mahajan	29-32
7	The Study of Role of Indian Government Policies in Rural Development Dr. Mahendra L. Vanjari	33-39
8	Role of Biomechanics in Sports and Physical Education Dr. Ravindra Uddhavrao Machale	40-46
9	Impact of Covid-19 on Right to Life and Personal Liberty Mr. Dharmender Dr. Anamika Yadav	47-53
10	Cement Industry in Rajasthan Dr. L. R. Patel Santosh Kumar Dhakar	54-58
11	Healthy Physical Activity's Benefits for Good Health Prof. Wazarkar U. M.	59-62
12	ThThe Level of Practice of Academic Freedom in Thamar University (A Field Study on the Opinions of a Sample of Faculty Members at Thamar University) Ebrahim Mohammed AL-Mogahed Dr. Pardeep Kumar	63-70

 **CONTENTS OF ENGLISH PART - I** 

Sr. No.	Name & Author Name	Page No.
13	Effective of Digital Marketing VS Internet Marketing Mr. Hivraj Isru Raut	71-78
14	E-Commerce in India: Challenges and Opportunities Dr. M. A. Koli	79-84
15	Journey Inward: The Need of Young Adult Soul Dr. Mrs. Patil Manik Shantinath	85-90
16	A Study of Demonetization and its Impact on Common Peoples Rajhans D. Wankhade Dr. M. S. Waghmare	91-96
17	Women Political Leadership in India (A Discriptive and Anytical Study) Virendra Murlidhar Gharde	97-100
18	The Relevance of the Concept of State Socialism of Dr. Babasaheb Ambedkar in Today's Indian Society Dr. R. G. Kumawat	101-109
19	Critical Study of the Effectiveness of the Laws on Money Laundering in India Deshaboina Raghu	110-116
20	High-Intensity Interval Training and Obesity Ulhas Vijay Bramhe	117-123
21	The Theme of Human Suffering in 'The Guide' by R. K. Narayan Dr. V. P. Shekokar	124-129

CONTENTS OF MARATHI PART - I

अ.क्र.	लेख आणि लेखकाचे नाव	पृष्ठ क्र.
१	जी. डब्ल्यु. लायबनीझचे द्रव्यविषयक विचार डॉ. सुनिल काळमेघ	१-६
२	महिलांच्या आर्थिक सक्षमीकरणात भारत सरकारच्या योजनांचे योगदान डॉ. गणेश जाणोजी मस्के	७-१२
३	दैनिक लोकसत्ता वृत्तपत्रातील अग्रलेखांचा इतिहास एक अभ्यास योगेश कचरु अंबिलढगे	१३-१६
४	भारतामध्ये महिला सबलीकरण: सामर्थ्य आहे कर्तृत्वाचे प्रा. मृणाल गुलाबराव भोसले	१७-२०
५	कामगार चळवळीतील बदलते समीकरण व असंघटित कामगारांच्या समस्या प्रा. डॉ. प्रविण व्ही. चव्हाण	२१-२५
६	महिलांच्या राजकीय आरक्षणाची वाटचाल व भारतीय संविधानात्मक तरतूदी प्राचार्य डॉ. सुभाष गवई महादेव उकडा सातव	२६-३१
७	भारतीय स्त्रीयांचे मध्ययुगीन व आधुनिक काळातील योगदान प्रा. डॉ. नारायण कारभारी म्हस्के	३२-३५

21. The Theme of Human Suffering in ‘The Guide’ by R. K. Narayan

Dr. V. P. Shekokar

Associate Professor, Department of English, Vidya Bharati Mahavidyalaya, Amravati.

Abstract

The use of words like suffering, pain, misery, and distress in literature indicate human being's suffering and so can only be fully understood from the accumulation of knowledge about its causes, contexts, and results from many disciplines, including the humanities, social sciences, biological sciences, and professional health care. The word, ‘Suffering’ is often used in the literary forms; Suffering refers to psychological or social hurt. Suffering also encompasses social affliction and stress, as well as the emotional component of pain. R.K. Narayan (1906-2001) was one of the prominent Indian novelists in Indian English Literature. He received the Sahitya Akademi Award for his experimental novel the Guide in 1960. Raju, Rosie and Marco suffered in the novel ‘The Guide’.

Keywords: *Human Suffering, Pain, Humiliation, Hurt, Stress.*

To consider suffering separately, the word ‘Suffering’ emerged from Middle English word *suffrir* and the Latin word *suffero*, both of which were defined as being ‘long-suffering’ or facing a burden of pain with patience. Over several hundred years, the word lost its reflection of endurance and remained a verbal representation of hardship, distress, and turmoil. Suffering is used in so many different ways that the very word might become a barrier rather than an aid to understanding once feelings and emotions.

R.K. Narayan (1906-2001) was one of the prominent Indian novelists in Indian English Literature. He received the Sahitya Akademi Award for his experimental novel the Guide in 1960 and won the Padmabhusahn Award in 1964. The Delhi University bestowed on him the honorary D.Litt. in 1973. As a story-teller, Narayan's belief was that the novel does not fit to deal with the social ills, and this point of view explains his aloofness from the socio-political and economic problems. However, he was not an unapproachable observer of Indian society. He has presented the Indian social concerns from the universal point of view. Compared to Leo Tolstoy, Anton Chekov, Dostoevosky and Henry James, Narayan is a readable novelist due to

his simple style and plain language. Like Jane Austin, his fictional world, restricted to Malgudi and populated with the middle class life, is the world of comedy.

The Guidewas published in 1959 A central theme of the novel is the transformation of Raju from his role as a tourist guide to that of a spiritual guide. The title of the novel, *The Guide*, has a double meaning, and Raju is in a sense a double character. As a tourist guide and lover, he is a spontaneous, unprincipled, and self-indulgent. After his imprisonment, and after his transformation as a holy man, he is careful, thoughtful, and self-disciplined.

The novel also tells two stories, that of Raju's relationship with Rosie, and Raju's relationship with the villagers as a holy man. The novel begins with Raju sitting beside the temple and meeting the villager named Velan, who mistakes him for a holy man. The novel then alternates between an account of Raju's career as a holy man, which is told in the third-person, and Raju's account to Velan of his previous career as a tourist guide and lover, which is told in the first-person. This dualism reflects the dualism in Raju's character. He is transformed from a sinner to a saint, though he is never truly a sinner, and never truly a saint. Because of his capacity for empathy, Raju is a sympathetic character throughout the novel.

Raju was born and brought up in very common and poor family. It indicates the Indian poverty and employment and Indian peoples' daily sufferings for livelihood. This was also the major reason of Raju to become guide of travellers and then entangles himself in the various predicaments. This evidence specifies when Velan raises his own problem about his father's last wife's youngest daughter, Raju talks marvellously like a holy man. Raju then tells him that everyone has a problem. Velan's problem was that his step sister does not show gratitude and disinclines to accept his plans for her marriage with his cousin's son. Raju tells Velan to bring her there to talk with Raju. He says:

“I shall be rewarded for this profound service to humanity. People will say, “Here is the man who knows the exact number of stars in the sky. If you have any trouble on that account, you had better consult him. He will be your right guide for the skies”. While counting the stars in the sky, he fell asleep under the open sky.

Next morning Velan comes to Raju with his cousin sister of fourteen years old. Velan also brings a basket filled with bananas, cucumbers, pieces of sugar-cane, fried nuts, and copper vessel brimming with milk. Raju sits in silence, eyeing the gift for a while and then picks up the

basket and goes into an inside the hut. He places the basket of edibles at the feet of the image. Here Narayan indicates the Indian peoples' Hindu religious acts to worship the god, in the following passage Narayan says:

“It's His first. Let the offering go to Him, first; and we will eat the remnant.”

Then he begins to narrate the story of Devaka, a man of ancient times who begs alms at the temple gate every day and never use any of his collections without first putting them at the feet of the god. This story was told him by his mother but he couldn't remember the whole story.

Indian person believes more in the existence of god. He attracts more towards an abstract ideas and unconscious thinks which make Indian an ideological slave of holy texts of religion. Indian common man becomes the slave of this tradition and custom. Behind this many more social and economical illegal thinks are going on which Narayan indicates in the form of Raju who become holy saint for the villager. The Indian illiterate people never think about the real background of the sanit. In the following given passage Narayan rightly says:

Suddenly Raju says to Velan, “I am not going to think of your problem, Velan, not now.” Velan retorts ‘why’. Raju says, “When the time is ripe for it”. He also tells Velan that he should think over the problem and further adds, “Whatever is written here will happen.....We may not change it, but we may understand it”..... ‘And to arrive at a proper understanding, time is needed’.

There is not any human being on the earth who has not any problem and free from the problem. The problem creates suffering to the human being which takes place in the difference forms such as economical, social and psychological. The suffering makes human being weak psychological and economical. Raju and Velan suffer socially but their nature of suffering was different. In this suffering mood, human sometimes express wisdom, in the following words Narayan indicates this in Raju's talking. He says:

Velan understands and appreciates his wisdom. Raju looks at Velan's difficult sister and says, ‘What must happen; no power on earth or in heaven can change its course, just as no one can change the course of that river’.

Actually, Raju was famous as ‘Railway Raju’ among the people but the ideal visitors begin to ask him about the well-known spots around Malgudi. His friend, Gaffur takes the visitors in his car to different places. Within a few days Raju becomes a full-blown tourist

guide. Occasionally, he asks the porter's son to look after his shop and he goes with tourists in Gaffur's taxi. At home, Raju's mother asks him to accept the proposal of Lalitha, the young daughter of her brother. In a few months Raju becomes a seasoned guide. He becomes a part-time shopkeeper and a full time tourist guide. Malgudi and its surroundings were his special show. His tourists are of many kinds and types and he tries to please them of all. One day a very strange tourist named Marco come to Malgudi along with his wife named Rosie from Madras. Raju makes their lodging provision at the Anand Bhavan Hotel. Rosie was a charming and attractive beautiful lady.

While Marco was busy in investigating carving episodes from the Ramayana on the stone wall in Iswara Temple in North Extension, Raju takes Rosie in Gaffur's taxi to watch a cobra dance at Nallappa's grove on the other side of the river. Rosie persuades her body in a dance giving the snake- girl performance. Rosie appears to be the greatest dancer of the century to Raju. In the following given words Narayan rightly explains the affair between Raju and Roise. He Says:

A slight and slender one beautifully fashioned with sparkled eyes and dusky complexion. As soon as she set foot in Malgudi, she asked Raju, 'Can you show me a cobra- a King Cobra it must be- which can dance to the music of a flute'.

Symbolically cobra indicates attractiveness to fulfil sexual satisfaction; Roise's husband does not much indulge to fulfil sexual satisfaction of his wife and so that's the illegal love affair take place between Raju and Roise. It was also the major reason behind the Rosie's husband; Marco was an extraordinary hateful fellow. When Raju tells his mother about their visit to a snake charmer, she doubts about the girl as a snake dancer but Raju said to his mother that she is a good girl, not a snake-worshipper. She is a dancer'. Next day, Raju goes upstairs to Room 28 on the second floor of the hotel. Marco, the strange man wants to study the friezes and also wants to study the cave-paintings. Rosie does not want to come with Marco to see caves. But Raju goes back to Room no. 28 and persuades Rosie to come along with them. Here Narayan indicates that male does praise of female to keep love affair with lady, in the following passage it indicate:

He appreciates her dance, form and figure and introduces himself in these words, 'My name is Raju'. Then he asks her to be ready and remarks, 'Who would

decorate a rainbow?’ Yet she is not willing to join them but Raju says, ‘Because life is so blank without your presence’. She responds, ‘Wait a minute, then’.

The real suffering point between pair starts from the attractiveness to each other. Rosie, Raju and Marco go to the Peak House in Gaffur’s car. The Peak House was situated on the uppermost cliff on Mempi Hills. The river Sarayu sparks at a distance. Joseph was its caretaker. He was nearly sixty years old man. He was converted into a Christian by the missionaries. They have to stay there for a night.

After several years Men and women are busy worshipping to Raju as a saint. His disciples bring him special gifts according to seasons and festivals of the year. So he does not require a calendar. His beard embraces his chest and his hair wraps his back. He put on a necklace of prayer-beads around his neck. His eyes fill with the light of wisdom. Whatever gifts his disciples brought him; he offers those all back to the women and children at the end of each day. This pretending nature of Raju, the author Narayan says in following given passage:

He protested to Velan and said, ‘I am a poor man and you are poor men; why do you give me all this? You must stop it.’ But people called him Swami. In the first half of the year there were rains; but in the second half of the year there were no rains. The summer seemed to continue. Raju asked, ‘Where are the rains?’ The millet crop is all scorched on the stalks. A thousand banana seedlings are dead.

Conclusion

R. K. Narayan’s art of characterization in *The Guide* is very mature and perfect, in the sense that Raju, Rosie and Marco are portrayed as mature characters. Raju, Rosie, and Marco have been drawn with passion keeping the inner and outer facts of human personality in their portrayal. They breathe the modern air. The characters have been set both in the traditional and conservative mould, but in their matrix they shine with definite individuality. They surprise by their action and not by the turn of the mannerism which is seen in modern novels.

Raju’s character is as flat as it can be, but Rosie is depicted as a subtle soul. A few characters have been drawn cursorily. Psychological realism is not seen in the characterization of *The Guide*. R. K. Narayan has not drawn three dimensional characters in the novel. He has portrayed two dimensional characters in Rosie. His characters on the whole are of single

dimensional, yet they are not wooden. They are very convincing and life-like. After all, the characters in this novel are not types but individuals.

Rosie and Marco are the memorable characters with their peculiar traits. Marco and Rosie suffer from their maladjustment. Raju is materialistic. His materialism is seen in his excessive love for money even at the cost of true love for Rosie. Velan is superstitious and Gaffur and Joseph have the dryness of life. Raju's mother and maternal uncle are tradition-bound and the District Superintendent of Police is a mere mechanical force.

References

1. Abbas, Asma. (2010). *Liberalism and Human Suffering Materialist Reflections on Politics, Ethics, and Aesthetics*. PALGRAVE MACMILLAN® in the United States.
2. Iyengar, K.R. Srinivasa (2004). *Indian Writing in English*. New Delhi: Sterling Publishers.
3. Jeff Malpas (2012). *Perspectives on Human Suffering*. Springer Dordrecht Heidelberg New York London.
4. Naik, M.K. (1982). *A History of Indian English Literature*. New Delhi: SahityaAkademi.
5. Naik, M.K. (1985). *Perspectives on Indian Fiction in English*. New Delhi: abhinav publications Book Publishers.
6. Narayan, R. K.(2013). *The Guide*. Chennai: Indian Thought Publication.
7. Ronald E. Anderson (2014). *Human Suffering and Quality of Life Conceptualizing Stories and Statistics*. Springer Dordrecht Heidelberg New York London
8. Sawant, Patil, Jadhav (2010). *Literature in English Novel*. YCMOU, Developed by Shivaji University, Kolhapur.
9. Naik, M.K. ed. (1983), *the Ironic Vision: A Study of the Fiction of R.K. Narayan*. New Delhi: Sterling Publishers Pvt. Ltd.
10. Mahod, Vina. (1982), *Social Realism in R.K. Narayan's Novels*
11. Biswal, Jayant K. (1987), *A Critical Study of the Novels of R.K. Narayan*. New Delhi.
12. Srinath, C.N. (2000), *R.K. Narayan: An Anthology of Recent Criticism* edited by, Delhi Penoraft International.
13. Singh, R. A. (2002), *Critical on R.K. Narayan's Novels*. Book Enclave, Jaipur, India.



Bulk viscous LRS bianchi type-I cosmological model of the inflationary universe in general relativity

A P Wasnik¹, P P Khade²

¹ Department of Mathematics, Bhartiya Mahavidyalaya, Amravati, Maharashtra, India

² Department of Mathematics, Vidyabharati Mahavidhyalaya, Amravati, Maharashtra, India

Abstract

We discussed Bianchi type-I bulk viscous cosmological model of inflationary universe in GR. To get deterministic model of the universe, it has been considered a flat region in which potential $V(\phi) = \text{Constant}$ where ϕ is Higg's field and also considered coefficient of bulk (ξ) is inversely proportional to the expansion (θ). The physical and geometrical aspects of the model in the context of inflationary scenario are also discussed.

Keywords: LRS Bianchi type-I, inflationary scenario, bulk viscosity

Introduction

The basic idea behind inflation is that a repulsive form of gravity caused the universe to expand. General relativity from its inception predicted the possibility of repulsive gravity; in the context of general relativity you basically need a material with a negative pressure to create repulsive gravity. According to general relativity it's not just matter densities or energy densities that create gravitational fields; it's also pressures. A positive pressure creates a normal attractive gravitational field of the kind that we're accustomed to, but a negative pressure would create a repulsive kind of gravity. It also turns out that according to modern particle theories, materials with a negative pressure are easy to construct out of fields which exist according to these theories. By putting together these two ideas — the fact that particle physics gives us states with negative pressures, and that general relativity tells us that those states cause a gravitational repulsion, we reach the origin of the inflationary theory.

By answering the question of what drove the universe into expansion, the inflationary theory can also answer some questions about that expansion that would otherwise be very mysterious. There are two very important properties of our observed universe that were never really explained by the Big Bang theory; they were just part of one's assumptions about the initial conditions. One of them is the uniformity of the universe — the fact that it looks the same everywhere, no matter which way you look, as long as you average over large enough volumes. It's both isotropic, meaning the same in all directions, and homogeneous, meaning the same in all places. The conventional Big Bang theory never really had an explanation for that; it just had to be assumed from the start. The problem is that, although we know that any set of objects will approach a uniform temperature if they are allowed to sit for a long time, the early universe evolved so quickly that there was not enough time for this to happen. To explain, for example, how the universe could have smoothed itself out to achieve the uniformity of temperature that we observe today in the cosmic background radiation, one

finds that in the context of the standard Big Bang theory, it would be necessary for energy and information to be transmitted across the universe at about a hundred times the speed of light.

In the inflationary theory this problem goes away completely, because in contrast to the conventional theory it postulates a period of accelerated expansion while this repulsive gravity is taking place. That means that if we follow our universe backwards in time towards the beginning using inflationary theory, we see that it started from something much smaller than you ever could have imagined in the context of conventional cosmology without inflation. While the region that would evolve to become our universe was incredibly small, there was plenty of time for it to reach a uniform temperature, just like a cup of coffee sitting on the table cools down to room temperature. Once this uniformity is established on this tiny scale by normal thermal-equilibrium processes and I'm talking now about something that's about a billion times smaller than the size of a single proton inflation can take over, and cause this tiny region to expand rapidly, and to become large enough to encompass the entire visible universe. The inflationary theory not only allows the possibility for the universe to be uniform, but also tells us why it's uniform: It's uniform because it came from something that had time to become uniform, and was then stretched by the process of inflation.

In terms of the evolution of the universe, the fact that the universe is at least approximately flat today requires that the early universe was extraordinarily flat. The universe tends to evolve away from flatness, so even given what we knew ten or twenty years ago. We know much better now that the universe is extraordinarily close to flat, we could have extrapolated backwards and discovered that, for example, at one second after the Big Bang the mass density of the universe must have been equal, to an accuracy of 15 decimal places, to the critical density where it counterbalanced the expansion rate to produce a flat universe. The conventional Big Bang theory gave us no reason to believe that there was any

mechanism to require that, but it has to have been that way to explain why the universe looks the way it does today. The conventional Big Bang theory without inflation really only worked if you fed into it initial conditions which were highly finely tuned to make it just right to produce a universe like the one we see. Inflationary theory gets around this flatness problem because inflation changes the way the geometry of the universe evolves with time. Even though the universe always evolves away from flatness at all other periods in the history of the universe, during the inflationary period the universe is actually driven towards flatness incredibly quickly. If you had approximately 10^{-34} seconds or so of inflation at the beginning of the universe, that's all you need to be able to start out a factor of 105 or 1010 away from being flat. Inflation would then have driven the universe to be flat closely enough to explain what we see today.

At the present time this inflationary theory, which a few years ago was in significant conflict with observation now works perfectly with our measurements of the mass density and the fluctuations. The evidence for a theory that's either the one that I'm talking about or something very close to it is very, very strong.

Bulk viscous cosmology is also an alternative to gravity modifying theories (Nojiri and Odintsov) [1] in that it alters the right hand side of Einstein's field equations instead of the left hand side. In this situation based on the Eckart theorem (Eckart) [2], the consideration of the DE fluid with viscous is important. The evolution of universe involves sequence of dissipative process. In isotropic and homogeneous model, the process of dissipative is modeled as a bulk viscosity (Ren and Meng [3]; Hu and Meng [4]; Meng and Duo [5]). Brevik *et al.* [6] investigated the overall cases of viscous cosmology in early and late time universe. Norman and Brevik [7] investigated the properties of the characteristic of two different viscous cosmological models for the future universe. Norman and Brevik [8] derived observance of the bulk viscous and approximated the bulk viscosity of the cosmic fluid. The finding of viscosity dominance by late epoch of the universe with accelerated expansion was studied by Padmanabhan and Chitre [9]. Velten *et al.* [10] have investigated phantom DE as an effect of bulk viscosity. It is illustrated by Brevik and Gorbunova [11] that the fluid which lies in the quintessence region can minimize its pressure and cross the barrier, and behaves like a phantom fluid which leads to the inclusion of large bulk viscosity in a sufficient way.

The characterization by shear and bulk viscosities leads to the effect of dissipation at microscopic interactions. As expansion fluid leaves its equilibrium state, the energy density and the pressure decrease. If there is no bulk viscosity, then, the fluid relaxes instantaneously with pressure and density related by. Bulk viscosity slightly met this nature by leading a certain relaxation time scale, but producing a shift between the equation of state pressure and the absolute pressure. Bulk viscosity becomes essential only for such effects where fluid compressibility is essential. Researchers have contributed a necessity role to bulk viscous fluid matter which is different from the traditional case. The effects of both shear and bulk viscosity were explained by Hoogeveen *et al.* [12] using kinetic theory during early time. The detailed implement for the origin of bulk viscosity in the universe is not correctly understood yet.

Based on the hypothetical view, the bulk viscosity can be derived from local thermodynamic equilibrium, the manifestation as an

effective pressure to bring back the system to its thermal equilibrium, which was broken when the cosmological fluid expands. The bulk viscosity pressure thus generated ceases as soon as the fluid reaches equilibrium condition. Very recently, the concept of bulk viscosity is introduced into a DE study. It is important to develop a cosmological model. The concept of viscosity has come from fluid mechanics, and it is related to the velocity gradient of the fluid. Misner [13] revealed that during cosmic evolution when neutrinos decouple from the cosmic fluid, bulk viscosity could arise and lead to an effective mechanism of entropy production. The isotropic homogeneous spatially flat cosmological model with bulk viscous fluid was discussed by Murphy [14]. Bulk viscosity related to the grand unified-theory phase transition (Langacher) [15] may lead to explain the cosmic acceleration. The presence of bulk viscosity leads to an inflationary-like solution in Friedmann-Robertson-Walker (FRW) space time obtained by Padmanabhan and Chitre [16]. Guth [17] introduced the idea of early inflationary phase the context of grand unified field theories. Panchapakshan and Sethi [18] have discussed inflationary scenario and large scale structure of the universe. Bali and Jain [19] discussed inflationary scenario in Bianchi type- I inflationary universe in general relativity. Bali and Singh [20] have investigated Bianchi Type-V viscous fluid string dust cosmological model assuming the condition that the bulk coefficient (ξ) is inversely proportional the expansion (θ) in the model.

In the present work, we have studied Bulk viscous LRS Bianchi Type-I cosmological model of the inflationary universe. The geometrical and physical aspects of the model have been studied. We consider string cosmology far the spherically symmetric is homogeneous anisotropic space-time with metric ansatz.

$$d^2s = -dt^2 + a^2 dx^2 + b^2 (dy^2 + dz^2) \tag{1}$$

Where $a=a(t)$, $b=b(t)$

The Lagrangian will be that of gravity minimally coupled to a scalar field $V(\phi)$ (Stein-Schalers, [21])

$$s = \int -g \left[R - \frac{1}{2} g^{\mu\nu} \partial_\mu \phi \partial_\nu \phi - V(\phi) \right] d^4x \tag{2}$$

[Notations have their usual meaning and units are taken so that $8\pi G = 1 = c$].

Now the variation of S with respect to the dynamical field, leads to the Einstein field equation in presences of bulk viscosity

$$R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} - \xi \Theta u_\mu u_\nu + g_{\mu\nu} \xi = -T_{\mu\nu} \tag{3}$$

Where u^μ the four velocity vector, $\theta = u^\mu_{;\nu}$, is the scalar of expansion and ξ the coefficient of bulk viscosity. Here

$$T_{\mu\nu} = \partial_\mu \phi \partial_\nu \phi - \left[\frac{1}{2} \partial_\mu \phi \partial^\mu \phi + V(\phi) \right] g_{\mu\nu} \quad ; \quad u_\mu u^\mu = -1 \tag{4}$$

And

$$\frac{1}{\sqrt{-g}} \partial_{\mu} [\sqrt{-g} \partial^{\mu} \phi] = -\frac{dV(\phi)}{d\phi} \tag{5}$$

The Einstein field equation (3) from the metric (1) are given by,

$$\frac{\dot{b}}{b^2} + 2\frac{\ddot{b}}{b} = -\frac{1}{2}\dot{\phi}^2 + V(\phi) + \xi \tag{6}$$

$$\frac{\ddot{a}}{a} + \frac{\dot{b}}{b} + \frac{\dot{a}\dot{b}}{ab} = -\frac{1}{2}\dot{\phi}^2 + V(\phi) + \xi \tag{7}$$

$$2\frac{\dot{a}\dot{b}}{ab} + \frac{\ddot{b}}{b} = \frac{1}{2}\dot{\phi}^2 + V(\phi) \tag{8}$$

And the equation for the scalar field ϕ leads to

$$\ddot{\phi} + \left(\frac{\dot{a}}{a} + 2\frac{\dot{b}}{b}\right)\dot{\phi} + \frac{dV}{d\phi} = 0 \tag{9}$$

Solution of the field equations:

We are interested in inflationary solutions, the flat region is considered where Potential is constant. Also to get determinant model we assume that coefficient of bulk viscosity (ξ) is proportional to expansion θ (Kandalkar *et al.*)^[22], thus we have

$$V(\phi) = \text{constant} = c_0 (sa)^{\alpha} \text{ and } \xi\theta = \text{constant} = c_1 (sa)^{\beta} \tag{10}$$

Thus Equation (9) leads to

$$\dot{\phi} = \frac{\alpha}{a^2} \tag{11}$$

Where α is the constant of integration.

From eq. (9) and (7), we have

$$\frac{\ddot{a}}{a} - \frac{\ddot{b}}{b} + \frac{\dot{a}\dot{b}}{ab} - \frac{\dot{b}^2}{b^2} = 0 \tag{12}$$

Which leads to

$$\left(\frac{\dot{a}}{a} - \frac{\dot{b}}{b}\right) + \left(\frac{\dot{a}}{a} - \frac{\dot{b}}{b}\right)\left(\frac{\dot{a}}{a} - \frac{2\dot{b}}{b}\right) = 0 \tag{13}$$

Which on integration leads to

$$\frac{\dot{a}}{a} - \frac{\dot{b}}{b} = \frac{k}{a^2} \tag{14}$$

Now equations ^{(7)+ $\frac{1}{2}$ (6)- $\frac{1}{2}$ (8)} together with eq. (11) leads to

$$\frac{\ddot{a}}{a} + 2\frac{\ddot{b}}{b} = V_0 - \frac{\alpha^2}{a^2 b^4} \text{ where } \alpha_0 + c_1 = V_0 \tag{15}$$

$$\left(\frac{\dot{a}}{a} + \frac{\dot{b}}{b}\right) + \frac{\dot{a}}{a^2} + \frac{2\dot{b}}{b^2} = V_0 - \frac{\alpha^2}{a^2 b^4} \tag{16}$$

Using $ab^2 = \eta$ in eq. (16), we get

$$\left(\frac{\dot{\eta}}{\eta}\right) + \left(\frac{\dot{a}^2}{a^2}\right) + 2\frac{\dot{b}^2}{b^2} = V_0 - \frac{\alpha^2}{\eta^2} \tag{17}$$

Now, $ab^2 = \eta$ leads to

$$\frac{\dot{\eta}}{\eta} = \frac{\dot{a}}{a} + 2\frac{\dot{b}}{b} \tag{18}$$

From eq.(14), we have

$$\frac{\dot{a}}{a} - \frac{\dot{b}}{b} = \frac{k}{\eta} \tag{19}$$

$$3\frac{\dot{a}}{a} = \frac{\dot{\eta}}{\eta} + \frac{2k}{\eta} \tag{20}$$

And

$$3\frac{\dot{b}}{b} = \frac{\dot{\eta}}{\eta} - \frac{k}{\eta} \tag{21}$$

Using eq.(20) and eq.(21) we have

$$\left(\frac{\dot{\eta}}{\eta}\right) + \frac{1}{9}\left(\frac{\dot{\eta} + 2k}{\eta}\right)^2 + \frac{2}{9}\left(\frac{\dot{\eta} - k}{\eta}\right)^2 = V_0 - \frac{\alpha^2}{\eta^2} \tag{22}$$

Which leads to

$$\eta\dot{\eta} - \frac{2}{3}\dot{\eta}^2 - V_0\eta^2 + \frac{2}{3}k^2 + \alpha^2 = 0 \tag{23}$$

Using suitable the transformation

$$\dot{\eta} = f(\eta) \tag{24}$$

So that

$$\ddot{\eta} = f \frac{d}{d\eta} \tag{25}$$

eq. (23) Now becomes

$$\frac{d^2 f}{d\eta^2} - \frac{4}{3\eta} f^2 = 2V_0\eta - \frac{4/3 k^2 + 2k^2}{\eta} \tag{26}$$

Which leads to

$$f^2 \eta^{-4/3} = 3V_0\eta^{2/3} + \left(k^2 + \frac{3}{2}\alpha^2\right)\eta^{-3/4} + m \tag{27}$$

We had four eq. (6)-(9) in five unknown a, b, ϕ, ξ and V , the condition $V = const. = c_0$ and $\xi\theta = const = c_1$, gives us four equations in three unknown which is over determined set. It is easy to verify that these equations are consistent and have a solution when $m=0$. Thus eq.(27) lead to

$$\dot{\eta} = \sqrt{3V_0\eta^{2/3} + \left(k^2 + \frac{3}{2}\alpha^2\right)} \tag{28}$$

Which leads to

$$\frac{d\eta}{\sqrt{3V_0\eta^2 + \left(k + \frac{3}{2}\alpha^2\right)}} = d \tag{29}$$

And this leads to

$$\eta = m s \quad \hbar \sqrt{3V_0} (t + \beta) \tag{30}$$

Where,

$$\frac{k^2 + \frac{3}{2}\alpha^2}{3V_0} = m^2 \tag{31}$$

And β is a constant.

Thus from (20) and (21), we have

$$3 \frac{\dot{a}}{a} = \frac{\dot{\eta} + 2k}{\eta} = \frac{m\sqrt{3V_0} c \text{ ohs}\sqrt{3V_0} (t+B) + 2k}{m s \text{ i h}\sqrt{3V_0} (t+B)} \tag{32}$$

And

$$3 \frac{\dot{b}}{b} = \frac{\dot{\eta} - k}{\eta} = \frac{m\sqrt{3V_0} c \text{ ohs}\sqrt{3V_0} (t+B) - k}{m.s \text{ i h}\sqrt{3V_0} (t+B)} \tag{33}$$

Equation (32) and (33), on integration lead to

$$a = \gamma^{1/3} \text{ s i n h}^{1/3} \sqrt{3V_0} (t+B) \left[\text{t a n h} \left\{ \frac{1}{2} \sqrt{3V_0} (t+B) \right\} \right]^{-\frac{2k}{3m\sqrt{3V_0}}} \tag{34}$$

$$b = \delta^{1/3} \text{ s i n h}^{1/3} \sqrt{3V_0} (t+B) \left[\text{t a n h} \left\{ \frac{1}{2} \sqrt{3V_0} (t+B) \right\} \right]^{-\frac{k}{3m\sqrt{3V_0}}} \tag{35}$$

Thus,

$$\dot{\phi} = \frac{\alpha}{a^2 b} = \frac{\alpha}{\eta} = \frac{\alpha}{m s \text{ i h}\sqrt{3V_0} (t+B)} \tag{36}$$

Using suitable transformation

$$t + \beta = T, \quad \gamma^{1/3} x = X, \quad \delta^{1/3} (y + z) = Y + Z$$

The metric (1) leads to

$$ds^2 = -dT^2 + \sinh^{2/3}(\sqrt{3V_0} T) \tanh^{2/3} \sqrt{3V_0} \left(\frac{1}{2}\sqrt{3V_0} T\right) dX^2 + \sinh^{2/3}(\sqrt{3V_0} T) \tanh^{-k/3} \sqrt{3V_0} \left(\frac{1}{2}\sqrt{3V_0} T\right) (dY^2 + dZ^2) \tag{37}$$

The model (37) represents LRS Bianchi type I bulk viscous inflationary universe in general relativity.

Some Physical and Geometrical Aspects

The scalar of expansion (θ) and the shear (σ) for the model (37) are given by

$$\theta = \frac{\dot{a}}{a} + \frac{2\dot{b}}{b} = \frac{\dot{\eta}}{\eta} = c \text{ o}(\sqrt{3V_0} T) \tag{38}$$

$$\sigma = \frac{\sqrt{2}}{\sqrt{3}} \left(\frac{\dot{a}}{a} - \frac{\dot{b}}{b} \right) = \frac{\sqrt{2}}{\sqrt{3}} \frac{k}{m s \text{ i h}(\sqrt{3V_0} T)} \tag{39}$$

The rate of expansion H; (Hubble parameters) in the direction of X, Y, Z are given by

$$H_1 = \frac{m\sqrt{3V_0} c \text{ oh}(\sqrt{3V_0} T) + 2k}{3m s \text{ i h}(\sqrt{3V_0} T)} \tag{40}$$

And

$$H_2 = H_3 = \frac{m\sqrt{3V_0} c \text{ oh}(\sqrt{3V_0} T) - k}{3m s \text{ i h}(\sqrt{3V_0} T)} \tag{41}$$

The Higg's field (ϕ) is given by Eqⁿ (36)

$$\dot{\phi} = \frac{\alpha}{m s \text{ h}(\sqrt{3V_0} T)} \tag{42}$$

The spatial volume V is given by

$$V = s \quad \hbar \sqrt{3V_0} \text{ i n} T \tag{43}$$

The coefficient of bulk viscosity is $\xi \propto \frac{1}{\theta}$

$$\xi = c_0 t^a \left(\sqrt{3V_0 T} \right) \quad (44)$$

Conclusion

The model (37) has no initial singularity. From equation (31), we

have, $\frac{\alpha}{m} = \sqrt{2V_0 - \left(\frac{2k^2}{3m^2} \right)}$, α is constant of integration. Clearly

$V_0 > \frac{k^2}{3m^2}$. Thus the quantity $\frac{k}{m}$ has physical relevance. The

quantity $\frac{\alpha}{m}$ appears in Higg's field. The model vanishes for large T and it diverges when $T=0$.

The coefficient of bulk viscosity ξ is vanishes for $T=0$ and tends to constant as $T \rightarrow \infty$.

Hence the universe remains anisotropy throughout the evolution of the universe.

The expansion $\theta \rightarrow \infty$, when $T \rightarrow 0$ and $\theta \rightarrow \text{finite quantity}$ when $T \rightarrow \infty$. The $\frac{k}{m}$ measures anisotropy in the model. The shear value of $\sigma \rightarrow 0$ when $T \rightarrow \infty$ and $\sigma \rightarrow \infty$ when $T \rightarrow 0$

The spatial volume increases with time. When $T \rightarrow \infty$ then spatial volume $V \rightarrow \infty$. Hence inflation is exist in Bulk viscous LRS Bianchi type I model with a mass less scalar field in the potential which has flat space region.

References

- Nojiri S, Odintsov SD. "Unified cosmic history in modified gravity: from theory to Lorentz non-invariant models," Physics Reports,2011:505(2-4):59-144.
- Eckart C. "The thermodynamics of irreversible processes III. Relativistic theory of the simple fluid," Physical Review,1940:58(10):919-924.
- Ren J, Meng XH. "Cosmological model with viscosity media (dark fluid) described by an effective equation of state," Physics Letters B,2006:633(1):1-8.
- Hu MG, Meng XH. "Bulk viscous cosmology: state finder and entropy," Physics Letters B,2006:635(4):186-194.
- Xin-He M, Xu D. "Friedmann cosmology with bulk viscosity: a concrete model for dark energy," Communications in Theoretical Physics,2009:52(2):377-382.
- Brevik I, Grøn O, Haro J de, Odintsov SD, Saridakis EN. "Viscous cosmology for early- and late-time universe," International Journal of Modern Physics D,2017:26(14):1730024.
- Normann BD, Brevik I. "Characteristic properties of two different viscous cosmology models for the future universe," Modern Physics Letters A,2017:32(4):1750026.
- Normann BD, Brevik I. "General bulk-viscous solutions and estimates of bulk viscosity in the cosmic fluid," Entropy,2016:18(6):215.
- Padmanabhan T, Chitre SM. "Viscous universes," Physics Letters A,1987:120(9):433-436.
- Velten H, Wang J, Meng X. "Phantom dark energy as an effect of bulk viscosity," Physical Review D, 2013:88(12):123504.
- Brevik I, Gorbunova O. "Dark energy and viscous cosmology," General Relativity and Gravitation, 2005:37(12):2039-2045.
- Hoogeveen F, Van Leeuwen WA, Salvati GAQ, Schilling EE. "Viscous phenomena in cosmology," Physica A: Statistical Mechanics and its Applications,1986:134(2):458-473.
- Misner CW. "The isotropy of the universe," Astrophysical Journal,1968:151:431-458.
- Murphy GL. "Big-bang model without singularities," Physical Review,1973:8(12):4231-4233.
- Langacher P. "Grand unified theories and proton decay," Physics Reports,1981:72(4):185-385.
- Padmanabhan T, Chitre SM. "Viscous universes," Physics Letters A,1987:120(9):433-436.
- Guth AH. "Inflationary universe: A possible solution to the horizon and flatness problem", Phys. Rev,1981:D23:347.
- Panchapakshan N, Sethi SK. "Inflationary scenario and large scale structure of the universe", Int. J. Mod. Phys,1992:A7:3769.
- Bali R, Jain VC. "Bianchi type-I inflationary universe in GR", Pramana,2002:59(1):1-7.
- Bali, Singh." Bianchi Type-V viscous fluid string dust cosmological model in GR", Astrophys. Space Sci,2005:300:387.
- Stein-Schables JA." Inflationary in spherically symmetric inhomogeneous models", Phys. Rev,1987:D35:2345.
- Kandalkar SP, Wasnik AP, Gawande SP." Bianchi type-V string dust cosmological models with bulk viscous magnetic field, IJSER,2013:4:3.

INTERNATIONAL RESEARCH FELLOWS ASSOCIATION'S
RESEARCH JOURNEY

International E-Research Journal

PEER REFREED & INDEXED JOURNAL

March 2021

Special Issue 261(A)

**Use of ICT in Teaching - Learning
 Opportunities and Challenges**



Guest Editor of the Issue :-

Dr. P. D. Hudekar
 Act. Principal
 Vidarbha Mahavidyalaya, Buldana
 Dist - Buldana [M.S.] INDIA

Executive Editor of the Issue :-

Prof. Sangita K. Pawar
 Vidarbha Mahavidyalaya, Buldana

Chief Editor : Dr. Dhanraj T. Dhangar

Co Editors of the Issue :-

Prof. Vaibhao G. Vaghmare
 Prof. M. T. Jamalwar
 Prof. S. H. Dandade
 Prof. N. D. Raut
 Prof. L. F. Shirale
 Prof. S. V. Kalne
 Shri. S. G. Pande



This Journal is indexed in :

- Scientific Journal Impact Factor (SJIF)
- Cosmos Impact Factor (CIF)
- Global Impact Factor (GIF)
- International Impact Factor Services (IIFS)



अनुक्रमणिका

अ.क्र.	शीर्षक	लेखक/लेखिका	पृष्ठ क्र.
1	Information Communication and Technology (ICT) Integration in Teacher Education: A Necessity for Sustainable Educational Growth	Dr.G.V. Sreenivasamurthy	13
2	ICT in Sports and Physical Education	Dr. Sangita Khadse	20
3	Need of Innovations in ICT Teaching Learning Process	Dr. Hari Kale	22
4	E-Resources, Concepts and Use : An Overview	Dr. Harshal Nimbhorkar	26
5	Role of ICT in Home Economics : Enhancement in Teaching and Learning	Ms. Pratibha S. Katkar	30
6	Emphasis on ICT / Digital Technology in National Education Policy 2020	Dr. Durga Pande	34
7	Application of ICT in Academic Library : An Overview	Dr. Pranali Pete	40
8	Use of ICT in World of Sports : Development in Physical Education	Dr. Balasaheb Paul	45
9	Need of Library Automation in Current Era: A Study	Prof. Swapnil Dandade	50
10	Use of Internet in Language Acquisition	Pundlik Nalinde	53
11	Role of Information and Communication Technologies (ICT) in Physical Education and Sports	Prof. Sunil B. Chordiya	57
12	Use of RFID Technology In Libraries	Dr. Prashant Pagade	61
13	Use of E-Learning in Teacher Education Institutions	Mrs. Sharmila Kerkar	64
14	Usage of ICT Based Library Services.	Dr. Dattatray Dhumale	70
15	ICT Enabled English Language Teaching-Learning in Rural India	Dr. Archana Deshmukh	73
16	A Study of Various Parameters of ICT and Teaching Aids used in the Higher Education in Gadchiroli District	Ganesh Dandekar	80
17	Impact of ICT in Higher Education in India	Dr. Pawan Naik	85
18	Role of ICT in Quality Enhancement in Higher Education in India	Dr. Archana Patki (Kahale)	90
19	IT and Teaching	Dr. Mahesh Joshi	95
20	Role and Significance of ICT in Rural Higher Education	Dr. Dadarao Upase	97
21	Impact and Suitability of Use of Information Technology with Reference to Physical Education in the Context of Pandemic of Covid-19	Dr. Pratap Chauhan	100
22	Suitability of ICT in Teaching Learning Process	Dr. Bhavesh Bhuptani	104
23	Innovations in Physical Education and Sports: Technological Tools Developing Sports	Dr. Devendra Gawande	107
24	Information and Communication Technology Uses in Teaching - Learning: with Respect to Opportunities and Challenges	Dr. Umesh Rathi	111
25	Technology in Physical Education and Sports	Sanjay Kale	116
26	Use of IT in Education	Dr. D. S. Wankhade	120
27	Social Media as Modern Tool	Dr. R. G. Suralkar	123
28	Expressionistic Technique in Tennessee Williams' 'The Glass Menagerie'	B. W. Somatkar	126
29	Use of ICT in Developing Academic Libraries	Dr. Sanjay Shenmare	129



Use of IT in Education

Dr. D. S. Wankhade

Director

Dept. of Physical Education.

Vidya Bharati Mahavidyaalaya, Amravati

Introduction:

Information Technology (IT) may be defined as the use of all conceivable digital media in managing and processing information. It provides any time any where access the reliable knowledge. Earlier traditional education system is far away from IT and based on class room teaching and learning methods, which has its own merits and demerits. Recently with the innovation in the field of computational sciences it is need of time to revolutionaries the education system with addition of IT resources which will be beneficial to both teachers and students. With the passage of time at every minute huge data/information is added to every field and to cope with all the information one has to incorporate IT tools to modernize education system. Worldwide communication provides instant access to vast array of data which is found to be improving the performance of students as it will help them to get exact information and right time and right place.

To keep the pace with ever advancing education field, IT skills and IT understanding are equally important for the students/teachers and researchers also.

Requisite of IT in Education:

- I. To lower down cost of education to face the challenges of illiteracy and poverty. IT can be master stroke.
- II. Every day huge amount of scientific and academic information is added to earlier information
- III. With advancement in Education system which is dynamic as ever
- IV. Traditional system can be reach to limited people in masses.

Implications of IT in Education:

- Online libraries made literature searching easy and time saving option.
- Recent information is available at tip of finger.
- Valuable educational information can be stored in soft copy.
- Access to variety of learning resources instantly, anytime and anywhere
- Use of multiple communication channels like e-mail, chat, forum, blogs, and etc. improving social communication skills.

Significance of IT in Education

Efficient use of learning resources.

In this advance technology world IT aids to utilize variety of learning resources. With advancement in education system now a day teachers are routinely using audio visual aids to impart more details of their subject to students. Researchers can able to present their data/information more efficiently as compare to earlier. Now a day's students from all disciplines are encouraged to undertake basic IT education in addition to regular studies which help them to communicate ideas, describe projects, and other information in their work.



Without Time barrier learning:

With the help of web networks exchanges of information is increased to a great extent. One can access the information at any time and there is no barriers like day, night, etc.

Role of Multimedia to improve education:

Audio-Visual Education, Planning, preparation, and use of devices and materials that involve sights, sound, or both for educational purposes. They includes photographs, movies, television, transparencies, audiotapes, records, teaching, machines, computers, and videodisc. Among the devices used are still and motion pictures, filmstrips etc. studies in the psychology of learning suggest that the use of audio-visual in education has several advantages. All leaning is based on perception, the process of memory and concept formation cannot occur without prior perception. People can attend to only a limited amount of information at a time. Their selection and perception of information is influenced by past experience.

Distance learning:

Distance learning, method of learning at a distance rather than in a class-room. Late 20th century communication technologies, in their most recent phases multimedia and interactive, open up new possibilities. Both individual and institutional, for an unprecedented expansion of home-based learning, much of its part-time. The term distance learning was coined within the context of a continuing communication revolution, largely replacing a hitherto confusing mixed nomenclature-home study, independent study, external study, and most common though restricted in pedagogic means, correspondence study. Whatever the reasoning, distance learning widens access for students unable for whatever reason to study alongside others.

Change in approach of learning:

Advancement in IT had brought about people for various part of world together. With this the collaborative and co-operative approach had improved. Peoples from diverse part of world are communication with each other using IT tools and improving the academics and research quality by collaborating each other.

Online library/Information resources:

Earlier access to the academic literature is very tedious and time consuming job earlier. With advancement and providing online resources it information it is now turn in t baby game. Now a day's maximum libraries are turn to online libraries with free access to journals and books. Online Data bases are available which made the information handy.

Immediacy to information:

IT has added immediacy to education. One can access the information from any site of the world. The transmission and transport of information is so fast that made possible to share, distribute information fastly.

Availability of Recent information:

Many things are happening in and around our world. It is not possible to get acquainted with all the news and information. It tools provide and additional to provide up to date and recent information.

Limitations:

1. Cost:

Purchasing and use of IT resources is costly affair and many times because of it poor people are remaining away from its use.



2. Authenticity:

The information and the available data are not always authentic, before using one has to try to check its authenticity from other resources.

Reference:

1. Koehler M (2005), what happens when teachers design Educational Technology.
2. Educational Computing Research, Vol 32 (2)
3. Kumar K.L. (1996), Educational Technology, New Delhi, New age international Pvt. Ltd.
4. Steketee C. (2005), Integrating ICT as an Integral teaching and learning tool into preservice teacher training course.
5. Vermal L (2004), ICT in Teacher Education: A case study, University News, Punjab University Chandigarh, vol 42 (39).



(SJIF) Impact Factor-7.675

ISSN-2278-9308

B.Aadhar

Peer-Reviewed & Refreed Indexed

Multidisciplinary International Research Journal

March -2021

ISSUE No- CCLXXVIII (278) C

**'CONTEMPORARY APPROACHES AND APPLICATIONS IN
PHYSICAL EDUCATION & SPORTS SCIENCES'**



Chief Editor

Prof. Virag S. Gawande
Director
A. S.R.& D. T.
Institute Amravati

Editors

Dr. Vijay Dhote and Dr. A P Joshi

The Journal is indexed in:

Scientific Journal Impact Factor (SJIF)

Cosmos Impact Factor (CIF)

International Impact Factor Services (IIFS)

Aadhar International Publication

For Details Visit To : www.aadharsocial.com

**INDEX**

No.	Title of the Paper	Authors' Name	Page No.
1	Fitness and Immunity Important for Corona Virus	Prof.D.P. Girhepuje	1
2	Spinal Injuries In Athletes And There Rehabilitation	Dr.Digambar Singh	3
3	Stress Management by Yoga During The Covid-19 Outbreak	Dr.Sharda Lokesh Naidu	7
4	Yoga For Health In Modern Era	Dr.Archana Madhukar Falke	10
5	Sports Psychology	Dr. B. A. Sarpate/ Dr. Bavikar Samir	13
6	Role And Importance Sports Psychologist	Dr. Babulal S. Dhotre /Dr. Aditya Kishor Sarwe	17
7	Role of Nutrition in Sports Performance and Fitness	Dr. Dnyaneshwari S. Wankhade	20
8	Study Of The Factors Necessary For The Physical And Mental Fitness Of Human Being	Asst Prof. Dr. Gunaji Pandurang Nalge	24
9	Leadership in the Organization, Sports Management	Dr.Indrajit Basu	27
10	Psychological And Sociological Preparation Of Sports Personnel For Performance	Dr. Jitendra Kumar Thakur	34
11	Significant Role Of Yogic Practices And Physical Activity In Present Era	Asst. Prof. Dr. Jaikumar G. kshirsagar	40
12	Importance Of Yoga In Modern Life	Dr. Madan Ingle	45
13	Role of Nutrition in sports	Dr. Manojkumar Varma	47
14	Significant Impact Of Covid-19 On The Global Sports	Dr. Meenakshi Pahuja	50
15	Nutrition and Sport Performance	Prof. Dr. Naresh P. Borkar	54
16	Empowerment of Women Through sports: A Social Aspect.	Ms. Nilima Waghmare Sonkamble	63
17	Effects of Dance Aerobics	Dr.Prashant Bambal	69
18	Effect of static stretch training on Kabaddi women players in Maharashtra.	Dr. Sheetal Laxman Shendkar	76
19	Sociological Aspects Of Sports: Society And Sports	Dr. Subhash S. Dadhe	78
20	The Physical Fitness and its Mental Benefits	Dr. Bhairat Suhas Nivrutti	83

**Role of Nutrition in Sports Performance and Fitness****Dr. Dnyaneshwari S. Wankhade**

Associate Professor

Vidyabharti Mahavidyalaya, Amravati.

Abstract:

This article highlights the Role of Nutrition in Sports performance and Fitness. Games Nutrition applies nourishment standards to brandish with the goal of expanding exhibition. Diets have been endorsed for diverse classifications of games relaying on the form weight and vigor needs. This article outlines the current energy, nutrient, and fluid recommendations for sports person. The performance of athletes and sports person can be enhanced by well-chosen nutrition strategies. It outlines the stance on nutrition factors that have been determined to influence sporting performance and emerging trends in the field of sports nutrition.

Introduction :

Games nourishment gathers basic vitality in light of the fact that long before inadequacy manifestations begin seeming, physical exhibition decays. It might not be reasonable to think regarding least ought to keep the blood levels or chemical levels at ordinary points of confinement. Rather endeavours ought to be made to figure out the level beneath which physical exhibition begins appearing. The level, which allows the jok to attain the greatest conceivable physical exhibition, ought to be the base level pointed in the games nourishment.

An optimal eating methodology may be described as one in which the supply of needed supplements is satisfactory to blanket vigor use, for tissue support, repair and development. The wholesome needs vary from single to single dependent upon age, sex, form measure and arrangement, occupation, physiological condition and so forth. Numerous mentors make dietary suggestions dependent upon their particular "sentiments" and past encounters as opposed to depend on accessible experimental confirmation the way that players frequently have either lacking or erroneous informative content concerning judicious dietary rehearses and the part of particular supplements in the eating regimen.

Meaning of Energy Requirement :

It is a created certainty that sustenance plays an imperative part in physical exhibition, separated from preparing and other identified segments. In India, games sustenance is yet to be distinguished as a paramount part of preparing project, and seems to be expansively dismissed.

Sports Authority concerning India (SAI) stresses every now and then on games nourishment and readiness of menus for the diverse games disciplines. As Nutrition goals and requirements are not static. Athletes undertake a periodized program in which preparation for peak performance in targeted events is achieved by integrating different types of workouts in the various cycles of the training calendar. Nutrition plans need to be personalized to the individual athlete to take into account specificity & uniqueness of the event, performance goals, practical challenges, food preferences, responses to various strategies. Energy



availability, which considers energy intake in relation to the energy cost of exercise, sets an important foundation for health and the success of sports nutrition strategies. The achievement of the body composition associated with optimal strategies. The achievement of the body composition associated with optimal performance is now recognized as an important but challenging goal that needs to be individualized and periodized. Training and Sports nutrition guidelines should also consider the importance of the timing of nutrient intake and nutritional support over the day and in relation to sport rather than general daily targets. Competition nutrition should target specific strategies that reduce or delay factors that would otherwise cause fatigue in an event; these are specific to the event, environment / scenario in which it is undertaken and the individual athlete. A pragmatic approach to advice regarding the use of supplements and sports foods is needed for high prevalence of interest and use by athletes and the evidence that some products can usefully contribute to a sports nutrition plan and / or directly enhance performance.

General Dietary Guidelines :

A well-designed diet that meets energy intake needs and incorporates proper timing of nutrients is the foundation upon which a good training program can be developed. Maximum research has clearly shown that not ingesting a sufficient amount of calories and/or enough of the right type of macronutrients may impede an athlete's training adaptations while athletes who consume a balanced diet that meets energy needs can augment physiological training adaptations. Incorporating good dietary practices as part of a training program is one way to help optimize training adaptations and prevent overtraining.

- **Energy Intake**

To optimize training and performance through nutrition is to ensure the athlete is consuming enough calories to offset energy expenditure. For elite athletes, energy expenditure during heavy training or competition may be enormous. Maintaining energy deficient diet during training often leads to significant weight loss, illness, onset of physical and psychological symptoms of over training, and reductions in performance. Nutritional analyses of athletes' diets have revealed that many are susceptible to maintaining negative energy intakes during training. Female athletes have been reported to have a high incidence of eating disorders. Some athletes do not like to exercise within several hours after eating because of sensations of fullness and predisposition to cause gastrointestinal distress. Care should be taken to plan meal times in concert with training, as well as to make sure athletes have sufficient availability of nutrient dense foods throughout the day for snacking between meals.

- **Carbohydrate:**

Athletes must consume proper amounts of carbohydrate, protein and fat in their diet. However, athletes involved in moderate and high volume training need greater amounts of carbohydrate and protein in their diet to meet macronutrient needs. Many sports nutrition specialists recommend that athletes consume concentrated carbohydrate juices/drinks and consume high carbohydrate supplements to meet carbohydrate needs. Care should be taken to consider the type of carbohydrate to ingest prior to, during, and following intense exercise in order to optimize carbohydrate availability.



- **Protein :**

If insufficient amount of protein is obtained from the diet, an athlete will maintain a negative nitrogen balance, which can increase protein catabolism and slow recovery which way lead to muscle wasting and training intolerance. Care should be taken to ensure that athletes consume a sufficient amount of quality protein in their diet to maintain nitrogen balance. Proteins differ based on the source that the protein was obtained, the amino acid profile of the protein, and the methods of processing or isolating the protein and anabolism.

- **Vitamins:**

Vitamins are essential organic compounds that serve to regulate metabolic processes, energy synthesis, neurological processes, and prevent destruction of cells. 2 primary classifications of vitamins are fat and water soluble. Some vitamins may help athletes tolerate training to a greater degree by reducing oxidative damage (Vitamin E, C) and help to maintain a healthy immune system during heavy training (Vitamin C). Since dietary analyses of athletes have found deficiencies in caloric and vitamin intake, many sports nutritionists' recommend that athletes consume a low dose daily multivitamin and / or a vitamin enriched post workout carbohydrate / protein supplement during periods of heavy training.

- **Minerals**

Minerals are essential inorganic elements necessary for a host of metabolic processes. Minerals serve as structure for tissue, important components of enzymes and hormones, and regulators of metabolic and neural control. Some minerals have been found to be deficient in athletes or become deficient in response to training and / or prolonged exercise. When mineral status is inadequate, exercise capacity may be reduced. Dietary supplementation of minerals in deficient athletes has generally been found to improve exercise capacity. Additionally, supplementation of specific minerals in non-deficient athletes has also been reported to affect exercise capacity. Iron supplementation in athletes prone to iron deficiencies and anemia has been reported to improve exercise capacity. Sodium phosphate loading has been reported to increase maximal oxygen uptake, anaerobic threshold, and improve endurance exercise capacity by 8 to 10%. Increasing dietary availability of salt (sodium chloride) during the initial days of exercise training in the heat has been reported to help maintain fluid balance and prevent dehydration. There is no benefit of mineral supplementation for athletes and it is unethical for a sports nutrition specialist to recommend that their clients take minerals for health and / or performance benefit is not consistent with current available literature.

- **Water :**

Most important nutritional ergogenic aid for athletes is water. Exercise performance can be significantly impaired when 2% or more of body weight is lost through sweat. Weight loss of more than 4% of body weight during exercise may lead to heat illness, heat exhaustion, heat stroke, and possibly death, for this reason, it is critical that athletes consume a sufficient amount of water/sports drinks during exercise in order to maintain hydration status. Athletes should not depend on thirst to



prompt them to drink because people do not typically get thirsty until they have lost a significant amount of fluid through sweat, they should weight themselves prior to and during exercise training to ensure that they maintain proper hydration and make sure that they consume more fluid in hotter/humid environments. Sports nutrition specialists can play an important role in educating athletes and coaches about proper hydration methods and supervising fluid intake during training and competition.

- **Dietary Supplements :**

Dietary supplements can help athletes to consume proper amount of calories carbohydrate, and protein in their diet. However, they should be viewed as supplements to the diet, not replacements for a good diet. Most dietary supplements available for athletes have little scientific data supporting their potential role to enhance training and performance. As, number of nutrients and dietary supplements have shown to help in improving performance and recovery. Supplementation with these nutrients can help to augment normal diets which help to optimize performance. Also, convenience supplements which are meal replacement powders, ready to drink supplements, energy bars, and energy gels currently represent the largest segment of the dietary supplement. They are typically fortified with vitamins and minerals and differ on the amount of carbohydrate, protein, fat they contain. Use of these types of products can be particularly helpful improving carbohydrate, protein, and other nutrients prior to exercise. Care should also be taken to make sure they do not contain any banned or prohibited nutrients.

Reference:

1. Beals, K. and Manore, M. (2007), "Nutritional considerations for the female athlete. In: Advances in Sports and Exercise Science Series". Philadelphia, PA: Elsevier; Pp. 187-206.
2. Brown, RC. (2002), "Nutrition for optimal performance during exercise: Carbohydrate and fat" Current Sports Medicine Reports 2002; Vol. 1, No.4, Pp 222-229.
3. Coyle, E.F., (2004), "Fluid and fuel intake during exercise", *Journal of Sports Sciences* 2004, Vol. 22, No. 1, Pp. 39-55.
4. Conningham, J.J. (1980). A reanalysis of the factors influencing basal metabolic rate in normal adults. *The American journal of Clinical Nutrition*. Vol. 33 (11), Pp: 2372-2374.
5. Jones, A.M. (2014). Influence of dietary nitrate on the physiological determinants of exercise performance: a critical review. *App. Phys. NutrMetab*. Vol. No. 39(9), Pp: 1019-1028.
6. Lambert, C.P., Frank, L.L. and Evans, W.J. (2004). "Macronutrient considerations for the sport of bodybuilding", *Sports Medicine* Vol. 34, No.5, Pp. 317-327.
7. Loucks, A.B. (2013). Energy balance and energy availability. In :Maughan RJ, ed. *Sports Nutrition. The Encyclopaedia of Sports Medicine, an IOC Medical Commission Publication*. West Sussex, UK: John Wiley & Sons, Ltd. Pp: 72-87.
8. Lukaski, H. C. (2001). Magnesium, zinc, and chromium nutrition and athletic performance. *Cand Jour. App. Phy.* Vol. No. 26 (Suppl). Pp: 13-22.

IFSIJ IMPACT FACTOR: 5.565

E-ISSN: 2455-1511

**SANSKRUTI INTERNATIONAL MULTIDISCIPLINARY
RESEARCH JOURNAL**

PEER REVIEWED, REFEREED & INDEXED JOURNAL

Special Issue: 009

April - 2021

**G. S. Tompe Mahavidyalaya Sarvajanik Trust's
G. S. Tompe Arts, Commerce & Science College**

Chandur Bazar, Dist. Amravati, Maharashtra (India)

Affiliated to Sant Gadge Baba Amravati University, Amravati.

NAAC Re-accredited B++ CGPA 2.89

On the Occasion of International Women Day

Held by

WOMEN CELL IN ASSOCIATION WITH IQAC

ORGANIZED

ONE DAY INTERNATIONAL E- CONFERENCE

QUEENS OF ERA – 2021

23rd March 2021



CONFERENCE THEME :
DRIVE WOMEN TO THRIVE SCIENCE,
TECHNOLOGY, COMMERCE AND HUMANITIES

Special Issue Editors:

Dr. P. S. DEOLE

Dr. Y. M. RAJGURE

Dr. J. J. CHORE

Dr. P. K. SHIRKE



17	Dr. Babasaheb Ambedkar As an Emancipator of Women <i>Dr. Ravikant N. Mahindkar</i>	105
18	Female Roles in Utopian Literature with Special Reference to Bellamy's 'Looking Backward" <i>Dr. Savita D. Thakare</i>	111
19	Gender Inequality <i>Dr. D. S. Wankhade</i>	115
20	Mahatma Jotiba Phule: An emancipator of Indian Women <i>Mr. Devendra Sarvdas Sandmare</i>	118
21	Women and Health/ Stress <i>Vaishali Sambre, Dr. Rajshri Meshram</i>	125
22	Gender equality and violence against women <i>Sarandha Sharma</i>	129
23	Women in 2021 Era <i>Prof. Dr. Sangita A. Jawanjil</i>	135
24	Exposure to Women's in Sports: Need Gender Equality by the Society <i>Dr. Rajendra S. Ramteke, Mr. Vinay S. Deole</i>	139
25	Gender Equality and Violence against Women <i>Dr. Sunil S. Bidwaik</i>	144
26	Role of Information Communication Technology in Libraries and its Services <i>Usha Kumari Sharma</i>	152
27	Lata Mangeshkar: The Queen of Melody <i>Dr. Shaheda Munaf</i>	157
28	E-Business: Management & Strategy <i>Mr. Suresh Namdeo Gawai</i>	160
29	Impact of Social Media on E-Commerce <i>Dr. B. S. Sawant</i>	166
30	Women Empowerment and Technology <i>Dr. Jahangeer Ahamd Bhat, Irfan Hassan Ganaie</i>	171
31	Poetry of Rabindranath Tagore and Aesthetics <i>Showkat Hussain Wani</i>	175
32	Continually fluctuating GVA hampers the sectoral growth <i>Dr. Sonal Santosh Chandak</i>	180
33	संस्कृत साहित्यातील स्त्रियांचे योगदान <i>विषय विशेषज्ञ प्रा. डॉ. पूनम गहुकर</i>	186
34	श्री. क्षेत्र चिंतामणी कळंब पर्यटन केंद्र म्हणून विकासाच्या उपाय योजना <i>प्रा. एन. व्ही. नरुले</i>	191
35	अमरावती जिल्ह्यातील पिक प्रारूप व उत्पादन वाढीचे भौगोलिक अध्ययन <i>प्रा. सचिन एन. भोंबे</i>	197



Gender Inequality

Dr. D. S. Wankhade

Associate Professor, Vidya Bharati Mahavidyalaya, Amravati
dnyaneshwariingle18@gmail.com

Abstract:

Gender Inequality refers to unequal treatment or perception of Individuals based on their sex. It reflects and manifests in humorous dimensions of our life. Gender Inequality and resultant discrimination in varying degrees on the grounds of sex are commonly witnessed admitted and even justified in India despite the policy measures to overcome gender discrimination gender in equalities within the family still exist. Most of the forms of discrimination against women have their roots in patriarchal system and its values. Women are dominated over their male members in their own family they have little power in making decision.

An attempt has been made to find out those factors which are responsible for this problem in India. The present study have fired to suggest some relevant strategies and implication for reducing this gender inequality and to promote the dignified position for children women.

Keywords: gender inequality social and cultural issues.

Introduction:

Gender inequality refers to unequal treatment or perceptions of individuals based on their gender it arises from differences in socially constructed gender roles. Gender Inequality in the family has mostly, but not exclusively been studied by social scientists in the context of married better sexual couples in which men/husband have power than women/wives as manifest in the division of house hold labor, family decision-making and in more extreme. Gender inequality in India refers to health education, economics and political inequalities between men and women in India. Various international gender inequality indices rank India differently on each of these factors as well as on a composite basis, and these indices and controversial.

Gender inequalities, and their social causes, impact India's sex ratio, women's health over their lifetimes their educational attainment and economic conditions, In India, discriminatory attitudes towards either sex have existed for generations and affect the lives of both sexes. Although the constitution of India grants men and women equal rights, gender disparities.



Objectives of gender inequality are intrinsically linked to sustainable development and vital to the realization of human rights for all. The overall objective of gender inequality is a society in which women and men enjoy the same opportunities, rights and obligations in all spheres of life.

▪ **Reasons for gender inequality**

The root cause of gender inequality in Indian society lies in the patriarchy system. According to the famous sociologist "Sylvia Walby patriarchy is a system of social structure and practices in which men dominate and exploit women. Exploitation is an old cultural phenomenon of Indian Society. One of the causes for gender inequality within employment is the division of jobs in most societies. There is an inherent belief that men are simply better equipped to handle certain jobs that pay the best. This discrimination results in lower income for women. Some more aspects are as follows.

- Lack of employment equality.
- Job segregation
- Poor medical care
- Lack of religious freedom
- Lack of bodily autonomy
- Lack of legal protections.
- Uneven access to education
- Poverty
- Illiteracy
- Lack of employment facilities
- Social customs, beliefs and practices
- Social altitude
- Lack of awareness of women

▪ **Two main types of inequality**

- Economics inequality
 - ✓ Income Inequality: income inequality is the extent to which income is distributed unevenly in a group of people
 - ✓ Pay Inequality: A persons pay is different to their income pay refers to payment form employment only.
 - ✓ Wealth inequality



- Social Inequality
 - ✓ Health: Poor health and poverty do go hand-in-hand but high levels on inequality the epidemiological research shows negatively affect the health of even the affluent mainly because, researchers contend, inequality reduces social cohesion that leads to more stress, fear and insecurity for everyone.
- **Inequality in Society today**
 - ✓ The major example of social inequality includes income gap, gender inequality, health care and social class. In health care, some individuals receive better and more professional care compared to others. They are also expected to pay more for these services.
- **Importance of gender awareness**
 - ✓ Gender awareness rising plays an important role in informing women and men about gender equality, the benefits of a more gender equal society and the consequences of gender inequality. Gender awareness rising intends to change attitude, behaviors between women and men.
- **Conclusion**

Through the gender gap has reduced, it is not that significant government should strive to increase and promote the inclusion of women into labor force of this country. There are number of specific actions that can be carried out at global national and institutional level to eliminate gender inequality or discrimination especially at workplace. Since the position of women and their status in any society is usually considered to be an index of its civilizations, the betterment of the women for the country by eliminating gender discrimination may be of the way to achieve growth of whole community as well as nation.

References:

1. www.india.celebrating.com/s
2. <https://en.m.wikipedia.org>
3. <http://www.vic.gov.au>
4. <http://www.nebi.nim>
5. *Esteve-Volart, B (January 2004) gender discrimination and growth: theory & evidence from India 1-61*

RESEARCH NEBULA

*An International Refereed, Peer Reviewed & Indexed Quarterly Journal in
Arts, Commerce, Education & Social Sciences*

DOI PREFIX 10.22183

JOURNAL DOI 10.22183/RN

IMPACT FACTOR 7.399

ONE DAY

INTERNATIONAL INTERDISCIPLINARY E-CONFERENCE

On

ROLE OF PHYSICAL ACTIVITIES, HEALTH AND FITNESS IN TODAY'S CRISIS

16th October, 2021

Jointly Organized by



**IQAC, AND DEPARTMENT OF PHYSICAL EDUCATION & SPORTS
MAHATMA JYOTIBA FULE MAHAVIDYALAYA, AMRAVATI. (M.S.)**



**LATE DATTATRAYA PUSADKAR ARTS COLLEGE, NANDGAON PETH,
AMRAVATI. (M.S.)**

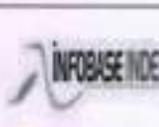


NARAYANRAO RANA MAHAVIDYALAYA, BADNERA, AMRAVATI. (M.S.)

Special Issue on 16th October, 2021

www.ycjournal.net

67.	DR. ANIL S. VAIDYA Director of physical education and sports Late N. A. Deshmukh Arts and Commerce College, Chandur Bazar, Dist - Amravati	THE IMPACT OF SURYANAMASKAR ON ATHLETES' FLEXIBILITY, EXPLOSIVE STRENGTH, AND CARDIO RESPIRATORY ENDURANCE	222
68.	DR. D.S. WANKHADE Director Physical Education & Sports, Vidya Bharti Mahavidyalaya, Amravati	A ROLE OF DIET AND NUTRITION IN SPORTS PERFORMANCE & FITNESS	226
69.	DR. ATUL BIJWE PG department of Physical Education Sant Gadge Baba Amravati University, Amravati atul.bijwe@rediffmail.com	STUDY OF IMAGERY INTERVENTION ON THROWING SKILL PERFORMANCE OF SOFTBALL PLAYERS AT THE UNIVERSITY LEVEL	229
70.	MR. SURENDRA TULSHIRAMSINGH CHAUHAN Director of Physical Education, Art's & Science College KamargaonTq. Karanja, District Washim (M.S.)	A COMPARATIVE STUDY OF SPORTS COMPETITIVE ANXIETY AND AGGRESSION BETWEEN JUDO AND BOXING PLAYERS	235
71.	PROF. MANJUSHA J. DESHMUKH Asstt.Prof. Physical Education, Post Graduate Institute, Dr. P. D. K. V. Akola mjdeshmukh20@gmail.com	STATISTICAL ANALYSISOF SELECTED COLLEGE-LEVEL KABADDI AND KHO- KHO PLAYERS' PHYSICAL FITNESS COMPONENT	238
72.	DR.P.M. DESHMUKH Shri Shivaji College Of Physical Education Amravati	THE IMPORTANCE OF ENVIRONMENTAL ON PROTECTION AND SUSTAINABLE DEVELOPMENT	240
73.	DR.ULHAS V DESHMUKH Shri Shivaji College of Physical Education Amravati	THE IMPACT OF COVID-19 ON PHYSICAL ACTIVITY AND PSYCHOLOGICAL WELL-BEING	243
74.	PROF. RAMKRUSHNA N. GAWANDE Asst.Prof. (English), B.L.D.College,Pinjar, Dist.Akola	ROLE OF MUSIC FOR MENTAL FITNESS AND HEALTH	245
75.	DR. PRAMOD N. HUMBAD Director of Sports & Physical Education MES Arts & Commerce College, Mehkar	IMPACT OF YOGIC PRACTICES ON MOTOR FITNESS COMPONENTS AMONG HIGH SCHOOL STUDENTS - A STUDY	248
76.	DR. CHHAYA B. JATKAR Librarian, Smt. V.N. Mahila Mahavidyalaya, Pusad chhayajatkar@gmail.com	STRESS MANAGEMENT IN LIS PROFESSIONALS	251
77.	PROF.DR. KADAM R.M. (D.P.E.), B.B. Arts, N.B.Commerce & B.P.Science College, Digras District- Yavatmal, Maharashtra rmkadam.judo@gmail.com	A STUDY OF PHYSICAL MOTION IN THE COURSE OF COVID-19 INDUCED LOCKDOWN HINTS	253
78.	DR. SAVITA M. KENE Associate Professor, Yuvashakti Coll.Of Phy Edu Amravati savita.kene2014@gmail.com	THE COMPARATIVE STUDY OF LIFESTYLE OF RURAL AND URBAN ATHLET	257
79.	PUSHPAK PRADIPRAO KHONDE Research Scholar, Shri H.V.P. Mandal's Degree College of Physical Education, Amravati (M.S.)	THE EFFECT OF SIX WEEK CALISTHENIC EXERCISE PROGRAM ON SELECTED PSYCHOMOTOR ABILITIE	259
80.	DR. KAMINI MAMARDE Director of Physical Education & Sports Raje Chatrapati Kala Mahavidhyalaya, Dhamangaon Badhe.Tq. Motala, Dist. Buldana.	IMPORTANCE OF GAME PLAN IN THE SPORTS	264



DR. D.S.
WANKHADE

Director Physical
Education & Sports,
Vidya Bharti
Mahavidyalaya,
Amravati

One Day International Interdisciplinary E-Conference On
ROLE OF PHYSICAL ACTIVITIES, HEALTH AND FITNESS IN TODAY'S CRISIS
On 16th October, 2021 @
Mahatma Jyotiba Fule Mahavidyalaya, Amravati, Late Dattatraya Pusadkar Arts College,
Nandgaon Peth, Amravati. & Narayanrao Rana Mahavidyalaya, Badnera, Amravati.

A ROLE OF DIET AND NUTRITION IN SPORTS PERFORMANCE & FITNESS

ABSTRACT
Nutrition is important for an athlete because it provides energy required to perform the activity. The foods they take leave an impact on strength, training performance & recovery. Not only the type of food is important for sports nutrition but also the time is equally important for what they eat throughout the day. Physical activity or exercise can improve your health & reduce the risk of developing several diseases like type 2 diabetes, cancer & cardio vascular disease. Physical activity & exercise can have immediate & long term health benefits most importantly, regular activity can improve your quality of life.
KEYWORDS: Nutrition, sports performance, intake, fitness.

Introduction:

Nutrition plays as essential role on sports performance following an adequate nutrition pattern determines winning the gold medal of failing in the attempt.

That is commonly referred to as "Invisible training".

However, regarding food & performance, it is not only referred to professional athletes. Now a day, a large number of amateur Athlete perform daily physical activity both recreationally & semi-professionally.

That, population also seeks to achieve an improvement in their follow proper nutritional guideline. In athlete population, nutrient requirements are incremented compared with non-athlete population. Therefore, it is essential to carry out a nutritioning approach adapted to the athlete & training sessions. 1

Nutrition is important for an athlete because it provides energy.

Good nutrition can enhance sporting performance. A well planned nutritious diet should meet most of an athlete's vitamin & mineral needs & provide enough protein to promote muscle growth & repair food rich in unrefined, carbohydrates, like whole grain breads & cereals should form the basis of the diet. At the basic level, nutrition is important for athletes because it provides a source of energy required to perform the activity. The food we eat impact on our strength, training, performance & recovery. 4

Good nutrition physical activity & a healthy body weight are essential parts of a person's overall

health & wellbeing. Together these can help decrease a person's risk of developing serious health conditions, such a high blood pressure high cholesterol, diabetes, heart disease, stroke & cancer.

Food provides, energy for physical activity as you get more activity & more fit & or as you lose weight your energy needs (how many calories you need) many change to get the energy you require, you need to get the proper amount of protein, which is needed to maintain & rebuild tissues such as muscles.

Nutrition & Physical Activity:

Eating a balanced diet being physically active are two of the most important things you can do to be stay healthy at any age.

A balanced diet includes eating the right amount of calories & nutrients to maintain a healthy weight. Physical activity is any form of movement that uses energy can benefit from being physically activities some physical activity is better than none & the more you do the more benefits you gain.

Eating smart & being active have similar effects on our health. This include

- 1) Reduce the risk of chronic disease, such as diabetes, heart disease, high blood pressure stroke, & cancer & associated disabilities.
- 2) Prevent weight gain & /or promote weight loss.
- 3) Improve overall wellbeing active can also improve your personal appearance, encourage with family & friends maintain the ability to live independently & enhance fitness for sports.

Sporting Performance & Food:

- Good nutrition can enhance sporting performance.
- A well planned nutritious diet should meet most of an athlete's vitamin mineral needs & provides enough protein to promote muscle growth & repair.
- Foods rich in unrefined carbohydrates like wholegrain breads & cereal, should form the basis of the diet.
- Water is great choice of fluid for athletes to help performance & prevent dehydration.

The role of diet & nutrition for sports performance while there are numerous general health benefits to a good diet there are 3 main purpose a nutrition plan has to fulfil in relation to sports performance.

Provide energy for training & compulsion:

This sounds quite oblivious but it is not only essential for athletes to meet their daily calories demand but also the manner by which those calories are consumed macronutrients, micronutrients, rations, meal timing & frequency hydration & supplementation are all factors that should be taken into consideration when crafting and optional diet plan.

Facilitate recovery after training & competition:

After an intense match or training session, glycogen (or energy storage) in the muscles are depleted & some proteins in the muscles are broken down & damaged by eating soon after (or even during) training or competition, those glycogen stores are replenished, which prevents muscle (protein) breakdown & accelerates recovery.

Achieve & maintain optimal body weight & composition:

It is important for athletes to achieve optimal body weight body fat levels to maximise their performance. Even more importantly a diet plan must develop in a way that the athlete can effortlessly where long time.

The role of diet & nutrition:

A healthy diet is essential for good health & nutrition, It protect you against many chronic non communicable diseases, such as heart disease diabetes & cancer, earn a variety of foods & consuming less salt, sugar & saturated & industrially produces trans fats are essential for healthy diet.

Eating a healthy balanced diet accompanied by regular exercise is essential in maintain physical & mental health & wellbeing. Not only are these effective in preventing excess weight loss but healthier lifestyle is also associated with improves sleep & mood.

Carbohydrates:

The key nutrients the most important macronutrient in any athlete's diet carbohydrates act

as the main energy source for any physical activity carbohydrates are the main source of glucose. Glucose is converted by the body in to glycogen & stored in the liver & muscles tissues stored glycogen is then used as energy to fuel athletes during physical activities.

Protein:

Protein is a critical part of a training diet & plays a key role in post exercise muscles recovery & repair generally, strength & endurance athletes should aim for 1.5 -2 g of protein per kilogram of body weight per day which is rather easy to achieve. This is because protein need is often already met by following a high & varied carbohydrate diet as many card source also contains a good amount of protein.

Fat:

Carbohydrates & protein are the most important macronutrients in the diet of an athlete but what role do fats play? Contra to old belief & misconception, fats actually play a crucial role in a diet. Hormonal production joint structure & cell membranes are all dependent on fats. Soluble which means they need fat to be fully absorbed in the body?

The optimal fat intake is generally around 10% of total daily calories intake. The most important factor to consider is the type of fat that is consumed.

Mono unsaturated & polyunsaturated fats are what is known to be healthy fats saturated fats also carry health benefits but should be consume in moderation while trans faces especially artificial ones, should be avoided as they have been shown t increase harmful cholesterol reduce the amount of beneficial cholesterol while also in increasing the risk of heart disease.

Minerals:

Minerals are essential inorganic elemental necessary for a host of metabolic process minerals serve as structure for tissue, important components of enzymes & hormones & regulators of metabolic neural control. Some minerals have been found t be deficient in athletes or became deficient in response to training &/ r prolonged exercise capacity may be reduced. Dietary supplementation of minerals in deficient athletes has generally been found to improve exercise capacity. Additionally, supplementation of specific minerals in non-deficient athletes has also been reported to affect exercise capacity. Iron deficiencies & anomie has been reported to improve exercise capacity. Sodium phosphate loading has been reported to increase maximal oxygen uptake anaerobic threshold & improve endurance exercise capacity by 8 to 10% increasing dietary availability of salt (Sodium Chloride) during the initial days of exercise training in the heat has been reported to help maintain fluid balance & prevent dehydration. There

is no benefit of minerals supplementation for athletes & it is unethical for a sport nutrition specialist to recommend that their client take mineral for health / or performance benefit is not considering with current available literature.

Water:

Most important nutritional cryogenic aid for athletes is water exercise performance can be significantly impaired when 2% or more of body weight is lost through sweat weight loss of more than 4% of body weight during exercise may lead to heart illness, heat, exhaustion, heat stroke and possibly death, for this reason, it is critical that athletes consume a sufficient amount of water / sports drink during exercise in order to maintain hydration status. Athletes should not depend on thirst to prompt them to drink because people do not typically get thirsty until they have lost a significant amount of fluid through sweat they should weight themselves prior to and during exercise training to ensure that they maintain proper hydration & make sure that they consume more fluid in hotter/ humid environment sports nutrition specialists can play an important role in educationally athletes & coaches about proper hydration method & supervising fluid intake during training & competition.

Dietary supplements: Nitrates, better almandine, Vitamin D:

Dietary supplements can help athletes to consumes proper amount of calories caffeine, beetroot juice, beta almandine centime, & bicarbonate comprehensive reviews on other supplements including caffeine creative & bicarbonate can be found elsewhere in recent yours research has focused on the role of nitrate BA, vitamin D & performance. During exercise nitric oxide potentially influences skeletal muscles function through regulating of flood flow & glucose homeostasis as well as mitochondrial respiration during endurance exercise, nitrate supplementation has been shown to increase exercise efficiency (4% - 5% reduction in VO2 at steady state, 0.9% improvement in time trails) reduce fatigue & attenuate oxidative stress vitamin D is essential for the maintenance of bone health & control of calcium homeostasis but is also important for muscles strength regulation of the imbue system & cardiovascular health. Thus inadequate vitamin D status has potential implication for the overall health of athletes of performance.

Reference:

1. <https://www.opensciencepublication>
2. www.betterhealth.vic.gov.au
3. www.spectramhealth.ie.blog
4. www.healthypeople.gov
5. www.healthmov.gov
6. www.spectrumhealth.ie
7. Jones, Am (2014) influence of dietary nitrate on the physiological determinations of exercise performance a critical review. App Phys. Nutri Metab Vol. No. (39), Pp 1019- 1028.

The Phenotypic Effect of Cardamom oil on the Developmental Stages of *Drosophila Melanogaster*

Dr. Y. D. Akhare, H. A. Patharikar

Department of Zoology, Vidya Bharati Mahavidyalaya, C. K. Naidu Road, Camp. Amravati, Maharashtra, India

ABSTRACT

Article Info

Volume 8, Issue 5

Page Number : 225-231

Publication Issue

September-October-2021

Article History

Accepted : 20 Sep 2021

Published : 30 Sep 2021

The fruit fly *Drosophila melanogaster* has been extensively studied as a model organism for genetic investigation. It also has many characteristics which make it an ideal organism for the study of animal development and behaviour, neurobiology and human genetic disease and condition. *Drosophila melanogaster* share several basic biological and chemical neurological and physiological similarities with mammals. In the present study, we noted the phenotypic effect of cardamom oil on the different stages of *Drosophila melanogaster*. The fruit flies were grown on 10-gram culture media supplemented with different concentration of cardamom oil (0.5µl, 1 µl, 2.5 µl). Further, the size and growth of different life stages of *Drosophila melanogaster* were observed and total protein estimated from it. The increase in the size and protein concentration in different life stages of controlled *Drosophila melanogaster* were recorded. Cardamom is a highly valued herbal spice used in tropical and subtropical Asia. cardamom is used as a flavouring and cooking spices in both food and drink and as a medicine.

Keyword: *Drosophila Melanogaster*, Cardamom Oil, Herbal Spice, Culture Media, Larval Stages, Protein

I. INTRODUCTION

Spices were among the most valuable items of trade in ancient and medieval times. As long ago as 3500 BC the ancient Egyptians were using various spices for flavouring food, in cosmetics, and for embalming their dead. The use of spices spread through the Middle East to the Eastern Mediterranean and Europe. Today, it is available in most tropical places in Asia, including India, China, Bhutan, Vietnam, Malaysia, Korea, and Japan. In botanical terms, it belongs to the family of Zingiberaceae and its scientific name is *Elettaria cardamomum*. There are two main types or

subspecies of cardamom. Their scientific names are *Elettaria*, which is called green or true cardamom, and *Amomum*, which stands for black white, or red cardamom (Nair KPP 2002). Cardamom has digestive, antispasmodic, carminative, anti-inflammatory, anti-microbial, aphrodisiac and diuretic properties.

Drosophila melanogaster, also known as the fruit fly, is an excellent model organism widely used in biological research that has made significant contributions to the greater scientific community over the last century. It is because fruit flies are inexpensive to maintain in the laboratory, have

simplified genetics, and short generation times allow for quick experiments with high sample numbers. *Drosophila melanogaster* exhibits complete metamorphosis, meaning the life cycle includes an egg, larval forms, pupa and finally emergence as a flying adults. Eco-friendly and easily biodegradable plant products with natural insecticidal activity has increased in recent years. To control pests without disturbing the environment, natural products have been screened for potential sources of insecticides. Plant materials with insecticidal properties have been employed to kill insects all over the world for generations. These materials are considered to be an alternative to conventional pesticides owing to their low toxicity to warm-blooded mammals, in addition to their high volatility. Botanical insecticides may be safer for the environment than synthetic insecticides, and they are usually easily processed and utilized by farmers and small industries (Belmain et al., 2001). The cardamom, *Elettaria cardamomum* (L.) Maton (Zingiberaceae), as also known as the “Queen of spices”, is a tall, perennial, reed-like herb growing wild. Some studies have shown that cardamom essential oils have inhibitory effects against fungus growth (Utta-Ur' et al., 2000), insecticidal activity (Abbasipour et al., 2011b) and marked antiapamodic, analgesic, as well as anti-inflammatory activity (Al-Zuhair et al., 1996). This study aims to investigate the effect of Cardamom Oil on the developmental stages of *Drosophila melanogaster*.

II. MATERIAL AND METHODS

Fruit flies were collected on ripe banana fruit. After that flies were cultured on potato, dextrose and agar culture media. Cardamom oil has monoterpenes, vitamins and minerals.

Preparation of culture media: Boil the potato and clean it then weight up to 100gm then add 1.09 gm agar, 100ml of H₂O added in it and 1 gm dextrose. Homogenize the mixture and make a fine paste. After

the formation of paste, autoclave it at 15 lb for 10-15 min. Cool the media up to 60°C. Then add 0.8 ml propionic acid put the culture media up to 40°C add 0.59gm yeast in it by dissolving it in distilled water then sterilized it in the autoclave at 15 IP

Composition of Cardamom Oil (Bitter): Essential oil of *Elettaria cardamom* leaves was analyzed by GC-MS Sixteen components constituting 93.62% of the total were identified The identified component were monoterpenes (27.37%), oxygenated monoterpenes 63%, acetates (0.63) sesquiterpenes (1:43) and fatty acids caster 1.17 % among this 4 terpineol 30.261 & and 1:8 cenol 25,74% were found as major components where as other components were found to be a terpinene (9,807%) p cymene (5.300%) terpinene (4.6759) tujene (1633%) la pinene (1.16596) sabinene (2.069%) terpieness (2.675%) linalool (2.6739) menth-2-en-1-01 0.754%), terpanol (3:44%) and endbory acetate(0.593)

Experimental set up: Take four culture vial, 1st having 0.5ul cardamom oil and 10gm of normal culture media, 2nd having 1ul cardamom oil and 10gm of normal culture media, 3rd having 2.5ul cardamom oil and 10gm of normal culture media and 4th having only 10 gm culture media without cardamom oil for normal growth of *Drosophila melanogaster*. *Drosophila* flies anaesthetized with diethyl ether and transfers into culture vial and each vial contains 10 lies (5 female and 5 male). The vial cover with cotton plug and place on the working table. The flies were laid eggs in culture, hatch and completed the life cycle within 10 to 15 days. Then collect the larval stages for measurement of total length and protein estimation by De Lowry's method. The measurement of the size of the various stages of *Drosophila melanogaster* was carried out with help of Oculometer. The total estimation of protein was done by the De' lowery method. Photographs were taken in the Carl Zeiss trinocular microscope and stereomicroscope. All results are presented as Mean ±

SD. Test of significance (t-test) was used to analyse the data collected.

III. RESULT AND DISCUSSION

The size and growth in the control *D. melanogaster* are normal but in experimental *D. melanogaster* size and growth rate was slightly higher. In the micrometre, the change was 1.5um, 3.3um, 0.3um, 2um, 0.5um, 0um of 1st, 2nd, 3rd instar larvae, prepupa, pupa and adult respectively as compared to controlled *D. melanogaster*.

The protein concentration in control *D. melanogaster* is normal but in experimental *D. melanogaster* level of protein concentration was slightly higher. The change protein concentration was 0.7ul, 0.5ul, 0.34ul, 0.11ul, 0.19ul, -0.2ul of 1st, 2nd, 3rd instar larvae, prepupa, pupa and adult respectively as compared to control *D. melanogaster*. In the present study we found that with the increased in concentrations the size of different development stages of *Drosophila*

melanogaster were increased as compared to control. This increase in size may be result due to the effect of cardamom oil.

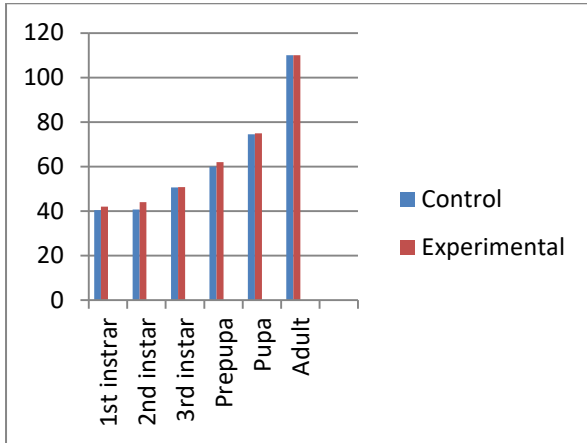
In other studies (Arslan et al. 2005; Al-Saleh et al. 2006). The higher bursa index observed in the chickens fed CEO-supplemented diets could be attributed to the antimicrobial and anti-inflammatory activities of active substances from cardamom, which may induce a positive effect in the activation and development of these organs. The inflammatory cytokines, IL-1b and IL-6, which are widely expressed in lymphoid tissue of birds, are the most powerful cognitive pathways in response to inflammatory processes.

Green cardamom supplement improved the grade of fatty liver, serum glucose indices, lipids, and irisin level among overweight or obese NAFLD patients. The changes in these biomarkers may yield beneficial effects on NAFLD (Daneshi-Maskooni et al., 2019).

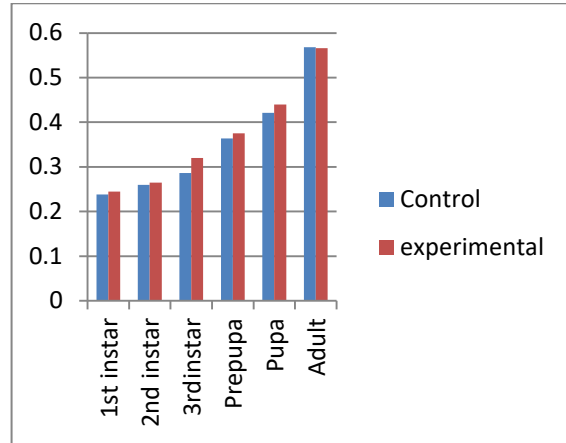
Table No. 1 : Measurement of size and protein from various larval stages of *Drosophila melanogaster*.

Sr. No.	Larval stage	Size of larval stage in micrometer (µm)		Protein concentration larval stage in microlitre (µl)	
		Control	Experimental	Control	Experimental
1	1 st instar	40.5	42	0.238	0.245
2	2 nd instar	40.7	44	0.260	0.265
3	3 rd instar	50.5	50.8	0.286	0.320
4	Prepupa	60	62	0.364	0.375
5	Pupa	74.5	75	0.421	0.440
6	Adult	110	110	0.568	0.566

Size of larval stage in micrometer (μm)



Protein conc. larval stage in microlitre (μl)



First Instar Larvae

Control



Experimental



0.5 μl

1 μl

2.5 μl

Second Instar Larvae

Control



Experimental



0.5 μl

1 μl

2.5 μl

Third Instar Larvae

Control



Experimental



0.5 μl

1 μl

2.5 μl

Pre-pupa

Control



Experimental



0.5µl



1µl



2.5 µl

Pupa

Control



Experimental



0.5µl



1µl



2.5 µl

Adult

Control



Experimental



0.5µl



1µl



2.5µl

Laarif et al. (2013) reported the chemical composition and insecticidal activity of essential oil from (*Citrus aurantium*) (Rutaceae) fruit peels against two greenhouse insects such as *Spodoptera littoralis* (Noctuidae) and *T. absoluta*. Moawad et al. (2015) reported that mixed clove, bitter orange and zinc sulfate collectively had the ability to cause highest mortality to *T. absoluta* larvae to a degree of 97.0%.

In another study, the insecticidal properties of essential oils extracted from *Citrus aurantium* were tested against third instar larvae of pest species, *T. absoluta* (Ebadah et al., 2016). LC50 values for 3rd larval instars (outside leaf and inside leaf) were 17.56 and 59.93 ml/L air, respectively. However, in recent studies, the LC50 values of cardamom essential oils in the leaves inside and outside of the larvae were 7.87,

1.54 ml/L air, respectively. Hence, the essential oil of cardamom is more effective than that of orange. Moawad et al. (2013) stated that clove, eugenol and iso eugenol caused highly reduced percentage of penetration and accumulative mortality of larvae and caused ovipositional deterrence reaction toward adult stage of *T. absoluta* under laboratory conditions. In addition, Perez et al. (2013) evaluated the effect of different cinnamates on the behavior of larval herbivory of the second stage of *T. absoluta* on tomato discs in different aqueous solutions of different concentration.

These results suggest that *E. cardamomum* oil has potential to be used in sustainable pest management in the greenhouse. It will purely safeguard the environment and health of the user especially when the application of synthetic insecticides gives rise to the development of resistance and pollution of the environment.

IV. REFERENCES

- [1]. Abbasipour H, Mahmoudvand M, Rastegar F, Hosseinpour MH. (2011b). Fumigant toxicity and oviposition deterrence of the essential oil from cardamom, *Elettaria cardamomum*, against three stored-product insects. *J Insect Sci* 11:165.
- [2]. Al-Saleh et al. 2006 Al-Saleh, Griselli Billedo & Innam I. El-Doush (2006), levels of selenium, DL- α -tocopherol, DL- γ -tocopherol, all-trans-retinyl, thymoquinone and thymol in different brands of *Nigella sativa* seeds, *Journal of Food Composition and Analysis*, Vol. 19, Issues 2-3, March-May 2006, 167-175.
- [3]. Al-Zuhair H, Al-Sayed B, Ameen HA, Al-Shoora H. (1996). Pharmacological studies of cardamom oil in animals. *Pharmacol Res* 34:79-82.
- [4]. Arslan H. , Ozlem Kurt Azap *, Onder Ergonul 2 and Funda Timurkaynak1 (2005) Risk factors for ciprofloxacin resistance among *Escherichia coli* strains isolated from community-acquired urinary tract infections in Turkey, *Journal of Antimicrobial Chemotherapy* (2005) 56, 914-918.
- [5]. Belmain SR, Neal GE, Ray DE, Golop P. (2001). Insecticidal and vertebrate toxicity associated with ethnobotanicals used as postharvest protectants in Ghana. *Food Chem Toxicol* 39:287-91.
- [6]. Daneshi-Maskooni, M., Keshavarz, S. A., Qorbani, M., Mansouri, S., Alavian, S. M., Badri-Fariman, M. & Sotoudeh, G. (2019). Green cardamom supplementation improves serum irisin, glucose indices, and lipid profiles in overweight or obese non-alcoholic fatty liver disease patients: a double-blind randomized placebo-controlled clinical trial. *BMC complementary and alternative medicine*, 19(1), 1-11.
- [7]. Ebadah IM, Shalaby SEM, Moawad SS. (2016). Impact of certain natural plant oils and chemical insecticides against tomato insect pests. *J Entomol* 13:84-90.
- [8]. Laarif A, Zarrad K, Tayeb W, et al. (2013). Chemical composition and insecticidal activity of essential oil from (*Citrus aurantium*) (Rutaceae) fruit peels against two greenhouse insects; *Spodoptera littoralis* (Noctuidae) and *Tuta absoluta* (Gelechiidae). *Adv Agric Sci Eng Res* 3:825-30.
- [9]. Moawad SS, Ebadah IM, Mahmoud YA. (2013). Biological and histological studies on the efficacy of some botanical and commercial oils on *Tuta absoluta* Meyrick (Lepidoptera: Gelechiidae). *Egypt. J Biol Pest Control* 23:301-8.
- [10]. Moawad SS, Sharaby A, Ebadah IM, El-Behery H. (2015). Efficiency of zinc sulfate and some volatile oils on some insect pests of the tomato crop. *Global Adv Res J Agric Sci* 4:182-7.

- [11]. Nair KPP (2002) "The nutrient buffer power concept" – a revolutionary soil management technique for sustainable agriculture. Invitational research paper presented at the 17th world soil science congress. Bangkok, August, p 2002.
- [12]. Perez ME, Haramboure M, Mirande L, et al. (2013). Biological activity of three Alkyl Cinnamates on young larvae of *Tuta absoluta*. *Comm Appl Biol Sci* 78:299–303.
- [13]. Utta-Ur' R, Choudhary MI, Ahmed A, Iqbal MZ. (2000). Demirci B, Demirci F, Baser KHC. Antifungal activity and essential oil constituent of some spices from Pakistan. *J Chem Soc Pakistan* 22: 60–5.

Cite this article as :

Dr. Y. D. Akhare, H. A. Patharikar, "The Phenotypic Effect of Cardamom oil on the Developmental Stages of *Drosophila Melanogaster*", *International Journal of Scientific Research in Science and Technology (IJSRST)*, Online ISSN : 2395-602X, Print ISSN : 2395-6011, Volume 8 Issue 5, pp. 225-231, September-October 2021. Available at doi : <https://doi.org/10.32628/IJSRST218534>
Journal URL : <https://ijsrst.com/IJSRST218534>

Variability of Nocturnal Insects In Sant Gadge Baba Amravati University Campus, Amravati

Dr. Y. D. Akhare., Mohd. Khadim, H. A. Patharikar

Vidya Bharati Mahavidyalaya, C. K. Naidu Road, Camp, Amravati, Maharashtra, India

ABSTRACT

Article Info

Volume 8, Issue 5

Page Number : 157-193

Publication Issue

September-October-2021

Article History

Accepted : 15 Sep 2021

Published : 23 Sep 2021

The insects are known to be the most successful and diverse animals on earth. Variability of nocturnal insects in Sant Gadge Baba Amravati University campus was recorded by using the light trap method for collection of insect. This trap consists of the light source and a big white sheet. The sheet was hung on the pillars. Insects were attracted to the light source settled on the sheet. The second sheet was also spread on the ground to catch the insect that fall. Collected insect in the killing bottle by beating tray aspirator, and forceps. Lightning hours were set for seven hours. During this study, total 6 orders of species of insects are found are Coleoptera, Lepidoptera, Orthoptera, Hemiptera, Diptera, and Hymenoptera, insects belonging to different families was collected in university campus by using light trapping methods respectively and moderate species diversity were observed. The abundance of species of most night insect in the study area maybe due to the loss of their natural habitat. This may be the virtue of anthropogenic activity such as fragmentation of habitat, replacement of the habitat for humane purpose or may be other due reasons that eventually lead to the decline of species diversity.

Keywords : Variability, Nocturnal insects, Sant Gadge Baba Amravati University.

I. INTRODUCTION

Insects have been around for more than 400 million years and it could be argued that they are the most successful and enduring life form that has ever arisen on this planet. Insects are abundant and ubiquitous. From the poles to the equator, from the surface of the sea to the highest peaks and from deserts to rain

forests. Diverse as well as abundant, insects comprise roughly half of the Earth's one and a half million known species. There are many more species than those to which we have given names and past estimates have been as high as 100 million. The majority view nowadays is that we share the planet with somewhere between 5-15 million species, of which insects will be a sizeable proportion. Insects

are the dominant component of biodiversity in terrestrial ecosystems and play important roles in ecosystem processes (Weisser and Siemann, 2004). Insects exhibit considerable variations in their season availability, size, trophic level, life history, mobility, strategy and habitats. Insect communities constitute an integral part of terrestrial ecosystems by the diversity of both the species and life forms (Adjaloo et al., 2012). Insects dominate in many food webs and food chain lengths (Sugihara et al., 1997) and have great importance because of their diversity, ecological roles and influence on the agriculture, natural resources and human health (Footitt and Adler, 2009). About one million species are extant and certainly, many more await discovery and description, especially in the tropics. The class includes 30 orders with the variable number of species: the less diverse include the Mantophasmatodea (gladiators, 24 spp.) (Zompro et al, 2002; Damgaard et al. 2008), Grylloblattodea (ice crawlers, 32 spp.) (Wipfler et al. 2014), and Zoraptera (angel insects, 30 spp.) (Mashimo et al. 2014); whereas the more diverse comprise the Dipteran (flies, 100,000 spp.). Hemiptera (bugs, 100,000 spp.), Hymenoptera (wasps, bees, ants, and sawflies; 120,000 spp.) Lepidoptera (butterflies and moths, 150,000 spp.) and of course, the Coleopteran (beetles, 370,000 spp.) (Capinera 2008). The Orthoptera senses triticois somewhat in the middle regarding the number of species 25,000 although this may be doubled with new species from poorly surveyed or unexplored regions. The order includes well-known insects such as grasshoppers, locusts, katydids and crickets but also, bush-crickets, wetas, mole-crickets, ant-inquiline crickets

Keeping in mind the above literature the present study aimed to know the different variety of insects from different order at selected places of Amravati university campus.

II. MATERIAL AND METHODS

In the present study, we have been used the light trap method for collection of insect. This trap consists of

the light source and a big white sheet. The sheet was hung on the pillars. Insects were attracted to the light source settled on the sheet. The second sheet was also spread on the ground to catch the insect that fall. Collected insect in the killing bottle by beating tray aspirator, and forceps. Lightning hours were set for seven hours.

A cotton swab dipped in benzene was put in killing bottle which used to anesthetize the insects. The insect was cached up in the following month (January, February, and March) at different interval of time.

The collected insects was put in the killing bottle to anesthetize and then removed to dry preservation for overnight by using of 160w mercury bright light. The average temperature ranged from 30–40 °C maximum and 15–25 °C minimum during the collection. The collected insects were then allowed for labelling, and sorting order wise.

Site of collection :

Insects are a remarkable group of animals. They occur virtually everywhere and make up more than half of all living things on earth. In the present study, we have selected the study area “Sant Gadge Baba Amravati University campus” for insect collection. The area was mostly vegetation and moderately temperate during the collection month. Most insects were noticed near the swampy area as well as the gardening site.

Preservation of collected insects :

The collected insects were transferred from killing bottle to preservation box. Smaller insects were preserved in 70% alcohols and large insects were preserved by pinning. Insects were pinned through properly thorax region about halfway between the two ends of the body.

Labelling and storage :

In order to have scientific values the collected specimens were labelled, in one side of the label, we noted locality, data, and last name of the collector in

an international convention writing procedure by a permanent marker pen.

Now, the labelled dried specimens were kept in a storage box with a small amount of flake naphthalene so that it may be protected from other museum insects.

III. OSERVATIONS AND RESULT

During this study, total 6 orders of species of insects are found are Coleoptera, Lepidoptera, Orthoptera, Hemiptera, Diptera, and Hymenoptera, insects belonging to different families were collected in

university campus by using light trapping methods respectively and moderate species diversity was observed. Temperature variations during the study period of different months. The highest number of Lepidoptera was collected during February and March 2016. Some habitats factors that influence the patterns of Lepidoptera and other orders diversity are vegetation including host plants, food availability, temperature and wind exposure. So the temperature plays a major role in distributions of night insects. Mostly the nocturnal insects identified and classified at genus level and the rest of few are pause at the family level below under classified are.

Classification	Figure 1	Figure 2	Figure 3	Figure 4	Figure 5	Figure 6	Figure 7	Figure 8	Figure 9	Figure 10
Order	<i>Lepidoptera</i>	<i>Lepidoptera</i>	<i>Lepidoptera</i>	<i>Lepidoptera</i>	<i>Lepidoptera</i>	<i>Lepidoptera</i>	<i>Lepidoptera</i>	<i>Lepidoptera</i>	<i>Lepidoptera</i>	<i>Lepidoptera</i>
Family	<i>Noctuidae</i>	<i>Noctuidae</i>	<i>Noctuidae</i>	<i>Aretidae</i>	<i>Arctidae</i>	<i>Natodontidae</i>	<i>Crambidae</i>	<i>Geometridae</i>	<i>Erebidae</i>	<i>Tineoidea</i>
Genus	<i>Luypurina</i>	<i>Sesamia</i>	<i>Orthosia</i>	<i>Apaidia</i>		<i>pheosia</i>	<i>Agriphilla</i>			<i>Tineola</i>
Species	<i>Testacea</i>	<i>nonagriodes</i>	<i>hebisci</i>	<i>mesogona</i>						<i>bisselliella</i>

Classification	Figure 11	Figure 12	Figure 13	Figure 14	Figure 15	Figure 16	Figure 17
Order	<i>Hemiptera</i>	<i>Hemiptera</i>	<i>Hemiptera</i>	<i>Hemiptera</i>	<i>Hemiptera</i>	<i>Hemiptera</i>	<i>Hemiptera</i>
Family	<i>Alydidae</i>	<i>Rhopalidae</i>	<i>Coreidae</i>	<i>Coreidae</i>	<i>Rhopalidae</i>	<i>Coreidae</i>	<i>Reduviidae</i>
Genus	<i>Alydis</i>	<i>Biosea</i>	<i>Leptoglossus</i>	<i>Leptoglossus</i>	<i>Halymortha</i>		
Species			<i>occidentalis</i>				



Figure No- 1



Figure No- 2



Figure No- 3



Figure No- 4



Figure No- 5



Figure No- 6



Figure No- 7



Figure No- 8

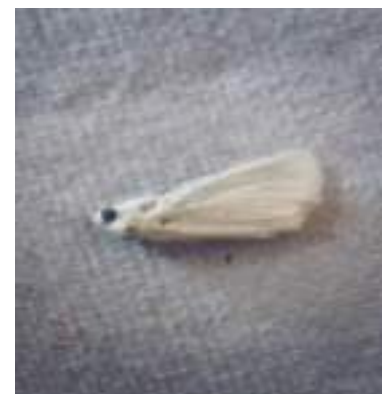


Figure No- 9



Figure No- 10



Figure No- 11



Figure No- 12



Figure No- 13



Figure No- 14



Figure No- 15



Figure No- 16



Figure No- 17



Figure No- 18



Figure No- 19



Figure No- 20



Figure No- 21



Figure No- 22



Figure No- 23



Figure No- 24

IV. DISCUSSION

Based on the survey which was carried out in the present study a total of 41 species belonging to 12 families were found. This study was mainly carried

out to elucidate the biodiversity of moth fauna that has not been studied previously. It was observed that the number of moth species belonging to family Erebidae, was found more than other families viz., Noctuidiae, Crambidae, Arctiidae, Geometridae,

Sphingidae, Lymantriidae, Saturniidae, and Lasiocampidae. The collection was more mainly in August. Similar studies were carried out at 16 sites in southern Korea to determine the patterns of diversity for moths in this area. A total of 975 moth species were recognized in the 6 month collection periods (May to Oct) between 2001 and 2007. Species diversity and seasonal abundance of fruit piercing moth were carried out from different localities in Tamil Nadu. They observed five species of fruit piercing moth belonging to two genera (Ramkumar 2010). Comprehensive surveys of moth diversity have been done in Hawaii (Zimmerman 1948) and on larger continental islands New Zealand (Hudson 1928), and (Holloway 1976).

Orthoptera and coleopteran showed fewer species in the study area. The presence of less food and shelter on the restored site could be the reason for the less diversity of these insect species on that site. Most species of belongs to these three order just visit the ecosystem for sake of food and shelter as "Visitors" and come from the nearby agricultural fields on which they act as pests (Ahmed et al., 2004; Sultana and Wagan, 2010). Dipteran communities were more clearly differentiated by habitat as defined by dominant vegetation rather than geographic proximity between sites. This result does not mean that spatially autocorrelated factor does not influence these communities, however. In regions such as our study area that incorporate many habitat type distance may not be as important as habitat type (Harrison et al. 1992). Lepidopterans were probably one of the common and abundant among other insect species in the study site. This could be many reasons elaborated by Holloway (1980, 1984 and 1985), especially their abundance, species richness, response to vegetation and climate.

In tropical zones, the most significant temperature is not seasonal, but rather diurnal and nocturnal. The differences in temperature between day and night

could also explain the diversity and species richness between diurnal and nocturnal insects, as the insects must explain more energy to adapt to the lower nocturnal temperature. Indeed, at the individual level the available heat, as indicated by body temperature, is the most significant variable that determines growth and activity (Huffaker and Gutierrez 1999). Metabolic activities essential for development, feeding, dispersal, reproduction, and survival may all are impeded by the decrease in nocturnal temperature, which likely results in great diversity, species richness, and abundance. The abundance of species of most night insect in the study area may be due to the loss of their natural habitat. This may be the virtue of anthropogenic activity such as fragmentation of habitat, replacement of the habitat for humane purpose or maybe other due reasons that eventually lead to the decline of species diversity.

V. REFERENCES

- [1]. Adajaloo, M., K., Oduru, W., and Mochian., M.B. (2012). Spatial distribution of insect assemblage in cocoa farm in relation to natural forest., J., appl., Biosci., 54: 3870-3879
- [2]. Ahmad, A., Suhail., A., Abdin., Z., Iftikhar, S. and Zahoor, K., (2004). Biodiversity of insects associated with sugarcane crop in Faisalabad. Pak. Entomol., 26: 65-69.
- [3]. Capinera and Vlisidou. (2008).The diversity of insect-bacteria interactions and its applications for disease control. Biotechnol Genet Eng Rev. 2008;25:203-43.
- [4]. Footit., R. G. & P. H. Adler Eds. (2009). Insect Biodiversity: Science and Society. ISBN 978-1-4051-5142-9
- [5]. Holloway., J. D. (1985). Moths as indicator organisms for categorizing rain forest and monitoring changes and regeneration processes.

- [6]. Hudson G.V. (1928). Butterflies and moths of New Zealand. Ferguson & Osborn. Wellington, N.Z.
- [7]. Huffaker., C. B. and A. P. Gutierrez., (1999). Ecological entomology. John Wiley and Sons, Inc., New York, USA.
- [8]. Mashimo, Y., et al. (2015). The morphology of the eggs of three species of Zoraptera (Insects), *Arthropod Structure & Development*, j.asd.2015.09.005
- [9]. Ramkumar, (2010). Species diversity and seasonal abundance of fruit piercing moths Tamil Nadu. *Journal of Biopesticides* 3 (1 Special Issue) 011 - 015 11
- [10]. Sugihara., Bersier., L.F and Schoenly., K. (1997). Effects of taxonomic and trophic aggregation on food web properties. *Oecologia*, 112: 272-284.
- [11]. Sultana. R. and Wagan., M.S., (2010). The effects of various food plants on nymphal development and egg production in *Hieroglyphusperpolita* (Uvarov 1933) (Acrididae Orthoptera) from Pakistan. *Pak. Trop. Zool.*, 23: 1-8.
- [12]. W.W. Weisser., E. Siemann Eds., Springer, Berlin (2004). *Insects and Ecosystem Function*, Ecological Studies, Vol. 173, (415pp., €155, ISBN 3-540-21672-3)
- [13]. Wipfler et al. (2014). Ice Crawlers (Grylloblattodea) – the history of the investigation of a highly unusual group of insect *Journal of Insect Biodiversity* 2(2): 1-25.
- [14]. Zompro., O., Kristensen., N.P. & Adis, J. Mantophasmatodea., (2002). a new insect order with extant members in the Afrotropics. *Science* 296(5572): 1456-1459.

Cite this article as :

Dr. Y. D. Akhare., Mohd. Khadim, H. A. Patharikar, "Variability of Nocturnal Insects In Sant Gadge Baba Amravati University Campus, Amravati", *International Journal of Scientific Research in Science and Technology (IJSRST)*, Online ISSN : 2395-602X, Print ISSN : 2395-6011, Volume 8 Issue 5, pp. 157-193, September-October 2021. Available at doi : <https://doi.org/10.32628/IJSRST218527>
Journal URL : <https://ijsrst.com/IJSRST218527>

Effect of *Ocimum sanctum* Extract on Developmental Stages of *Drosophila melanogaster*

Y. D. Akhare¹, H. A. Patharikar²

^{1,2}Dept. of Zoology, Vidya Bharati Mahavidyalaya, C. K. Naidu Road, Camp., Amravati, Maharashtra, India, 444602

ABSTRACT

The fruit fly *Drosophila melanogaster* has been extensively studied as a model organism for genetic investigation. It also has many characteristics which make it an ideal organism for the study of animal development and behavior, neurobiology, human genetic disease and condition. *Drosophila melanogaster* share several basic biological and chemical neurological and physiological similarities with mammals. In the present study, we noted the phenotypic effect of *Ocimum sanctum* extract on the different developmental stages of *Drosophila melanogaster*. The fruit flies were grown on 10-gram culture media supplemented with concentration of *Ocimum sanctum* (10µl). Further, the size and growth of different life stages of *Drosophila melanogaster* were observed and total protein estimated from it. The increase in the size and protein concentration in different life stages of controlled *Drosophila melanogaster* were recorded.

Keyword: *Drosophila melanogaster*, cardamom oil, herbal spice, culture media, larval stages, protein.

INTRODUCTION

Plants have always been an important source of drugs. A large number of the world's populations, especially in developing countries, depend upon medicinal plants as an alternative and complimentary drugs therapy for various ailments. Some of the most common practices involve the use of crude plant extracts, which may contain a broad diversity of molecules with often unknown biological effects (Yakob et al., 2012).

Since the medicinal plants are being used indiscriminately without notifying to their possible unhealthy or toxic effects, the World Health Organization has recommended that traditional plants used for the treatment of diseases need further scientific investigation their toxic side effects. Plants produce bioactive compounds which act as defence mechanisms against any disease, and at the same time, may be toxic in nature (Roch 2001). *Ocimum sanctum* Linn (Labiatae), known as holy basil, is a commonly used home remedy and has been advocated for various ailments like cold, fever, dysentery, hemorrhage and dyspepsia, glaucoma, cataract, chronic conjunctivitis, and other painful eye diseases, as well as gastric and hepatic disorders in indigenous system of medicine (Prakash & Gupta 2005).

The chemical composition of *Ocimum sanctum* L. is highly complex, containing many nutrients and other biologically active compounds. The quantity of many of these constituents are significantly affected by differing growing, harvesting processing and storage conditions that are not yet well understood.

The sugars are composed of xylose and polysaccharides. Phytochemical investigation of OS stem and leaves have shown constituents like saponins, flavonoids, triterpenoids and tannins. *Ocimum sanctum* L. contains Vitamin C, A and minerals like calcium, zinc and iron, as well as chlorophyll and many other phytonutrients. (Siva M et al., 2016).

In the last decade the fruit fly has gained popularity as a model system for human neurodegenerative diseases (Gistelinet et al., 2012). The genetic toolkit of the fly is extensive and flexible; furthermore it is not too different from the human genome. *D. melanogaster* contains around 13,767 protein-coding genes while humans have 19,599, as a result fairly 70% of human's disease-causing genes are preserved in the fly (Bonner and Boulianne, 2011; Gistelinet et al., 2012). In the present study, we evaluated the effect of *Ocimum sanctum* on development of *Drosophila* flies supplemented with an extract of this plant exhibited reduced mortality. Although our study does not reveal the causal mechanism behind our observed anti-aging effects, it does suggest that *Ocimum sanctum* species is worthy of continued investigation. "Chattopadhyay" 1993 observed the hyperglycemic effects of *ocimum sanctum* leaf extract in normal and streptozotocin diabetic rats. He noticed the oral administration of alcoholic extract of leaves of *Ocimum sanctum* led to marked lowering of blood sugar level in normal rat glucose fed hyperglycemic rats and streptozotocin induced diabetic rats.

MATERIAL AND METHODS

Fruit flies were collected on ripe banana fruit. After that flies were cultured on potato, dextrose and agar culture media. Cardamom oil has monoterpenes, vitamins and minerals.

Preparation of culture media: Boil the potato and clean it then weight up to 100gm then add 1.09 gm agar, 100ml of H₂O added in it and 1 gm dextrose. Homogenize the mixture and make a fine paste. After the formation of paste, autoclave it at 15 lb for 10-15 min. Cool the media up to 60°C. Then add 0.8 ml propionic acid put the culture media up to 40°C add 0.59gm yeast in it by dissolving it in distilled water then sterilized it in the autoclave at 15 IP.

Extraction of *Ocimum sanctum*: Fresh leaves of *Ocimum sanctum* were collected from local area. Specimen were deposited in the laboratory. Collected leaves were oven dried at oven at 40°C overnight and then crushed in a crucible and weighted separately. For methanolic extracts samples of the crush leaves (5gm tulasi) were soaked separately in 100 ml of methanol in a sealed container at 4°C for two days. The sample was then filtered and centrifuged for the aqueous extract.

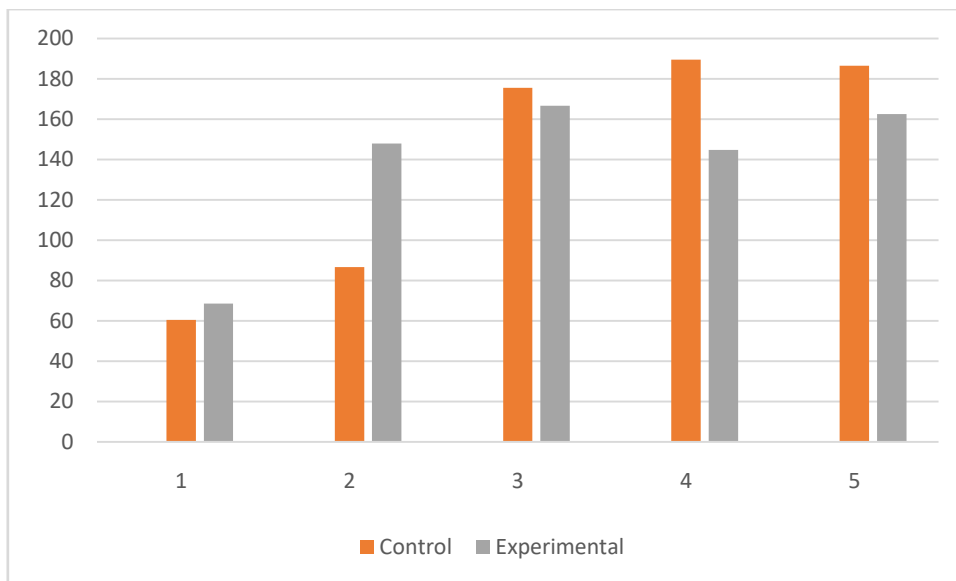
Experimental set up: Take four culture vial having 10 µl *Ocimum sanctum* extract and 10gm of normal culture media and 2nd having only 10 gm culture media without *Ocimum sanctum* extract for normal growth of *Drosophila melanogaster*. *Drosophila* flies anaesthetized with diethyl ether and transfers into culture vial and each vial contains 10 flies (5 female and 5 male). The vial cover with cotton plug and place on the working table. The flies were laid eggs in culture, hatch and completed the life cycle within 10 to 15 days. Then collect the larval stages for measurement of total length and protein estimation by De Lowry's method. The measurement of the size of the various stages of *Drosophila melanogaster* was carried out with help of Oculometer. The total estimation of protein was done by the De' lowery method. Photographs were taken in the Carl Zeiss trinocular microscope and stereomicroscope. All results are presented as Mean ± SD. Test of significance (t-test) was used to analyse the data collected.

RESULT & DISCUSSION

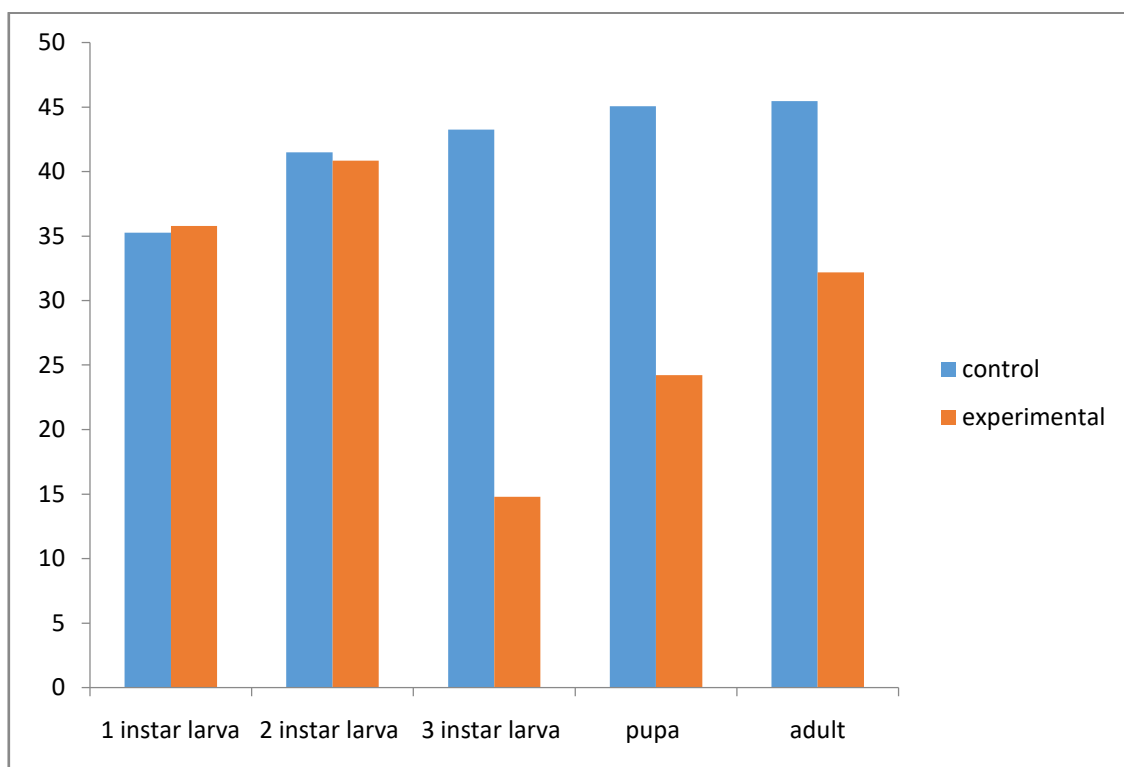
Table: Measurement of size (µm) & Protein (µg) from various larval stages of *Drosophila melanogaster*

		Size	Protein
1 st instar larva	Control	60.45 ± 8.20	35.256 ± 1.57
	Experimental	NS 68.5±12.70 (13.31%)	NS 35.776 ± 981.27 (1.474%)
2 nd instar larva	Control	86.7 ± 13.49	41.488 ± 0.98
	Experimental	NS 148±49.39 (68.75%)	NS 40.85±1.07 (1.537%)
3 rd instar larva	Control	175.6 ± 10.65	43.216 ± 0.50
	Experimental	NS 166.6±7.68 (5.12%)	NS 14.79±0.90 (65.77%)
Pupa	Control	189.5 ± 31.13	45.054 ± 0.26
	Experimental	NS 144.8±12.33 (23.58%)	NS 24.28±2.96 (46.27%)
Adult	Control	186.5 ± 15.64	45.458 ± 0.69
	Experimental	NS 162.5±8.79 (12.86)	NS 32.188±.098 (29.19%)

Mean ± S.D. of five flies P<0.05(*), P<0.01(**) and P<0.001(***) and NS= not significant. The value mention in parenthesis is the % change over the control.



Graphical representation of size



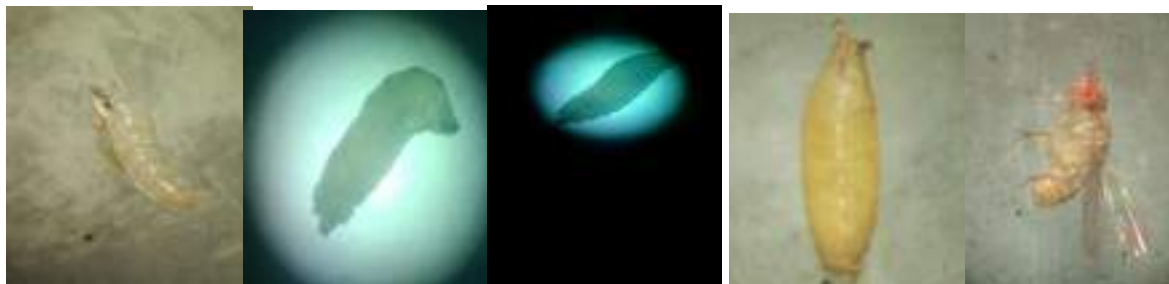
Graphical representation of concentration protein.

Control



1st instar larva 2nd instar larva 3rd instar larva Pupa Adult

Experimental



1st instar larva 2nd instar larva 3rd instar larva Pupa Adult

In the present study the Ist and IInd instar larvae increasing the protein concentration but IIIrd instar larvae, Pupa and Adult the amount of protein is decrease due to extract of *Ocimum sanctum*. Extract leaves of *Ocimum sanctum* have been shown to possess hypoglycemic effects in experimental animals (Sethi et al, 2004). The phenotypic character in different larval stages of *Drosophila melanogaster* is change due to effect of *Ocimum sanctum* extract. The variation in size and transparency in colour observed in the present study.

In vivo cytogenetic assay (Babu K. & Uma MaheswariK. C., 2006) in *Allium cepa* root tip cells has been carried out to detect the modifying effect of *O. sanctum* L. aqueous leaf extract against chromium (Cr) and mercury (Hg)-induced genotoxicity. It was observed that the roots post-treated with the leaf extract showed highly significant recovery in mitotic index (MI) and chromosomal aberrations. When compared to pretreated (Cr/Hg) samples, the lower doses of the leaf extract were found to be more effective than the higher doses. *Ocimum sanctum* L. extract (10 mg/kg body wt., PO) before and after mercury (HgCl₂) intoxication (5 mg/kg body wt.) showed a significant decrease in lipid peroxidation.(Sharma M. K. et. al., 2002) Serum glutamate pyruvate transaminase (SGPT) activities compared to HgCl₂ - induced values suggests that *O. sanctum* L. extract provides protection against HgCl₂ -induced toxicity in mice.

In the study of antimicrobial activity on extract of *Ocimum sanctum* based on the growth inhibition zone diameter obtained by 20 and 40 mg/ml *Ocimum* methanol extract concentration, bacterial strains were tested. This methods allows better diffusion of the extracts into the medium thus enhancing contact with organism. The antimicrobial activity of extracts of *Ocimum sanctum* was used against four pathogenic organism, Escherichia coil, Staphylococcus aureus , *Aeromonashydrophila* and *Eterococcusfaecalis*.

Benzene extract of *O. sanctum* L. leaves have a reversible antifertility (Ahmed M. et. al., 2002) effect, as *O. sanctum* L. extract (250 mg/kg body weight) for 48 days decreases the total sperm count, sperm motility and forward velocity. The percentage of abnormal sperm increased in caudal epididymal fluid and the fructose content decreased in the caudal plasma of the epididymis and the seminal vesicles. All these parameters returned to normal two week after the withdrawal of the treatment.

REFERENCES

- [1] Ahmed M., Ahamed RN, AladakattiRH, Ghosesawar MG. Reversible anti-fertility effect of benzene extract of *Ocimum sanctum* leaves on sperm parameters and fructose content in rats. *J Basic ClinPhysiolPharmacol* 2002; vol. 13:51-9.
- [2] Babu K. & Uma MaheswariK. C., In vivo studies on the effect of *Ocimum sanctum* L. leaf extract in mordifying the genotoxicity induced by chromium and mercury in *Allium* root meristems. *J Environ Biol*2006; vol. 27:93-5.
- [3] Bonner, J.M. and Boulianne, G.L. 'Drosophila as a model to study age-related neurodegenerative disorders: Alzheimer's disease ', *Experimental Gerontology*, 2011; vol. 46 (5), pp. 335-339.
- [4] Chattopadhyay R. R. &BandyopadhyayM. 1993 Effect of *Azadirachtaindica* leaf extract on serum lipid profile changes in normal and streptozotocin induced diabetic rats, *African Journal of Biomedical Research*, 2005; vol. 8; 101 – 104.
- [5] Gistelinck, M., Lambert, J., Callaerts, P., Dermaut, B. and Dourlen, P. 'Drosophila Models of Tauopathies: What Have we learned? ', *International Journal of Alzheimer's disease*, 2012; pp. 1-14.
- [6] Prakash P. &Gupta. Therapeutic uses of *Ocimum sanctum* Linn (Tulsi) with a note on eugenol and its pharmacological actions: a short review, *Indian J. PhysiolPharmacol*2005; vol. 49(2): 125-31.
- [7] Roch A.B.D.,R. M. Lopes, and G. Schwartzmann, "Natural products in anticancer therapy," *Current Opinion in Pharmacology*, 2001; vol. 1, pp. 364–369.
- [8] Sharma M. K., Kumar M, Kumar A. *Ocimum sanctum* aqueous leaf extract provides protection against mercury induced toxicity in Swiss albino mice. *Indian J ExpBiol* 2002; vol. 40:1079- 82.



- [9] SethiJyoti, ShushmaSood, Shashi Seth, AnjanaTalwar, Evaluation of hypoglycemic and antioxidant effect ofOcimum sanctum, Indian J Cli Biochem., 2004; vol. 19(2): 152–155.
- [10] Siva, M., Shanmugam, K. R., Shanmugam, B., Subbaiah, V., Ravi, S., Reddy, S., &Mallikarjuna, K.. Ocimum sanctum: a review on the pharmacological properties. *International Journal of Basic & Clinical Pharmacology*, 2016; vol. 5. pp. 558-565
- [11] Yakob K. H, ManafUyub A, FarizaSulaiman S. 2012 Toxicological evaluation of 80% methanol extract of Ludwigiaoctovalvis (Jacq.) P.H. Raven leaves (Onagraceae) in BALB/c mice. *J Ethnopharmacol.* 2012;vol.142 (3):663-8.

Fourier transform infrared spectroscopy analysis of *Sauromatum venosum* (Ait.) Schott. Tuber extract for alkaloids

M U Ghurde¹, S N Malode^{2*}

¹ Department of Botany, Vidya Bharati Mahavidyalaya, Camp, Amravati, Maharashtra, India

² Department of Botany, Govt. Vidarbha Institute of Science & Humanities, Amravati, Maharashtra, India

Abstract

Alkaloids are one of the largest groups of plant secondary metabolites; represent a highly diverse group of chemical entities. Plants are considered pharmacologically active when they contain alkaloids which could be of benefit to health. The FTIR spectroscopic analysis of *Sauromatum venosum* (Ait.) Schott. tuber extract for alkaloids revealed the presence of different characteristic peak values of alkaloids with various functional groups in the extracts. The FTIR analysis of *Sauromatum venosum* (Ait.) Schott. tuber extract confirmed the presence of Ar-OH stretching, -CH₃ stretching, Ar-H stretching, C-O stretching and C-O-C stretching which showed the major peaks for alkaloids. The FTIR, which was used to detect the characteristic peak values and their functional groups in tuber extract which may be used to prepare new drugs for the benefit of human health in future.

Keywords: *Sauromatum venosum*, tuber extract, secondary metabolites, alkaloids, FTIR

Introduction

Alkaloids are heterogeneous group compounds which contain one or more nitrogen atoms in acyclic system. These may contain one of the functional groups viz. carboxylic acid, alcohol, ether, esters amide, amine, alkene, aromatic ring and alkyl or arylhalide. The plants are considered as pharmacologically active contain polyphenolic compound such as alkaloids which exerts a biochemical and /or physiological effect on the cell (Ajibesin, 2005) [1], which could be of benefit to health. Moreover, many alkaloids revealed the significant bioactivities, such as the relieving action of ephedrine for asthma, the analgesic action of morphine, and the anticancer effects of vinblastine (Benyhe, 1994; Huang *et al.*, 2007; Wang, *et al.*, 2009; Lee, 2011; and Lu *et al.*, 2012) [4, 10, 24, 34, 46]. In fact, alkaloids are among the most important active constituents in natural herbs and some of these compounds have already been successfully developed into chemotherapeutic drugs, such as camptothecin (CPT), a famous topoisomerase I (Top I) inhibitor (Haug *et al.*, 2007) [16] and vinblastine, which interacts with tubulin (Li *et al.*, 2007) [15]. Alkaloids are imperative chemical compounds that give a rich reservoir for drug discovery.

Sauromatum venosum (Ait.) Schott. syn. *S. gutatum* (Wall.) Schott. (Araceae), is locally known as 'Saap ki Booti', 'Voodoo lily' or 'Snake Plant.' It is a shade loving perennial herb and found in Melghat region of Amravati District. Tuber/Corm is a condensed form of rhizome consists of solid, stout, fleshy underground stem. It contains heavy deposits of food material. Tuber size is upto 13 cm or more broad, globose and depressed. It is reported to contain the constituents like Lectins, Dimethyl sulphides, p-caryophyllene, indole, ammonia, trimethylamine, primary amines (Shinwari and Khan, 2000) [22]. Plant is reputed to be capable of neutralizing the action of snake venom and their antivenom activity has been related to certain chemical

compounds identified in the plants (Pereira *et al.*, 1994; Jain *et al.*, 2008) [19, 11]. The plant is used as antidote and antitumor (Choudhary *et al.*, 2008) [9].



Fig 1: A) *Sauromatum venosum* Plant B) Tuber/Corm

FT-IR spectroscopy is capable of providing strong insight into the structural and functional alterations induced by various factors due to its high sensitivity, the technique was successfully explored for the study of biological materials. IR spectroscopy has become a widely accepted tool for the characterization of biomolecules present in the sample (Margarita and Quinteiro, 2000) [17]. Fourier Transform Infrared Spectroscopy (FTIR) is the most powerful tool to identify the types of chemical bonds (functional groups) present in the compounds. FTIR spectroscopy is one of the most widely used methods to identify the chemical constituents and elucidate the compound structures to propose in medicinal purposes [Marimuthu and Gurumurthi (2013) [14], Ashokkumar and Ramaswamy, (2014) [1]. *Sauromatum venosum* are short lived found in rainy season's remains ethno-medicinally undocumented and have not been explored with respects to phyto-constituents and bioactive constituents. Hence, the present study was planned to investigate the alkaloids in *Sauromatum venosum* corm extract by FTIR Spectroscopy.

Materials and Methods

Selection and Collection of Plants: *Sauromatum venosum* tuber was selected on the basis of its medicinal value and collected from Semadoh region of Melghat of Amravati District in the month of June to September.

Identification and Authentication of plants: The plant was further identified with the help of standard floras (Sharma *et al.*, 1996; Almeida, 2001; Dhore, 2002) [2], [2, 6] and authenticated by Dr. S.P. Ro the Professor and Head, Department of Botany, Shri. Shivaji Science College, Akola.

Preparation of extract: The *Sauromatum venosum* tuber was cut into fine pieces and kept for shade-dried (at 25°C) and was powdered in the mechanical grinder. The extraction was done by using 20gm of tuber powder in 150ml of solvent by Soxhlet extraction method (Sadasivam and Manickam, 1996) [20] using different solvents such as distilled water, acetone, ethanol, petroleum ether and chloroform. The presence of alkaloids in the sample was detected by performing the qualitative tests (Khandelwal, 2007). The crude quantification of alkaloid was carried out by using precipitation method (Krishnaiah *et al.*, 2009) [13].

Fourier Transform Infrared Spectrophotometer analysis (FTIR): The Fourier-Transform Infrared Spectroscopy (FTIR) analysis of the crude extract of *Sauromatum* tuber was carried out from Sophisticated Analytical Instrumentation Facilities, Central Drug Research Institute Lucknow, India. The dried 10mg of powder was mixed with

KBr salt and encapsulated in 100 mg of KBr pellet in order to prepare translucent sample discs. The infrared spectrum of solid was recorded in the scan range from 4000-400 cm^{-1} on a FTIR spectrophotometer. Samples were introduced in sample chamber and spectra were taken in ATR mode. These spectra were recorded as absorbance values at each data point in replicate two times. Perkin Elmer spectrophotometer was used to detect characteristic peak and their functional group. Results obtained were plotted against wave number verses percent transmission.

Results and Discussion

The results revealed that the presence of various phytochemicals in crude extract *Sauromatum venosum* tuber. The FTIR spectrum was used to identify and detect the characteristic peaks and functional groups of the active components based on the peak value in the region of infrared radiation (Gupta *et al.*, 1977) [9] (Table 1; Fig. 1). Extract of *Sauromatum venosum* tuber was subjected to FTIR analysis and the functional groups of the components were separated based on their peaks. The results exhibited the presence of characteristic bands at 3019.79 cm^{-1} indicating the presence of Ar-OH stretching. Strong peak at 2925.88 cm^{-1} show -CH₃ stretching, at 2855.32 assign to C-H stretching shows some alkane group present in this rare medicinal plant. Besides Aromatic compounds with Ar-H stretching, at 1215.76 shows C-O stretching indicates presence to ketones, aldehydes, carboxylic acids, and esters bands and sensitive to peptide linkage. Peak at 1113.47 shows C-O-C strong stretching which can from two coordinate compounds (Fig. 1).



Fig 1: IR Spectra analysis of *Sauromatum venosum* corn for alkaloids

Table 1: IR spectral analysis of *Sauromatum venosum* (Al.) Schott. tuber for alkaloids

Sr. No.	Absorption Observed (cm^{-1})	Assignment	Absorption Expected (cm^{-1})
1.	3019.79	Ar-OH stretching	3500-3000
2.	2925.88	-CH ₃ stretching	3000-2800
3.	2855.32	Ar-H stretching	3000-2500
4.	1215.76	C-O stretching	1375-1200
5.	1113.47	C-O-C stretching	1150-1070

Fourier Transform Infrared Spectrophotometer (FTIR) is perhaps the most influential tool for identifying the types of chemical bonds (functional groups) present in compounds. The wavelength of light absorbed by the molecule is characteristic of the chemical bond as can be seen in the annotated spectrum. By interpreting the infrared absorption spectrum, the chemical bonds present in the molecule can be

determined. FT-IR spectroscopy is capable of providing strong insights into the structural and functional alterations induced by the various factors due to its high sensitivity (Toyran *et al.*, 2005; Dogan *et al.*, 2007) [23, 7]. FT-IR technique was used for assessment the type of organic and inorganic complexes in plants. The qualitative phytochemical analysis of aqueous extract of *Sauromatum venosum* tuber detected the presence of carbohydrate and glycosides, protein and amino acids, alkaloids, phenolic compounds & flavonoids, phytoosterols, saponins and terpenoids (Ghurde and Malode, 2018) [9]. The present study, the FTIR spectroscopic analysis showed the presences of phytoconstituents in the corn extract. Alkaloids are used as efficacious remedies in low concentrations serves as cardiotonics, potential oncological medicines, etc. Nature offers resources that include chemical compounds which can potentially solve many of

these problems. Investigation of natural products obtained from plants, the isolation of compounds and their modification, and the evaluation of their biological activities, represents an important field of biological and pharmaceutical research.

Conclusion

The Present study concluded that the *Sauromatum venosum* tuber has rich source of phytoconstituents like alkaloids, flavonoids, saponins, phenols and tannins. FTIR spectra showed the presence characteristic functional group of compound showing the strong peak values C-O stretching indicates presence to ketones, aldehydes, carboxylic acids, and esters bands and sensitive to peptide linkage as well as aromatic and peptide linkage with C-O-C indicate complex compounds used as a cardio tonic with their respective functional groups. Further study is needed with this corn extract to isolate, characterize and elucidate the structure of bioactive compound of this corn extract for industrial drug formulation in future.

Acknowledgement

Authors are thankful to Sophisticated Analytical Instrumentation Facility (SAIF) CDRI, Lucknow for helping to carry out FTIR analysis.

References

1. Ajibesin KK. Doctoral thesis: Chemical and antimicrobial study of the extracts and constituents of selected medicinal plant of Akwa Ibom State, Nigeria. Department of Pharmacology and Traditional Medicine, University of Uyo, Nigeria, 2005.
2. Almeida MR. The Flora of Maharashtra, (B), Orient Press, Shreeji Enterprises, Mumbai, 2001:3:97-99.
3. Ashok kumar A, Ramaswamy M. Phytochemical screening by FTIR spectroscopic analysis of leaf extracts of selected Indian Medicinal plants. Int. J. Cur. Micobi. Ap. Sci, 2014:3(1):395-406.
4. Benyhe S. Morphine: new aspects in the study of an ancient compound. Life Sciences, 1994:55(13):969-979.
5. Choudhary K, Singh M, Pillai U. Ethnobotanical survey of Rajasthan - An Update American-Eurasian Journal of Botany, 2008:1(2):38-45.
6. Dhore MA. Flora of Amravati District with special reference to the distribution of Tree species, 2002.
7. Dogan A, Ergen K, Budak F, Severcan F. Evaluation of disseminated candidiasis on an experimental animal model: A Fourier transform infrared study. Appl. Spectrosc, 2007:61:199-203.
8. Ghurde MU, Malode SN. Phytochemical screening and Thin layer chromatography of crude extract of *Sauromatum venosum* (Ait.)Schott. tuber. Proceedings of Multidisciplinary International Conference on Green Earth: A Paranomic View. ISBN:978-81-923628-5-4.
9. Gupta OP, Ali MM, Ray Ghatak BJ, Atal CK. Some pharmacological investigations of emblina and its semisynthetic derivatives. Indian J Physiol Pharmacol, 1977, 21(1).
10. Huang Min, Gao Heyong, Chen Yi, Zhu Hong, Chi Yujun, Xiongwen Zehong Zhang, Hualiang Jiang Miao, et al. Chimmitecan, a Novel 9-substituted Camptothecin, with improved anticancer pharmacologic profiles. In Vitro and In Vivo Clinical Cancer Research, 2007:13(4):1298-1307.
11. Jain Anita, Katewa SS, Sharma SS, Galav Pravin, Jain Vartika. Some therapeutic uses of biodiversity among the tribals of Rajasthan. Indian Journal of Traditional Knowledge, 2008:7(2):256-262.
12. Khandelwal KR. Practical Pharmacognosy, Techniques and Experiments. Nirali Prakashan, 2007:18:151-153.
13. Krishnaiah D, Devi T, Bano A, Sarbatly R. Studies on phytochemical constituents of six Malaysian medicinal plants. J. Medicinal Pl Research, 2009:3(2):67-72.
14. Lee MR. The history of *Ephedra* (ma-huang). J R Coll Physicians of Edinburgh, 2011:41(1):78-84.
15. Li WY, Shao HuL, Zhang X, Chen Y, Tong L. "BM6, a new semi-synthetic *Vinca* alkaloid, exhibits its potent *in vivo* anti-tumor activities via its high binding affinity for tubulin and improved pharmacokinetic profiles. Cancer Biology and Therapy, 2007:6(5):787-794.
16. Lu JJ, Bao JL, Chen XPM, Haung YT, Wang YT. Alkaloids isolated from natural herbs as the anticancer agents. Evid Based Complement Alternat Med, 2012:2012:485042.
17. Margarita P, Quintero R. Fourier Transform Infrared (FTIR) Technology for the identification of Organisms. Clinical Microbiology Newsletter, 2000, 8(22).
18. Marimuthu M, Gurumurthi P. Phytochemical screening and FTIR studies on wild and common south Indian legumens. Asian J Pharm Clin Res, 2013, 6(2).
19. Pereira NA, Ruppelt Pereira BM, Nascimento MC, Parente JP, Mors MB. Pharmacological screening of plants recommended by folk medicine as snake venom antidotes. IV: Protection against jararaca venom by isolated constituents. Plant Medica, 1994:60:90-100.
20. Sadasivam S, Manickam A. Biochemical Methods. New Age International Publishers, New Delhi, 1996:2:154-155.
21. Sharma BD, Kartikeyan S, Singh NP. Flora of Maharashtra State Monocotyledones, Botanical Survey of India, 1996, 220-221.
22. Shinwari MI, Khan MA. Folk use of medicinal herbs of Margala Hills National Park, Islamabad J Ethnopharmacol, 2000:69:45-56.
23. Toyran N, Zorlu F, Severcan F. Effect of stereotactic radiosurgery on lipids and proteins of normal and hypo perfused rat brain homogenates: a Fourier transform infrared spectroscopy study. Int. J. Radiat. Biol, 2005:81:911-918.
24. Wang ZT, Liang GY, Zhong, Yao Hua Xue. Shanghai Scientific & Technical, 2009.

ISSN 2319 - 359X
AN INTERNATIONAL MULTIDISCIPLINARY
HALF YEARLY RESEARCH JOURNAL

IDEAL

Volume - X

Issue - II

March - August - 2022

English / Marathi / Hindi Part - I

Peer Reviewed Refereed
and UGC Listed Journal No. 47026



IMPACT FACTOR / INDEXING
2019 - 6.601
www.sjifactor.com

❖ EDITOR ❖

Assit. Prof. Vinay Shankarrao Hatole
M.Sc (Math's), M.B.A. (Mkt), M B A (H R),
M.Drama (Acting), M.Drama (Prod & Dirl), M.Ed.

❖ PUBLISHED BY ❖



Ajanta Prakashan
Aurangabad. (M.S.)

The information and views expressed and the research content published in this journal, the sole responsibility lies entirely with the author(s) and does not reflect the official opinion of the Editorial Board, Advisory Committee and the Editor in Chief of the Journal "IDEAL". Owner, printer & publisher Vinay S. Hatole has printed this journal at Ajanta Computer and Printers, Jaisingpura, University Gate, Aurangabad, also Published the same at Aurangabad

Printed by

Ajanta Computer, Near University Gate, Jaisingpura, Aurangabad. (M.S.)

Printed by

Ajanta Computer, Near University Gate, Jaisingpura, Aurangabad. (M.S.)

Cell No : 9579260877, 9822620877 Ph. No. : (0240) 2400877

E-mail : ajanta6060@gmail.com, www.ajantapublishing.com

IDEAL - ISSN 2319 - 359X - Impact Factor - 6.601 (www.sjfactor.com)



EDITORIAL BOARD



Mehryar Adibpour
Faculty of Computing London
Metropolitan University
Holloway Road London

Dr. Ashraf Feroz Fata
College of Art's and Science,
Salman Bin Abdul Aziz University, KAS

Dr. Altaf Husain Paudi
Dept. of Chemistry University
of Kashmir, Kashmir, India.

Dr. Ramdas S. Wanare
Associate Professor & Head Accounts & Applied Stat
Vivekanand Art's Sandar Dalip Sing Commerce
& Science College Samant Nagar, Aurangabad (M.S.)

Dr. Prashant M. Dolia
Dept. of Computer Science and Applications,
Bhavnaragar University, India

Dr. P. A. Koli
Professor & Head (Retd)
Dept. of Economics, Shivaji University,
Kolhapur - (M.S.) India

Dr. Rana Pratap Singh
Professor & Dean School for Environment Science,
Dr. Babasaheb Bhimrao Ambedkar University
of Raebareilly, Lucknow- India

Dr. Joyanta Barborra
Head Dept. of Sociology University of
Dibrugarh- India

Dr. Jagdish K. Baheti
H.O.D. SNUB College of Pharmacy,
Nerthangar, Chauthvat, Nashik (M.S.) - India

Prof. P. N. Gajjar
Head, Dept. of Physics,
University of School of Sciences,
Gujarat University, Ahmedabad- India.

Dr. Memon Ubed Mohd Yusuf
Asst. Prof. Dept. of Commerce,
Sir Sayyed College Aurangabad (M.S.) - India



PUBLISHED BY



Ajanta Prakashan

Aurangabad. (M.S.)




CONTENTS OF ENGLISH PART - I


S.No.	Title & Author	Page No.
1	The Role of Exercise in Treating Obesity Dr. Madan B. Ingle	1-4
2	Effect of Yogasanas and Pranayam on the Physiological Variables of Senior Citizen during Covid-19 Pandemic Lockdown Dr. Pradeep H. Shende	5-8
3	A Study on Awareness of Soft Skills in Teacher Educators and Student Teachers M.U.K. Sudha Rani	9-14
4	A Study on the Effect of Reasoning Strategies on the Academic Achievement of IX Standard Students in Biology T. Lingaiah	15-21
5	Developing Creativity as an Employability Skill through ODL Dr. Rajendra Vadnere Dr. Anuradha Bhosale	22-29
6	Siderophore Producing Rhizobacteria Retarding <i>Alearophthora phoscolina</i> Growth S. G. Sangle U. P. Mogle	30-36
7	Basics of Open Access, Open Access Resources and Open-Source Tools Saungramsinh S. Pawar	37-41
8	Gopal Krishna Gokhale: Thoughts on Education Dr. Santosh Abhimanyu Jethithor	42-47
9	The Core Principles of Management for Business Management Dr. S. K. Rodde	48-51
10	Health, Welfare Provisions & Wages to Contract Employees - An Overview N. Rajashekar Rao	52-58

 **CONTENTS OF MARATHI PART - I** 

अ. क्र.	लेख आणि लेखकाचे नाव	पृष्ठ क्र.
१	छत्रपती संभाजीराजे यांच्या विचारनाल 'राज्यकर्ता' कुबेर भगत पांडुरंग	१-३



CONTENTS OF HINDI PART - I



अ. क्र.	लेख आणि लेखकाचे नाव	पृष्ठ क्र.
१	महिला सर्वांकुष और यास्तयता संदीप महादेवराव हाडोळे	१-५

9. The Core Principles of Management for Business Management

Dr. S. K. Rodde

Assistant Professor, Department of Commerce, Vidya Bharati Mahavidyalaya, Anuvati.

Abstract

Management as a subject of administering the resources is a modern era development but before the scientific management has emerged there was a technical management running from the time to time beneath the surface which made the nomadic life of human beings to revolutionised to live a life where they are free from any kind of threat.

The core principle of management lies managing self at the individual level towards ourselves at the higher level to ensure the personnel management for any business management to be successful at the larger plane.

Modern 21st century is a technological era where the ability to manage things are being done through technological management even though the nature of human effort has changed from physical to mind management the need for fundamental principles of core values of management are still a high valued principle.

Introduction

The system of management in the making of a successful business fundamentally lies in the hands of manager whose efficiency and ability to look or to forecast the needs of future makes the business to last for long to make at most profits with a fair service through the mechanism of principles of technical management.

The prospects of managerial methods to make some successful principles came out specially from the problems of working together for a same objective. through the system of training for specialisation, work division, and job timing at first laid a fundamental difference in output of production with the maximum satisfaction from the employees.

For to achieve the targeted results in the business and to function with the minimal efforts led a major difference in the strategy to push the targets in faster pace to get the maximum benefits from the business. With the growth in the industries and from the change towards

modern mechanical establishments the change in the behaviour led further division in the system of running business like enterprise.

Management and its 5M's

- 1 Men
- 2 Material
- 3 Money
- 4 Methods
- 5 Machinery

The above five terms are the much-needed inputs for to escalate any business to hold up in the market for running in profit orientation. The time of industrialisation and its results in the market economy laid the first foots steps of business to work for pro market needs and in those times of acquiring wealth to conquer power of one country over other led the human force to steep rise of hard working labour to stand as a monopoly not only in the market but in the reign of development at the cost of hard labourers work.

Systematic management and deeper exploitation of resources at the early stages of industrial world led the rise of transforming the traditional ways of living towards a fragmented way of life survival for to acquire the wealth to stand in the crowd of learned as a mere spectator of fast pace life with the work culture of employee and employers.

From the dispersed ways of exploitation of resources, the researches observations to make management studies for effective utilisation of resources led to aspire for studies of employees to make necessary changes in the industries to manage and make some principles for managing the industries. The core level management studies and principles laid out by **Henry Fayol's 14 principles of Management.**

The Core Principles of Management

According to Henry Fayol the principles of management are the true guiding factors for to manage any organisation or body of management

- 1 Division of work
- 2 Authority and Responsibility
- 3 Discipline
- 4 Unity of command
- 5 Unity of Direction

6. Subordination of individual
7. Remuneration
8. The degree of Centralization
9. Line of Authority/Scalar Chain
10. Order
11. Equity
12. Stability of Tenure of Personnel
13. Initiative
14. Esprit de corps Team spirit

All the above principles are the fundamental/core values for any organisation to grow strong for to achieve the targets set out for the management to do efficient business. To make any organisation the need for to earn profits with the decrease in the human efforts is the foremost value given by management scientists like Henry Fayol.

Modern Era Management

Modern management era is a technological era where the human hard work is being replaced by mind/smart work through the technological advancement, the mind management and its behaviour towards the work must be done with the at most care because of a slightest mistake will lead to collapse of the roots of managements

From the core principles of management for the new era managerial skills the most important one is **Discipline** management where by to make a disciplined mindset for rightful work management

Business Management in Globalised Era

The challenging era where the new fundamentals of business is drawing prominence the very transition of industrial labour to new minds of changing the society is a welcoming move from the management sciences and its efforts to improve human force and the nature of work

Management in the 21st century is less management and more like behavioural sciences for to ensure the smooth functioning of business and of personal life to fore see the greater possibility of human aspiration for to achieve the set targets for escalating life with the help of managerial sciences of management even in the business management

The true nature of managerial skills and abilities lies within the thoughts of improving the ways and means of working to improve the environment around which makes the employees to

output their production not only for profits but for the growth in the human effort for to reach a goal with collective work culture in management.

References

1. Fayol, H., **General and industrial Management**, London, Pitman, 1949, p. 6
2. Terry, G. R., **Principles of Management**, Illinois, Richard D. Irwin, Inc. 1956, p. 18
3. **Theories of Organization-[A] Henry Fayol's Contribution**
4. **Management Values of Management**
5. **Management Tests of Good Management.**
6. Hicks, Herbert, G., **Management of Organisations.**

ISSN 2319 - 359X
AN INTERNATIONAL MULTIDISCIPLINARY
HALF YEARLY RESEARCH JOURNAL

IDEAL

Volume - X

Issue - II

March - August - 2022

English Part - I / II

Peer Reviewed Refereed
and UGC Listed Journal No. 47026



IMPACT FACTOR / INDEXING
2020 - 6.008
www.sjifactor.com

❖ EDITOR ❖

Assit. Prof. Vinay Shankarrao Hatole
M.Sc (Math's), M.B.A. (Mkt), M B A (H R),
M.Drama (Acting), M.Drama (Prod & Dirl), M.Ed.

❖ PUBLISHED BY ❖



Ajanta Prakashan
Aurangabad. (M.S.)

The information and views expressed and the research content published in this journal, the sole responsibility lies entirely with the author(s) and does not reflect the official opinion of the Editorial Board, Advisory Committee and the Editor in Chief of the Journal "IDEAL". Owner, printer & publisher Vinay S. Hatole has printed this journal at Ajanta Computer and Printers, Jaisingpura, University Gate, Aurangabad, also Published the same at Aurangabad

Printed by

Ajanta Computer, Near University Gate, Jaisingpura, Aurangabad. (M.S.)

Printed by

Ajanta Computer, Near University Gate, Jaisingpura, Aurangabad. (M.S.)

Cell No : 9579260877, 9822620877 Ph. No. : (0240) 2400877

E-mail : ajanta6060@gmail.com, www.ajantapublishing.com

IDEAL - ISSN 2319 - 359X - Impact Factor - 6.008 (www.sjifactor.com)



EDITORIAL BOARD



Mehryar Adibpour
Faculty of Computing London
Metropolitan University
Holloway Road London

Dr. Ashraf Feroz Fata
College of Art's and Science,
Salman Bin Abdul Aziz University, KAS

Dr. Altaf Husain Paudi
Dept. of Chemistry University
of Kashmir, Kashmir, India.

Dr. Ramdas S. Wanare
Associate Professor & Head Accounts & Applied Stat
Vivekanand Art's Sandar Dalip Sing Commerce
& Science College Samanth Nagar, Aurangabad (M.S.)

Dr. Prashant M. Dolia
Dept. of Computer Science and Applications,
Bhavnaragar University, India

Dr. P. A. Koli
Professor & Head (Retd)
Dept. of Economics, Shivaji University,
Kolhapur - (M.S.) India

Dr. Rana Pratap Singh
Professor & Dean School for Environment Science,
Dr. Babasaheb Bhimrao Ambedkar University
of Raebareilly, Lucknow- India

Dr. Joyanta Barborra
Head Dept. of Sociology University of
Dibrugarh- India

Dr. Jagdish K. Baheti
H.O.D. SNUB College of Pharmacy,
Nerthangar, Chauthvat, Nashik (M.S.) - India

Prof. P. N. Gajjar
Head, Dept. of Physics,
University of School of Sciences,
Gujarat University, Ahmedabad- India.

Dr. Memon Ubed Mohd Yusuf
Asst. Prof. Dept. of Commerce,
Sir Sayyed College Aurangabad (M.S.) - India



PUBLISHED BY



Ajanta Prakashan

Aurangabad. (M.S.)




CONTENTS OF ENGLISH PART - I


S. No.	Title & Author	Page No.
1	The Role of Exercise in Treating Obesity Dr. Madan B. Ingle	1-4
2	Effect of Yogasanas and Pranayam on the Physiological Variables of Senior Citizen during Covid-19 Pandemic Lockdown Dr. Pradeep H. Shende	5-8
3	A Study on Awareness of Soft Skills in Teacher Educators and Student Teachers M.U.K. Sudha Rani	9-14
4	A Study on the Effect of Reasoning Strategies on the Academic Achievement of IX Standard Students in Biology T. Lingaiiah	15-21
5	Developing Creativity as an Employability Skill through ODL. Dr. Rajendra Vadnere Dr. Anuradha Bhosale	22-29
6	Siderophore Producing Rhizobacteria Retarding <i>Macrophomina phaseolina</i> Growth S. G. Sangle U. P. Mogle	30-36
7	Basics of Open Access, Open Access Resources and Open-Source Tools Sangramsinh S. Pawar	37-41
8	Gopal Krishna Gokhale Thoughts on Education Dr. Santosh Abhimanyu Jethithor	42-47
9	The Core Principles of Management for Business Management Dr. S. K. Rolde	48-51
10	Health, Welfare Provisions & Wages to Contract Employees - An Overview N. Rajashekar Rao	52-58
11	Sustainability and Innovations in Agriculture and Rural Development in India Dr. Sanjay B. Shinde	59-66

 **CONTENTS OF ENGLISH PART - I** 

S. No.	Title & Author	Page No.
12	Indian Socio-Ethnic Sense and Sensibility in the Verses of Sarojini Naidu Prof. Dr. Vinodkumar Chaudhari	67-70
13	The Administration of the Chairperson on Chiang Mai Province in Thailand Mr. Worayuth Sathapanasuphakul Dr. Subhas Morale	71-76
14	Classroom Participation of Secondary School Students in Relation to their Perceived Parental Behaviour N. Rajashekar	77-84
15	Corporate Philanthropy - An Ancient Analogy in Building Modern India Dr. Varsha Gondaliya	85-91
16	Marginal People in Tarashankar's Fiction Dr. Subrata Kumar Mandal	92-98
17	Impact of Covid-19 on Farmers' Income of Marathwada Region in Maharashtra Dr. Sanjaykumar P. Phad	99-106
18	Microfinance: A Way Forward for Women Development Miss. Sarviinder Kaur	107-113
19	Family Environment and Income as an Influential Factors for Participation of Females in Sports at College Level in Gondwana Region of Maharashtra Dr. Subhash M. Shekhar	114-119
20	Self Respect among Self Help Group Women, Belagavi City Corporation, Belagavi Dr. Ravi S. Dalawai	120-125
21	The Stress among Doctors in Hospitals Dr. L. Karthikeyan Ambily Das K.	126-133

 **CONTENTS OF ENGLISH PART - I** 

S. No.	Title & Author	Page No.
22	Impact of Constitutional Measures for Development of Scheduled Castes and Scheduled Tribes in Hyderabad Karnataka-Overview Sadashiva Dr. P. B. Rathod	134-140
23	Intervention Programmes for the Wellbeing of Elderly - An Analysis Prof. Padala Prasad	141-148

 **CONTENTS OF ENGLISH PART - II** 

S. No.	Title & Author	Page No.
1	Tourism Marketing and Sales Promotion Strategies Dr. Pratik Bhaskarrao Upase	1-8

1. Tourism Marketing and Sales Promotion Strategies

Dr. Pratik Bhaskarrao Upase

Assistant Professor, Department of Commerce, Vidya Bharati Mahavidyalaya, Amravati.

Abstract

Marketing is a subject of vital concern in tourism because it is the principal management influence that can be brought to bear on the size and behavior of this major global market. The figure shows the relationship between market demand, generated in areas of origin and product supply, mainly at visitor destinations. Marketing strategy is a powerful that gives an organization a competitive advantage in the marketplace. While just defining a marketing strategy will not automatically create a competitive advantage, it will allow the organization to concentrate its resources on the greatest opportunities to increase sales and achieve a sustainable competitive advantage. Sales promotion contributes in a number of ways to achieving overall promotional objectives. While it can be used merely to gain attention for a service, it is more likely to be employed as an incentive incorporating an offer which represents value to the target audience. It can also act as an invitation to engage in a transaction now rather than later. Sales promotion usually attracts brand switchers but is unlikely to turn them into loyal brand users without the use of other elements of the promotion mix. Many people would argue that marketing is the antithesis of sustainable tourism and that achieving the latter must inevitably involve reducing the power of marketing of tourism. However the author believes that this view is both unrealistic and incorrect. Marketing is now a very powerful force that influence everyone in every aspect of their lives, and there seems little evidence that this situation will change in the short term at least. Perhaps in the longer term there may be a consumer backlash against marketing and consumerism, but until then we need to see if we can use marketing to help advance the cause of sustainable tourism.

Key word : Tourism Marketing, Management & Sales promotion techniques

Introduction

Marketing is a subject of vital concern in tourism because it is the principal management influence that can be brought to bear on the size and behavior of this major global market. The figure shows the relationship between market demand, generated in areas of origin and product supply, mainly at visitor destinations. In particular, the model shows how the main sectors of the tourism industry -travel organizers, destination organizations,

transportation, various product suppliers- combine to manage visitors' demands a range of marketing influences. The marketing mix in the centre of the diagram, and it is discussed in detail in this book. However, it is important to note that the influence of this marketing activity is likely to vary according to visitors' interests and circumstances. For example, domestic visitors travelling by car to stay with friends or relatives may not be influenced by destination marketing in any way, whereas first-time buyers of package tours to exotic by the marketing decisions of the tour operator they choose. In between these two examples a business traveler will select his or her own destination according to business requirements, but may be influenced as to which airline or hotel he or she selects. Knowledge of the customer, and all that it implies for management decisions, is generally referred to as consumer characteristics and buying behavior is central to the activity.

Marketing Management

Many people would argue that marketing is the antithesis of sustainable tourism and that achieving the latter must inevitably involve reducing the power of marketing of tourism. However, the author believes that this view is both unrealistic and incorrect. Marketing is now a very powerful force that influence everyone in every aspect of their lives, and there seems little evidence that this situation will change in the short term at least. Perhaps in the longer term there may be a consumer backlash against marketing and consumerism, but until then we need to see if we can use marketing to help advance the cause of sustainable tourism.

It is against this positive, modern view of tourism marketing techniques may be used to create more sustainable forms of tourism. The classic marketing techniques could be utilized to help achieve more sustainable forms of tourism. Firstly, this means seeking to understand our customers in terms of their motivations and determinants, reflecting the consumer-led concept of modern marketing. You cannot influence a tourist's behavior unless you understand how they think, what they are looking for and the factors which influence their purchase decisions. There is thus a need for sophisticated marketing research on consumer attitudes towards sustainable tourism. Then, the organization, be it a destination marketing agency or a tour operator or hotel, needs to scan its business environment for relevant data that might determine the stance it takes towards sustainable tourism. This scanning could involve considering,

- Government legislation on environmental issues
- The economic climate
- The level public concern over the social and environmental impacts of tourism
- The potential influence of technological innovations such as virtual reality

Sustainable Tourism Involves Both

- Developing products which are more sustainable in nature
- Moving away from offering products which are intrinsically not sustainable

The Former Category Might Included

- Conservation holidays
- Vacation packages using public transport rather than private cars
- Small-scale rural community-based tourism initiatives such as those outlined in the chapter on rural areas.

The Latter Principle could Cover the Following

- Hunting trips
- Holidays packages to countries with a poor human rights record
- Destination with poor environmental standards where inappropriate development is taking place
- Holidays on which tourists consume too many local resources which may be in short supply, such as water.

Marketing Strategy

Marketing strategy is a powerful that gives an organization a competitive advantage in the marketplace. While just defining a marketing strategy will not automatically create a competitive advantage, it will allow the organization to concentrate its resources on the greatest opportunities to increase sales and achieve a sustainable competitive advantage. The word strategy comes from the Greek word strategies, meaning general. Strategy is what generals use to win battles. Thus properly understood, marketing strategy is a high-level exercise involving the "generals" of the organization in determining how to build on the firm's strengths while taking advantage of competitors' weaknesses. Marketing strategy is most effective when it is a vital component of corporate strategy, defining how the organization will engage customers, prospects and the competition in the market arena for consistent success. A strategy is different from a tactic. While it is possible to write a tactical marketing plan without a sound, well-considered strategy, it is not recommended. Without a sound marketing strategy, a marketing plan has no foundation. Marketing strategies serve as the fundamental underpinning of marketing plans designed to reach marketing objectives. It is important that these objectives have measurable results. A good marketing strategy should integrate an organization's marketing goals, policies, and action sequences into a cohesive whole. The

objective of a marketing strategy is to provide a foundation from which a tactical plan is developed. This allows the organization to carry out its mission effectively and efficiently.

The Role of Sales Promotion

Sales promotion contributes in a number of ways to achieving overall promotional objectives. While it can be used merely to gain attention for a service, it is more likely to be employed as an incentive incorporating an offer which represents value to the target audience. It can also act as an invitation to engage in a transaction now rather than later. Sales promotion usually attracts brand switchers but is unlikely to turn them into loyal brand users without the use of other elements of the promotion mix.

Sales Promotion Planning

As in the case of advertising, effective sales promotion involves an on-going process with a number of stages

- **Establishment of objectives:** Sales-promotion objectives vary according to the target market. If the target is the customer, objectives could include the encouragement of increased usage or the building of trial among non-users. For intermediaries, objectives could be to encourage off-season sales or offsetting competitive promotion activity could also be aimed at internal personnel, making up part of the reward system.
- **Selection of promotional tools:** Promotional objective form the basis for selecting the most appropriate sales-promotion tools. The cost and effectiveness of each tool must be assessed with regard to achieving these objectives in respect of each target market. The tools available to the service marketer are described in more detail in the next section.
- **Planning the sales promotion programme:** the major decision that need to be made when designing the sales-promotion programme relates to the timing of the promotion and how long this tool is to be used. Also important are the size of incentive, rules for eligibility and, of course, the overall budget for the promotion.
- **Pre-testing:** this needs to be undertaken to ensure that potentially expensive problems are discovered before the full launch of a promotion. Testing in selected market segments can highlight problems of ambiguity, response rates and give an indication of cost effectiveness.
- **Implementation:** the programme for implementation must include two important time factors. First, it must indicate the 'lead time' –the time necessary to bring the programme up to the point where the incentive is made available to the public.

Second, the 'sell-in time' which is the period of time from the date of release to when approximately 90-95 percent of incentive material has been received by potential customers.

- Evaluation - the performance of the promotion needs to be assessed against the objectives set. If objectives are specific and quantifiable, measurement would seem to be easy. However, extraneous factors could account for the apparent success of many sales-promotion activities for example, competitive actions or seasonal variations may have influenced customers' decision making. It can also be extremely difficult to separate out the effects of sales-promotion activity from other promotional activity or indeed from other marketing.

Promotional Objectives

The general purposes of promotion in services marketing are much the same as in other kinds of marketing to:

1. Build awareness and interest in the service product and the service organization
2. Differentiate the service offer and the service organization from competitors.
3. Communicate and portray the benefits of the services available
4. Build and maintain the overall image and reputation of the service organization.
5. persuade customers to buy or use the service

Promoting Services

The promotional elements of services are the under

- a. Advertising
- b. Personal Selling
- c. publicity
- d. Sales promotion

These four types of promotion are the traditional marketer-dominated ways of influencing sales for services Products. Between them they offer the organization a wide range of possibilities for promoting services in marketplace. Absolutely used they can be most effective tools of communication and influence. However two important qualifications should be noted. First, service organizations promote themselves and their service products in other ways than the four forms considered here. Though this chapter deals with 'personal salesman' as part of 'promotion', in many services markets the 'selling' of services and the 'performance' of services cannot be separated. Indeed there may not be a separate sales force as such. So this separation of the 'selling role' may be unrealistic in practice. The comments about people in this chapter should be seen alongside the 'other service personal' who can influence

promotional effort. Also physical evidence will influence customer response to service products. Second, there is increasing research evidence that in some kinds of services markets these marketer dominated ways of influencing sales for service products may not necessarily be the most effective. For example, a study covering eight different types of retail services (dry cleaning, hairdressing, car loan finance, dentistry, photography, carpet shampoo, car repair, gynecological services) showed that customers of retail services do not rely on the more traditional sources. Instead, they prefer to use more personal sources of information from friends and other contacts. Also location is a key factor in the final purchase decision. The importance of personal recommendations has been confirmed in the life insurance market too. Customers in this market have been shown to rely heavily on the opinions and influences of others through word of mouth communication.

Sales-Promotion Tools

A wide and ever increasing range of sales promotion tools is employed by service organizations. Some of the more commonly used tools aimed at the final consumer include the following:

- **Free samples, visits/ consultations.** these encourage trial of a service and can be valuable where consumers are loyal to an existing service supplier. It could, for example, be used by a video film-rental chain to entice potential customers into their branch so that they can learn something about the nature of the service on offer. In the case of new services that are perceived as being expensive and of poor value to a consumer, they can encourage trial –this has been used by satellite television companies.
- **Money-off price incentives:** these are used to stimulate demand during slack periods where price is considered to be a key element in customers' purchased activity. Price incentive can be used tactically in competitor activity. they can also be employed to stimulate sales of a new service shortly after launch. Price incentives tend to be expensive to the service provider, as the incentives is given to customers regardless of its motivational effect on individual customers. A leisure park reducing its prices for all is unable to extract the full price from those customers who may have otherwise been willing to pay that price.
- **Coupons, Vouchers:** These allow holders to obtain a discount off a future purchases and can be targeted at quiet specific groups of users or potential users. To encourage trial by potential new users, vouchers can be distributed to non-users who fit a specified profile. In this way the operator of a plaza might arrange for vouchers to be

given to customers of a cooperating hairdressing salon. To encourage repeat usage, vouchers can be given as a loyalty bonus. Voucher offers tend to be much more cost effective than straight price incentives because of their ability to segment markets. In this way, a leisure park operator is able to recognize that a visitor to the park (coupons usually take one of three forms: a straight price cut; a discount or fee waiver for one or more patrons accompanying the original purchaser

- Gifts offers: These allow an organization to augment its service offer with an additional gift which can satisfy a number of objectives. In order to promote initial enquiry and to give tangible cues of the service company's offering and image, many firms—especially in the financial services sector—offer a gift for merely enquiring about their service. Items offered by insurance companies for merely requesting a quotation often include pocket calculators and pens. A gift can also be used to bring about immediate action—for example, a free clock radio if a policy is taken out within a specialized period.
- Competitions: The inclusion of a competition in a service offer adds to the value of the total offer. Instead of simply buying an insurance policy, customers buy the policy plus a dream of winning a prize to which they attach significance. Competitions can be used both to create trial among non-user and to retain loyalty among existing users.
- Short-term discount: Price cuts that are promoted as being available for only a limited time period—such as any form of "sale" designed to boost business during slack periods.
- Sign-up rebates: May be offered by "membership" service organizations that charge a preliminary sign-up fee for applying, joining or making connections to a network.
- Short-term increases in sales commission: These, together with sales bonuses, can be used to stimulate sales during slack periods or to develop loyalty from intermediaries in the face of competitor activity.
- Point-of-sale material: To stimulate additional sales, a service principal can provide a range of incentives to help intermediaries.
- Cooperative advertising: A service principal often agrees to subscribe to local advertising by an intermediary—often in conjunction with a significant to turn to the chapters on the different sectors of tourism for more detail on this subjects.

Conclusion

The present study has covered the tourism marketing, Service marketing Service marketing and mix, promotional mix, sales promotion and tourism. The theoretical background helps to provide the base for the further research so the researcher has covered the background of sales promotion and tourism. Moreover, it is stated that theoretical background helps to clarify the key concept of research and understand the influence factors, which affected on the sales promotion policies and tourism industry. Furthermore, sales promotion policies are more useful and profitable for the tourist as well as tourism industries.

References

1. B .Dalaja - Services Marketing And Management - S Chand & Company Ltd. New Delhi .
2. Jagdish Goswami - International Tourism - ABD Publications Jaipur-New Delhi
3. John Swarbrooke-Sustainable Tourism Management Rawal Publications New Delhi
4. Ravi Shanker - Services Marketing ,Excel Books New Delhi
5. Services Marketing - ICFAI Center for Management Reaserch, Hyderabad.
6. Dr. S. shajahan - Services Marketing, Himalaya Publishing House New Delhi Edition.
7. S.M. Jha- Services Marketing , Himalaya Publishing House New Delhi .
8. Dr. N RAJAN NAIR, SANJITH R NAIR, SULTAN CHAND & SONS - 23, Daryaganj, New Delhi
9. R.S.N. PILLAI & BAGAVATHI, Marketing , S. CHAND & COMPANY LTD, RAM NAGA New Delhi.
10. RAVI SHANKAR, Services Marketing , the Indian perspective, EXCEL BOOKS New Delhi.
11. Marketing Management - ICFAI Center for Management Research, Road # 3 - Banjara Hills, Hyderabad
12. Joel R. Evans , Barry Bernan Marketing Management , CENGAGE Learning, 2007
13. Biplab S. Bose, Marketing Management , Himalaya Publishing House , Mumbai.
14. Ms.Ahana Charaboty & Dr. B K CHAKRAVARTI - Global Tourism - A P.U PUBLISHING CORPORATION ,ANSARI ROAD,DRYA GANJ New Delhi - 110002
15. Anand Sahgal -Bsics of Tourism Marketing - Ishika Publishing House jaipur-New Delhi, 2018-



IJRASET

International Journal For Research in
Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 10 **Issue:** VII **Month of publication:** July 2022

DOI: <https://doi.org/10.22214/ijraset.2022.45604>

www.ijraset.com

Call:  08813907089

E-mail ID: ijraset@gmail.com



Analysis of IoT Data Transfer Messaging Protocols on Application Layer

Ms. Shilpa B.Sarvaiya¹, Dr. D.N. Satange², Prof. Ather Iqbal³

¹Department of Computer Science Vidyabharati Mahavidyalaya, Amravati,

²Department of Computer Science Narsamma Hirayya Arts Commerce & Science, Amravati,

³HOD Department of Computer Science Vidyabharati Mahavidyalaya, Amravati

Abstract: Now on a daily basis during a smarter embedded world, have Internet of Things. IoT have lot of things for the embedded systems, and it's the potential to remodel our world with the assistance of it. Internet of Things (IoT) or Web of Things (WoT) is emerging technology and it wireless network between two or more objects or smart things connect via Internet. IoT classified in two types first is within IoT and second side is outside of IoT.

In inside IoT consider as protocols in IoT. In outside of IoT consider as sensor, actuators, etc., those are physically possible. In within IoT consider as protocol and IoT have own Protocol Stack. Protocol Stack has different layer like Application Layer, Transport Layer, Internet Layer and Physical/Link Layer. The judgmental role goal of IoT is to confirm effectual communication between two objects and build a sustained bond among them using different application. The application Layer liable for providing services and determining a group of protocol for message passing at the application Layer. This paper understands Application Layer Data Transfer Messaging Protocol like MQTT, AMQT, COAP, XMPP, DDS, HTTP, RESTFULL, and WEB-SOCKET. Also describe which sort of architecture (like Request/Response, Client/Server and Publish/Subscribe) and security (like DTLS, TCL/SSL and HTTPS) support in those protocols that decide upon appropriate protocol supported application needs.

Keywords: IoT, WoT, M2H, M2M, MQTT, AMQT, COAP, XMPP, DDS, HTTP, RESTFULL, WEB-SOCKET.

I. INTRODUCTION

IoT has significantly changed ones perspective of living style. It has enabled many non-living objects to behave smartly and intellectually according to the circumstance and environment. A growing number of physical objects are being connected to the Internet at an unprecedented rate realizing the idea of the Internet of Things (IoT).

The IoT envisions hundreds or thousands of end-devices with sensing, actuating, processing, and communication capabilities able to be connected to the Internet [1]. The captured data needs a direction to be Transferred, Modified, Controlled, Acknowledged, Stored or Exported to other devices. These tasks can be performed through suitable protocols. Typically, IoT is expected to offer advanced connectivity of devices, systems, and services that goes beyond, Human-to-Machine (H2M), Machine-to-Machine (M2M) communications. In section 2 studies the literature review, in section 3 explain the working of the data transfer protocols between the transmitter and the receiver is given. Section 4 Analyse the result and discuss, section 5 gives the conclusion of the work, the comparisons of the messaging protocols is important to choose a suitable design platform.

1) *Human To Machine Communications In IoT (H2M):* Human to Machine communication originally emerged from Telemetry technology, and its main aim was to measure data and automatically transmit it from remote sources typically by cable or a radio. Nowadays, plethora of sensors are being developed, which have better perceptual abilities than humans and can detect information that humans cannot.

Affordable electronic devices halved to an increasing number of them being connected to the Internet. The smart IoT devices open up the possibility to reduce the burden on the user end by equipping everyday objects. Human-to-Machine communication is a very important development in Internet of Things. Figure 1 shows the human to machine interaction and its relation to the IoT. Figure 1 shows the elaborate human to machine interaction which demonstrates the use of sensors, actuators, cognition unit and processing unit.

The data transfer in HMI model is based on the cognitive ability of the human. The machine sends the control data to the actuator unit to perform the required action.

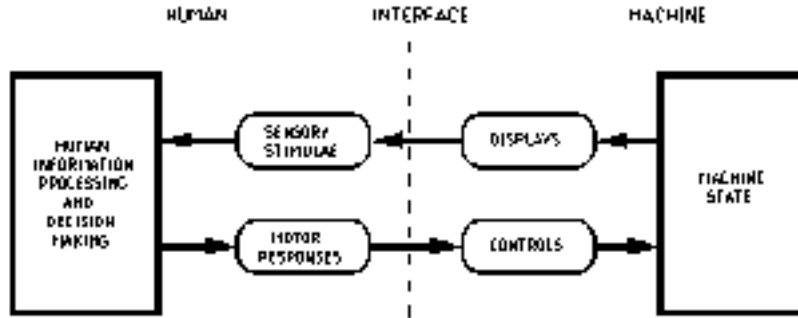


Figure 1: Human to Machine Interaction

2) *Machine To Machine Communications In Iot (M2m)*: Machine to Machine (M2M) is a broad term that can be used to describe any technology that enables networked devices to exchange information and perform actions without the manual assistance of humans. It forms the basis for a concept known as the Internet of Things (IoT). Key components of an M2M system include sensors, RFID, a Wi-Fi or cellular communications link and autonomic computing software programmed to help a networked device interpret data and make decisions. M2M communications expands telemetry's role beyond its common use in science and engineering and places it in an everyday setting. People already are using M2M, but there are many more potential applications as wireless sensors networks and computers improve, benefitting the concept to be amalgamated with other technology. The main goal of M2MC is to enable the sharing of information between electronic systems autonomously; Figure 2 shows the amalgamation of Machine to Machine and IoT in real time. In other words Machine to Machine refers to technologies that allow both wireless and wired systems to communicate with other devices. M2M application protocols take a fundamental role in communication efficiency: Protocol overheads, necessary number of management or control and information messages, reliability and security. All these impact the number and size of transmissions consequently, the energy and bandwidth consumptions in a mobile device. It finds its major application in protocols namely MQTT, AMQP, COAP, XMPP, DDS, HTTP, RESTFULL, and WEB-SOCKET.

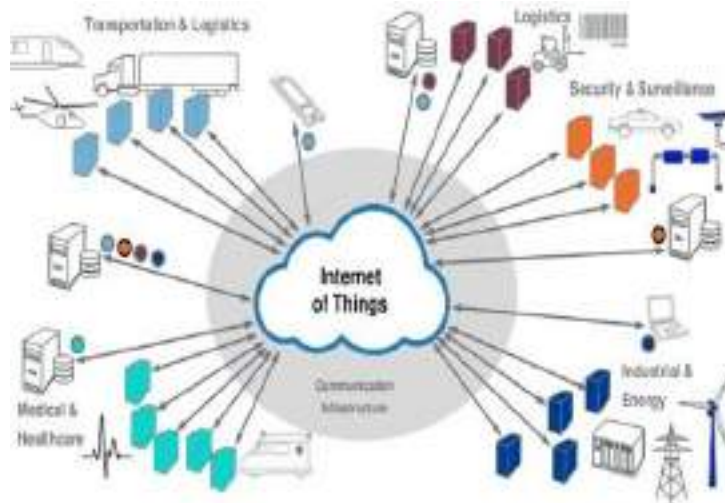


Figure 2: Amalgamation of M2M and IoT

II. RELATED WORK

Internet of Things (IoT) is emerging technology. Show in previous section inside of IoT describe as protocols. By studying paper regarding IoT and IoT protocol related IETF standards paper show Application Layer protocols focus basically on message exchange between applications and the Internet [2]. A Data Transfer Protocol is a standardised format for transmitting data between two devices. The type of protocol used can determine the variables. Application Layer work with other layer like Transport Layer, Internet Layer and Physical Layer. In this paper our aim to provide comprehensive survey to describe all main eight Application Layer Protocols.

A. IoT Ecosystem

Figure 3 shows a 7-Layer model of IoT Ecosystem. At the bottom Layer is the market or application domain. The second layer consists of sensors that enable the application. The third layer consists of interconnection layer that allows the data generated sensors to be communicated, usually to a computing facility, data centre, or a cloud. Finally the top layer consists of services that enable the market and may include energy management, health management, education, transportation etc. [3].

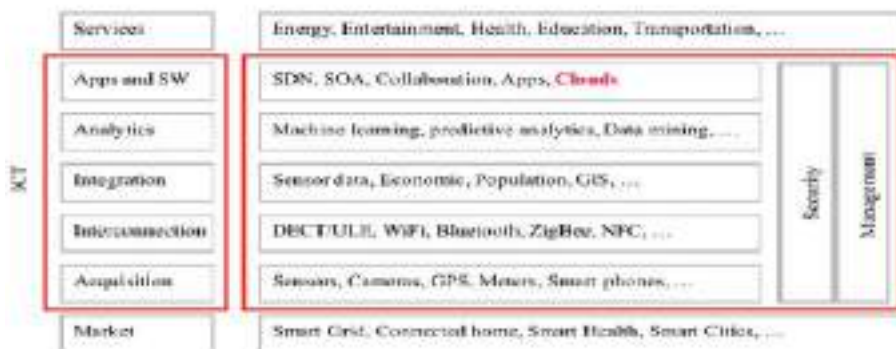


Figure 3: IoT Ecosystem

In this paper, we concentrate on the Session Layer. This layer itself can be shown in a Multi-Layer stack as shown in figure 4. We have shown only the Data link, Network, and Transport or Session Layers. The data link layer connects two IoT elements which generally could be two sensors or the sensor and the gateway device that connects a set of sensors to the Internet. Often there is a need for multiple sensors to communicate and aggregate information before to the Internet. Specialized protocols have been designed for routing among sensors and part of the routing layer. The Session Layer protocols enable messaging among various elements of the IoT communication subsystem. A number of security and management protocol have also been developed for IoT as shown in the figure.

Session		MQTT, SMQTT, CoRE, DDS, AMQP, XMPP, CoAP, ...	Security	Management
Network	Encapsulation	6LoWPAN, 6TISCH, 6Lo, Thread, ...	TCG, Coth 2.0, SMACK, SASL, ISAsecure, ace, DTLS, Dice, ...	IEEE 1905, IEEE 1451, ...
	Routing	RPL, CORPL, CARP, ...		
Datalink		WiFi, Bluetooth Low Energy, Z-Wave, ZigBee Smart, DECT/ULE, 3G/LTE, NFC, Weightless, HomePlug GP, 802.11ah, 802.15.4e, G.9959, WirelessHART, DASH7, ANT+, LTE-A, LoRaWAN, ...		

Figure 4: Protocols for IoT

III. IOT APPLICATION LAYER MESSAGING PROTOCOLS

A communication protocol is nothing but a language that is used by objects to interact among them. In simple words, a protocol is a set of rules that must be obeyed by the communicating objects. Communication Protocols are extremely essential in heterogeneous systems; where the objects interacting may be heterogeneous in nature, needing a common framework for them to interact. This section reviews standards and protocols for message passing in IoT Application Layer proposed by different standardization [2]. All most Web-based application and IoT application are IP based and they use TCP and UDP for transport. However, there are several message distribution functions that are common among many IoT applications; it is desirable that these functions be implemented in an interoperable standard ways by different applications. Those protocols are:

- 1) *MQTT (Message Queuing Telemetry Transport Protocol)*: MQTT (Message Queue Telemetry Transport) was developed by and introduced by IBM in 1999 and standardized by OASIS in 2013 to target come up with lightweight M2M communication [3]. It is Publish/Subscribe Protocol architecture similar to Client/Server Protocol shown in figure 5 below. The importance of MQTT Protocol is due to its simplicity and the no need of high CPU and memory usage (reason is the lightweight protocol) [4]. MQTT

supports a wide range of different devices and mobile platforms. At Transport Layer TLS/SSL security provides to MQTT. Show above figure 5 there three components are there Publishers, a Broker and Subscribers. Publishers are generally lightweight sensors that connect with a broker and send data to a broker and go back to sleep. Subscribers are IoT applications that interested in data send by sensors and also connect with a broker, so broker send interested data to subscribers. The brokers classify sensor data in Topics and send them to subscribers interested in the Topics. This all thing is on IoT point of View [5].

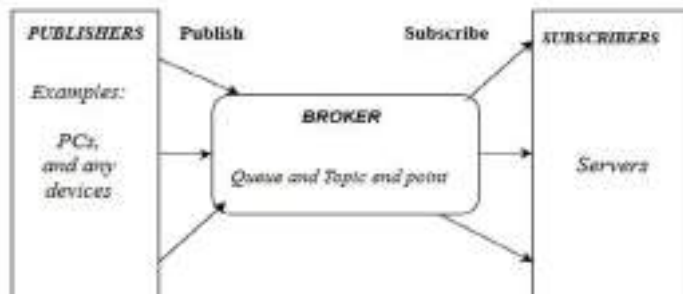


Figure 5: Working of MQTT Protocol

- 2) *CoAP (Constrained Application Protocol)*: CoAP (Constrained Application Protocol) is used for low power and low memory embedded devices where it can be used for communication instead of HTTP. Currently there is HTTP Protocol available with Request/Response paradigm but HTTP has many features and more footprint [5]. HTTP runs over TCP where TCP will need more resources due to three way handshake and many more complex mechanism. Now for low power embedded devices, there is no need of this heavy protocols and we can optimize it to run over TCP. As CoAP is a Restful web Transfer Protocol for use with constrained networks. CoAP uses Client/Server model of approach same as HTTP. It is designed for constrained networks with low overhead and lower footprint [6].

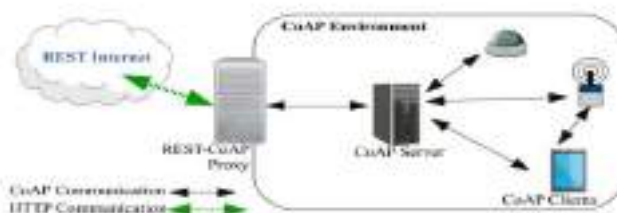


Figure 6: Constrained Application Protocol

- 3) *AMQP (Advanced Message Queuing Protocol)*: The Advanced Message Queuing Protocol (AMQP) is a protocol that across from the financial industry. Security is managed with the use of the TLS/SSL Protocols. Its run over TCP. AMQP is follow Publish/Subscribe communication Protocol for messaging [6,7]. AMQP is same like MQTT but AMQP have advantage its store data then forward it, and this features used at when network disruptions that time ensures reliability. Show in figure 7 below a broker divide into two part Exchange and Queue. Exchange responsibility to receive Publishers messages and distribute to Queue. A queue is based on Pre-define Roles and Condition and it's basically send message to subscribers who subscribe those data.

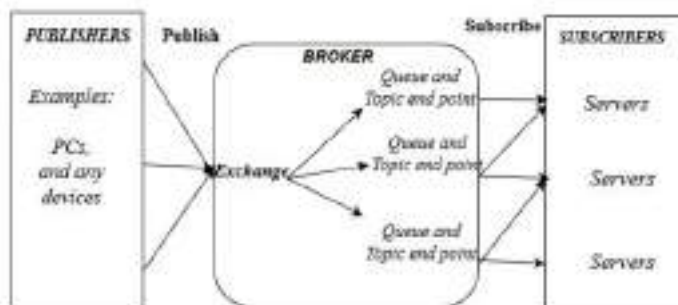


Figure 7: Advanced Message Queuing Protocol

- 4) *XMPP (Extensible Messaging and Presence Protocol)*: Extensible Messaging and Presence Protocol (XMPP) is a Messaging Protocol that was designed originally for chatting and message exchange applications. It was standardized by IETF in 1999 named as jabber. In all Application Layer Protocols only XMPP Protocol support Publish /Subscribe and Request/Response model and it's depend on application developers to develop application which model they use [8]. It does not provide any quality of service guarantees and, hence, is not practical for M2M communications. XMPP is IP based communication protocol with Extensible Mark-up Language (XML) support [9].

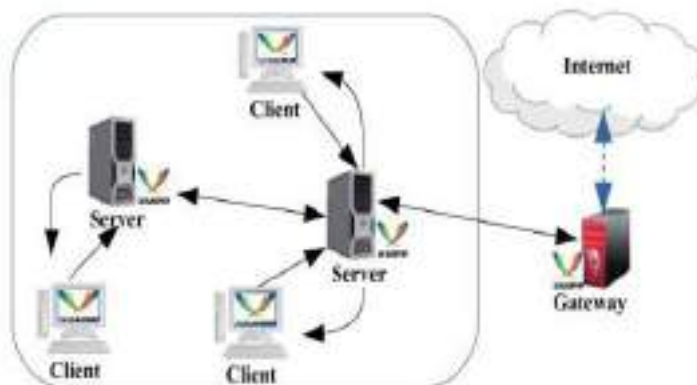


Figure 8: Extensible Messaging and Presence Protocol

- 5) *DDS (Data Distribution Service)*: Data Distribution Service (DDS) is another Publish/Subscribe Protocol that is designed by the Object Management Group (OMG) for M2M communications. It defines two sub layers: Data-Centric Publish-Subscribe (DCPS) which disseminates information to subscribers and Data-Local Reconstruction Layer (DLRL) which is an optional and is an interface to the DCPS functionalities. It shares data among distributed objects [9].

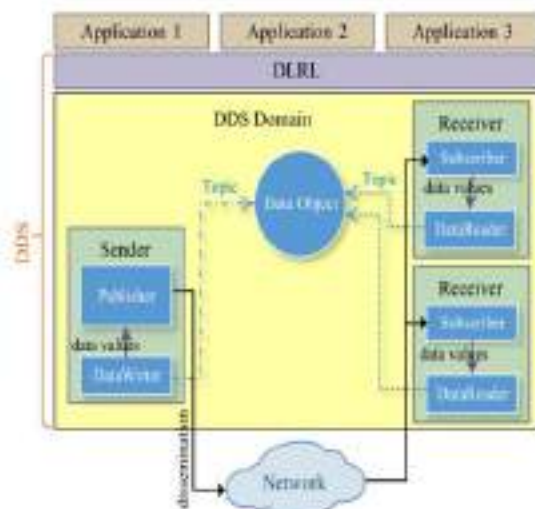


Figure 9: Data Distribution Service

- 6) *HTTP (Hyper Text Transfer Protocol)*: HTTP is specially designed for Internet. It was developed by Tim Berners Lee and later standardized by IETF in 1997. Though HTTP uses Request/Response architecture, it doesn't use topics. HTTP is based on Representational State Transfer (REST), an architectural style that makes information available as resources identified by URIs [8]. HTTP is simple text based protocol where no fixed header size is defined. It has features on Persistent and Non-Persistent connections. By default TCP is used as HTTP's Transport Protocol, but HTTP doesn't have any QoS support [9]. HTTP is very powerful protocol, but it's relatively expensive in implementation and network resource usage. This makes it difficult to adopt HTTP as it is for IoT networks. HTTP transfers a large number of small packets over web but overhead of HTTP causes many problems, such as consumption of network resources and large delays [10].

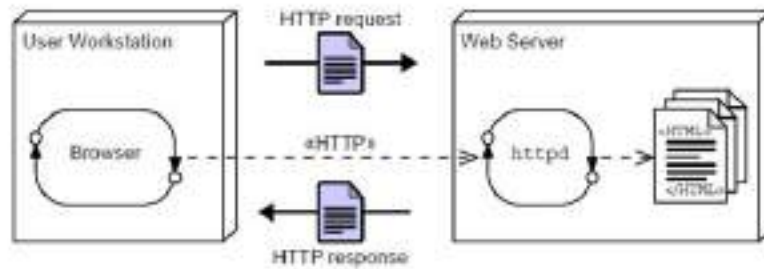


Figure 10: Working of Hyper Text Transfer Protocol

7) **RESTFULL (REpresentational State Transfer):** Representational State Transfer (RESTFUL Services) is an engineering that gives web administrations which permit correspondence and information trade between various gadgets utilizing HTTP in IoT condition [5]. REST utilizes the HTTP strategies GET, POST, PUT, and DELETE to give an asset arranged informing framework where all activities can be performed essentially by utilizing the synchronous request/response HTTP commands. RESTFUL services use the secure and reliable HTTP which is the proven worldwide Internet language. It can make use of TLS/SSL for security [9].



Figure 11: Working of RESTFULL Protocol

8) **WEB-SOCKET:** The WEB-SOCKET Protocol provides two ways for communication between clients and a remote server. Web-Socket provides security similar to the security model used HTTPS Protocol. For browsing Application Layer used and web-Socket work on TCP Transport Layer Protocol, so they need to interact and communicate with host those who connect with remote. Web-Socket is a Web-based Protocol that works on the single TCP channel and provides full duplex communications. Web-Socket starts session without Publish/Subscribe and Request/Response models like previous protocols [9, 10].

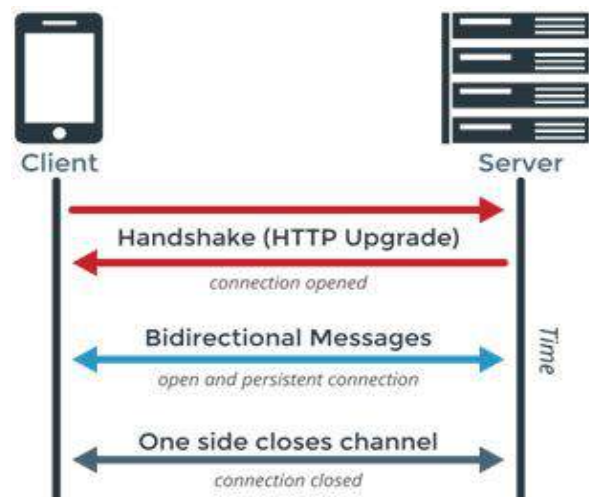


Figure 12: Working of WEB-SOCKET Protocol



IV. RESULT AND DISCUSSION

To interconnect objects to transfer data and to publish information over a network without requiring Human-to Human or Human-to-Computer interaction is now possible. That is why the IoT application programmers are faced with the challenges of choosing an appropriate communication protocols for their resource-constrained applications.

Performance Evaluation: In The process of Subscription-Publication, the challenge is to accomplish the message delivery with a high efficiency, a low latency and a low packet loss rate using one of reliability and QoS level. Otherwise, it is up to the application to select the appropriate QoS level for its publications and subscriptions, thus the decision to use one of these levels impacts on the application performance as well as on the use of bandwidth and battery life of devices. However, although the traditional protocol’s effectiveness, the need for a suitable protocol for IoT applications involving constrained devices is necessary because the biggest obstacle in the functioning of these devices is energy consumption. Both MQTT and CoAP protocols are being implemented for Mesh-Networking applications in networks in order to allow inter-standard communication between lightweight end nodes. Data transfer Protocols are rapidly emerging and integrating the IoT market as leading lightweight messaging protocols for constrained devices.

Each protocol offers unique benefits and each poses challenges and tradeoffs. The strengths and issues of these eight protocols are summarized in Table 1 [9, 10]. The purpose of this evaluation is to choose the use of either MQTT or CoAP according to the best throughput and lowest latency resulted in the presence of different criteria. AMQP can be integrated with TLS in order to ensure secure communication. DDS, RESTFUL, and WEB-SOCKET is an excellent quality of service levels and reliability guarantees. XMPP is very secure protocol which supports encryption, authentication, and access control. MQTT and CoAP use different transmission protocols of TCP and UDP. CoAP present the best performances in terms of both throughput and latency.

Table 1: Comparative Analysis of IoT Data Transfer Messaging Protocols

Sr.No	Criteria	MQTT	CoAP	AMQP	XMPP	DDS	HTTP	RESTFUL	WEBSOCKET
1	Year	1999	2010	2003	1999	2004	1997	2000	2011
2	RESTFUL	No	Yes	No	No	No	Yes	Yes	yes
3	Transport Protocol	TCP	UDP	TCP	TCP	TCP/UDP	TCP	TCP/UDP	TCP
4	Publish/Subscribe Model	Yes	Yes	Yes	Yes	Yes	No	No	No
5	Request/Response	No	Yes	No	Yes	No	Yes	Yes	YES
6	Security	SSL	DTLS	SSL	SSL	SSL/DTLS	SSL	SSL/TLS	TLS
7	QoS	yes	Yes	Yes	No	Yes	No	NO	NO
8	Header Size	2	4	8	-----	-----	-----	-----	-----
9	XML Support	No	No	No	Yes	No	Yes	Yes	Yes
10	Encoding Format	Binary	Binary	Binary	Character	Binary	Text	Text	Text
11	Default Port	1883/8883	5683/5684	5671/5672	5222/5223	7400/7401	80/443	23450/tcp	RFC 6455
12	Proxy Support	Partial	Yes	Yes	Yes	Yes	Yes	Yes	Yes

V. CONCLUSION

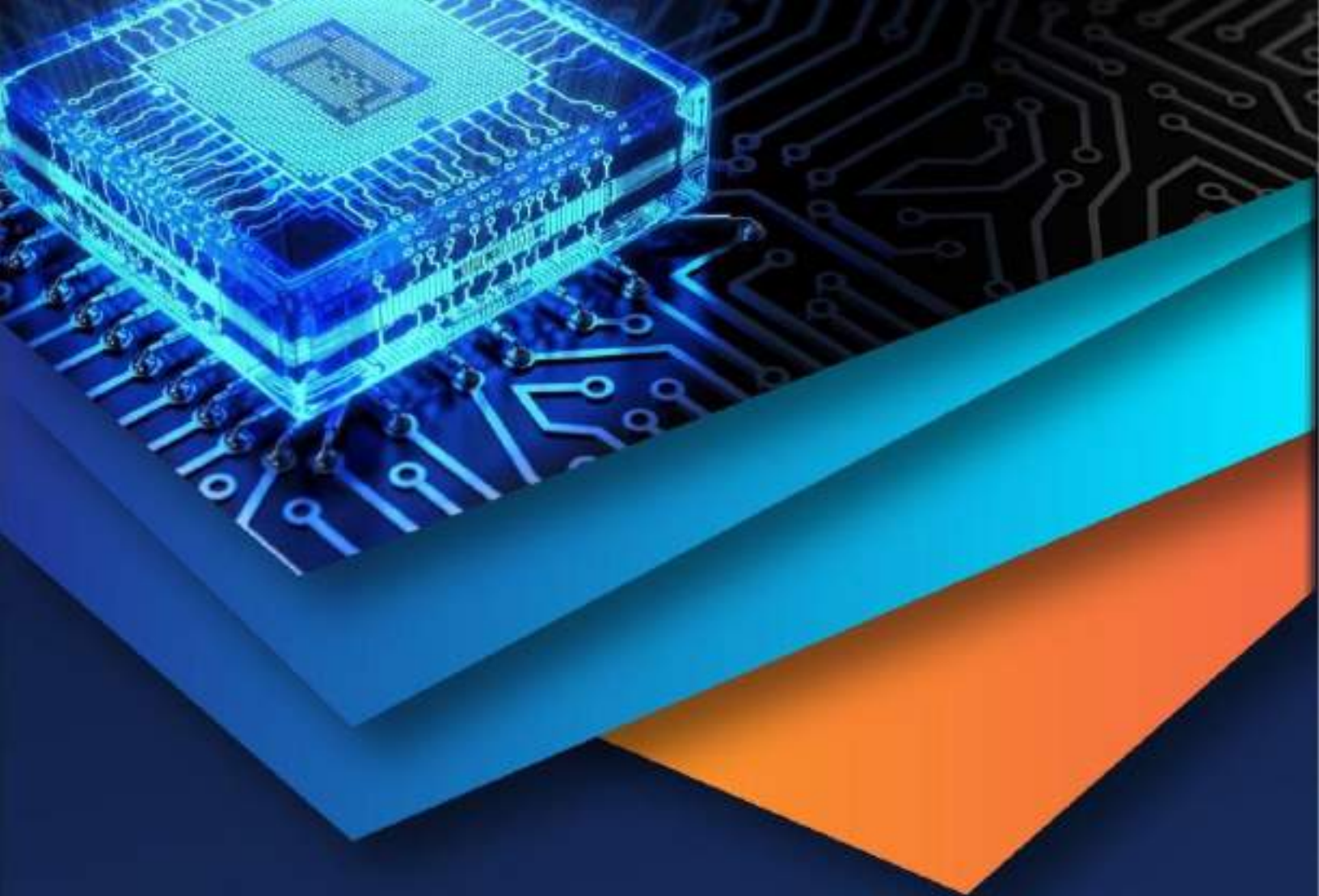
In this paper, we made a shot to supply survey on the embedded world and/or Machine to Human (M2H), Machine to Machine (M2M) communication around Data transfer Messaging Protocols. During this paper we have analyzed and compared messaging protocols for IoT systems. Started our discussion with MQTT then discussed, MQTT, AMQT, COAP, XMPP, DDS, HTTP, RESTFULL, and WEB-SOCKET. Each of these protocols has their different pros, cons and is meant for particular scenarios. Also we performed comparative analysis of this protocol which is in a position to assist us to choose appropriate messaging protocol depending upon application requirements.

There are three major components for implementing IoT on different applications: Security, Privacy and Trust. While increasing the expansion of IoT, Security is more important for reliable data transferred among the billions of smart objects. This paper concentrates on Application Layer Data Transfer Messaging Protocols on IoT devices. CoAP having light weight and consume low energy; CoAP is utilized on many applications of IoT. To secure data transferred, CoAP combined with DTLS Protocols named as Datagram Transport Layer Security Protocols because the safety agent. So, in future we concentrate on Security in Application Layer Protocols.



REFERENCES

- [1] Ms .B.Sarvaiya, Dr.S.S. Sherekar, Dr.V.M.Thakare, "Study of Security Challenges in Multi-layered Structure and Various Attacks on IOT", AIC 2K18 Annual IETE Convention International Journal of Electronics, Communication And Soft Computing Science & Engineering (IJECSCSE),Impact Factor- 4.526, ISSN :2277-9477, 29 and 30 September-2018.
- [2] Jayavardhana Gubbi et al. "Internet of Things (IoT): A Vision, Architecture Elements, and Future Directions", In Future Generation Computer Systems, PP.1645-1660, and ISSN: 0167-739X, 2013 IEEE.
- [3] Kushal Reshamdalal, "A Comprehensive Study of Application Layer Protocols (ALP)", International Journal of Innovations & Advancement in Computer Science (IJACS), ISSN 2347-8616, Volume 6, Issue 10 October 2017 IEEE.
- [4] Tanya Mohan Tukade, R.M.Banakar, "Data Transfer Protocols in IoT-An Overview", International Journal of Pune and Applied Mathematics, ISSN: 1311-8080, Volume 118, No.16 2018 IEEE.
- [5] Fathia Ouakasse, Said Rakrak, "A Comparative Study of MQTT and CoAP Application Layer Protocols via. Performances Evaluation", Journal of Engineering and Applied Sciences. ISSN: 6053-6061, Volume 13, No 15.2018 IEEE.
- [6] Yuang Chen,Thomas Kunz, "Performance Evaluation of IoT Protocols under a Constrained Wireless Access Network",2016 International Conference on Selected Topics in Mobile & Wireless Networking(MoWNeT).ISSN:5090-1743,2016 IEEE.
- [7] Danish Bilal Ansari, Atteeq-Ur-Rehman, Rizwan Ali Mughal, "Internet of Things (IoT) Protocols: A Brief Exploration of MQTT and CoAP", International Journal of Computer Applications.ISSN: 0975-8887, Volume 179, No.27, March 2018 IEEE.
- [8] Suvam Monanty, Sagar Sharma, Vaibhav Vishal, "MQTT-Messaging Queue Telemetry Transport IoT based Messaging Protocol", International Research Journal of Engineering and Technology (IRJET).ISSN:2395-0056, Volume 03, Issue 09, SEP 2016 IEEE.
- [9] Makkad Asim, "A Survey on Application Layer Protocols for Internet of Things (IoT)", International Journal of Advanced Research in Computer Science, ISSN: 0976-5697, Volume 8, No.3, March-April 2017 IEEE.
- [10] Sagar P Jaikar, Dr.Kamatchi R.Iyer, "A Survey of Messaging Protocols for IoT Systems", International Journal of Advanced in Management Technology and Engineering Sciences, ISSN: 2249-7455, Volume 8, Issue II, FEB 2018 IEEE.



10.22214/IJRASET



45.98



IMPACT FACTOR:
7.129



IMPACT FACTOR:
7.429



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089  (24*7 Support on Whatsapp)

Trustworthy IoT Traditional Network Security to Authentication and Access Control Model for Heterogeneous Devices

Ms.Shilpa B.Sarvaiya¹, Dr.D.N.Satange²

¹*Department of Computer Science, Vidyabharati Mahavidyalaya, Amravati*

²*Director, Students' Development, S.G.B.A. University, Amravati*

Abstract-Security is the basic requirement of any user for Internet of Things (IoT) traditional network. An internet user will not share his confidential and important data on the network unless the traditional network is trusted. IoT is considered as a collection of heterogeneous devices, such as, radio frequency identification, sensors and actuators, which form a huge traditional network, enabling not connected to computer in the network to produce a trustworthy world of services. Security and privacy are the two most important aspects of the IoT network, which includes authentication, authorization, data protection, network security, and access control. Additionally, traditional network security cannot be directly used in IoT traditional networks due to its limitations on computational capabilities and heterogeneous devices storage capacities. Authentication and Access control is the mainstay of the IoT traditional network, as all components undergo an authentication process before establishing communications between heterogeneous devices therefore, securing authentication and access control is essential to ensure that resources are only granted to the authorized users. With authentication and access control information, it sets the access rights of the subject to the object and protects heterogeneous devices from unauthorized access to ensure confidentiality and integrity of the system resources in the send and receive data signal is one of the basic security services. Current access control technology can be divided into Role-based Access Control (RBAC) and Resource Role Hierarchy Based Access Control (RRBAC). The first kind of access control model is RBAC, which is widely used in traditional networks and second is suitable for multiple security domains with different applications. In this paper focused on IoT security particularly on their authentication and access control model. Also, studies on existing evaluation schemes of IoT authentication and access control.

Keywords- Access Control, Authentication, Attribute-Based Access Control (ABAC), Role-Based Control (RBAC), Resource Role Hierarchy Based Access Control (RRBAC).

1. INTRODUCTION

The fundamental question that needs to be answered is how we can trust the validity of the data being generated in the first place. IoT therefore needs to improve its trustworthiness before it can be used to solve challenging economic and environmental problems tied to our social lives.

Due to huge number of IoT devices and machine to machine communication feature of IoT, legacy authentication and authorization techniques are not viable for it. Devices must authenticate each other before exchanging any information between them (M2M communication) which is a challenge for researcher due to massive number of heterogeneous devices. IoT is focusing on Machine to Machine (M2M) mode of communication. For such communication nodes authentication is very important for insuring security and privacy. When two or more nodes are communicating with each other for a common objective they should authenticate each other first in order to block fake node attack. However, there is no efficient authentication mechanism for massive number of IoT devices. Authentication and access control mechanisms are capable of preventing unauthorized users from accessing the data of sensor nodes on the IoT perception layer and guaranteeing the data security effectively. User authentication is to allow legitimate user to access resources as well as to decline malicious person or attacker [1]. After authentication, access control is to restrict authenticated user to access the only data that have the privileges. However, due to the characteristics of

wireless sensor network, secure access is faced with more severe challenges. The trustworthiness to heterogeneous device authentication and access control model in IoT traditional network are discussed here. In this paper, authors focus various evaluation techniques with their parameters and supporting equations. This paper presents an overview of the existing work on trust authentication, access control models in IoT. The first access control model is role-based access control (RBAC), which is widely used in traditional networks. Adopt ABAC-based authorization method in order to access various resources and data in this type of model, users require certain certificate information that falls into ABAC. If a user has some special attributes in ABAC, it is possible to access a particular resource or piece of data. ABAC is a more flexible and scalable than abstract identity, role, and resources information of the traditional access control into entity attributes. Additionally, ABAC can support either fine-grained access control in the complex system or dynamic extension of large-scale users. The second access control model RRBAC is suitable for multiple security domains with different applications [2, 3].

The paper focuses on building an access control model and system based on trust computing, which is a new field of access control techniques that includes Access Control, Trust Computing, Internet of Things, network attacks, and cheating detection technologies. Because target access control systems can be very complex to manage, there has been substantial research in this domain, most of which has been related to attacks like self-promotion and ballot stuffing where a node falsely promotes its importance and boosts the reputation of a malicious node (by providing good recommendations) to engage in a collusion-style attack. The traditional trust computation model is inefficient in differentiating a participant object in IoT, which is designed to win trust by cheating. There is an urgent need to put forward more suitable and effective methods to ensure the security of IoT

This paper is organized as follows. Section II describes the authentication model for IoT security. Section III presents Access control model for IoT security. In section IV Authentication Evaluation Techniques for IoT security, Section V presents Access Control Evaluation Techniques for IoT security finally, section VI concludes the paper and future research.

2. AUTHENTICATION MODEL FOR IOT SECURITY

Authentication allows communicating entities to convince the identity of each other and exchange session keys. In wireless sensor network, user and terminal nodes in the communication process require mutual authentication to ensure network security, while terminal nodes require authentication mutually to prevent malicious nodes attacks. Encryption mechanism ensures confidentiality to prevent data from being stolen during communication process via encoding the data. Usually, the authentication is divided into two parts [4, 5].

(1) Authentication: authentication between user and terminal nodes ensures only the legitimate User can access the network.

(2) Key establishment: session keys should be created between the user and nodes for secure Communication.

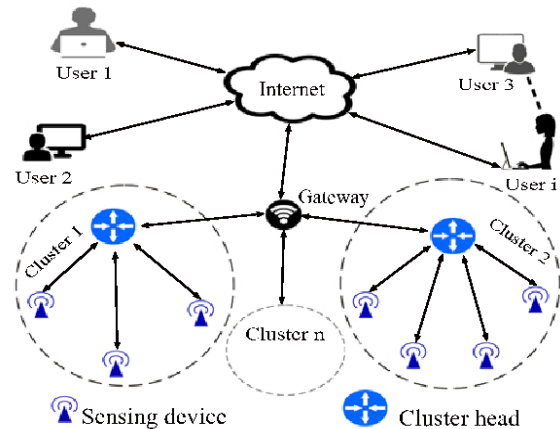


Figure 1: Authentication Model for IoT security

3. ACCESS CONTROL MODEL FOR IOT SECURITY

As discussed in Section 3, access control is the most fundamental component in trustworthy security. Many access control models have been developed in the past three decades and among all these models, Role-Based Access Control (RBAC) models [6] are most widely used in enterprises and other organizations. Role-based models can greatly cut down the cost for policy specification. Also, Role hierarchy in RBAC provides a natural representation (role hierarchy) of the structure of the users in an organization. Role faithfully describes the responsibility and authority of

the user in the position represented by the role. The RBAC model focuses on building a hierarchy of the subjects to reduce the overhead in access right specification and management but does not consider the same for the objects (i.e., the resources to be accessed). In IoT security, there are an enormous number of resources. If permissions have to be assigned for individual IoT resources to roles, permission assignment and management can have a very high complexity, likely to be infeasible [7]. RBAC model also has limitations in highly open environment where no role hierarchy can be formulated [8] In RBAC, the only alignment required for interoperability is to map the roles from one domain to another and role mapping techniques has been well explored [9]. For interoperability in ABAC, need to align the attributes as well as the values for the attributes. If two systems do not have equivalent attributes, it is impossible to align them. Extended the RBAC model and created the RRBAC (Resource and Role hierarchy Based Access Control) model [10] to circumvent the problems in RBAC and ABAC discussed above. Similar to role hierarchy, IoT resources can be organized in a hierarchy and permissions can be assigned based on the resource hierarchy. By providing resource hierarchy as a part of the access control model, we can greatly simplify access rights assignments using the resource groups and privilege inheritance concept on the resource hierarchy. The high level RRBAC model is formally specified in Section 3.1. For the dynamic and open IoT systems develop a “resource role hierarchy” based access control model to support easy policy specification. An entity in the system can build a resource role hierarchy to specify its view of the other entities in the system without knowing the specific entities. Integrate the RBAC model with RRBAC so that access control policies can be specified based on the relative role hierarchy and resource hierarchy [11]. When a dynamic IoT network is formed, the other entities are mapped to the relative role hierarchy of entity based on their attributes. The attribute values are obtained by mining the societal databases and social networks. The resource role hierarchy concept is presented in Section 3.2.

3.1 Role-Based Access Control Model (RBAC)

Role-Based Access Control approach (RBAC), a policy mechanism defined roles and privileges. This approach scales better than other models. However,

when talking about a huge amount of devices, managing roles for individual entities the possibility of grouping sensors and assigning roles to those that have the same rights is a good solution for this problem. For providing access rights to user, it is important to know the user’s responsibilities assigned by the organization. RBAC try to reduce the gap by combining the forced organizational constraints with flexibility of explicit authorizations [12]. RBAC mostly used for controlling the access to computer resources. RBAC is very useful method for controlling what type of information users can utilize on the computer, the programs that the users execute, and the changes that the users can make. In RBAC roles for users are assigned statically, which is not used in dynamic environment. It is more difficult to change the access rights of the user without changing the specified roles of the user. RBAC is mostly preferable access control model for the local domain. Due to the static role assignment, it does not have complexity. Therefore, it needs the low attention for maintenance [13, 14]. Role is nothing but the abstractions of the user behaviour and their assigned duties [15].

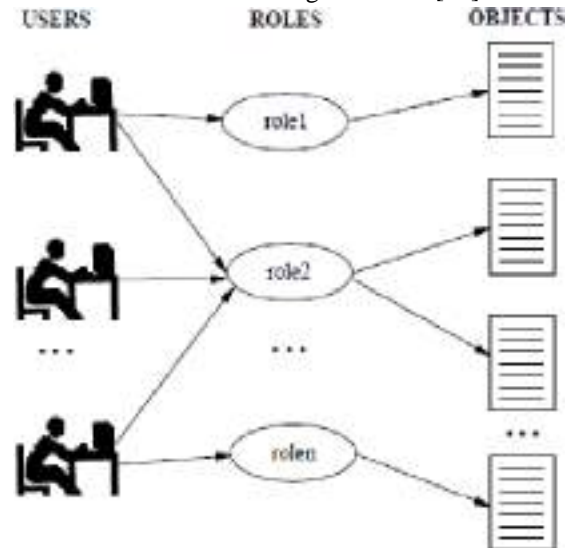


Figure 2: Role-Based Access Control Model

Essentially, in role-based access control policies need to identify the roles in the system, a role can be defined as a set of responsibilities and actions associated with a particular working activity. In an Access control security model, a role is considered as a job-related access right which can be given to the authorized users within an organization. It allows authorized user to achieve its associated responsibilities [14, 15].

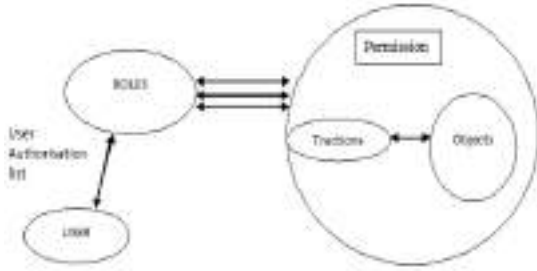


Figure 3: User-Role-Permission Mapping

A permission p is a pair $\langle \text{trans}, \text{objset} \rangle$, where trans represents the transaction that executes on the set of objects that is objset . Consider P indicate the universal set of permissions, Trans indicate the universal set of transactions, and Obj indicates the set of objects.

3.2 Resource and Role Hierarchy Based Access Control (RRBAC)

The big difference between RRBAC and (ABAC). RRBAC is suitable for multiple security domains with different applications. Figure 4 is the structure graph of RRBAC model. From Figure 4, the users are distributed anywhere, in a school, in a company etc. In every security domain, the administrator is charge of managing the sessions and roles. Usually, the session IDs are randomly generated as a procedure for a user to perform actions. The roles are man-made according to the registration of the resources. The resources are also distributed. After a resource registers and passes the examination, it can become a legitimate resource. Surely, a valid resource is treated as a part of the domain [16].

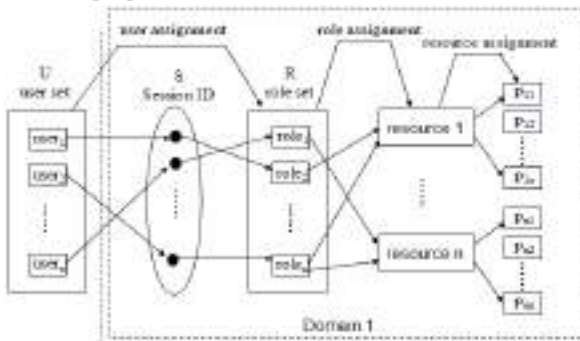


Figure 4: The structure graph of RRBAC Model [16]

4. IOT AUTHENTICATION EVALUATION TECHNIQUES

As the new and challenging authentication techniques are necessary to protect the IoT environment from

various emerging attacks, evaluation of those proposed schemes are equally important to check their potency. In this section, discuss several evaluation techniques with their parameters and supporting equations [17].

A. Average Response Time

Response time is assumed to be the time taken by the server to result in the response of a request to the client. This can be affected by few factors, such as, number of users, number of request, type of requests, think time, network bandwidth, and server configuration. First response time can be executed by the time of client request and the time of first response, which is defined in equation 1.

$$T_{res} = t_{res} - t_{req} \quad (1)$$

Here T_{res} , t_{res} , t_{req} are response time, time of client request and time of first response respectively. Average response time is calculated by the mean of all response time, which is determined in equation 2.

$$T_{avg_res} = (n/r) - T_{think} \quad (2)$$

Where T_{avg_res} is the average response time, n is the number of concurrent users. r is the number of requests per second the server receives. T_{think} is the average think time (in seconds). However, to obtain an accurate response time result, a user should always include think time in the equation.

B. Impact on Throughput

Throughput (TP) can be described as the amount of data passes through a system in a unit of time. In the traditional network of IoT, find out the total number of transmitted data preserved in a second. The TP can be defined in equation 3.

$$TP = \sum (Q_i^f * l_i) / T_w \quad (3)$$

Here, TP denotes throughput, while Q_i^f is the Quantity and l_i is the length of the i^{th} kind, and T_w denotes as the whole time of the simulation.

C. Packet Delivery Ratio

Packet Delivery Ratio is calculated based on the number of packets sent by the sender and the number of packets successfully received at the receiver end. However, it depends on several factors like network configuration, device capabilities, and bandwidth; therefore, it is difficult to test the network performance. Equation 4 can be used to calculate the packet delivery ratio.

$$PDR = N_{rp} / N_{sp} \quad (4)$$

Where PDR is Packet Delivery Ratio; N_{sp} is the total

number of sent packets, and N_{rp} is the total number of received packets. It has been identified that throughput falls when the number of nodes increases in a network. In the wireless sensor network, packet-sending circumstances are defined in the energy model, like that; energy is consumed when a packet is sent over the network. Therefore, more packet transfer cost core energy consumption. Ultimately, the packet can be discarded due to less energy or long-distance travel [17].

D. Handshake Duration

Handshaking is the process of negotiation between two network parties in the IoT network. These parties can be user, sensor, actuator, server or other nodes. As shown in Figure 5, handshaking takes place by completing the two-roundtrip message, whereas, client's discovery offers by the server and again the client's request acknowledges by the server. Duration to a handshake T_{hs} is computed at the client-end using equation 5.

$$T_{hs} = T_s + T_{res} + T_p \quad (5)$$

Where T_s is the time taken by whole session request, T_{res} is client response time and T_p denotes as processing time at the server. However, to calculate the handshake duration, a user must perform several random numbers of handshakes between the client and the server.

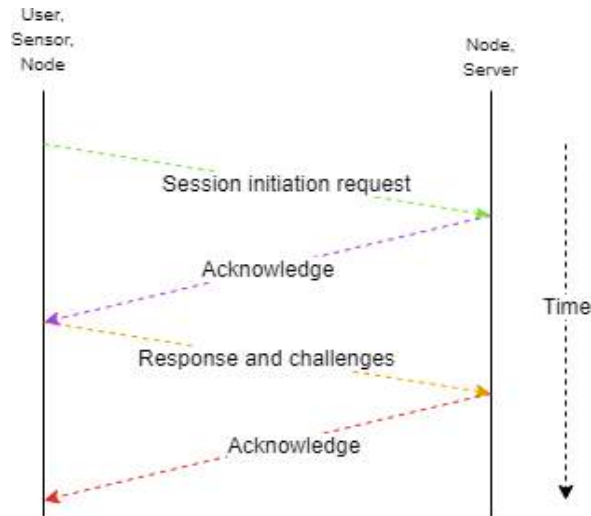


Figure 5: Four Way Authentication Handshaking.

E. End-To-End Delay

End-to-End Delay or E2ED denotes the average time to deliver packets from sender to receiver. E2ED can be calculated by using equation 6.

$$E2ED = \sum (T^r_i - T^s_i) / n \quad (6)$$

Here, i is the number of packets and n is the number of received packets, while T^r_i is the received and T^s_i denotes the sent timestamp for i^{th} packet. E2ED is proportional to the number of sensors in the IoT network. Therefore, an increased number of nodes put up the congestion in the network.

5. IOT ACCESS CONTROL EVALUATION TECHNIQUES

Access Control technologies are known as the main elements to address the security and privacy issues in the Internet of Things. Any effective access control system should satisfy the main security properties of confidentiality (preventing unauthorized divulgence of resources), integrity (preventing resource to be modified without authorization resources), and availability (assuring access to resource by legitimate users when needed). In addition, the classification of IoT heterogeneous devices by assigning them a particular class according to their evaluation techniques. These evaluation techniques will help in building an adequate access control framework to achieve the required security level for each domain application.

A. Quantitative and Qualitative Evaluation techniques
In this section, we evaluate in both quantitative and qualitative way towards access control in IoT and their versatility for preserving security and privacy by referring to the above Role-Based Access Control Model (RBAC) and Resource Role Hierarchy Based Access Control (RRBAC) Model based on the described legend below this evaluation is highlighted as follow [18].

Legend: VH= Very High=> 5, H= High =>4, M=Medium =>3, L= Low => 2, VL= Very Low =>1, No= Null => 0

Each quality criteria is assigned a value ranging from 0 to 5, where 0 signifies not defined, 1 signifies a very low quality sufficiency, 2 signifies a low quality sufficiency, 3 signifies a medium quality sufficiency, 4 signifies a high quality sufficiency while 5 signifies a very high quality sufficiency. These values are then used to indicate the sufficiency of each access control model's interoperability between heterogeneous devices for each security and privacy preserving analysis for IoT.

B. Evaluation of Access Control based on RBAC and RRBAC Model

The web service technology is known to provide great interoperability between heterogeneous devices. For this reason, we classify all the devices that adopt the web of think approach (based on web service) as high-quality sufficiencies in term of interoperability based on RBAC model have the following issues [18].

(1) Interoperability: the difficulty to approve a real consensus regarding the meaning of role to be shared with different applications, platforms, domains and enterprises.

(2) Role explosion: The role explosion issue justifies the critical dynamicity aspect of RBAC. Actually, RBAC defines access permissions in a static and fixed manner without taking the context of the access into consideration. As a result, a pure RBAC solution may be inappropriate for defining fine-grained access permissions based on context, and dynamics of IoT environment.

(3) Critical scalability: policies cannot evolve easily. In fact, the creation of new roles can lead to rebuilding the entire model.

(4) Nonsupport of delegation: a subject cannot grant access rights to another subject, as well as grant the right to further delegate all or part of the granted rights.

6. ANALYSIS AND DISCUSSION

The current concept of network and connectivity is going to be changed in the next few years. As it is predicted that the number of connected heterogeneous devices in the world will take over the headcount of human beings soon, this can be possible because of the expansion of the authentication and access control evaluation techniques in Internet of Things. However, security on IoT is still searching for its way to improve so that it can provide reliability and protection against threats. Again, suitable selection of authentication and access control model is one of the main important parts in security, because it is the gateway of a user or device to introduce in a network. In addition, a proper selection of authenticate devices to protect the network from attacks. Due those issues, basic RBAC model is not really a suitable solution to perform authorization functions in IoT domain applications requiring high level of interoperability/scalability, such as smart grids and smart cities. Authentication and access control is the process for giving the authority to access the

specific resources, applications and system. Access control defines a set of criteria to access the heterogeneous devices of the IoT system and its resources. In Role Based model creates different authorities permissions by assigning access rights to specific roles or jobs within the IoT system then role based access control assigns these roles to users, It is effectively implemented in an environment because command and resources are assigned according to the roles.

7. CONCLUSION

The data can be perceived from any device at any moment. The powerful IoT heterogeneous devices authentication and Access Control is needed to make sure connected devices on the IoT can be trusted and access to the device resources to be what they intend to be. Accordingly, each IoT device needs a distinctive identity that can be authenticated when the device connected to the traditional network, it can track every device communicate securely with it, and inhibit it from executing detrimental processes. If a device shows unforeseen behaviour can simply revoke its privileges and compulsory to verify that the data do not change during transit. Then the network checks if this information is correct or not and ensures that the device is connected to the right network or not. It can be done using authentication and access control mechanism.

This paper is to propose a flexible Resource and Role Hierarchy Based Access Control (RRBAC) model for the dynamic IoT environment. Working from the RBAC access control model, the RRBAC access control model is more extendable, flexible and trustworthy. The main contribution of this paper is to understand the trust models enable the owners and roles to determine the trustworthiness of individual roles and users in the RBAC system respectively. RRBAC allow the data owners to use the trust evaluation techniques to decide to save their data in the IoT environment. RRBAC model provides a flexible approach for many security domains. Authentication and Access Control models supports different types of resources sharing to reduce the impact of any wrong data signal on the IoT traditional network. As a consequence, the RRBAC model provides a self-adaptive framework which is reliable enough too attached to any heterogeneous devices in the IoT environments.

REFERENCE

- [1] Ms.S.B.Sarvaiya, Dr.S.S.Sherekar, Dr.V.M .Thakare,” Taxonomy of Authentication Techniques in Security Attacks of Internet of Things”, NCETS “Research Journey” International E- Research journal, Impact Factor 6.261 ISSN: 2348-7143,February-2019.
- [2] Inayant Ali, Sonia, Sabir, Zahid Ullah," Internet of Things Security, Device Authentication and Access Control: A Review”, International Journal of Computer Science and Information Security (IJCSIS), vol. 14, no. 8, August 2016.
- [3] Sowmya Ravidas, et al. " Access Control in Internet-of-Things: A Survey”, ResearchGate March 29, 2019.
- [4] Antonio L.Maia Neto et al., " AoT: Authentication and Access Control for the Entire IoT Device Life-Cycle” November 2016.
- [5] Yungpeng Zhang,Xuqing Wu, " Access Control in Internet of Things: A Survey”, Asia-Pacific Engineering and Technology Conference, ISBN: 978-1-60595-443-1, 02 October 2018.
- [6] R. Sandhu, E. Coyne, H. Feinstein and C. Youman, "Role- based access control models," IEEE Computer, vol. 29, no. 2, pp. 38-47, 1996.
- [7] E. Yuan and J. Tong, "Attribute based access control (ABAC) for web services," in IEEE International Conference on Web Services, 2005.
- [8] C. Hu, D. Ferraiolo, D. Kuhn, A. Schnitzer, K. Sandlin, R. Miller and K. Scarfone, "Guide to attribute-based access control (abac) definition and considerations," in NIST Special Publication 800-162, 2014.
- [9] B. Shafiq, J. Joshi, E. Bertino and A. Ghafour, "Secure interoperation in a multidomain environment employing RBAC policies," IEEE TKDE, vol. 17, no. 11, pp. 1557- 1577.
- [10]N. Solanki, Y. Huang, I.-L. Yen, F. Bastani and Y. Zhang, "Resource and role hierarchy-based access control for resourceful systems," in CompSAC, 2018.
- [11]Xingdong Li, Zhengping Jin “Resource and Role Based Access Control Model”, 3rd International Conference on Mechatronics and Industrial Informatics (ICMII 2015).
- [12]Bokefode Jayant.D., Ubale Swapnaja A,Apte Sulbha S,Modani Dattatray G,” Analysis of DAC MAC RBAC Access Control based Model for Security”, International Journal of Computer Applications (0975-8887) Volume 104-No.5,October 2014.
- [13]Zhuo Tang, Juan Wei, Ahmed Sallam, Kenli Li, and Ruixuan Li,” A New RBAC Based Access Control Model for Cloud Computing Springer-Verlag Berlin Heidelberg 2012.
- [14]Yizhu Zhao, Yanhua Zhao, Hongwei Lu,” A flexible role-and resource-based access control model”, International Colloquium on Computing, Communication, Control, and Management 2018 ISECS
- [15]H.L.F.Ravi Sandhu, Edward J.Coyne, C.E. Youman. Role-based Access Control Models. IEEE Computer, 29 February 1996.
- [16]Nidhiben Solanki,Yongtao Huang,I-Ling Yen,Farokh Bastani,Yuqun Zhang,”Resource and Role Hierarchy Based Access Control for Resourceful Systems, International Conference on Computer Software and Applications 2018 42nd IEEE.
- [17]Tarak Nandy,Rafidah MD Noor,et al. " Review on Security of Internet of Things Authentication Mechanism ",IEEE,vol.7, August 26,2019.
- [18]Aafaf Ouaddah, et al.," Access control in The Internet of Things: Big challenges and new opportunities”, ScienceDirect, ISSN:1389-1286 17 January 2018.

Higher Education in Commerce is aligned with Employability Skills: A myth or a reality

Dr. Devendra Shivdas Rangacharya
Assistant Professor in Economics, Vidyabharti Mahavidyalaya
Camp Amravati.

Abstract

In India, a rapidly developing country, enrolment in higher education is accelerating. It is believed that higher education will translate into better skill enhancement and so, jobs. However, the real scenario is starkly different. In this research paper, the researcher gathered primary data of 250 students all of them pursuing higher education. The parameters of the questionnaire included items like skill-building, alignment with industry needs, the need for professional certification, and topicality of subjects, and so on. According to 149 students, the commerce syllabus is not related to the current economic situation, 131 students opine that the syllabus did not give the knowledge of current market conditions, 142 students disagree that commerce graduate gets a handsome salary. According to 136 students colleges have not signed MOU with nearby companies to give training to the students, 147 students agree that professional certification is mandatory for better job prospects. Findings reveal some positive responses such as higher education in commerce is job oriented, industrial visits are organized, practicals for certain subjects are taken in college they have to write the practical in the book. Overall analysis reveals, that there is a scope to improve the commerce syllabus as there is a gap between industry needs and higher education inputs.

Keywords: Alignment, Commerce, Industry, Job, Higher education.

Introduction

Commerce deals with trading things of economic value between two or more entities that is exchange of goods and services. The field of commerce is the most influential and strongest social institution these days.

Commerce education is given to the students to provide in-depth knowledge of the business and its different functional areas. While pursuing higher education in commerce the students acquire knowledge of business trade, fluctuations in the market, and basics of economics, industrial policies, and fiscal policies, and so on. Commerce education helps support the increasing needs of business houses. Commerce education equips students with specialized skills that come useful in tackling problems in different functional areas of commerce, Industry, and Trade.

There are several factors to take into consideration when deciding whether obtaining a higher education degree can be a positive decision or a negative decision. The advantages of pursuing higher education range from guaranteed employment, improved healthy lifestyles, higher income, to social recognition. To validate this objective, a questionnaire was designed. The parameters of the

26. Cashless Transactions: Challenges and Remedies

Dr. D. S. Rangacharya

Department of Economics, Vidyabharti Mahavidyalay, Amravati.

Abstract

The real aim of demonetization was to change the mode of Indian Economy from cash to cashless economy. The biggest problem in this was that 86% of Indian economy was cash based. Taking the situation under consideration Indian Government has launched many a programs to promote cashless transaction, such as BHIM App. BHIM (Bharat Interface for Money) app is a simple, quick and easy digital payment solution. BHIM app is interoperable with other Unified Payment Interface (UPI) applications and bank accounts for quick money transfer. Infact BHIM is a mobile app developed by National Payments Corporation of India (NPCI) based on the Unified Payments Interface. This app was launched by the Prime Minister at a Digi Dhan Mela at Talkatora Stadium I new Delhi on Dec, 30, 2016. It was named after Dr Bhimrao Amedkar and intended to facilitate E-payment directly through banks as part of the 2016 Indian Bank Note demonetization and drive towards cashless transactions. The app supports to all Indian Bank which use this platform for the immediate payment service infrastructure and allows the user to instantly transfer money between bank accounts of any two parties. This app can be used on all mobile devices. The major concern in this cashless economy is security. Any cashless transaction can't be done without internet, so cyber security, cyber crime and online frauds are some major concern in this regards. In India, there is not a single law which deals with cyber security or cyber crime. So there is a need to make laws for this concern so that cashless transactions can be made safe and interests of our consumers can be secure. This threat is not only for India but also the World. Protection of consumer is a big challenge in the E-commerce. United Nation, The International Consumer Protection and Enforce Network and other organizations are working to protect consumers rights and making policies for their protection. To tackle this problem Indian government has launched many a program like "Digital India" and "Vittiya Saksharta Abhiyan" to increase awareness among the people.

4. Financial Inclusion: Solution for Banking Shadow

Dr. D. S. Rangacharya
Department of Economics, Vidyabharti Mahavidyalay, Amravati.

Introduction

The dark and enormous banking shadow depriving the poor of economic opportunities in India can only be removed through financial inclusion. Rain shadow is a dry area, where because of its topography, there is very little rainfall while areas around it get abundant rain. Similarly, there are a very large number of low-income families, particularly in rural areas, who are a banking shadow, i.e., they are not served at all or served well by the formal banking system. They are thus deprived of the opportunities of economic advancement enjoyed by the middle-income and upper-income families. The disparities caused by this banking shadow are increasing the income gap between the rich and the poor, which is likely to exacerbate social tensions. It is estimated that nearly half of our population does not have operating bank accounts or access to bank credit. Only 10% of our six lakh villages have a bank branch. Despite extensive efforts by the government over the past several decades, the availability of financial services to the poor families has not improved.

Financial Inclusion

Financial inclusion refers to delivery of banking services and credit at an affordable cost to a vast section of disadvantaged and low income groups. Rangarajan's committee on financial inclusion defines it as, "The process of ensuring access to the financial services and timely and adequate credit where needed by vulnerable groups such as weaker sections and low income groups at an affordable cost."

Financial Exclusion

Financial exclusion could be looked at in two ways:

1. Lack of access to financial services which could be due to several reasons such as
 - Lack of sources of financial services in our rural areas, which are popular for the ubiquitous money lenders but do not have (safe) saving deposit and insurance services.



EMPIRICAL STUDY OF PREDICTION MODEL IN EDUCATION DOMAIN

¹K. P. Raghuvanshi, ²Dr. V. R. Dhawale, ³N. P. Pardesi, ⁴P. S. Selokar, ⁵A. M. Chopade

¹Professor, Department of MCA, Vidyabharati Mahavidyalaya, Amravati, India

²Professor, K. K. Wagh Institute of Engineering Education and Research, Nashik, India

^{3,4,5}Student, Department of MCA, Vidyabharati Mahavidyalaya, Amravati, India

Abstract: Data mining has been used in many different sectors to extract relevant data from the large data sets for analytical purposes. In recent years, several research studies have proven the significance of using data mining techniques in education. With the use of various data mining techniques, relevant data can be efficiently extracted from the large sets of data present in the education system. By analyzing the extracted data, many different aspects regarding education can be predicted with a significantly high accuracy rate. This includes predicting aspects such as newly emerging fields, performance of students and instructors, learning pace of the students, chances of employment after graduation and dropout rate in a particular field etc. This paper presents an empirical study on various papers which show how prediction models can be used to extract the relevant data from the large data sets present in the education system for analyzing several important aspects.

Keywords: Data Mining, Prediction Model, Data Analysis, Big Data.

I. INTRODUCTION

Data mining is an exploratory data analysis, an art of extracting information from data, a process of discovering patterns in data, and a task of discovering meaningful data from big data with the aim of obtaining clear and useful results. There are many goals of data mining such as to reduce large data into useful knowledge, to find useful patterns from large data, to identify important trends and previously unknown behavior pattern to find patterns in historical data, and to uncover previously unknown patterns.

II. DATA MINING TECHNIQUES

1. Association

Association analysis is the finding of association rules showing attribute-value conditions that occur frequently together in a given set of data.

2. Classification

Classification is the process of finding a set of models that describe and distinguish data classes or concepts, for the purpose of being able to use the model to predict the class of objects whose class label is unknown.

3. Prediction

Prediction uses a combination of other data mining techniques such as trends, clustering, classification, etc. It analyzes past events or instances in the right sequence to predict a future event.

4. Clustering

Clustering is the task of dividing the data points into a number of groups such that data points in the same groups are more similar to other data points in the same group and dissimilar to the data points in other groups.

5. Regression

Regression can be defined as a statistical modeling method in which previously obtained data is used to predict continuous quantity for new observations.

6. Sequential Patterns

The sequential pattern is a data mining technique specialized for evaluating sequential data to discover sequential patterns.

III. LITERATURE REVIEW

Sr. No.	Prediction Model	Authors	Techniques	Conference/Journal and Year	Conclusion
1	College Students' Employment	Houwen Fan	Decision Tree	13th International Conference on Intelligent Computation Technology and Automation, 2020	Based on the enrollment, student status management and employment data of a university over the years, this paper preprocesses and analyzes the employment related data, applies the classification algorithm to generate the decision tree model of employment prediction, extracts the prediction rules, and evaluates the accuracy of the employment data Through analyzing and obtaining the regular potential information, the prediction model of student employment and employment unit type is generated. In the process of the experiment, based on the research of predicting the employment of college students, we realized the construction of the decision tree model of employment prediction through programming.
2	Predicting Student Performance	Wan Fairos Wan Yaacob, Syerina Azlin Md Nasir, Wan Faizah Wan Yaacob, Norafefah Mohd Sobri	K-NN Naives Bayes DT – Information Gain DT – Gini Logistic Regression (LR)	Indonesian Journal of Electrical Engineering and Computer Science, Dec 2019	In this paper, four supervised data mining algorithms were applied on the students performance data to predict student performance either excellent or non-excellent based on predictive accuracy. The results indicate that the Naïve Bayes classifier outperformed other algorithm compared to Decision Tree, k-NN, and Logistic Regression with accurate and comprehensive classifier.
3	Predict Student's Performance	Annisa Uswatun Khasanah1), Harwati 2)	Bayesian Network & Decision Tree	IOP Conference Series, 2017	Some high influence attributes in student prediction that can be considered by university management to minimized student drop out based on the result of this study are First Semester GPA, First Semester Attendance, Senior High School Department, Gender, Father Occupation, Mother Occupation, Mother Education, and Origin. From this study that feature selection can increase the accuracy rate of the prediction, and Bayesian Network is outperforming Decision Tree since it has higher accuracy rate.
4	Students' Employability Prediction Model	Tripti Mishra., Dharminder Kumar.,	J48 Random forest Random Tree SMO	International journal of applied engineering research, Nov 2016	The aim of this paper was to apply various classifiers to find the employability of students and develop employability model based on the suitable classifier. It was found that J48 algorithm which is

	through Data Mining	Sangeeta Gupta	Multilayer Perceptron Naïve Bayes		implementation of pruned C4. 5 Decision Tree algorithm of WEKA is most suitable for the employability prediction.
5	Predict Slow Learners	Parneet Kaura ,Manpreet Singhb ,Gurpreet Singh Josanc	Multilayer Perception 75%, Naïve Bayes 65.13%, SMO 68.42%, J48 69.73% , REPTree	3 rd International Conference on Recent Trends in Computing, 2015	Among all data mining classifiers Multi Layer Perception performs best with 75% accuracy and therefore MLP proves to be potentially effective and efficient classifier algorithm. Also comparison of all 5 classifiers with the help of WEKA experimenter is also done, in this case also MLP proves to be best with F-measure of 82%. Therefore, performance of MLP is relatively higher than other classifiers. A model performance chart is also plotted. This research help the institutions to identify students who are slow learners which further provide base for deciding special aid to them.
6	Students' Academic Performance	Fadhilah Ahmad, Nur Hafieza Ismail and Azwa Abdul Aziz	Navie Bayes Decision Tree, RB	Applied Mathematical Sciences, Vol. 9, 2015	The amount of data stored in an educational database at IHL is increasing rapidly by the times. In order to get the knowledge about student from such large data and to discover the parameter that contributed to the students' success, the classification techniques are applied to the students' data. This study also conducts a comparative analysis of three classification techniques; DT, NB, and RB using WEKA tool. The experimental result shows that the RB has the best classification accuracy compared to NB and DT.
7	A Study Model On The Impact Of Various Indicators In The Performance Of Students In Higher Education	Jai Ruby1 , K. David2	Chi-square, Information Gain, Gain Ratio, Linear Regression, Correlation.	International Journal of Research in Engineering and Technology	This study paper on performance analysis of student data help the institution to decide on the factors to concentrate for the better performance of the academic results of the students.
8	Performance of Students in Higher Education	Jai Ruby 1 , Dr. K. David 2	ID3, J48, NB Tree, MLP, Simple Cart, Rep Tree, Decision Table	International Journal for Research in Applied Science and Engineering, 2014	This study paper helps the institution to know the academic status of the students in advance and can concentrate on weak students to improve their academic results. All the classification algorithms MLP, ID3, J48, REPTree, NBTree, SimpleCart and Decision Table considered in the study shows prediction accuracy above 68% for the student dataset.
9	Mining Students' Data for Performance Prediction	Tripti Mishra, Dr. Dharminder Kumar, Dr. Sangeeta Gupta	Random Tree, J48	Fourth International Conference on Advanced Computing & Communication Technologies, 2014	This paper focused on identifying attributes that influenced students 'third semester performance. Effect of emotional quotient parameters on placement has been established. Random tree gave higher accuracy of prediction than J 48.
10	Prediction of Student Performance	M. Durairaj #1, C. Vijitha *2	Naïve Bayes Decision Tree	International Journal of Computer Science and Information Technologies, 2014	Using K-Means clustering algorithm, we predicted the pass percentage and fail percentage of the Overall students appeared for a particular examination. The results show the students' performance and it is seems to be accurate. The comparison between Naviebays algorithm and

					decision stump tree technique shows that the Navie bayes techniques produce accurate result than the other and it is measured using confusion matrix. The results are predicted within 0 seconds.
11	Student Performance	Dorina Kabakchieva	OneR Rule Learner, Decision Tree, Neural Network and K-Nearest Neighbour	International Journal of Computer Science and Management Research, 2012	The classification models, generated by applying the selected four data mining algorithms – OneR Rule Learner, Decision Tree, Neural Network and K-Nearest Neighbour, on the available and carefully pre-processed student data, reveal classification accuracy between 67.46% and 73.59%. The highest accuracy is achieved for the Neural Network model (73.59%), followed by the Decision Tree model (72.74%) and the k-NN model (70.49%). The Neural Network model predicts with higher accuracy the “Strong” class, while the other three models perform better for the “Weak” class.
12	Academic Performance for First Year Student	Ernesto Pathros Ibarra García ¹ Pablo Medina Mora ²	Naïve Bayes	10th Mexican International Conference on Artificial Intelligence, 2011	Predictions for the low and high groups have significant percentage accuracy in some cases, exceeding 70% if the naïve Bayes classifier is used. This shows that it is possible to obtain a good prediction model. It can be used to detect low performing students and high performing students take appropriate decisions even before the courses start and, hence, to revert their academic standing.
13	Early Prediction of Student Success	Zlatko J. Kovačić	CHAID Tree Cart Tree	Informing Science & IT Education Conference, 2010	This study examines the background information from enrolment data that impacts up on the study outcome of Information Systems students at the Open Polytechnic. Based on results from feature selection (Figure 2 and Table 3), the CHAID tree presented in Table 4 and Figures 3 and 4, the CART trees presented in Table 5 and Figures 5 and 6 it was found that the most important factors that help separate successful from unsuccessful students are ethnicity, course programme and course block. Demographic data such as gender and age though significantly related to the study outcome, according to the feature selection result, were not used in the classification trees. Unfortunately the classification accuracy from the classification trees was not very high. In the case of the CHAID tree the overall classification accuracy was 59.4% and in the case of the CART tree slightly higher at 60.5%.

V. CONCLUSION

Prediction models are very important in any domain. We use the prediction model in the Education domain. Using this model we can predict the College Students' Employment, Early Prediction of Student Success, Academic Performance for First Year Student, Performance of Students in Higher Education, Students' Academic Performance and Students' Employability.

IV. FUTURE SCOPE

Data mining techniques are already used for the prediction of success rate of students and chances of employment after graduation. By extracting the relevant data from the database of a particular university or a college, the average number of Start-ups in upcoming years from the respective university or college can be predicted. These techniques can also be used for the prediction of average placement rate of a university or a college in upcoming years.

REFERENCE AND BOOKS

- [1] "Data Mining: The Textbook", Charu C. Aggarwal
- [2] "Data Mining: Concepts and Techniques", Jiawei Han
- [3] "Introduction to Data Mining", Michael Steinbach, Pang-Ning Tan, and Vipin Kumar
- [4] "Data Mining Techniques", Arun K. Pujari
- [5] Kovačić, Z. (2010), "Early Prediction of Student Success: Mining Students Enrolment Data", Paper presented at Proceedings of Informing Science & IT Education Conference (InSITE) ,Casinio Italia, June, 19-24,2010.
- [6] Dorina Kabakchieva, "Student Performance Prediction by Using Data Mining Classification Algorithms", International Journal of Computer Science and Management Research Vol 1 Issue 4 November 2012.
- [7] Ramaswami, M., Bhaskaran, R. (2010), "A CHAID Based Performance Prediction Model in Educational Data Mining", IJCSI International Journal of Computer Science Issues, Vol. 7, Issue 1, No.1, January 2010, pp.10-18.
- [8] Cortez, P., Silva, A. (2008), "Using Data Mining to Predict Secondary School Student Performance", EUROSIS, A. Brito and J. Teixeira (Eds.), 2008, pp.5-12.
- [9] Vandamme, J., Meskens, N., Superby, J. (2007), "Predicting Academic Performance by Data Mining Methods", Education Economics, 15(4), pp405-419.
- [10] M. Durairaj, C. Vijitha (2014), "Clustering Algorithm for Educational Data Mining: A Systematic Review of Literature and Techniques", 53 International Journal of Scientific Research in Computer Science Applications and Management Studies, Volume 3, Issue 4 (July 2014).
- [11] Md. Hedayetul Islam Shovon, "Prediction of Student Academic Performance by an Application of K-Means Clustering Algorithm", International Journal of Advanced Research in Computer Science and Software Engineering, Vol. 2(7), July 2012.
- [12] Sajadin Sembiring, "Prediction Of Student Academic Performance by an Application of Data Mining Techniques", International Conference on Management and Artificial Intelligence IPEDR, IACSIT Press, Vol.6,2011.
- [13] Tripti Mishra, Dr. Dharminder Kumar , Dr. Sangeeta Gupta , "Mining Students' Data for Performance Prediction", 2014 Fourth International Conference on Advanced Computing & Communication Technologies, 2014 IEEE.
- [14] R. R. Kabra, R.R. Bichkar , "Performance Prediction of Engineering Students using Decision Trees", International Journal of Computer Applications, Volume 36, No.11, 2011.
- [15] Jai Ruby, K. David, "A Study Model On The Impact Of Various Indicators In The Performance Of Students In Higher Education", International Journal of Research in Engineering and Technology, Volume: 03 Issue: 05 May-2014.



Journal of Emerging Technologies and Innovative Research

An International Open Access Journal Peer-reviewed, Refereed Journal

www.jetir.org | editor@jetir.org An International Scholarly Indexed Journal

Certificate of Publication

The Board of

Journal of Emerging Technologies and Innovative Research (ISSN : 2349-5162)

Is hereby awarding this certificate to

Kunal. P. Raghuvanshi

In recognition of the publication of the paper entitled

EMPIRICAL STUDY OF PREDICTION MODEL IN EDUCATION DOMAIN

Published In JETIR (www.jetir.org) ISSN UGC Approved (Journal No: 63975) & 7.95 Impact Factor

Published in Volume 9 Issue 11 , November-2022 | Date of Publication: 2022-11-25

Parisa P

EDITOR

JETIR2211411

[Signature]

EDITOR IN CHIEF

Research Paper Weblink <http://www.jetir.org/view?paper=JETIR2211411>

Registration ID : 504717





An Empirical Study on Virtual Private Network

¹Prof. Sunita Totade, ²Dr. V. R. Dhawale, ³Sakshi U. Choudhary, ⁴Gauri P. Alone, ⁵Falguni D. Joshi

¹HOD, Department of MCA, Vidyabharti Mahavidyalaya, Amravati, India

²Assist.Prof., K.K. Wagh Institute of Engineering Education and Research, Nashik, India

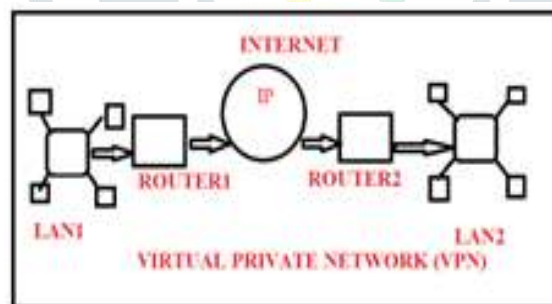
^{3,4,5}Student, Department of MCA, Vidyabharti Mahavidyalaya, Amravati, India

Abstract: Now a days, network security becomes an essential part of human life. Everyone wants that his or her network should be secure and free from every kind of malicious activities. Network security is important because it keeps sensitive data safe from cyber-attacks and ensure the network is usable and trustworthy. Virtual Private Network provides us a protected network connection, whenever we use public network connection Virtual Private Network encrypt our internet traffic and hide our online identity. Due to this, third-party applications cannot get our data. Virtual Private Network connects our mobiles or PCs to server computer so that internet connection of server computer can be easily used. Virtual Private Networks are legal networks. Our data remains secure from hackers, public networks and government and third-party applications. The intention of this paper is to enlighten the structure of Virtual Private Network and how it is easy to use for individual and other sectors as well. At the beginning we have introduced about Virtual Private Network and then working of VPN.

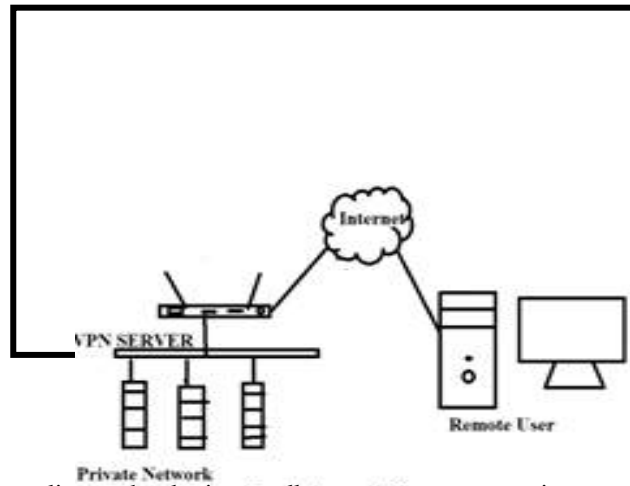
Keywords: Tunnel, Encryption, Remote Access VPNs, Personal VPNs, Mobile VPNs, Site-to-Site VPNs, WAN

I. INTRODUCTION

A Virtual Private Network (VPN) is a part of data security which provides security to its users from hackers, government, public networks and third-party applications. Virtual Private Network provides users an encrypted server and conceal user's IP address from corporations, government agencies and hackers. Virtual Private Network protects user identity even if they are using public or shared Wi-Fi and keep user's data private and safe from any prying internet eyes. Virtual Private Network or VPN is a technology which used for connecting the components and resources of one network over another. In common usage, a Virtual Private Network (VPN) is a private corporate network whose Wide Area Network (WAN) connections are made over a shared public network, usually the Internet.



VPN was firstly developed by Microsoft in 1996. The purpose behind it is that to provide secure access of Internet network of company to the remote employees of the company, here remote employees are the employee who works from outside location. By doing this productivity of company get increased and observing this other companies also started using VPN services. A Virtual Private network extends a private network across public network and enables users to send and receive data across shared or public networks as their devices are directly get connected to the private network. The benefits of VPN consist of increase in functionality, security and management of the private network. A VPN is created by establishing a virtual point-to-point connection through the use of dedicated circuits or with tunneling protocols over existing networks. A VPN available from the Public Internet can provide some of the benefits of a Wide Area Network (WAN).



II. WORKING OF VPN

Virtual Private Networks use tunneling technologies to allow users to access private network resources through the Internet or other public network. When user get connected to secure Virtual Private Networks then their internet traffic passes through encrypted tunnel which is not seen to any entity such as hackers, government and not even by Internet Service Provider of user. Due to use of VPN data is not readable to these entities. To understand how VPN exactly works let us take two different situations first is without using VPN and second is with using VPN-

2.1. Without using VPN

When without using VPN, we access internet then Internet Service Provider (ISP) also connect on the site which is success by us as they are providing Internet Service to us. Internet Service Provider gives us unique IP address, as ISP handle our internet data, they can get to know what we are browsing on internet so in that case our privacy not remains secure. Internet Service Provider can see what we are browsing over the Internet.

2.2. With using VPN

When we get connected to Internet through VPN then VPN established secure connection through VPN server this works when we have proper application installed in our devices. Application of VPN which present in our device for usage of VPN service is known as VPN client. As our internet traffic is still passing through ISP, but ISP is unable to see the final destination of the traffics. And websites we visit cannot see our original IP address as VPN masked our IP address. A VPN creates a private tunnel within a public connection (ex. the internet). VPN software allows it users to send and receive data transfers securely. VPN use different types of VPN Protocols to encrypt web connections and make them private. Furthermore, different versions of VPN software exist.

III. TERMINOLOGIES

1.Virtual: - It means existing, seen or happening online or on a computer screen, rather than in person or in the physical world.

2.Private: - It means belonging to or intended for one particular person or group and not to be shared by others.

3.Network: - A network consists of two or more computers that are linked in order to share resources, exchange files, or electronic communications.

4.Encryption: - Encryption is a method by which information is converted into secret code that hides the information's true meaning.

5.WAN (Wide Area Network): - A wide area network (WAN) is a geographically distributed private telecommunication network that interconnects multiple local area networks (LANs).

6.ISP (Internet Service Provider): - Internet Service Provider (ISP) refers to a company that provides access to the internet to both personal and business customers.

7.IP Address: -An IP address is a unique address that identifies a device on the internet or a local network.

8.Tunnel: - It is a kind of passage, similarly in networking, tunnels are a method for transporting data across a network using protocols that are not supported by that network. Tunneling works by encapsulating packets: wrapping packets inside other packets.

VI. SECURITY AND PRIVACY OF VPN

VPN provides you online privacy and anonymity by creating a private network from a public internet connection. VPN mask your IP address so that your online activities are virtually untraceable. VPN service establishes secure and encrypted connections to provide greater privacy than a secured Wi-Fi hotspot. A virtual private network is a key privacy tool that should be used whenever logging onto the internet from a public place or any other spot that offers access to free public Wi-Fi. A VPN creates a type of tunnel that hides your online activities. Virtual Private Network can hide a lot of information that can put your privacy at high risk. Here are five of them-

- Your Browsing History
- Your IP Address
- Your location for streaming
- Your devices
- Your web activity – to maintain internet freedom

A virtual private network is an internet security service that allows users to access the internet as though they were connected to a private network. VPNs use encryption to create a secure connection over unsecured internet infrastructure. Encryption is a way of scrambling data so that only authorized parties can understand the information. It takes readable data and alters it so that it appears random to attackers or anyone else who intercepts it; so, we can say that encryption is "secret code". VPNs are one way to protect corporate data and manage user access to that data. VPNs protect data as users interact with apps and web properties over the Internet, and they can keep certain resources hidden.

V. TYPES OF VIRTUAL PRIVATE NETWORK

Virtual Private Network services mainly falls in four types, viz., Remote Access VPNs, Personal VPNs, Mobile VPNs and Site-to-Site VPNs.

5.1. Remote Access VPNs Service

A Remote Access VPN lets us use the internet to connect to a private network. The internet is an untrusted link in the communication. VPN encryption is used to keep the data private and secure as it travels to and from the private network. User can connect to another network using private encryption tunnel. User can get connect to public network or company network through this VPN. Remote Access VPN are also sometimes called client-based VPNs or client-to-server VPNs.

5.2. Personal VPNs Service

A Personal VPN service connects you to a VPN server which act as intermediate between your device and the online services which you want to access. The Personal VPN is sometimes called as a consumer or commercial VPN which encrypts our connection, hides your identity online and lets you spoof your geographic location. A Personal VPN service is differing from a Remote Access VPN as it doesn't give you access to a private network instead of that a Personal VPN works by giving you access to the public internet but over encrypted connection.

5.3. Mobile VPNs Service

While remote access VPNs let you connect to a local network from anywhere, they assume that the user will stay in one location. If the user disconnects the IP tunnel closes. A Mobile VPN is better option than a Remote Access VPN if the user is unlikely to have a stable connection, on the same network, for the entire session. With a mobile VPN, the VPN connection persists even if the user switches Wi-Fi or cellular network, loses connectivity, or switches their devices off for a while. A Mobile VPN can be used with any device and any connection: it doesn't have to be a mobile phone on a mobile network.

5.4. Site-to-Site VPNs Service

Whereas a Remote Access VPN is designed to let individual users to connect to a network and use its resources, a Site-to-Site VPN joins together two networks on different sites. It is also known as Router-to-Router VPN. It is mainly used in corporate environment which have their offices and headquarters at different locations, in this case Site-to-Site VPN creates a closed internal network where each and every location get connected with each other and connect them into a single network. Site-to-Site VPN is also known as intranet. Site-to-Site VPNs are also sometimes known as network-based VPNs.

Depending on who owns the networks being joined, there are generally two different forms of Site-to-Site VPN:

- **Intranet-Based VPN:** When the network being connected belong to a VPN single company, the combined VPN is known as an intranet-based VPN. This enables a company to establish a single wide area network (WAN) that spans two or more of its offices. Users in the company can access resources from other sites as easily as if they were on their own site.

- **Extranet-Based VPN:** When the networks being connected belong to different companies, the combined VPN is known as extranet-based VPN. An extranet VPN is used, for example, when a company wants to connect to its supplier's, so they can trade more efficiently.

VI. USES OF VIRTUAL PRIVATE NETWORK

- VPN is used whenever user feel that privacy is most vital factor for users.
- Users can use VPN while Travelling, Streaming, while using public Wi-Fi, Playing Games, Online Shopping. VPN can be used in that devices which can access internet.
- Devices like Laptops, Tablets, Smart phones, Voice assistant, Smart applications, Smart TV can access internet through VPN Services. VPN provides service on multiple platforms.
- As VPN are legal and it is used by the companies to protect their data from hackers and it is also used by individual.
- VPN are also used in the countries having highly respected government.
- VPN can be used for one's own security over public network.

VII. CONCLUSION

Security of data is very important for everyone, as there are many different sources and services are available in the market but among them Virtual Private Networks services provides better privacy and security to the users. VPN hides our original online identity and activities or work we perform on internet from hackers and third-party applications. User can switch as often as he or she lies with no limits using different locations; using VPN one can download or access content from anywhere; VPN can access on various platforms like Windows, Mac, iOS, Android, Linux, routers, etc. Virtual Private Networks make users relax in terms of online security and privacy.

REFERENCE

1. Kanuga Karuna Jyothi, Dr. B. Indira Reddy "Study on Virtual Private Network (VPN), VPN's Protocols And Security", Int © 2018 IJSRCSEIT | Volume 3 | Issue 5.
2. Komalpreet Kaur, Arshdeep Kaur "A Survey of Working on Virtual Private Network" © 2019 IRJET | Volume 6 | Issue 9.
3. <https://scholar.google.com/citations?hl=en&user=OOI01CwAAAAJ>
4. <https://www.servercake.blog/types-virtual-private-network-vpn/>
5. <https://www.geeksforgeeks.org/types-of-virtual-private-network-vpnand-its-protocols/>
6. D. Simion, M.F. Ursuleanu, A. Graur, A.D. Potorac, A. Lavric "Efficiency Consideration for Data Packets Encryption with in Wireless Tunneling for Video Streaming" INT J COMPUT COMMUN 8(1):136-145
7. <https://whatismyipaddress.com/vpn-comparison>
8. <https://scholar.google.com/citations?hl=en&pli=1&user=ks9yhS0AAAAJ>
9. Charlie Scotte et al., "Virtual Private Network" Second Edition, O'Reilly, January



Journal of Emerging Technologies and Innovative Research

An International Open Access Journal Peer-reviewed, Refereed Journal

www.jetir.org | editor@jetir.org **An International Scholarly Indexed Journal**

Certificate of Publication

The Board of

Journal of Emerging Technologies and Innovative Research (ISSN : 2349-5162)

Is hereby awarding this certificate to

Prof. Sunita Totade

In recognition of the publication of the paper entitled

An Empirical Study on Virtual Private Network

Published In JETIR (www.jetir.org) ISSN UGC Approved (Journal No: 63975) & 7.95 Impact Factor

Published in Volume 9 Issue 11 , November-2022 | Date of Publication: 2022-11-19

Parisa P

EDITOR

JETIR2211292

[Signature]

EDITOR IN CHIEF

Research Paper Weblink <http://www.jetir.org/view?paper=JETIR2211292>

Registration ID : 504484



Cyber Security Awareness

¹Dr. V.R.Dhawale, ²Prof. Rana Afreen Sheikh, ³Rajendra Dahare, ⁴Nirav Girnari, ⁵Pratik Perkawar, ⁶Palash Sahare

¹Professor, K. K. Wagh Institute of Engineering Education and Research, Nashik, India

²Professor, Department of MCA, Vidyabharati Mahavidyalaya, Amravati, India

^{3,4,5,6}Student, Department of MCA, Vidyabharati Mahavidyalaya, Amravati, India

Authors E-mail: 2ranafreen.sheikh@gmail.com, 3mrrajendradahare@gmail.com, 4niravgirnari91@gmail.com,
5pratikperkawar@gmail.com, 6palashsahare4444@gmail.com

Abstract - In this research paper we will talk about cyber security awareness. What is cyber security and why we need awareness of cyber security. Today we all are using internet and sharing our data and information on internet so it can be theft or steal so we need to have its security by all manner. We will also cover market of cyber security. We will also know that what are threats that we will face on internet. We will also talk on why awareness of cyber security is required. Why it is needed, and what benefits we will get after knowing it. This whole research paper is based on awareness of security of our data on internet. We will also talk how to get awareness of cyber security. And we will discuss what information of ours may be stolen via internet.

Keywords: Cyber Security, Awareness, data, information, internet, threats.

I. INTRODUCTION

Today we are all using internet and sharing our data via internet. We are using lots of web applications like Gmail, maps, digital locker, etc. So, we need security to keep our data securely because some hackers can steal our data and they can communicate to world by using our identity. They can steal any type of our data.

Like they can also view our bank details and they can also read our messages etc. We need to have awareness regarding our digital data in today's world. Because data is our most precious thing in this world. There are various ways that we can spread awareness of cyber security. We must think on every part of security. So, users get proper awareness of cyber security. There are also threats in the IT world. There is various type of users on internet so we have to give different type of awareness.

II. METHODOLOGY

We had read different types of research paper on cyber security awareness. So, we got much information regarding cyber security awareness. We knew that what are methods applied by some organizations to spread awareness among

users. We have read articles on cyber security awareness so we got knowledge that how to get awareness and what are key points we need to consider in this topic. We collected our points on bases of research we have done. We had read on many types of points for this paper and we selected some necessary points to prepare this paper.

What is cyber security awareness?

In today's world we live in the age of information technology we have our data online or on the servers. So, we must be aware that this information can be stolen by hackers and they can use it for they own interest. So, need it to keep our data secure. Some of our data is very important like bank details. If hacker has all details of our bank account and we had interaction with hacker then they can easily transfer our cash to their account. It is very risky. If we are chatting with anybody then hacker can hack our phone and read our messages. They can also send messages from our behalf.

Why awareness of cyber security is required?

We have lots of data in our devices. If we do not use antivirus software then our data can be stolen by hackers and they can have access of our data. Likewise, if have stored our data on servers then servers need security like we can add firewall to servers so we can keep our data secure. We all have some sensitive information like bank details so these types information we must keep securely. Because if it is stolen then we may be lost our cash. Some people called as hacker can steal our data and they can use it on their interest. They do not care about our identity or security.

III. THREATS OF IT

Hacking is the main threat to the IT field. We must be aware that our data can be stolen by hackers whether it is available online (servers) or offline (devices). We must use antivirus software in our devices so hackers cannot steal our data. In today's world email is primary tool for communication and hackers are sending emails to hack our data so we must be aware of that and we must not click on untheorized links that we get in our email account. There are

some viruses called malware and ransomwares they can hack our devices if they enter on our device by email. Web browser is hot target for hackers because it is directly connected to the internet and we visit every website by it. We have our data regarding to our activity in the browser.

IV. CYBER SECURITY AS AN OPPORTUNITY

We can run training program on awareness of cyber security so user can get knowledge regarding it. So, people also can earn from it. We can find jobs in the field of cyber security and earn money. Cyber security is very important field in today's world because we have seen data theft and hacking so we must keep our data securely. There are varieties of work in cyber security we can learn anything that comes in our interest area. We can start spreading awareness from school and colleges because as we know that today's students are spending so much their time on social media. Security of their data is very important. Security is also very important to servers so we can find jobs to companies to maintain security to their servers.

How to get awareness?

We can participate in different programs of cyber security awareness. We can learn about cyber security on internet by doing some certification course on awareness. We must know that hackers can hack our devices by sending emails so we must be aware on that we should not click any unauthorized link of emails. We should read daily news on cyber security and about hackers that how they are hacking our devices. What are tricks they are using to hack our devices. We must not click on unrecognized links which is received by messages. And most important do not trap in "lottery scheme" notifications which we receive in our phone. It is also a cyber-crime.

What information of ours can be stolen?

Hackers can steal our bank details, emails, messages, contacts, images, videos, etc. So, we must keep these things securely. We must not keep password in written in any file. And if we doing this then do not keep these types of files on internet. We must use antivirus software in our devices so we can keep them secure. If we are putting some type of data on the internet then we must have trust on that site. Our name, mobile number and occupation is very important because some organizations are purchasing this information to sell their products / services. Like we have seen that banks are calling us to buy their credit cards and for giving loans to us. So, from somewhere they are getting this information.

V. CONCLUSION

In this paper we have discussed on various types of points regarding to cyber security awareness. Like what is cyber security awareness? how to get awareness? threats of IT etc. By this paper user can get awareness on cyber security. And they can learn lots of things from it. We discussed on some opportunity that we can get from cyber security.

REFERENCES

- [1] <https://arxiv.org/abs/1901.02672>
- [2] <https://www.tandfonline.com/doi/abs/10.1080/0144929x.2012.708787>
- [3] <https://www.tandfonline.com/doi/abs/10.1080/08874417.2020.1712269>
- [4] <https://journals.co.za/doi/abs/10.10520/EJC154952>
- [5] <https://www.worldscientific.com/doi/abs/10.1142/S0219649216500076>
- [6] <https://www.kaspersky.com/resource-center/preemptive-safety/7-ways-to-cyberattack-vulnerability>
- [7] <https://www.ogl.co.uk/the-importance-of-cyber-security-awareness>

AUTHORS BIOGRAPHY



Rajendra Dahare, Student, Department of MCA, Vidyabharati Mahavidyalaya, Amravati, India.



Nirav Girnari, Student, Department of MCA, Vidyabharati Mahavidyalaya, Amravati, India.



Citation of this Article:

Dr. V.R.Dhawale, Prof. Rana Afreen Sheikh, Rajendra Dahare, Nirav Girnari, Pratik Perkawar, Palash Sahare, “Cyber Security Awareness” Published in *International Research Journal of Innovations in Engineering and Technology - IRJIET*, Volume 6, Issue 10, pp 116-118, October 2022. Article DOI <https://doi.org/10.47001/IRJIET/2022.610020>



ISSN(online): 2581-3048

Impact Factor : 5.87

CERTIFICATE OF PUBLICATION

INTERNATIONAL RESEARCH JOURNAL OF INNOVATIONS IN ENGINEERING AND TECHNOLOGY

Is Hereby Awarding this Certificate to

Prof. Rana Afreen Sheikh

**Professor, Department of MCA,
Vidyabharati Mahavidyalaya, Amravati, India**

In Recognition of the Publication of Manuscript Entitled

Cyber Security Awareness

*Published in International Research Journal of Innovations in
Engineering and Technology (IRJIET)*

Volume 6, Issue 10, pp 116-118, October 2022

<https://doi.org/10.47001/IRJIET/2022.610020>

Manuscript ID : IRJIET610020

Date of Issue : November 14, 2022



Editor-In-Chief
IRJIET

Managing Editor
IRJIET

Mail us at: editor@irjiet.com / irjietjournal@gmail.com

Journal Website : www.irjiet.com

Role of Web Technology in Primary Education

¹Prof. S.B. Bele, ²Aarti Parve, ³Linta Nandgaonkar, ⁴Hema Wankar, ⁵Vaishnavi Rajas

¹Assisnant Professor, Department of MCA, Vidyabharati Mahavidyalaya, Amravati, India

^{2,3,4,5}Student, Department of MCA, Vidyabharati Mahavidyalaya, Amravati, India

Abstract - World is changing day by day using technologies. Web technology plays vital role in educational field. Technology is progressing rapidly. Great amount of educational information is readily available and easily accessible throughout the web. There are lots of technologies like internet, email, online teaching platforms, etc. used by the teachers and students to communicate. In this research paper we are specially focusing on the 'Role of Web Technology in Primary Education'.

Keywords: Primary Education, Technology in Education, Student Life-changing Technology.

I. INTRODUCTION

Now days there are different types of teaching methodologies in primary education. Various online educational platforms are open to all students. Use of web technology in education sector has some advantages and disadvantages too. In this research paper, we are going to know about not only Web Technology but also the changing life of students due to this technology.

II. WEB TECHNOLOGY

Web technology is defined as the various tools and technologies that are used as a medium of communication between server and user/client using internet. Web Technology is the combination of World Wide Web (WWW), web browser, web server, web pages. WWW is the collection of documents, connected through links, spread out over thousands of computer over the internet and these pages over the web can be accessed using a web browser. It includes building, creating, and maintaining of websites. Furthermore it contains aspects such as web design, web publishing, web programming, and database management.



Primary Education

In India, Primary education is of eight years. Students' ages from 6 to 14 comes under the primary education. Mainly primary education has two stages first one is 'primary stage' from class 1 to 5 and second one is 'upper primary stage' from class 6 to 8. It is a time of remarkable growth with brain development.

For primary students, technology can be used to build fundamental skills to prepare them for future independent learning. Students are using interactive games to reinforce math, spelling, phonetic, and reading skills. Sites like Spelling training permit students our teachers to upload their own word lists to practice word pronunciation and create interactive games. Parents can also use these sites to exercise fundamental skills beyond the walls of the classroom.

In primary students, Technology help to teach Math and English and indeed other curriculum in a fun manner and put these subjects into environment making them easier to digest and more understandable to younger primary age pupils.

Technology in Teaching and Learning

It is very easy to know or learn something. Today we all are using technology and so all the educational information is easily available throughout the web. Just one click on Google, we come to know the answer to our questions. Also on you tube there are lot of tutorials to watch and study. The teacher will simply switch on or switch off radio programs, television programs, play DVDs or CDs that contain educational programs. Technology is a teaching tool, but can never replace a teacher. It is being used to facilitate and lightens the work of the teacher. Technology has modernized the teaching-learning environment.

While the teacher utilize technology as the tool for learning, likewise it is an effective tool, it makes learning interesting, easy and effective. Even the teachers who are teaching can utilize similar tools for learning. Students are using these learning tools and completing their academic task effectively using the technology. The topics they are studying are also becoming increasingly clear to them because of availability of information in images, videos and animation format.

III. IMPACT OF TECHNOLOGY ON EDUCATION

Technology is changing in such way that we learn from zero to hero. Education is one of the most impacted areas of technology. According to our research, there is some positive impact as well as some negative impact of technology on education.

Positive Impact

1) Easier Access to Information: The books can be stored in the classroom because the information that they need is easily accessed on a computer.

2) Broaden the mind: Before the internet, children knew what their family thought them. This caused their political and religious views to be the only ones that the children knew. Having access to the technology will expose them to things outside of their parents' interest and help them to form their own opinion.

3) Prepare student for the future: If the students are well-versed on using technology to collaborate and communicate as early as now. They will not have trouble fitting in, competing and finding jobs in the future.

4) Promotes exciting way to educate students: Since there are lots of images, videos and other graphics and text that may be found in your computer, more students would feel the excitement in studying through the use of the gadget, this is very important in order to arouse their interest in studying.

5) Make students challenging: Education is becoming challenging day by day. Educational websites, applications, games like word puzzle, sudoku, logical games make students' mind challenging that will be better for their future life. If students learn how to face challenges today, it will become easy for them to face challenges in their next educations.

Negative Impact

1) Expensive: Technology required in the education includes internet services, computer hardware, printers, internet browser etc. Acquiring all of these needs a lot of money to invest in technology at your school.

2) Distraction: Cartoons are becoming very popular among the children. They are spending most of their time, which is having a bad effect on their studies.

3) Students acquiring wrong information: Many websites and other online platforms are providing wrong information that users require. Many websites post deceptive material online in order to increase traffic, even if the information is inaccurate.

4) Health issue: Media devices may lead to psychological and physical issues, such as eyestrain and difficulty focusing on an important task. They may also contribute to more serious health conditions, such as depression.

Improvement of students due to technology



IV. CONCLUSION

Technology is being used more and more in education. Students are progressing highly using these technologies. Getting the prior knowledge is very helpful for students in their education. Students can get prior knowledge of their higher education so that whenever they will go for their higher education, it will not be difficult to learn and understand. Technology is very helpful for not only students but also teachers. Teachers are using new concepts for teaching using the technology. In some schools, teachers are providing 3D animated concepts so that students will visually understand all the concepts easily. But technology can not replace a teacher. They both together will make a bright student.

REFERENCES

- [1] https://en.wikipedia.org/wiki/Primary_education
- [2] https://www.researchgate.net/publication/303703009_Web_technology_based_innovations_in_education_sector
- [3] <https://www.dezzain.com/technology/technology-effects-on-education/>
- [4] <https://happynumbers.com/blog/list-of-teaching-methodologies-primary-school/>
- [5] https://prezi.com/t7_x60fsmas4/a-summary-of-the-role-of-web-technology-in-education/

Citation of this Article:

Prof. S.B. Bele, Aarti Parve, Linta Nandgaonkar, Hema Wankar, Vaishnavi Rajas, “Role of Web Technology in Primary Education” Published in *International Research Journal of Innovations in Engineering and Technology - IRJIET*, Volume 6, Issue 10, pp 136-138, October 2022. Article DOI <https://doi.org/10.47001/IRJIET/2022.610026>



ISSN(online): 2581-3048

Impact Factor : 5.87

CERTIFICATE OF PUBLICATION

INTERNATIONAL RESEARCH JOURNAL OF INNOVATIONS IN ENGINEERING AND TECHNOLOGY

Is Hereby Awarding this Certificate to

Prof. S.B. Bele

**Assisant Professor, Department of MCA,
Vidyabharati Mahavidyalaya, Amravati, India**

In Recognition of the Publication of Manuscript Entitled

Role of Web Technology in Primary Education

*Published in International Research Journal of Innovations in
Engineering and Technology (IRJIET)*

Volume 6, Issue 10, pp 136-138, October 2022

<https://doi.org/10.47001/IRJIET/2022.610026>

Manuscript ID : IRJIET610026

Date of Issue : November 24, 2022



Editor-In-Chief
IRJIET

Managing Editor
IRJIET

Mail us at: editor@irjiet.com / irjietjournal@gmail.com

Journal Website : www.irjiet.com



G. S. Tompe Mahavidyalaya Sarvajani Trust's
G. S. Tompe Arts, Commerce & Science College

Chandur Bazar, Dist. Amravati, Maharashtra
Affiliated to Sant Gadge Baba Amravati University Amravati. NAAC Re-accredited B++ CGPA 2.89 ISO 9001:2015 Certified
DST-SERB, New Delhi Sponsored

2nd INTERNATIONAL CONFERENCE
on Recent Advances in Material Science & Nanotechnology

(RAMAN-22)

12th to 14th May 2022

CERTIFICATE

This is to Certify that Dr. G. T. Lamdhade of Dept. of Physics

Vidya Bharati Mahavidyalaya, Amravati has actively Participated/Presented paper/Delivered keynote address/Plenary Talk/Invited Talk/Chaired session in an 2nd International Conference on Recent Advances in Material Science & Nanotechnology (Raman-2022) on the topic: 1) Investigation of Nonlinear.. 2) I-V chara. of CuO.. 3) One Step synt.., 4) To study thermal stability of nanocrystalline TiO₂. Organized by Department of Physics G.S. Tompe Arts Commerce & Science College, Chandur Bazar & In Collaboration with I.Q.A.C., Sant Gadge Baba Amravati University, Amravati.


Chairman
Dr. R. S. Ramteke
Principal
G. S. Tompe Mahavidyalaya


Chairman
Prof. Dr. S. A. Waghuley
Director, IQAC SGBAU, Amravati


Convener
Dr. D. R. Bijwe
Dept. of Physics
G. S. Tompe Mahavidyalaya


Organizing Secretary
Dr. P. S. Deole
Head Dept. of Physics
G. S. Tompe Mahavidyalaya

RAMAN-115**Investigation of Nonlinear I-V characteristics of nanocrystalline tin oxide**

[#]Balkhande V.M., Abdul Tantray, Raulkar K.B., Lamdhade G.T.

[#]Dept. of Physics, Prof. Ram Meghe Institute of Tech. & Research, Badnera-Amravati
M.S. India-444 701

Dept. of Physics, Vidya Bharati Mahavidyalaya, Amravati M.S. India-444 602

*Corresponding Author Email: vidya_balkhande@yahoo.com,oumgajanan@gmail.com

Abstract:

Current versus voltage characteristics (I-V) of nanocrystalline tin oxide (SnO₂) has been investigated at various temperatures (from 50°C to 350°C) in air. The characteristics were measured by using Keithley 6487 voltage source cum picoammeter. The nanocrystalline powder of SnO₂ was prepared by the liquid phase method and samples were prepared via spray pyrolysis technique in the form of thin films on an optically plane and clean glass surface. X-ray diffraction studies showed a tetragonal rutile structure for the samples. Microstructural studies were performed with scanning electron microscopy. The nanocrystalline tin oxide exhibited nonlinear I-V characteristics of the current-controlled negative resistance type (NTC). The results show that the threshold field (break down) voltage is higher due to the formation of a tin oxide layer over the crystalline tin. It is also found that the threshold field increases with the decrease in grain size.

Keywords: I-V characteristics, nanocrystalline tin oxide, liquid phase, XRD, SEM

RAMAN-116**I-V characteristics of Cupric Oxide (CuO) thin film**

[#]Balkhande V.M., Abdul Tantray, Raulkar K.B., Lamdhade G.T.

[#]Dept. of Physics, Prof. Ram Meghe Institute of Tech. & Research, Badnera-Amravati M.S.
India-444 701

Dept. of Physics, Vidya Bharati Mahavidyalaya, Amravati M.S. India-444 602

*Corresponding Author Email: vidya_balkhande@yahoo.com,oumgajanan@gmail.com

Abstract:

Cupric oxide (CuO), having a narrow bandgap of 1.2 eV and a variety of chemophysical properties, is recently attractive in many fields such as energy conversion, optoelectronic devices, and catalysts. Compared with bulk material, the advanced properties of CuO nanostructures have been demonstrated. Current versus voltage characteristics (I-V) of nanocrystalline cupric oxide (CuO) has been investigated at various temperatures (from 50°C to 350°C) in air. The characteristics were measured by using Keithley 6487 voltage source cum picoammeter. The nanocrystalline powder of cupric oxide (CuO) was prepared by the liquid phase method and samples were prepared via spray pyrolysis technique in the form of thin films. X-ray diffraction studies showed a monoclinic structure of CuO and no other peaks are observed. Microstructural studies were performed with scanning electron microscopy. The nanocrystalline cupric oxide exhibited nonlinear I-V characteristics of the current-controlled negative resistance type (NTC).

Keywords: I-V characteristics, nanocrystalline cupric oxide, liquid phase, XRD, SEM



2nd INTERNATIONAL CONFERENCE ON RECENT ADVANCES IN MATERIAL SCIENCE AND NANOTECHNOLOGY

SOUVENIR



RAMAN 2022

12th - 14th May 2022



Organized by

Department of Physics

G. S. Tompe Arts, Commerce & Science College,

Chandur Bazar

In Collaboration with

I.Q.A.C.,

Sant Gadge Baba Amravati University, Amravati. (M. S.)

In Association with

International Journal of Scientific Research in Science and Technology

Online ISSN : 2395-602X | Print ISSN : 2395-6011

2ND INTERNATIONAL CONFERENCE

**RECENT ADVANCES IN MATERIAL SCIENCE
AND NANOTECHNOLOGY**

RAMAN 2022

Organized by
Department of Physics
G. S. Tompe Arts, Commerce & Science College, Chandur Bazar
In Collaboration with
I.Q.A.C.,
Sant Gadge baba Amravati University

● Editors ●

Dr. P. S. Deole
Dr. D. R. Bijwe
Prof. M. N. Pawar
Dr. A. V. Rajgure

● Publisher ●

Dr. R. S. Ramteke

Principal

G. S. Tompe Arts, Commerce
& Science College,
Chandur Bazar, Dist.- Amravati (M. S.)

● Date of Publication : 12th May 2022 ●

RAMAN 2022

2ND INTERNATIONAL CONFERENCE

RECENT ADVANCES IN MATERIAL SCIENCE
AND NANOTECHNOLOGY

RAMAN 2022

Organized by
Department of Physics
G. S. Tompe Arts, Commerce & Science College, Chandur Bazar
In Collaboration with
I.Q.A.C.,
Sant Gadge baba Amravati University

• Patron •



Hon. Mr. Keshavdada Tompe
Chairman, G. S. Tompe Mahavidyalaya
Sarvajanic Trust's, Chandur Bazar

• Co-Patron •



Hon. Mr. Bhaskardada K. Tompe
Secretary, G. S. Tompe Mahavidyalaya
Sarvajanic Trust's, Chandur Bazar



Dr. Vijay K. Tompe
Member, G. S. Tompe Mahavidyalaya
Sarvajanic Trust's, Chandur Bazar

• Chairman •



Dr. R. S. Ramteke
Principal
G. S. Tompe Mahavidyalaya

• Chairman •



Prof. Dr. S. A. Waghuley
Director, IQAC SGBAU, Amravati

• Convener •



Dr. D. R. Bijwe
Dept. of Physics
G. S. Tompe Mahavidyalaya

• Org. Secretary •



Dr. P. S. Deole
HOD Dept. of Physics
G. S. Tompe Mahavidyalaya

Co-Convener



Dr. A. V. Rajgure
Dept. of Physics
G. S. Tompe Mahavidyalaya



Ms. M.N Pawar
Dept. of Physics
G. S. Tompe Mahavidyalaya



Dr. A. R. Bansod
Dept. of Physics
Dr. Ambedkar College Nagpur



Dr. U. P. Manik
Dept. of Physics
Sardar Patel Mahavidyalaya,
Chandrapur

CONTENTS

Sr. No	TITLE	AUTHORS	Page No.
	<u>KEYNOTE ADDRESS</u>		
1	Raman spectroscopy for characterisation of materials	Dr. V. Natranjan	1
	<u>PLANARY TALKS</u>		
1	Holographic acoustooptic display	Dr. Vladimir Petrov,	2
2	Phosphors and there applications	Dr. S. J. Dhoble	2
3	Dielectric relaxation in conducting polymers	Prof. Dr. S. P. Yawale	3
4	Oxide Nanomechanics	Dr. Luca Pellegrino	4
5	Perovskite: fullerene bulk heterojunction using microemulsion scheme: Modelling, Simulation and experimental studies	Dr. Jaydeep V. Sali	5
6	Nanostructured Metal Oxide Thin Films for Supercapacitor Applications	Dr. Abhijit A. Yadav	6
7	Sol gel derived advanced materials aeroagels, coatings and gas sensors	Dr. Digambar Y. Nadargi	6
8	Elastic, Mechanical and Thermal Properties of Rare-earth Materials by Ultrasonic Analysis	Dr. Devraj Singh	7
9	Tailoring of Perovskites films for Efficient and Stable Building Integrated Perovskite Solar cell devices	Dr. Dharendra Kumar Chaudhary	8
	<u>INVITED TALKS</u>		
1	Carbon Rich Materials: A Possible Source of Various Forms of Carbon for futuristic applications	Dr. D. E. Kshirsagar	9
2	Commercialization of Academic Research: Business perspective	Dr. Abhay D. Deshmukh	10

114	Synthesis and characterization of $KAlF_4: Ho^{3+}/Yb^{3+}$ upconversion phosphor	K. S. Janbandhu, V. B. Pawade, S. J. Dhoble	72
115	Investigation of Nonlinear I-V characteristics of nanocrystalline tin oxide	Balkhande V.M., Abdul Tantray, Raulkar K.B., Lamdhade G.T.	73
116	I-V characteristics of Cupric Oxide (CuO) thin film	Balkhande V.M., Abdul Tantray, Raulkar K.B., Lamdhade G.T.	73
117	One Step Synthesis of Nano crystalline MgO	A.S. Daware *, Abdul Tantray, Raulkar K.B., G. T. Lamdhade	74
118	To study thermal stability of nanocrystalline Titanium oxide	Balkhande V.M., Abdul Tantray, Raulkar K.B., Lamdhade G.T.	74
119	Investigate transport properties of metal oxide doped polyindole composites.	A. A. Dubey, G. R. Dhokane	75
120	Photovoltaic Response of TiO_2 Nanoparticle doped materials	D. J. Bhagat , G. R. Dhokane	75
121	Combustion synthesis of $CaAlBO_4: Eu^{3+}$ phosphor for solid state lighting	R. T. Maske, A.N. Yerpude, S. J. Dhoble	76
122	Ultrasonic Velocity, Adiabatic Compressibility, Intermolecular Free Length and Other Acoustical Parameters of Leaf Extract Solution of Cymbopogon	S. S. Kamble, S. R. Aswale, S. S. Aswale	77
123	Synthesis and characterization of pure and Mg^{2+} doped ZnS nanoparticles	Sandeep A. Waghuley Nikita S. Korde	78
124	Nurturing Interdisciplinary Competencies for Innovations in Science	Dr. U. R. Kanerkar, Dr. Y. M. Rajgure Dr. D. R. Bijwe	79



G. S. Tompe Mahavidyalaya Sarvajanic Trust's
G. S. Tompe Arts, Commerce & Science College

Chandur Bazar, Dist. Amravati, Maharashtra
Affiliated to Sant Gadge Baba Amravati University, Amravati. NAAC Re-credited B++ CGPA 2.89 ISO 9001:2015 Certified

DST-SERB, New Delhi Sponsored

2nd INTERNATIONAL CONFERENCE
on Recent Advances in Material Science & Nanotechnology

(RAMAN-22)

12th to 14th May 2022

CERTIFICATE

This is to Certify that Dr. K. B. Raulkar of Dept. of Physics

Vidya Bharati Mahavidyalaya, Amravati

has actively Participated/Presented
paper/Delivered keynote address/Plenary Talk/Invited Talk/Chaired session in an 2nd International Conference on Recent
Advances in Material Science & Nanotechnology (Raman-2022) on the topic: 1) Investigation of Nonlinear, 2) I-V chara. of CuO, 3) One Step synt., 4) To study thermal stability of nanocrystalline TiO₂. Organized by Department of Physics G.S. Tompe
Arts Commerce & Science College, Chandur Bazar & In Collaboration with I.Q.A.C., Sant Gadge Baba Amravati University, Amravati.

Chairman
Dr. R. S. Ramteke
Principal
G. S. Tompe Mahavidyalaya

Chairman
Prof. Dr. S. A. Waghuley
Director, IQAC SGBAU, Amravati

Convener
Dr. D. R. Bijwe
Dept. of Physics
G. S. Tompe Mahavidyalaya

Organizing Secretary
Dr. P. S. Deole
Head Dept. of Physics
G. S. Tompe Mahavidyalaya

RAMAN-115**Investigation of Nonlinear I-V characteristics of nanocrystalline tin oxide**

[#]Balkhande V.M., Abdul Tantray, Raulkar K.B., Lamdhade G.T.

[#]Dept. of Physics, Prof. Ram Meghe Institute of Tech. & Research, Badnera-Amravati
M.S. India-444 701

Dept. of Physics, Vidya Bharati Mahavidyalaya, Amravati M.S. India-444 602

*Corresponding Author Email: vidya_balkhande@yahoo.com,oumgajanan@gmail.com

Abstract:

Current versus voltage characteristics (I-V) of nanocrystalline tin oxide (SnO₂) has been investigated at various temperatures (from 50°C to 350°C) in air. The characteristics were measured by using Keithley 6487 voltage source cum picoammeter. The nanocrystalline powder of SnO₂ was prepared by the liquid phase method and samples were prepared via spray pyrolysis technique in the form of thin films on an optically plane and clean glass surface. X-ray diffraction studies showed a tetragonal rutile structure for the samples. Microstructural studies were performed with scanning electron microscopy. The nanocrystalline tin oxide exhibited nonlinear I-V characteristics of the current-controlled negative resistance type (NTC). The results show that the threshold field (break down) voltage is higher due to the formation of a tin oxide layer over the crystalline tin. It is also found that the threshold field increases with the decrease in grain size.

Keywords: I-V characteristics, nanocrystalline tin oxide, liquid phase, XRD, SEM

RAMAN-116**I-V characteristics of Cupric Oxide (CuO) thin film**

[#]Balkhande V.M., Abdul Tantray, Raulkar K.B., Lamdhade G.T.

[#]Dept. of Physics, Prof. Ram Meghe Institute of Tech. & Research, Badnera-Amravati M.S.
India-444 701

Dept. of Physics, Vidya Bharati Mahavidyalaya, Amravati M.S. India-444 602

*Corresponding Author Email: vidya_balkhande@yahoo.com,oumgajanan@gmail.com

Abstract:

Cupric oxide (CuO), having a narrow bandgap of 1.2 eV and a variety of chemophysical properties, is recently attractive in many fields such as energy conversion, optoelectronic devices, and catalysts. Compared with bulk material, the advanced properties of CuO nanostructures have been demonstrated. Current versus voltage characteristics (I-V) of nanocrystalline cupric oxide (CuO) has been investigated at various temperatures (from 50°C to 350°C) in air. The characteristics were measured by using Keithley 6487 voltage source cum picoammeter. The nanocrystalline powder of cupric oxide (CuO) was prepared by the liquid phase method and samples were prepared via spray pyrolysis technique in the form of thin films. X-ray diffraction studies showed a monoclinic structure of CuO and no other peaks are observed. Microstructural studies were performed with scanning electron microscopy. The nanocrystalline cupric oxide exhibited nonlinear I-V characteristics of the current-controlled negative resistance type (NTC).

Keywords: I-V characteristics, nanocrystalline cupric oxide, liquid phase, XRD, SEM



2nd INTERNATIONAL CONFERENCE
ON
**RECENT ADVANCES IN MATERIAL SCIENCE
AND NANOTECHNOLOGY**

SOUVENIR



RAMAN
2022

12th - 14th May 2022



Organized by

Department of Physics

G. S. Tompe Arts, Commerce & Science College,

Chandur Bazar

In Collaboration with

I.Q.A.C.,

Sant Gadge Baba Amravati University, Amravati. (M. S.)

In Association with

International Journal of Scientific Research in Science and Technology

Online ISSN : 2395-602X | Print ISSN : 2395-6011

2ND INTERNATIONAL CONFERENCE

**RECENT ADVANCES IN MATERIAL SCIENCE
AND NANOTECHNOLOGY**

RAMAN 2022

Organized by
Department of Physics
G. S. Tompe Arts, Commerce & Science College, Chandur Bazar
In Collaboration with
I.Q.A.C.,
Sant Gadge baba Amravati University

● Editors ●

Dr. P. S. Deole
Dr. D. R. Bijwe
Prof. M. N. Pawar
Dr. A. V. Rajgure

● Publisher ●

Dr. R. S. Ramteke

Principal

G. S. Tompe Arts, Commerce
& Science College,
Chandur Bazar, Dist.- Amravati (M. S.)

● Date of Publication : 12th May 2022 ●

RAMAN 2022

2ND INTERNATIONAL CONFERENCE

RECENT ADVANCES IN MATERIAL SCIENCE
AND NANOTECHNOLOGY

RAMAN 2022

Organized by
Department of Physics
G. S. Tompe Arts, Commerce & Science College, Chandur Bazar
In Collaboration with
I.Q.A.C.,
Sant Gadge baba Amravati University

• Patron •



Hon. Mr. Keshavdada Tompe
Chairman, G. S. Tompe Mahavidyalaya
Sarvajanic Trust's, Chandur Bazar

• Co-Patron •



Hon. Mr. Bhaskardada K. Tompe
Secretary, G. S. Tompe Mahavidyalaya
Sarvajanic Trust's, Chandur Bazar



Dr. Vijay K. Tompe
Member, G. S. Tompe Mahavidyalaya
Sarvajanic Trust's, Chandur Bazar

• Chairman •



Dr. R. S. Ramteke
Principal
G. S. Tompe Mahavidyalaya

• Chairman •



Prof. Dr. S. A. Waghuley
Director, IQAC SGBAU, Amravati

• Convener •



Dr. D. R. Bijwe
Dept. of Physics
G. S. Tompe Mahavidyalaya

• Org. Secretary •



Dr. P. S. Deole
HOD Dept. of Physics
G. S. Tompe Mahavidyalaya

Co-Convener



Dr. A. V. Rajgure
Dept. of Physics
G. S. Tompe Mahavidyalaya



Ms. M.N Pawar
Dept. of Physics
G. S. Tompe Mahavidyalaya



Dr. A. R. Bansod
Dept. of Physics
Dr. Ambedkar College Nagpur



Dr. U. P. Manik
Dept. of Physics
Sardar Patel Mahavidyalaya,
Chandrapur

CONTENTS

Sr. No	TITLE	AUTHORS	Page No.
	<u>KEYNOTE ADDRESS</u>		
1	Raman spectroscopy for characterisation of materials	Dr. V. Natranjan	1
	<u>PLANARY TALKS</u>		
1	Holographic acoustooptic display	Dr. Vladimir Petrov,	2
2	Phosphors and there applications	Dr. S. J. Dhoble	2
3	Dielectric relaxation in conducting polymers	Prof. Dr. S. P. Yawale	3
4	Oxide Nanomechanics	Dr. Luca Pellegrino	4
5	Perovskite: fullerene bulk heterojunction using microemulsion scheme: Modelling, Simulation and experimental studies	Dr. Jaydeep V. Sali	5
6	Nanostructured Metal Oxide Thin Films for Supercapacitor Applications	Dr. Abhijit A. Yadav	6
7	Sol gel derived advanced materials aeroagels, coatings and gas sensors	Dr. Digambar Y. Nadargi	6
8	Elastic, Mechanical and Thermal Properties of Rare-earth Materials by Ultrasonic Analysis	Dr. Devraj Singh	7
9	Tailoring of Perovskites films for Efficient and Stable Building Integrated Perovskite Solar cell devices	Dr. Dharendra Kumar Chaudhary	8
	<u>INVITED TALKS</u>		
1	Carbon Rich Materials: A Possible Source of Various Forms of Carbon for futuristic applications	Dr. D. E. Kshirsagar	9
2	Commercialization of Academic Research: Business perspective	Dr. Abhay D. Deshmukh	10

114	Synthesis and characterization of $KAlF_4: Ho^{3+}/Yb^{3+}$ upconversion phosphor	K. S. Janbandhu, V. B. Pawade, S. J. Dhoble	72
115	Investigation of Nonlinear I-V characteristics of nanocrystalline tin oxide	Balkhande V.M., Abdul Tantray, Raulkar K.B., Lamdhade G.T.	73
116	I-V characteristics of Cupric Oxide (CuO) thin film	Balkhande V.M., Abdul Tantray, Raulkar K.B., Lamdhade G.T.	73
117	One Step Synthesis of Nano crystalline MgO	A.S. Daware *, Abdul Tantray, Raulkar K.B., G. T. Lamdhade	74
118	To study thermal stability of nanocrystalline Titanium oxide	Balkhande V.M., Abdul Tantray, Raulkar K.B., Lamdhade G.T.	74
119	Investigate transport properties of metal oxide doped polyindole composites.	A. A. Dubey, G. R. Dhokane	75
120	Photovoltaic Response of TiO_2 Nanoparticle doped materials	D. J. Bhagat , G. R. Dhokane	75
121	Combustion synthesis of $CaAlBO_4: Eu^{3+}$ phosphor for solid state lighting	R. T. Maske, A.N. Yerpude, S. J. Dhoble	76
122	Ultrasonic Velocity, Adiabatic Compressibility, Intermolecular Free Length and Other Acoustical Parameters of Leaf Extract Solution of Cymbopogon	S. S. Kamble, S. R. Aswale, S. S. Aswale	77
123	Synthesis and characterization of pure and Mg^{2+} doped ZnS nanoparticles	Sandeep A. Waghuley Nikita S. Korde	78
124	Nurturing Interdisciplinary Competencies for Innovations in Science	Dr. U. R. Kanerkar, Dr. Y. M. Rajgure Dr. R. Bijwe	79



Online ISSN : 2395-602X

Print ISSN : 2395-6011

[UGC Journal No : 64011]

International Journal of Scientific Research in Science and Technology

Certificate of Publication

Ref : IJSRST/Certificate/Volume 9/Issue 13/9680

15-May-2022

This is to certify that **A. O. Chauhan** has published a research paper entitled '**Synthesis and Characterization of $\text{LiBaB}_9\text{O}_{15}$: Gd^{3+} Phosphor by Recrystallization Method**' in the International Journal of Scientific Research in Science and Technology (IJSRST), Volume 9, Issue 13, May-June-2022 .

This Paper can be downloaded from the following IJSRST website link

<https://ijsrst.com/IJSRST22291325>

DOI : <https://doi.org/10.32628/IJSRST.RAMAN22913>

IJSRST Team wishes all the best for bright future

Editor in Chief
IJSRST

website : <http://ijsrst.com>

Scientific Journal Impact Factor = 7.214

Peer Reviewed and Refereed International Journal





2nd International Conference on Recent Advances in Material Science and Nanotechnology (RAMAN-2022)

12th to 14th May 2022

Organized By

Department of Physics

G. S. Tompe Arts, Commerce & Science College,
Chandur Bazar, Amravati, Maharashtra, India

In Collaboration with

Internal Quality Assurance Cell (IQAC)

Sant Gadge Baba Amravati University, Amravati, Maharashtra, India

Certificate of Participation

Ref : RAMAN-2022/Certificate/9680

15-May-2022

This is to certify that **A. O. Chauhan** has presented a research paper entitled '**Synthesis and Characterization of $LiBaB_9O_{15} : Gd^{3+}$ Phosphor by Recrystallization Method**' in the RAMAN-2022 held on 12th to 14th May 2022, Department of Physics, G. S. Tompe Arts, Commerce & Science College, Chandur Bazar, Amravati, Maharashtra, India In Collaboration with Internal Quality Assurance Cell (IQAC), Sant Gadge Baba Amravati University, Amravati, Maharashtra, India.

Editor in Chief
IJSRST

Dr. D. R. Bijwe
Convener

Dr. R. S. Ramteke
Chairman





Synthesis and Photoluminescence study of $\text{LaF}_3:\text{Gd}^{3+}$ phosphors for Phototherapy Application

A. O. Chauhan^{1*}, C. B. Palan², N. S. Sawala³, S. K. Omanwar⁴

^{1*}Department of Physics, Vidya Bharati College, Amravati, Maharashtra, India

²Department of Physics, Bapuniya Sirajoddin Patel Arts, Commerce and Science College, Pimpalgaon Kale, Tq. Jalgaon (Jamod), Dist: Buldhana, Maharashtra, India

³Science and Humanities Department, Government Polytechnic, Arvi, India

⁴Department of Physics, Sant Gadge Baba Amravati University, Amravati, India

*(Corresponding author email: abhi2718@gmail.com)

Abstract

The narrowband UVB emitting phosphor LaF_3 doped with various concentration of Gd^{3+} ions were successfully synthesized by wet chemical synthesis method with RAP (Reactive Atmospheric process). The crystal structure and the phase purity of samples were characterized using powder X-ray diffractometer. Crystallographic data obtained for LaF_3 phosphor matches well with the reported data. Gd^{3+} ions were successfully incorporated in the LaF_3 host lattice separately and their photoluminescence properties were analyzed at room temperature. The Phosphor shows sharp and intense emission at 311 nm under the excitation of 275 nm due to ${}^6\text{P}_J \rightarrow {}^8\text{S}_{7/2}$ transition of Gd^{3+} ions. Optimum concentration of Gd^{3+} ions in the prepared phosphor was found to be 0.04 mol. For this concentration the critical distance R_0 was calculated to be 13.78 Å.

Keywords: Wet Chemical Synthesis, Photoluminescence, XRD, Narrowband UVB

1. Introduction

The UV radiation differentiated as per their Biological and Physical characteristic and divides into three parts such as UV-C: the rays that do not pass through the earth's atmosphere (200-290 nm) UV-B: the rays responsible for nearly all biological effects following sun light exposure including tanning, burning and skin cancer, (290-320 nm) and UV-A: those rays closest to the visible spectrum that pass through glass and are the least harmful to the skin (320-400nm).

Recently the treatment of skin diseases using artificial sources of ultraviolet (UV) radiation in controlled condition is well established. UV therapy is useful for treating more than 40 types of skin diseases and disorders such as psoriasis [1], or vitiligo [2], which could be treated by UV-B radiation, and lichen sclerosus [3], morphea [4] scleroderma [5], cutaneous T-cell lymphoma, lupus erythematosus [6], which could

be treated by UV-A radiation. In the treatment of hyperbilirubinemia [7], commonly known as infant jaundice. The basis of phototherapy is believed to be the direct interaction of light of certain frequencies with tissues to cause a change in the immune response. In the treatment of skin diseases, few methods are currently used: phototherapy with narrowband UVB (310 - 313 nm) and photochemotherapy PUVA, with UVA (365 nm) and psoralens as photosensitizers. In 1976, Fischer published a study wherein he evaluated the “healing” effects of 313, 334, 365, and 405 nm light on psoriasis. He concludes that, 313 nm light exerted a potent clearing effect on psoriasis than the longer wavelength [8].

Over the last decade, lanthanide doped fluorides (LnF_3) have been intensively studied as promising luminescent materials. LnF_3 are featured with low phonon energy of the crystal lattice (e.g. $\text{LaF}_3 = 350 \text{ cm}^{-1}$), resulting in a relatively high QY of luminescence and diminished nonradiative relaxation of their excited states. Fluoride materials captured the attention due to their properties such as reasonably high thermal conductivity, good enough mechanical hardness, and high chemical stability, multicolor luminescence [9], long radiative lifetimes (several ms) [10], invariable luminescence in time [11], low cytotoxicity, well defined crystal structures and possibilities of easy surface modifications. Zhang et al., has reported that the UC emission of Ho^{3+} was changed from green to red in cubic phase $\text{NaYF}_4: \text{Yb}^{3+}/\text{Ho}^{3+}$ NCs by introducing Ce^{3+} [12].

The Yttrium fluoride (YF_3) phosphor has been studied since the 1969's, that shows YF_3 is an ideal host for optically active rare earth's (RE's) mostly because of its low phonon energies (530 cm^{-1}) and a wide band gap (10.5 eV), which diminish the probability for non-radiative relaxation [13, 14]. M. Domineberges et al. reported the optical study of Ce^{3+} ion in YF_3 phosphor [15]. Pankratov *et al.*, has reported the intrinsic spectroscopic properties of YF_3 phosphor have been investigated for scintillation applications [16]. Yan et al. have successfully synthesized Eu^{3+} doped yttrium hydroxide fluoride mesocrystals by facile hydrothermal route to explore the possible application in optical thermometry [17].

The Lanthanum fluoride phosphor possesses a large band gap (10.1 eV), is an ideal host for studying Ce^{3+} fluorescence for scintillators because the 4f and 5d levels of cerium are located in the gap of the host lattice [18]. R.E. Kroon et al. has reported the decay study and energy transfer mechanism of Ce^{3+} and Tb^{3+} ions in LaF_3 phosphor [19]. Guss *et al.* [20] fabricated $\text{LaF}_3: \text{Ce}^{3+}$ nanoparticles mixed with oleic acid and characterized their optical, physical, and radiation detector properties.

2. Materials and Methods

2.1. Synthesis Method

The phosphors $\text{LaF}_3:\text{Gd}^{3+}$ were successfully prepared by wet chemical synthesis method with RAP (Reactive Atmospheric process). The precursors chemicals La_2O_3 (99.99 %, AR), Gd_2O_3 (99.90 %, AR) and Hydrofluoric acid used for synthesis of phosphor. The detail of molar ratio of each precursor used for phosphor synthesis is given in Table 1.

The starting materials were taken in a proper stoichiometric ratio mixed together in a Teflon beaker. A small quantity of double distilled water (D.W.) was added and paste was formed. After that, the solution of HNO_3 was added drop by drop and mixture was simultaneously heated slowly at 80°C , till the completely

clear homogeneous solution was obtained. The solution was further heated to remove the excess of nitric acid. Little quantity of double distilled water was again added. The resulting solution was considered as $\text{La}(\text{NO}_3)_3:\text{Gd}^{3+}$. Later, the Hydrofluoric acid (HF) was added drop by drop in solution using syringe to get precipitate. Filtered the precipitate and washed by DDW few times and then dried under an infrared (IR) lamp. The dried powder was finally heated at 500°C for 1 hr.

Table 1. Molar Ratio of each precursor used for synthesis

Sr. No.	Products	Corresponding reaction with balanced molar ratios of precursors
1	$\text{La}_{(1-x)}\text{F}_3: x\text{Gd}^{3+}$	$(1-x)\text{La}_2\text{O}_3 + \text{HF} + x\text{Gd}_2\text{O}_3$ {In stock solution form 1 gm = 100ml} $\xrightarrow{\Delta}$ $\text{La}_{(1-x)}\text{F}_3: x\text{Gd}^{3+} + \text{Gaseous products}$ (H_2O , fluorine acid and NO_2) ($x = 0.002, 0.005, 0.01, 0.03$ and 0.04) Δ (Heating) at 500°C for 1 h.

2.2. Characterizations

The phase purities of $\text{LaF}_3:\text{Gd}^{3+}$ samples were studied using Rigaku miniflex II X-ray Diffractometer with scan speed of $4.0^\circ/\text{min}$ and Cu K α ($k = 1.5406 \text{ \AA}$) radiation. Photoluminescence properties were measured on (Hitachi F-7000) fluorescence spectrophotometer at room temperature. The parameters such as spectral resolution, width of the monochromatic slits (1.0 nm), photomultiplier tube (PMT) detector voltage and scan speed were kept constant throughout the analysis of samples.

3. Results and Discussion

3.1. XRD Analysis

The diffraction pattern is usually used to identify the crystal structure and the phase purity of the sample. Fig. 3.1 show the powder X-Ray Diffraction patterns of $\text{LaF}_3:\text{Gd}^{3+}$ sample prepared by using wet chemical synthesis method. The XRD pattern for $\text{LaF}_3:\text{Gd}^{3+}$ agrees well with the ICDD file no. (00-032-0483). This agreement indicates that the material was successfully prepared using the wet chemical synthesis method. The crystal structure of the prepared materials can be refined to be Hexagonal, with lattice parameter $a = 7.187 \text{ \AA}$, $b = 7.187 \text{ \AA}$, $c = 7.350 \text{ \AA}$.

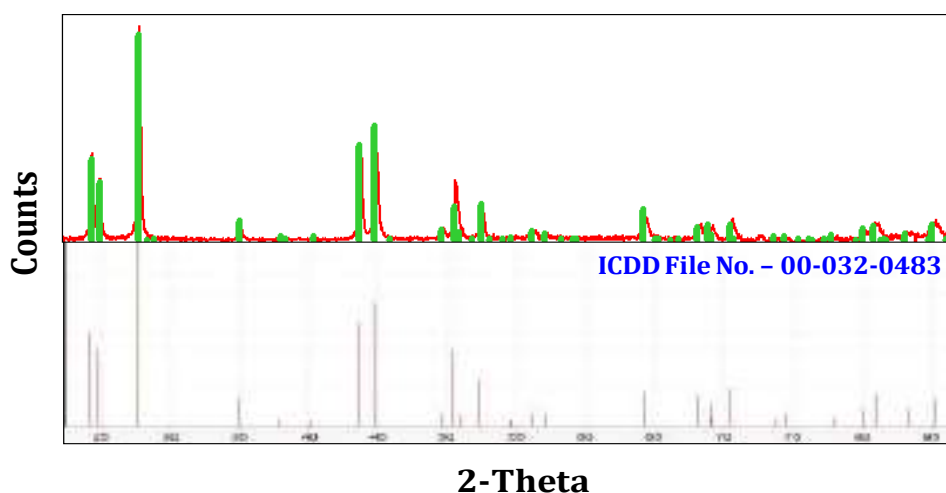


Fig. 1. X-ray diffraction pattern of LaF_3 phosphor.

3.2. Photoluminescence Spectra

3.2.1. PL measurement of LaF₃:Gd³⁺ phosphors

Fig.2. represented the excitation and emission spectra of LaF₃:_xGd³⁺ (x = 0.005, 0.01, 0.02, 0.03 and 0.04) phosphors. The double shoulder excitation peak observed at 272 nm - 282 nm maximum at 275 nm corresponding to ⁸S_{7/2} → ⁶I_J transition of Gd³⁺ ions. Under the excitation of 275 nm phosphor shows intense and sharp Narrow Band UVB (NB-UVB) emission at 311 nm corresponds to the ⁶P_{7/2} → ⁸S_{7/2} transitions of the Gd³⁺ ions. There was weak line was observed at 307 nm due to the ⁶P_{5/2} → ⁸S_{7/2} transition of the Gd³⁺ ions.

In addition, the emission intensity of phosphor increases with increasing concentration of an activator (Gd³⁺ ions) and achieve a maximum intensity for the concentration of 0.04 moles of Gd³⁺ ions. For LaF₃ system the quenching phenomena were not observed upto the 0.04 mol concentration of Gd³⁺ ion.

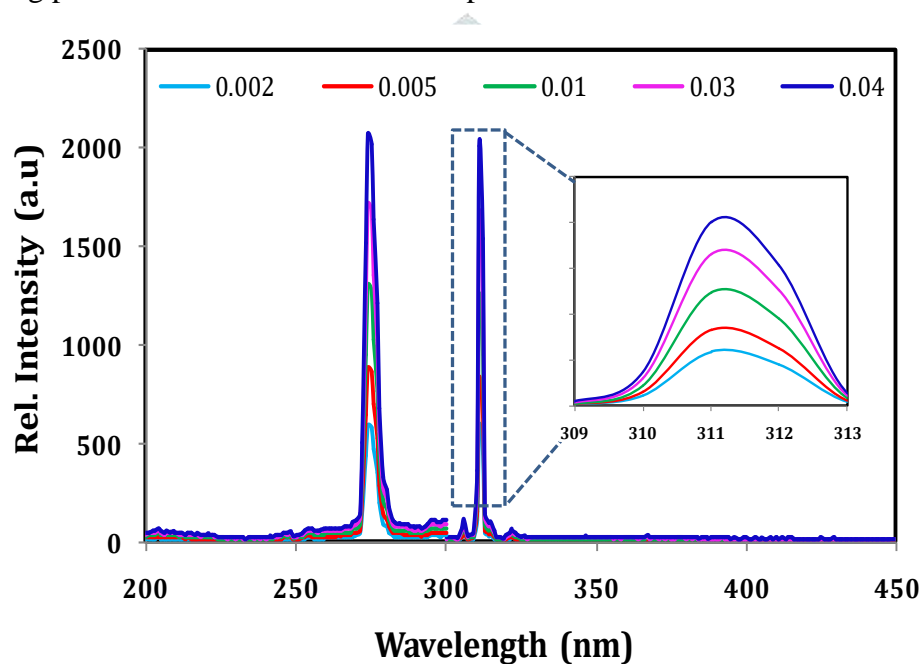


Fig. 2. Excitation and Emission spectra of LaF₃:_xGd³⁺ (x = 0.002, 0.005, 0.01, 0.03, 0.04) phosphors.

According to Blasse, the average shortest distance between nearest activator ions is equal to the critical transfer distance (R_C). On other hand the critical distance (R_C) was nothing but the critical separation between donor (activator ion) and acceptor (quenching ion), at which the non-radiative rate equals that of the internal single ion relaxation. Hence the critical distance was calculated using the following equation,

$$R_C \approx 2 \times \left(\frac{3V}{4\pi x_c N} \right)^{\frac{1}{3}}$$

Where, V is the volume of the unit cell (in Å³), χ_c is critical concentration and N is the number of La³⁺ ions in the unit cell. By taking the value of above said parameters from the experimental results and the crystal structure of the compound LaF₃, the value of $\chi_c = 0.04$, N = 6 and V = 328.80 Å³, the critical distance R_C of LaF₃ doped with Gd³⁺ phosphor is calculated to be about 13.78 Å.

4. Conclusions

The inorganic narrow UVB emitting LaF₃:Gd³⁺ phosphor was intentionally and successfully prepared by wet chemical synthesis method with RAP method which is low cost, low temperature and not required any other additive for initiation synthesis process. Photoluminescence properties in the UV region, which is used for phototherapy lamps, are studied. The XRD pattern of prepared sample found in agreements with the respective ICDD file no. (00-032-0483) and found to be in complete crystalline nature.

The photoluminescence spectra specify that the LaF₃:Gd³⁺ gives sharp narrow UVB emission i.e. 311 nm under the 275 nm excitation attributed to ⁶P_{7/2}→⁸S_{7/2} optical transition of Gd³⁺ ion. This narrow band UVB emission is of importance in medical phototherapy, biological agent detection, sterilization, and covert communication. The study of spectroscopic data has demonstrated that, with an increase in the concentration of gadolinium, the luminescence intensity of Gd³⁺ ions increase upto the 0.04 mol concentration of Gd³⁺ ion. Hence, we can conclude that LaF₃:Gd³⁺ phosphor is useful for phototherapy lamp application because of its intense emission at NB-UVB region (311±2 nm).

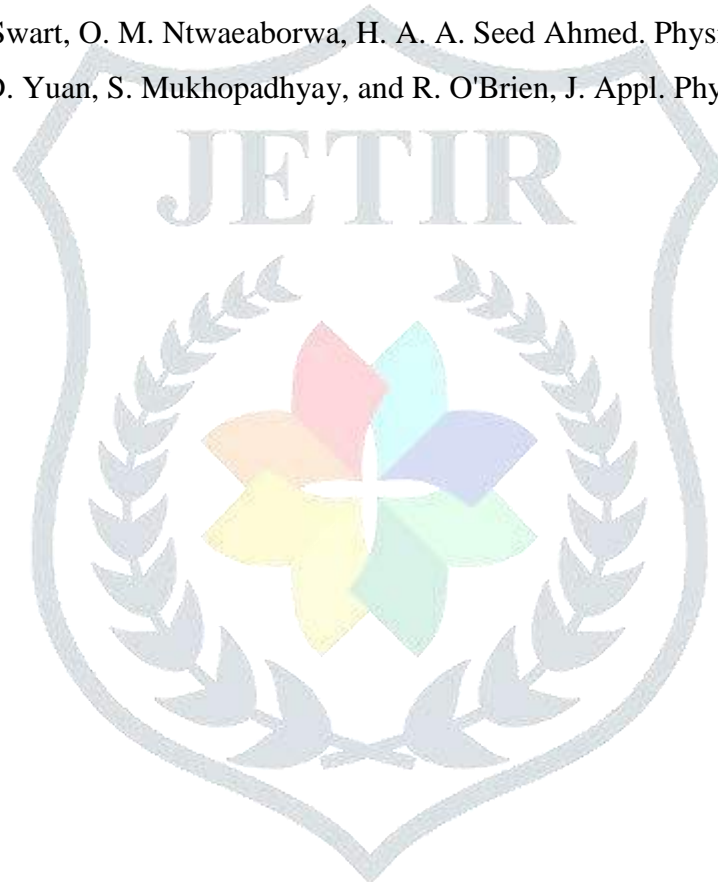
5. Acknowledgement

One of the authors Dr. Abhijit O. Chauhan is thankful to the Head of Department, Department of Physics, SGBA University, Amravati (MH), PIN- 444602 INDIA for providing XRD facility for this work.

6. References

- [1] Honigsmann H, Brenner W, Rauschmeier W. Photochemotherapy for cutaneous T cell lymphoma. A follow-up study. *Am. Acad. Dermatol.*, 1984, **10**: 238.
- [2] Scherschun L, Kim J J, Lim W H. Narrow band ultraviolet B is a useful and well tolerated treatment for vitiligo. *J. Am. Acad. Dermatol.*, 2001, **44**: 999.
- [3] Kreuler A, Jansen T, Stucker M, Herde M, Hoffmann K, Altmeyer P, Vonkobyletzki G. Low-dose ultraviolet-A1 phototherapy for lichen sclerosus et atrophicus., *Clin. Exp. Dermatol.* 2001, **26**: 30.
- [4] Hawk J L M, Sunbeds. *Radiat. Prot. Dosim.*, 2000, **91**: 143.
- [5] Morita A, Kobayashi K, Isomura I, Tsuji T, Krutmann J. Ultraviolet A1 (340–400 nm) phototherapy for Scleroderma in systemic sclerosis . *J. Am. Acad. Dermatol.*, 2000, **43**: 670.
- [6] Millard T P, Hawk J L M. Ultraviolet therapy in lupus. *Lupus.*, 2001, **10**: 185.
- [7] Dani C, Martelli E, Reali M F, Bertini G, Panin G, Rubaltelli F. Fiberoptic and conventional phototherapy effects on the skin of premature infants. *J. Pediatr.*, 2001, **138**: 438.
- [8] Fischer T. UV-light treatment of psoriasis. *Acta. Dermatovener (Stockh).*, 1976, **56**: 473.
- [9] N. Guo, Y. Song, H. You, G. Jia, M. Yang, K. Liu, Y. Zheng, Y. Huang, H. Zhang *Eur. J. Inorg. Chem.* 72 (2010) 4636.
- [10] Pankratov, V., Kirm, M. and von Seggern, H., 2005. Intrinsic luminescence in yttrium trifluoride. *J. Lumin.* 113, 143-150.
- [11] R. E. Kroon, H. C. Swart, O. M. Ntwaeaborwa, H. A. A. SeedAhmed. *Physica B* 439 (2014) 83–87.
- [12] G. Y. Chen, H. C. Liu, G. Somesfalean, H. J. Liang, Z. G. Zhang, *Nanotechnology.* 20 (2009) 385704.

- [13] Rast, H.E., Caspers, H.H. and Miller, S.A., 1969. Lattice vibrations and infrared properties of yttrium fluoride. *Phys. Rev.* 180, 890
- [14] Krupa, J.C. and Queffelec, M., 1997. UV and VUV optical excitations in wide band gap materials doped with rare earth ions: 4f-5d transitions. *J. Alloys Compd.* 250, 287-292.
- [15] Domineberges, M. and Loriers, J., 1985. Luminescence of trivalent cerium in yttrium fluoride YF₃. *Comptes Rendus De L Academie Des Sciences Serie II* 301, 915-918.
- [16] Pankratov, V., Kirm, M. and von Seggern, H., 2005. Intrinsic luminescence in yttrium trifluoride. *J. Lumin.* 113, 143-150.
- [17] G. Murali, B. Lee, R. Mishra, J. Lee, S. Nam, Y. Suh, D. Lim, J. Lee, S. Lee, Synthesis, luminescence properties and growth mechanisms of YF₃: Yb³⁺/Er³⁺ nanoplates. *J. Mater. Chem. C* 3, 10107 (2015)
- [18] A. Canning, A. Chaudhry, R. Boutchko, and N. Grønbech- Jensen, *Phys. Rev. B* 83, 125115 (2011).
- [19] R. E. Kroon, H. C. Swart, O. M. Ntwaeaborwa, H. A. A. Seed Ahmed. *Physica B* 439 (2014) 83–87.
- [20] P. Guss, R. Guise, D. Yuan, S. Mukhopadhyay, and R. O'Brien, *J. Appl. Phys.* 113, 0643031 (2013).





Synthesis and luminescence properties of $\text{MSO}_4:\text{Eu}$ (M=Ca, Ba) phosphors for radiation dosimetry

C. B. Palan^{1*}, A. O. Chauhan², N.S. Sawala³, S. K. Omanwar⁴

^{1*}Department of Physics, Bapumiya Sirajoddin Patel Arts, Commerce and Science College, Pimpalgaon Kale, Tq. Jalgaon (Jamod), Dist:-Buldhana, Maharashtra, India

²Department of Physics, VidyaBharati College, Amravati, Maharashtra, India

³Science and Humanities Department, Government Polytechnic, Arvi, India

⁴Department of Physics, Sant Gadge Baba Amravati University, Amravati, India

*(Corresponding author email: chetanpalan27@gmail.com)

Abstract

The goal of this work was to investigate the relevant dosimetric and luminescent properties of $\text{MSO}_4:\text{Eu}$ (M=Ca, Ba) phosphor for radiation dosimetry. The structural property was studied through X-ray diffraction technique. Additionally, the photoluminescence (PL), thermoluminescence (TL) and optically stimulated luminescence (OSL) behaviors of $\text{MSO}_4:\text{Eu}$ (M=Ca, Ba) phosphor was studied. The XRD pattern of $\text{MSO}_4:\text{Eu}$ (M=Ca, Ba) phosphor fully matched with the International center for diffraction data (ICDD) file. The PL spectra of as-prepared $\text{MSO}_4:\text{Eu}$ (M=Ca, Ba) phosphor showed characteristic emission in near-UV region, The TL glow curve of $\text{MSO}_4:\text{Eu}$ (M=Ca, Ba) phosphors were consist overlapping peaks in temperature 50–300°C. Also $\text{MSO}_4:\text{Eu}$ (M=Ca, Ba) phosphors show excellent CW-OSL response under β irradiation.

Keywords: Radiation dosimetry; Luminescence properties; X-ray diffraction; Thermoluminescence; Optically Stimulated Luminescence; $\text{MSO}_4:\text{Eu}$ (M=Ca, Ba).

1. Introductions

Radiation dosimetric methods are used for the estimation of dose absorbed by radiation in a detector material. These methods are required for estimation of absorbed dose in various applications of radiation, such as personnel and environmental dosimetry, retrospective, accidental, dosimetry and medical applications of radiation. The use of Thermoluminescence (TL) as a method for radiation dosimetry of ionizing radiation has been established for many decades and has found many useful applications in various fields, such as personnel, environmental, medical, archaeological, geological dating and space dosimetry. Several high

sensitive TL phosphor materials and thermoluminescent dosimeters (TLDs) are now commercially available in different physical forms. There are many commercial TLD systems which are being used for various dosimetric applications and even presently, TL is a popular technique in the field of radiation dosimetry, particularly in personnel monitoring [1-4].

Thermoluminescent materials based on rare earth (RE)-doped alkaline earth sulphates have been studied since 1970 [5]. Sulphates are known to be good Thermoluminescent materials. Sulphate based RE doped phosphors, particularly alkaline earth sulphates, because of its high sensitivity, ease of preparation and stability of response in adverse climates, have already been very popular for use in radiation dosimetry, using thermoluminescence [6]. The main characteristics of some Eu activated sulphate TL materials are tabulated in **Table 1**.

Table 1. The main characteristics of some Eu activated sulphate based TL material

S.N	Material	Z _{eff}	Glow peak Temp (°C)	Synthesis Method	Application	Ref.
1.	BaSO ₄ :Eu	30.3	190	Precipitation method	Radiation dosimetry using OSL as well as TL	[7]
2.	BaSO ₄ : Eu	30.3	-	Recrystallization method	Dosimetric applications	[8]
3.	SrSO ₄ :Mn	46.5	346	Wet chemical precipitation method	High temperature radiation dosimetry applications	[9]
4.	SrSO ₄ :Eu	46.5	231	Acid evaporative method		[10]
5.	SrSO ₄ :Eu	46.5	312	Chemical precipitation method	Research in mixed fields detection dosimeters	[11]
6.	SrSO ₄ :DyTb	46.5	217	Co-precipitation method	Estimating high dose of gamma rays	[12]
7.	CaSO ₄ : Eu	15.3	180	-	TLD	[13]
8.	CaSO ₄ : Eu	15.3	173	Chemical co-precipitation method	High dose in space craft payload	[14]

In the last two decades an alternative technique namely OSL has been developed and widely used. The OSL technique is the best of all the known techniques for measurement of radiation exposure since the out process does not involve problems of blackbody radiation and thermal quenching as in TL [15-18].

Many studies have searched for new OSL materials with suitable properties for dosimetry, including phosphate, borate, sulfate and halides [19-34]. Rare earth activated MSO_4 is an interesting material for TL/OSL, it has a low effective atomic number with hydrated Orthorhombic phases [35]. Rare earths (RE) activated sulfate based phosphors are widely used for radiation dosimetry.

The purpose of this work was to investigate the dosimetric and luminescence properties of a $\text{MSO}_4:\text{Eu}$ ($\text{M}=\text{Ca}, \text{Ba}$) phosphor, this phosphor is developed via co-precipitation method for radiation dosimetry application. However, to the best of our knowledge, there are no reports on the Comparative studies of TL/OSL properties of $\text{MSO}_4:\text{Eu}$ ($\text{M}=\text{Ca}, \text{Ba}$) phosphor.

2. Experimental details

MSO_4 ($\text{M}=\text{Ca}, \text{Ba}$) phosphors activated with Eu were prepared by co-precipitation method described in our earlier works [36]. The stoichiometry of the reaction is maintained by formula $\text{M}_{1-x}\text{SO}_4:x\text{Eu}^{2+}$ ($\text{M}=\text{Ca}, \text{Ba}$). The nitrate precursor of Calcium/ Barium was dissolved in 100 ml of double-distilled water with addition of the stock solution prepared for Eu_2O_3 . Solution in glass beaker under stirring to form a homogeneous aqueous solution and it was confirmed that precursor was dissolved in water. 10 ml H_2SO_4 solution was added drop by drop into the mixed aqueous solution of $\text{M}_{1-x}(\text{NO}_3)_2: x\text{Eu}$ ($\text{M}=\text{Ca}, \text{Ba}$) under rigorous stirring at room temperature and white precipitation formed. After that, the MSO_4 ($\text{M}=\text{Ca}, \text{Ba}$) precipitate was centrifuged and washed several times by water to remove the excess residual salts. The precipitate was dried at 60°C for 2 hr by optical heating. The dried sample was annealed at 900°C for 1 hr to get white crystalline powder of $\text{MSO}_4:\text{Eu}^{2+}$ ($\text{M}=\text{Ca}, \text{Ba}$). The complete process involved in the reaction was represented as a flow chart in **Fig. 1**.

1.

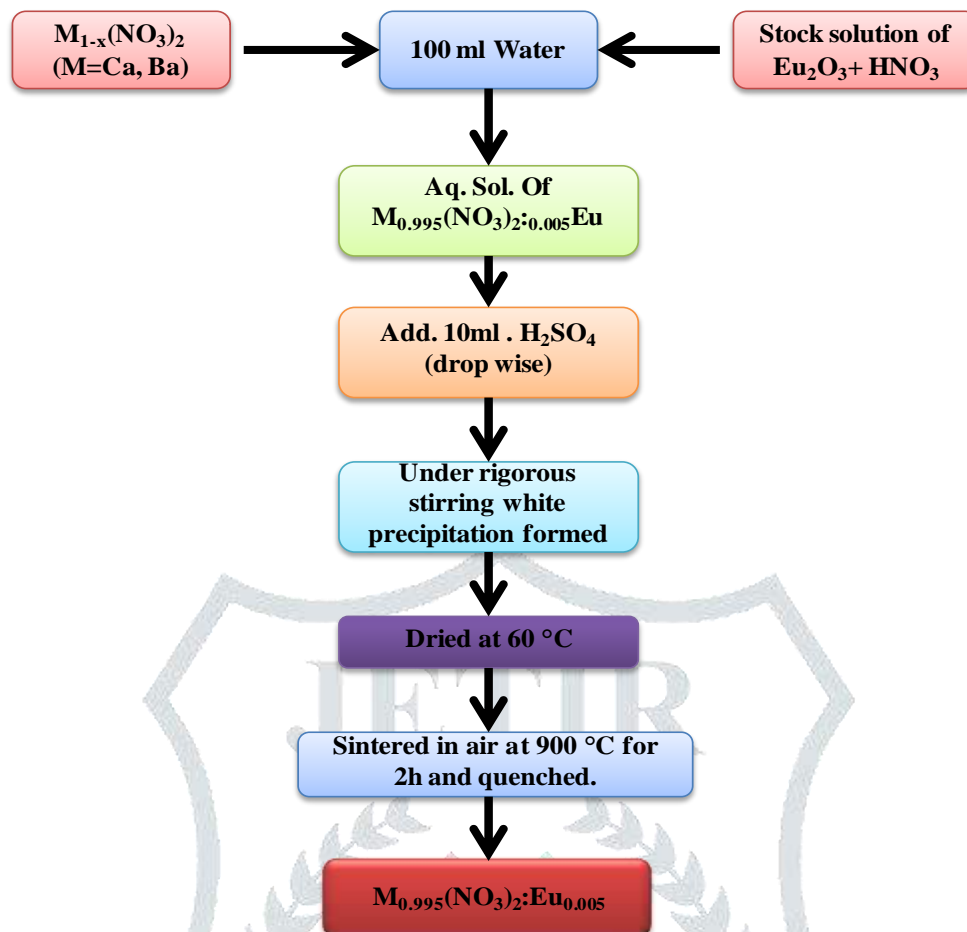


Fig. 1 Flow chart of $M_{(1-x)}SO_4:xEu^{2+}$ (M=Ca, Ba) synthesized by Co-Precipitation method.

3. Results and discussion

The structure of the as-prepared samples were analyzed by Rigaku Miniflex X-ray diffractometer, using monochromatic $CuK\alpha_1$ ($\lambda = 1.5405 \text{ \AA}$) radiation in the 2θ range of $10-60^\circ$. Photoluminescence was studied on a Hitachi F-7000 fluorescence spectrophotometer. Emission and Excitation spectra were recorded using a spectral slit of 2.5 nm for each window. For studying the TL and OSL response, all the samples were irradiated using $^{90}Sr/^{90}Y$ beta source with the dose rate of 20mGy per minute. All OSL measurements were carried out using an automatic Risø TL/OSL-DA-15 reader system which can accommodate up to 48 discs at IGCAR . Blue LEDs emitting at 470 nm (FWHM = 20 nm) are arranged in four clusters each containing seven individual LEDs. The total power from 28 LEDs at sample position is $w80 \text{ mW/ cm}^2$. A green long pass filter (GG-420) is incorporated in front of each blue LED cluster to minimize the amount of directly scattered blue light from reaching the detector system. The standard PMT used in the Risø TL/OSL luminescence reader is a bialkali EMI 9235QA, which has an extended UV response with maximum detection efficiency

between 300 and 400 nm. To prevent scattered stimulation light from reaching the PMT, the Risø reader is equipped with a 7.5 mm Hoya U-340 detection filter, which has a peak transmission around 340 nm (FWHM ~80 nm)

3.1. X-ray diffraction pattern

The structure of $\text{MSO}_4:\text{Eu}^{2+}$ ($\text{M}=\text{Ca}, \text{Ba}$) phosphor was orthorhombic.

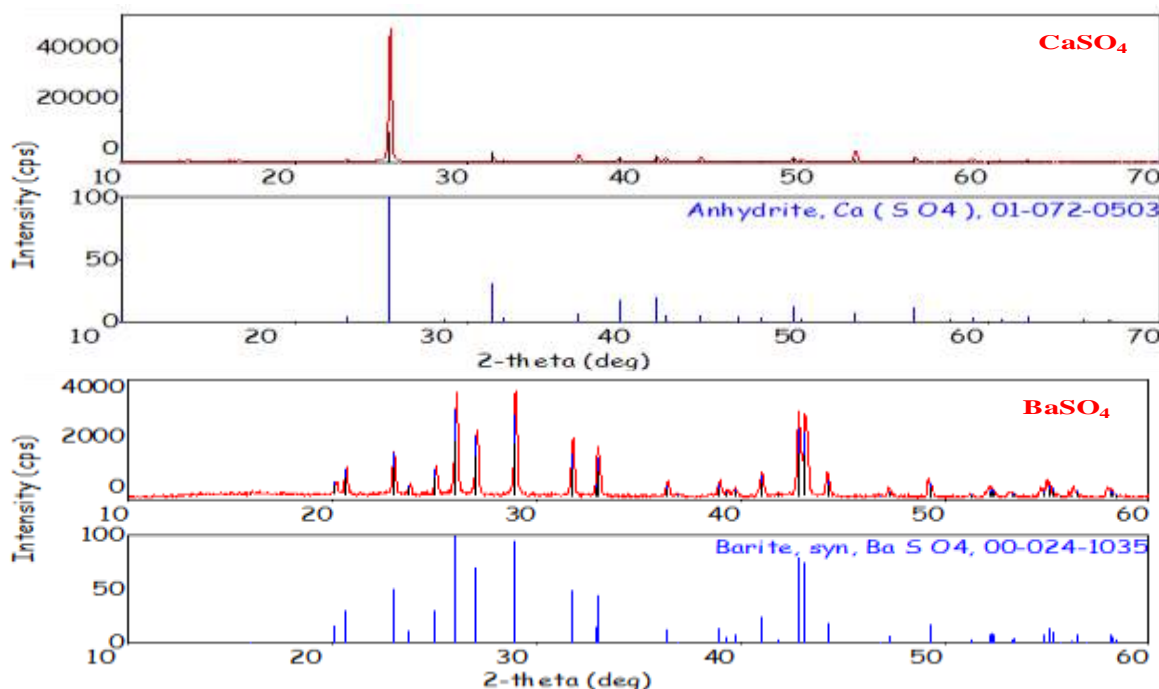


Fig. 2 X-Ray diffraction patterns of $\text{MSO}_4:\text{Eu}$ ($\text{M}=\text{Ca}, \text{Ba}$) phosphors with ICDD patterns

In order to determine the phase purity, chemical nature of the phosphor, X-ray diffraction (XRD) analysis was carried out. **Fig. 2** show the XRD pattern of $\text{MSO}_4:\text{Eu}^{2+}$ ($\text{M}=\text{Ca}, \text{Ba}$) phosphor along with the standard XRD pattern (ICDD Card No. 01-072-0503 and 00-024-1035). The XRD pattern shows the formation of pure CaSO_4 and BaSO_4 phase. The addition of the dopant (Eu) does not seem to have effect on the XRD pattern which suggests that the dopant was incorporated in the lattice.

3.2 Photoluminescence Properties (PL)

The combine excitation and emission spectra of $\text{MSO}_4:\text{Eu}$ ($\text{M}=\text{Ca}, \text{Ba}$) phosphors were give in **Fig. 3**. The excitation and emission spectra of $\text{BaSO}_4:\text{Eu}$ phosphor was observed at 377 nm and 326 nm respectively. The excitation and emission spectra of $\text{CaSO}_4:\text{Eu}$ phosphor was observed at 384 nm and 320 nm respectively.

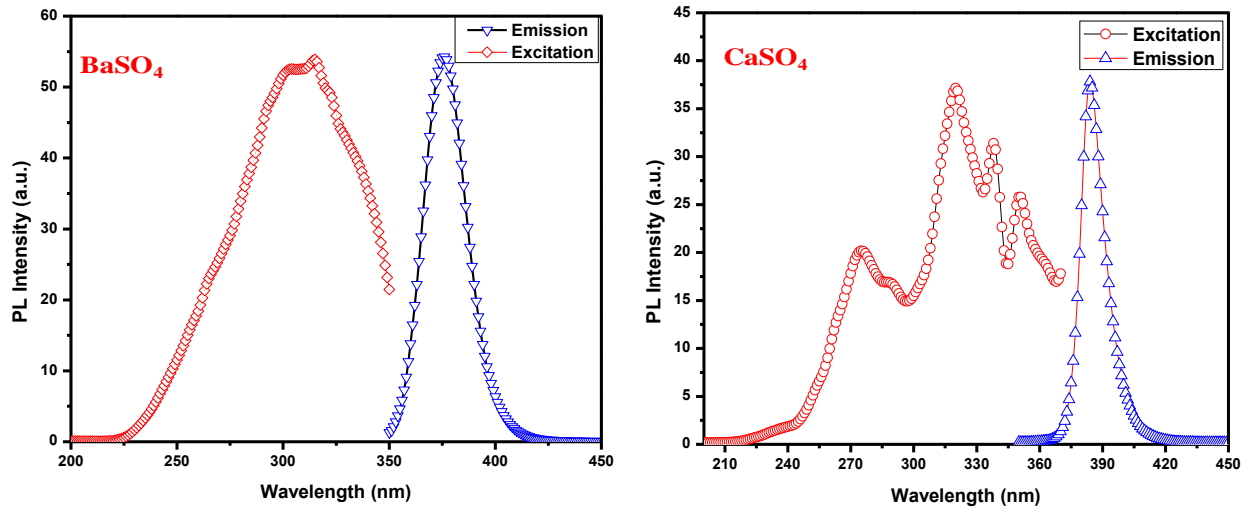


Fig. 3 Excitation and emission spectra of $\text{MSO}_4: \text{Eu}$ ($\text{M}=\text{Ca}, \text{Ba}$) phosphors

3.3 Thermoluminescence (TL)

Thermoluminescence glow curve of $\text{MSO}_4: \text{Eu}$ ($\text{M}=\text{Ca}, \text{Ba}$) phosphors under β irradiation as shown in Fig. 4. The TL glow curve of the $\text{MSO}_4: \text{Eu}$ ($\text{M}=\text{Ca}, \text{Ba}$) phosphor consists of overlapping peaks in temperature rang 50-300°C. The TL sensitivity of $\text{BaSO}_4: \text{Eu}$ phosphor was 5 time sensitive than TL sensitivity of $\text{CaSO}_4: \text{Eu}$ phosphor.

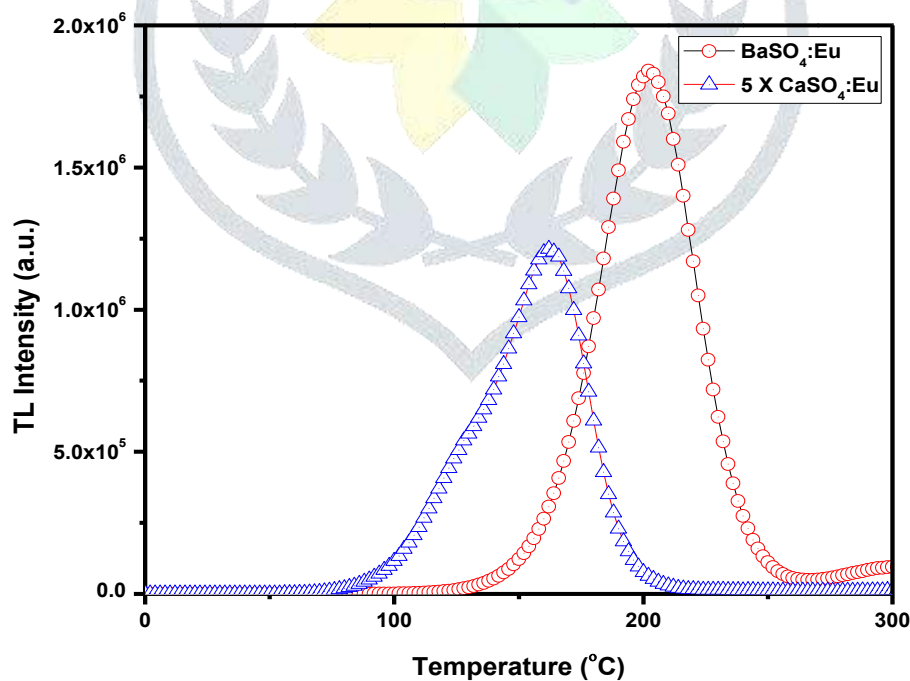


Fig. 4 Comparison of TL responses $\text{MSO}_4: \text{Eu}$ ($\text{M}=\text{Ca}, \text{Ba}$) phosphors under beta irradiation

3.4 Optically stimulated luminescence

The CW-OSL response of prepared $\text{MSO}_4:\text{Eu}$ ($M=\text{Ca}, \text{Ba}$) phosphor under 100mGy of β irradiation as shown in **Fig. 4**. The OSL sensitivity of compared with two different [37]. In the first method, The OSL sensitivity of $\text{BaSO}_4:\text{Eu}$ phosphor was found to be about 40 times to that of $\text{CaSO}_4:\text{Eu}$ phosphor. OSL sensitivity of $\text{CaSO}_4:\text{Eu}$ phosphor was more than $\text{BaSO}_4:\text{Eu}$ phosphor by using second method. The difference in the sensitivities in the two methods is because of the fact that the OSL decay in the $\text{BaSO}_4:\text{Eu}$ based phosphors is faster as compared to $\text{CaSO}_4:\text{Eu}$ phosphor.

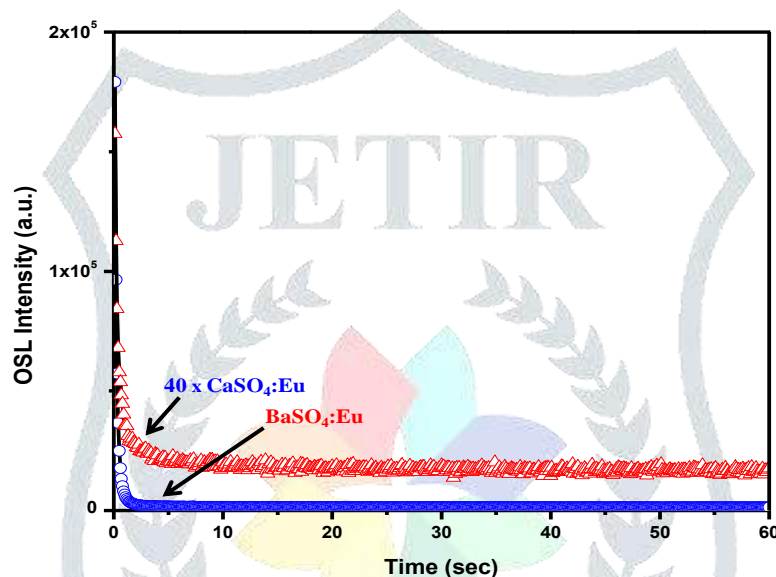


Fig. 4 Comparison of OSL responses $\text{MSO}_4:\text{Eu}$ ($M=\text{Ca}, \text{Ba}$) phosphors under beta irradiation

4. Conclusions

In this reports $\text{MSO}_4:\text{Eu}$ ($M=\text{Ca}, \text{Ba}$) phosphor was synthesized via Coprecipitation method. The XRD pattern of prepared $\text{MSO}_4:\text{Eu}$ ($M=\text{Ca}, \text{Ba}$) phosphor was perfectly matched with ICDD file. The TL glow curve consist overlapping peaks in temperature 50-300°C range. The emission spectra of prepared phosphor was observed in 350 – 450 nm range. The CW-OSL decay pattern of prepared $\text{MSO}_4:\text{Eu}$ ($M=\text{Ca}, \text{Ba}$) phosphor was same than CW-OSL decay pattern of commercial available $\alpha\text{-Al}_2\text{O}_3:\text{C}$ phosphor. The OSL sensitivity of $\text{BaSO}_4:\text{Eu}$ phosphor was found to be about 40 times to that of $\text{CaSO}_4:\text{Eu}$ phosphor. The OSL decay in the $\text{BaSO}_4:\text{Eu}$ based phosphors is faster as compared to $\text{CaSO}_4:\text{Eu}$ phosphor. The effective atomic number of prepared $\text{MSO}_4:\text{Eu}$ ($M=\text{Ca}, \text{Ba}$) phosphor is nearly 15. Although $\text{MSO}_4:\text{Eu}$ ($M=\text{Ca}, \text{Ba}$) is not a

material equivalent to tissue but, due to its high TL and OSL sensitivities, this phosphor can be proposed as a suitable candidate for radiation dosimetry, of course, after further progress in the studies.

Acknowledgement

One of the authors CBP is very much thankful to Head, RPAD, Radiological Physics and Advisory Division, Bhabha Atomic Research Centre, Mumbai-400085, India for providing the necessary facilities for the analysis of TL/OSL results and also to the Principal, Bapumiya Sirajoddin Patel Arts, Commerce and Science College, Pimpalgaon Kale for constant encouragement and guidance.

References

- [1].S. W. S. McKeever, M. Moscovitch, Radiat. Prot. Dosim. 104 (2003) 263.
- [2].S. W. S. McKeever, M. Moscovitch, P. D. Townsend Nuclear Technology Publishing, Ashford (1995).
- [3].S. W. S. McKeever, Radiat. Prot. Dosim. 27 (2002) 100.
- [4].S. W. S. McKeever, Radiat. Prot. Dosim. 109 (2004) 269.
- [5].O. Annalakshmi, M. T. Jose, U. Madhusoodanan, Radiat. Prot. Dosim. 150 (2012) 127.
- [6].K. Vohra, R. Bhatt, C Bhuwan, A. Pradhan, A. Lakshmanan, S. Shashtry, Health Phys. 38 (1980)193.
- [7].A.Patle, R. R. Patil, M. S. Kulkarni, B. C Bhatt S. V. Moharil, Conference: ADVANCED MATERIALS AND RADIATION PHYSICS (AMRP-2015) 1675 (2015) 020026 doi: 10.1063/1.4929184
- [8].M. S. Nikam, Y. K. More, B. Butey, M. Jog, S. P.Wankhede, S. B. Kondawar, S. V. Moharil J. Phys. Conf. Ser.1913 (2021) 012040.
- [9].S Jayasudha, K. Madhukumar, C. M. K. Nair, R. G. Nair, V. M. Anandakumar, Int. J. Eng. Res. 3 (2015) 1.
- [10]. Q. Tang, C. X. Zhang, D. L. Luo, P. L. Leung, Z. Y. Xiong, Radiat. Prot. Dosim. 119 (2006) 238.
- [11]. S. Jayasudha, K. Madhukumar, C.M.K. Nair, Resmi G. Nair, V.M. Anandakumar, T. S. Elias Spectrochim. ActaA Mol. Biomol. Spectrosc155 (2016) 21.
- [12]. K. R. E. Saraee, A. A. Kharikey, M. Erfani, J. Rare Earths 32 (2014) 1003.
- [13]. V. Guckan, V. Altunal, N. Nur, T. Depci, A. Ozdemir, K. Kurt, Y. Yu, I.Yegingil, Z. Yegingil, Nucl Instrum Methods Phys Res B 407 (2017) 145.
- [14]. N. Mandlik, V. N. Bhoraskar, B. J. Patil, S. S. Dahiwale, P. D. Sahare, S. D. Dhole, Indian J. Pure Appl. Phys 55 (2017) 413.
- [15]. B. C. Bhatt, M. S. Kulkarni Defect Diffusion Forum 347 (2014) 179.
- [16]. B. G. Markey, L. E. Colyott, S. W. S. McKeeverRadiat. Meas. 24 (1995) 457.
- [17]. S. W. S. McKeeverNucl. Instrum. Methods Phys. Res., Sect. B 184 (2001) 29.
- [18]. A.S. Pradhan, J. I. Lee, J. L. Kim J. Med. Phys. 33 (2008) 85.
- [19]. C.B. Palan, N.S. Bajaj, K.A. Koparkar, A. Soni, S.K. Omanwar J Inorg Organomet Polym 26 (2016) 845.
- [20]. C.B. Palan, N.S. Bajaj, S.K. Omanwar Bull. Mater. Sci., 39 (2016) 1619.
- [21]. C.B. Palan, K.A. Koparkar, N.S. Bajaj, A. Soni, S.K. Omanwar, Res Chem Intermed, 42 (2016) 7637.
- [22]. C.B. Palan, S.K. Omanwar Res Chem Intermed (2017) DOI 10.1007/s11164-017-2861-3.
- [23]. C.B. Palan, N.S. Bajaj, A. Soni, S.K. Omanwar J. Lumin. 176, (2016) 106
- [24]. C.B. Palan, N.S. Bajaj, A. Soni, M.S. Kulkarni, S.K. Omanwar Bull. Mater. Sci., 38 (2015) 1527.
- [25]. C.B. Palan, K.A. Koparkar, N.S. Bajaj, A. Soni, S.K. Omanwar Appl. Phys. A 122 (2016) 703.
- [26]. C.B. Palan, N.S. Bajaj, S.K. Omanwar AIP Conf. Proc. 1728 (2016) 020474.
- [27]. C.B. Palan, S.K Omanwar J. Lumin.178 (2016) 340.
- [28]. C.B. Palan, N.S Bajaj, SK Omanwar Mater. Res. Bull. 76 (2016) 216.

- [29]. C.B. Palan, K.A Koparkar, N.S. Bajaj, A. Soni, S.K. Omanwar J. Mater. Sci.: Mater. Electron. 27 (2016) 5600.
- [30]. C.B. Palan, A.O. Chauhan, N.S. Sawala, N.S. Bajaj, S.K. Omanwar Optik Int. J. Light Electron Opt. 127 (2016) 6419.
- [31]. C.B. Palan, N.S. Bajaj, S.K. Omanwar St. Petersburg Polytech. Univ. J. Phys. Math. 1 (2015) 410.
- [32]. M.S. Kulkarni , R.R. Patil, A. Patle , N.S. Rawat , P. Ratna , B.C. Bhatt , S.V. Moharil Radiat. Meas. xxx (2014) 1.
- [33]. B.C. Bhatt , A. Soni, G.S. Polymeris, D.K. Koul, D.K. Patel, S.K. Gupta, D.R. Mishra, M.S. Kulkarni, Radiat. Meas. 64 (2014) 35.
- [34]. L. C. Oliveira, E. G. Yukihara, O. Baffa Scientific Reports 6 (2016) 24348 DOI:10.1038/srep24348
- [35]. A. Muley, R.R. Patil, S.V. Moharil J. Lumin.128 (2008) 509.
- [36]. C.B. Palan , N.S. Bajaj , S.K. Omanwar DOI: [10.1016/j.spjpm.2015.12.007](https://doi.org/10.1016/j.spjpm.2015.12.007)
- [37]. B. Dhabekar, S.N. Menon, E. Alagu Raja, A.K. Bakshi, A.K. Singh, M.P. Chougankar, Y.S. Mayya Nuclear Instruments and Methods in Physics Research B 269 (2011) 1844–1848





Synthesis and Photoluminescence properties of Eu³⁺ ions doped LiBaPO₄ phosphor

N. S. Sawala^{1*}, S. R. Jaiswal,² A. O. Chauhan³, C. B. Palan⁴, S. K. Omanwar⁵

Corresponding author: nssawala@gmail.com

¹ Science and Humanities Department, Government Polytechnic Arvi, Dist- Wardha, MS, India

²Department of Physics, Shri. R.L.T. College of Science, Akola, India

³Department of Physics, Vidya Bharati College, Amravati, India

⁴Department of Physics, B.S. Patel Arts, Commerce and Science College, Pimpalgaon Kale,
Ta- Jalgaon Jamod, Dist- Buldhana, Maharashtra, India

⁵Department of Physics, Sant Gadge Baba Amravati University, Amravati, India

Abstract

The red orange emitting LiBa_(1-x)PO₄:xEu³⁺ phosphor was successfully synthesized by combustion method. The structure of the prepared LiBaPO₄: Eu³⁺ phosphor was confirmed by powder XRD analysis and its photoluminescence properties has been studied in the range of 200 – 400 nm. The XRD pattern of prepared LiBa_{0.98}PO₄:0.02Eu³⁺ phosphor is well matched with available standard ICDD file. The effect of different concentrations of Eu³⁺ ions in LiBaPO₄ phosphor was studied and optimum PL intensity was obtained at 0.01 mol of Eu³⁺ ion. The photoluminescence (PL) emission spectra of prepared LiBaPO₄:Eu³⁺ phosphor consist of series of sharp line peaks at 596 nm, and 615 nm under the excitation of 393 nm.

Keywords: Photoluminescence, Combustion method, LiBaPO₄:Eu³⁺, SSL

1.0 Introduction

The phosphates with chemical composition ABPO₄ (A and B are mono and divalent cations, respectively) are in a large family of monophosphates with the different structures rigorously depending on the relative ionic size of the A and B ions [1, 2]. These compounds have been considered to be efficient luminescent materials due to their excellent thermal and hydrolytic stability [3, 4]. Recently, Eu²⁺-doped ABPO₄ phosphates have received much attention because of its potential applications as new white light emission diodes (W-LEDs) phosphors, such as KCaPO₄:Eu²⁺ [5], KSrPO₄:Eu²⁺, [6] NaCaPO₄:Eu²⁺, [7] and ABaPO₄:Eu²⁺ (A = Na, K) [8]. The use of combustion synthesis has been developed looking at experiences of last few years in the preparation of luminescent materials, specifically borates, silicates, aluminates, and some oxides. The advantages of using the combustion method are low cost, highly effective as well as time saving approach to produce highly stable particles. Furthermore, under some specific conditions, uniform and narrow distribution of particle with regular size crystallite could be obtained by the combustion method [9]. Inspiring from the above discussion, it was decided to attempt the combustion method for the synthesis of LiBaPO₄ activated with Eu³⁺ as activator.

2.0 Experimental

2.1 Synthesis of Materials and Characterizations

The reputation on combustion synthesis was only due to drawbacks of solid state diffusion method [9]. The AR grade constituent raw materials were used such as lithium nitrate (LiNO_3), barium nitrate (BaNO_3)₂, ammonium dihydrogen orthophosphate ($\text{NH}_4\text{H}_2\text{PO}_4$), ammonium nitrate (NH_4NO_3) and urea ($\text{CO}(\text{NH}_2)_2$). The precursors were mixed thoroughly in China clay basins. The stock solution of the dopant Eu_2O_3 was then added into the mixture and later little amount of double distilled water was added into the prepared mixture. The mixture slowly converts into the thick paste. The thick paste was then heated for 5–10 min on hot plate, till it turns into a homogeneous clear solution. The clear solution obtained was then transferred directly to preheated furnaces maintained at temperature of 550°C . The clear solution was warmed and boiled till the evolution of lots of fumes, for 5 min. As the fuel reaches the required temperature for burning, the self-heat generating redox reaction was started. The temperature of environment during the reaction was may be around $1100\text{--}1200^\circ\text{C}$. This sudden high temperature generation from the fuel and oxidizer reaction increases the chances of the formation of fine and stable product. The fine powder of $\text{LiBa}_{(1-x)}\text{PO}_4\text{:}_x\text{Eu}^{3+}$ ($x = 0.001, 0.002, 0.005, 0.01$ and 0.02) were then finally obtained. The powder obtained from the reaction was then calcinated in the microwave furnace at 800°C for 2 hr and suddenly quenched at room temperature. This step was introduced to remove excess of organic impurities and carbon traces remaining in the material after reaction. Also the calcining gives the particles with fine structure due to sudden quenching.

The detail of stoichiometric molar amount of each precursor used for phosphors synthesis is given in Table 1.

Table 1. Molar ratio of precursors used for material preparation and corresponding chemical reaction.

Compounds	Corresponding reaction with balance molar ratios of precursors
$\text{LiBa}_{(1-x)}\text{PO}_4\text{:}_x\text{Eu}^{3+}$	$\text{LiNO}_3 + (1-x)\text{Ba}(\text{NO}_3)_2 + x(\text{Eu}_2\text{O}_3 + \text{HNO}_3) + \text{NH}_4\text{H}_2\text{PO}_4 + 5\text{CO}(\text{NH}_2)_2 + 7.5\text{NH}_4\text{NO}_3 \rightarrow \text{LiBa}_{(1-x)}\text{PO}_4\text{:}_x\text{Eu}^{3+} + (\text{Gaseous product like } \text{NO}_3, \text{N}_2, \text{CO}_2, \text{ and } \text{H}_2\text{O})$

The structural confirmation of as prepared material was done by XRD method by using Rigaku miniflex II X-ray diffractometer with scan speed of $2.000^\circ/\text{min}$ and $\text{Cu K}\alpha$ ($\lambda = 1.5406 \text{ \AA}$) radiation in the range 10° to 70° . The PL excitation and PL emission spectra were recorded at room temperature on (Hitachi F-7000) fluorescence spectrometer associated with 450W Xenon discharge lamp in range of wavelength 200 – 700 nm. The measuring parameter such as width of monochromatic slit (1.0 nm), photomultiplier tube (PMT) detector voltage, scan speed (240 nm/min), spectral resolution were kept constant during the entire analysis of materials.

2.2 Structural Confirmation

For the phase confirmation the sample with maximum dopant concentration chosen. The formation of the $\text{LiBa}_{0.98}\text{PO}_4\text{:}_{0.02}\text{Eu}$, sample in the crystalline phase synthesized by combustion method was confirmed by XRD pattern as shown in Fig.1. The XRD pattern of $\text{LiBa}_{0.98}\text{PO}_4\text{:}_{0.02}\text{Eu}$ agreed well with the standard data from ICDD file having file number (00-014-0270) of LiBaPO_4 . Also the XRD pattern show that the formed material was absolutely crystalline and in single phase. The XRD data can be refined to be hexagonal structure and crystallized in the P63 space group. Based on the effective ionic radii it was assumed that the Ba^{2+} was more preferably replaced by Eu^{3+} ions [10].

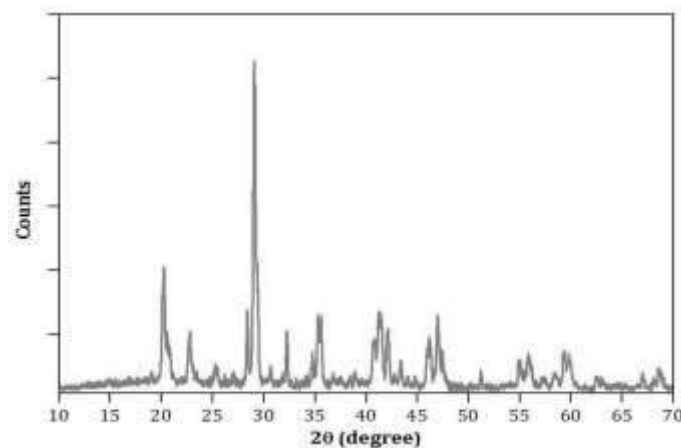


Fig. 1: XRD pattern of the $\text{LiBa}_{0.98}\text{PO}_4:0.02\text{Eu}^{3+}$ phosphor.

2.3 Photoluminescence Properties

The excitation and emission spectra of $\text{LiBa}_{(1-x)}\text{PO}_4:x\text{Eu}^{3+}$ phosphor is as presented in the Fig. 2 and Fig. 3 respectively. The excitation and emission spectra of $\text{LiBa}_{(1-x)}\text{PO}_4:x\text{Eu}^{3+}$ ($x = 0.001, 0.002, 0.005, 0.01$ and 0.02) phosphor, were monitored at wavelength 615 nm and 393 nm respectively. The excitation spectra consists of series of lines at $321, 365, 378$ and 393 nm correspond to the ${}^7\text{F}_0 \rightarrow {}^5\text{H}_3, {}^7\text{F}_0 \rightarrow {}^5\text{D}_4, {}^7\text{F}_0 \rightarrow {}^5\text{L}_7$ and ${}^7\text{F}_0 \rightarrow {}^5\text{L}_6$ transitions respectively. Among these peaks, the excitation at 394 nm is the dominant, as shown in Fig. 2. On the other hand emission spectra consist of two sharp peaks 596 nm and 613 nm with one weak peak at 653 nm . The emission at 596 nm corresponds to the ${}^5\text{F}_0 \rightarrow {}^7\text{D}_1$ (magnetic dipole), the emission at 613 nm corresponds to the ${}^5\text{F}_0 \rightarrow {}^7\text{D}_2$ (electric dipole) and weak emission at 653 nm corresponds to the ${}^5\text{F}_0 \rightarrow {}^7\text{D}_3$ transition of Eu^{3+} ions [11]. The PL emission intensity was studied as a function of concentration of Eu^{2+} ($0.1 \text{ mol}\%$ to $2.0 \text{ mole}\%$). It is observed that 0.01 mol is the optimum concentration of Eu^{3+} ions for the prepared phosphor. The luminescence intensity is found to be decreased when the concentration of Eu^{3+} ions is increased beyond the optimum value, due to well known consequence of concentration quenching.

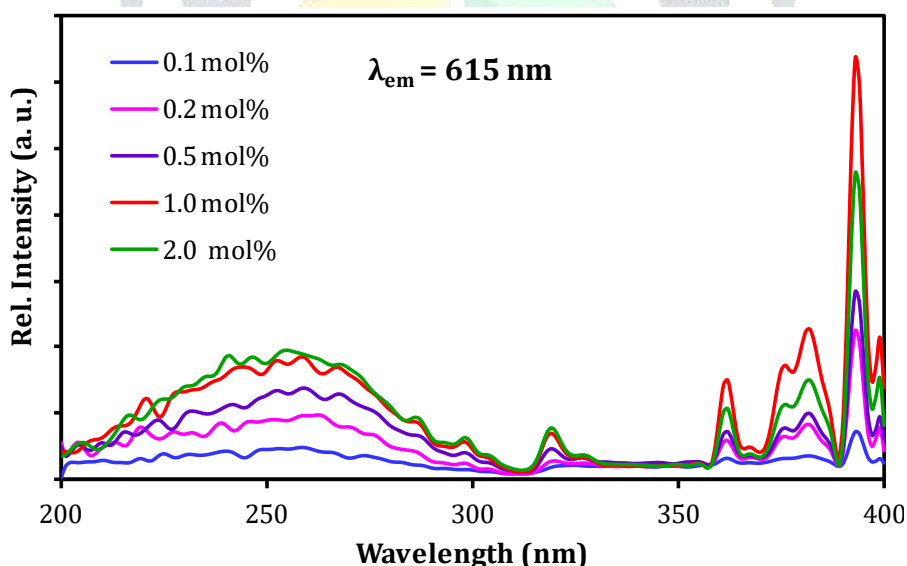


Fig. 2: PL excitation spectra of $\text{LiBa}_{(1-x)}\text{PO}_4:x\text{Eu}^{3+}$ ($x = 0.001, 0.002, 0.005, 0.01$ and 0.02) phosphors

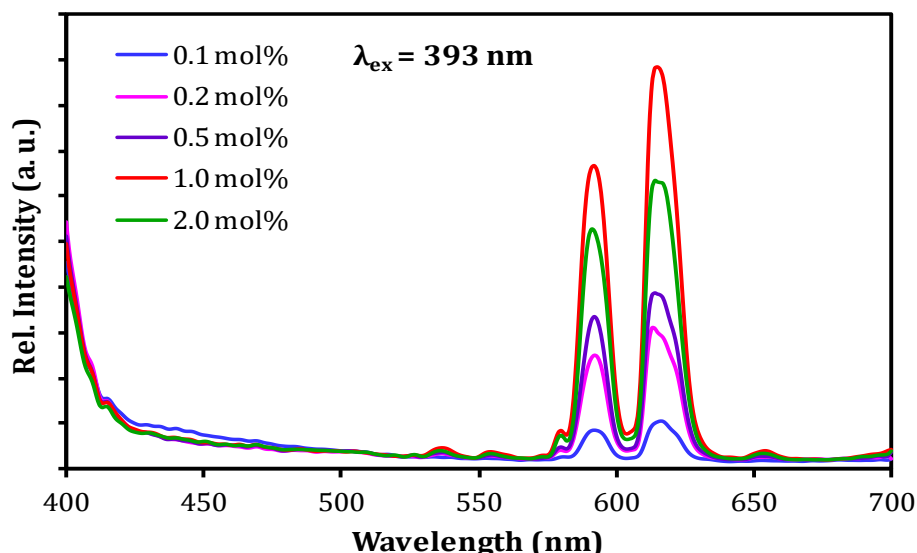


Fig. 3: PL emission spectra of $\text{LiBa}_{(1-x)}\text{PO}_4:x\text{Eu}^{3+}$ ($x = 0.001, 0.002, 0.005, 0.01$ and 0.02) phosphors

3.0 Conclusions

The polycrystalline $\text{LiBa}_{(1-x)}\text{PO}_4:x\text{Eu}^{3+}$ phosphor was effectively prepared by combustion method. The XRD pattern of prepared $\text{LiBa}_{0.98}\text{PO}_4:0.02\text{Eu}^{3+}$ phosphor is good settlement with the ICDD standard file with card no. 00-014-0270. The PL excitation spectra, was monitored at 615 nm and emission spectra was monitored at 393 nm. The outcome of this phosphor provides the support for the probable and potential application in red-orange emitter in many applications in fluorescent area including solid state lighting (SSL).

References

- [1]. W. J. Tang, D. H. Chen, J. Am. Ceram. Soc. 92 (2009), 1059.
- [2]. Y. Chen, J. Wang, M. Zhanga, Q. Zenga RSC Adv., 7, (2017), 21221.
- [3]. R. Yuan Yang, Y. Ming Peng, Y. Kuin Su J. Elect. Mat., 42, (2013) 129.
- [4]. C. C. Lin, Z. R. Xiao, G. Y. Guo, T. S. Chan, R. S. Liu, J. Am. Chem. Soc. 132 (2010), 3020.
- [5]. T. S.Chan, R. S. Liu, I. Baginskiy, Chem. Mater. 20 (2008), 1215.
- [6]. C. B. Palan, K. A. Koparkar, N. S. Bajaj, S. K. Omanwar, Res. Chem. Intermed. 42 (2016), 7637.
- [7]. Y. S.Tang, S. F. Hu, C. C. Lin, N. C. Bagkar, R. S. Liu, Appl. Phys. Lett. 90, (2007), 151108.
- [8]. C. X. Qin, Y. L. Huang, L. Shi, G. Q. Chen, X. B. Qiao, H. J. Seo, J. Phys. D: Appl. Phys. 42 (2009), 185105.
- [9]. M. Ben Amara, M. Vlasse, G. Le Flem, P. Hagenmuller, Acta Crystallogr., Sect. C 39 (1983), 1483.
- [10]. C. B. Palan, N. S. Bajaj, A. Soni, M. S. Kulkarni, S. K. Omanwar, Bull. Mater. Sci., 38 (2015), 1527.
- [11]. N. S. Bajaj, S. K. Omanwar, Opt. Mater. 35, (2013) 1222.
- [12]. L. Elammari, M. El Koumiri, I. Zschokke-Gr€anacher, B.Elouadi, Ferroelectrics 158 (1994), 19.

स्थानिक स्वराज्य संस्था आणि महिला नेतृत्व

अमित पुरोषोत्तम इंगोले

राज्यशास्त्र विभाग प्रमुख
 विद्याभाषात्री महाविद्यालय, अमरावती

प्रा.डॉ. विनायक

राज्यशास्त्र विभाग
 श्री. भिसनहाल नरपन्थ
 कला व वाणिज्य महाविद्यालय, अमरावती

प्रस्तावना

'स्थानिक स्वराज्य संस्था' हा राज्य व्यवस्थेचा प्रमुख आधारस्तंभ मानला जातो. प्रामुख्याने भारतातील राज व्यवस्था ग्रामीण विकासाच्या दृष्टीने महत्त्वाची प्रशासकीय यंत्रणा आहे. पंचायत राज व्यवस्थेचे जिल्हा पातळीवरील जिल्हा परिषद, विकास गट पातळीवर पंचायत समिती आणि पायाभूत समजल्या जाणाऱ्या ग्राम किंवा खेडे पातळीवरील ग्रामपंचायत अशी प्रशासकीय रचना आढळून येते. यातील ग्रामपंचायत म्हणजे प्राचीन पंचायतीचे आधुनिक स्वरूप. ग्रामीण भागात राहणाऱ्या जवळपास ७० टक्के लोकसंख्येची निगडित ग्रामपंचायतीची भूमिका असते. कालखंडापासून पंचायतीतील महिलांचा सहभाग अपवादात्मक राहिला आहे. मात्र १९९२ च्या ७३ व्या घटनासह महिलांना ग्रामपंचायतीतील ३३ टक्के आरक्षण देऊन निर्णय निर्धारण प्रक्रियेतील त्यांच्या राजकीय प्रवेशाचा मार्ग केला. अर्थात ग्रामपंचायतला दीर्घ इतिहास लाभला असला तरी त्यातील महिला सहभाग ही नवीन संकल्पना आहे.

सत्याचे विकेंद्रीकरण हा केंद्रबिंदू म्हणून महाराष्ट्राचे शिल्पकार कै. यशवंतराव चव्हाण यांनी पंचायत व्यवस्थेची मुहूर्तमेढ रोवली. महाराष्ट्रातील पंचायत राज व्यवस्था ही आदर्श व्यवस्था म्हणून संपूर्ण देशात मान्यता आहे. ७३ व्या घटना दुरुस्तीमुळे पंचायत राज संस्थांना विशेष करून ग्रामपंचायत व ग्रामसभा या शेवटच्या स्तरावरील संस्थांना घटनात्मक अधिकार प्राप्त झालेले आहेत. त्यामुळे राज्यातील पंचायत राज व्यवस्थेला बळकटी प्राप्त झाली आहे. देशातील महिलांना सत्तेत प्रत्यक्ष सहभाग देण्यात महाराष्ट्र अग्रेसर राहिला आहे. महिलांना सबळ कर्तव्य पथावरील ग्रह करावे घटना दुरुस्ती हे महत्त्वाचे पाऊल आहे. याची जाणीव असल्यामुळे प्रथम जिल्हा पातळीवरील पातळीवर महिला व बालविकास समितीची स्थापना करण्यात आली आणि पंचायतराज व्यवस्थेत ३३ टक्के मुलांसाठी राखून ठेवण्यात आल्या.

संशोधनाची उद्दिष्ट्ये-

महिलांचा राजकीय सहभागाचा अर्थ अभ्यासणे

स्थानिक स्वराज्य संस्थेतील महिलांची भूमिका अभ्यासणे

संशोधन पध्दती

प्रस्तुत शोधनिबंधासाठी संकलित तथ्यांचे संकलन हे प्रामुख्याने दुय्यम स्रोतांच्या माध्यमातून केले जाणारे आहे. यासाठी प्रामुख्याने स्थानिक स्वराज्य संस्थेच्या संदर्भात पुस्तके, महिलांचे नेतृत्व या संबंधीत साहित्याचा आढावा घेतला आहे.

महिलांचा राजकीय सहभागाचा अर्थ

राजकारणातील सहभाग ही लोकशाही राजकीय व्यवस्थेच्या यशस्वीतेतील अत्यावश्यक व महत्त्वाची

look healthier and freour.

Conclusion

Recognizing women's and men's distinct roles in family nutrition is a key to improving food security at the household level. To tackle this issue, FAO bases its approach to nutrition on the economic and cultural context of the area, on the economic and cultural context of the area concerned, and considers that food security depends not only on the availability of food, but also on access to food, as well as on food adequacy and acceptability to consumers. Other underlying causes of malnutrition must also be addressed. These include dietary intake and diversity, health and disease, and maternal and child care – areas in which women play decisive roles.

Recommendations

1. Olive oil can help reduce risk of certain cancers, blood pressure problems and heart disease, but there are 100 calories in each tablespoon so it is crucial to use it sparingly. One can either add a little in salad or use it in cooking.
2. Water-soluble vitamins include the vitamin B-complex and vitamin C, and are essential nutrients needed daily by the body in very small quantities.

References

Hotz Christine, González-Cossío (2016). The American Society for Nutritional Sciences, The Effect of Micronutrient Deficiencies on Child Growth: A Review of Results from Community-Based Supplementation Trials. <http://jn.nutrition.org/content/133/11/4010S.full>.

Hector, Cori; Kumar (2014), "Fortification of cereals, rice, oil and sugar" Proceeding of the International Conference on Micronutrient Fortification of Food Science, application and management held at Delhi", pg 35-36.

Orphan Nutrition (2016). An International Organization. Report: An Initiative of a child's best start to improve nutrition and feeding orphaned children Micronutrient Malnutrition. What is micronutrient deficiency? What is Vitamin- D Deficiency? http://www.orphannutrition.org/understanding-malnutrition/micronutrient-malnutrition/Vitamin_D.



Misconceptions of Covid-19 Vaccine among College Students

S. D. Wakode

Head, Department of Psychology, Vidya Bharati Mahavidyalaya, Amravati

Ms. Vidhya Ambhore

Department of Psychology, Mahatma Fule Mahavidyalaya, Amravati

Abstract: The study aimed to reveal vaccine related misconceptions among college students. A questionnaire based on seven misconceptions identified by UNICEF was used. Online survey was conducted. Nine hundred eighty nine respondents replied the questionnaire. The study revealed that majority of respondents (80.50%) believed in natural immunity. Study stream was a partial factor in vaccine related misconceptions. Lower income group (0-1 Lakhs) exhibited significantly higher vaccine related misconceptions. It is also evident that gender is not a crucial factor in vaccine related misconceptions.

Introduction

The emergence of a novel Corona virus indicates that our understanding of its mutations and risk factors for infection is still limited. Therefore, the tragedy that we are facing today makes us search for answers and corrective actions urgently for the survival of humans. But due to the information boom caused by internet and social media there are enormous misconceptions about this disease. Misconceptions related to corona virus disease-2019 (COVID-19) has been spread out broadly and the World Health Organization declared these as a major challenge to fight against the pandemic. Therefore,



Misconceptions of Covid-19: The Role of Social Media in its Spread

S. D. Wakode (1)

Head, Department of Psychology
Vidya Bharati Mahavidyalaya, Camp, Amravati
Mo- 9421829222
E- Mail : shankarwakode04@gmail.com

Dr. G. R. Ratnaparkhi (2)

Head, Department of Psychology
Mahatma Fule Mahavidyalaya, Amravati
Mo - 9850158676
E- Mail: gratnaparkhi@gmail.com

Abstract:

The study aimed to reveal Covid-19 related misconceptions and role of social media in spreading rumours and misconceptions. Online survey was conducted. Nine hundred eighty nine respondents replied the questionnaire. The study revealed study stream was a partial factor in Covid-19 related misconceptions. The factors of income and gender was found non significant. It is also evident that social media plays a partial role in spreading Covid related misconceptions.

Introduction:

In modern era, social media is one of the most used sources of information around the World. It is considered as a two edge sword. However, studies related to the role of social media in spread of rumors and correct information are equivocal. In the Covid -19 crisis, González-Padilla and Tortolero-Blanco (2020) reported that social media has become an important tool for individuals to communicate with friends and family during quarantine periods. It functioned as a tool to minimize the negative effect of isolation that turned into to anxiety, stress, and fear. On the other hand, Torales, O'Higgins, Castaldelli-Maia and Ventriglio (2020) found that during the pandemic, social media platforms spread misleading rumours, misinformation, life-endangering consequences of supposed cure of the disease, aetiology, preventions, vaccinations, and conspiracy theories about the origin of the virus. Moreover, the dangerous issue is that misinformation and rumours spread on social media are faster than reliable information, damaging the authenticity, balance of the news system, in particular health systems (Tasnim, Hossain and Mazumde, 2020).

Literature Review:

In the Cross-sectional online survey of Saudi population Mukhtiar Baig, et al. (2020) explored the predictors of misconceptions, knowledge, attitudes, and practices related to the COVID-19 pandemic. The sample comprised of 2006 participants; (47.5%) females and (52.5%) males. A majority of respondents (43.9%) have reported that social media platform was the leading source of information. The authors concluded that about one-third of participants (31.7%) had self-reported disturbed social, mental, and psychological wellbeing due to the pandemic and many participants became more religious during the pandemic.

Recently, Md. Sayeed Al-Zaman (2021) has reported that India has emerged as the biggest source of Covid information, with 1 in 6 pieces of fake information is coming out of the country. The author analyzed 9,657 pieces of misinformation that originated in 138 countries. Of all countries, India (18.07%) produced the highest amount of social media misinformation. The author concluded that social media (84.94%) produces the highest amount of misinformation, internet (90.5%) and Facebook alone produces 66.87% misinformation among all social media platforms.

Continuous holding of some misconceptions about the reality of the virus may create further problems for humankind. In this context Aminu (2020) has documented some of the

Effect of Endosulfan on Intestine and Pancreas of frog *Rana tigrina*

Rathod S. H, N. R, Thorat

Department of Zoology, Vidya Bharati Mahavidalaya, Amravati, Maharashtra, India

ABSTRACT

Article Info

Volume 9, Issue 1

Page Number : 256-260

Publication Issue

January-February-2022

Article History

Accepted : 08 Feb 2022

Published : 17 Feb 2022

Now a day's use of pesticides is the matter of great discussion for environmentalist to gain more relevant information on tolerance of organism to the pesticides. *Rana tigrina* were exposed to sub lethal concentration of Endosulfan (0.073 ml/lit) for 7 days regularly. The goblet cells of intestine was found to be swollen, enlargement of cells, cell shrinkage, damage of cell membrane, vacuoles formation and pancreas exocrine secretory acini cells was found to be damaged and connective tissue was loosely bounded also observed on exposure to Endosulfan.

Keywords : Endosulfan, Intestine, Pancreas.

I. INTRODUCTION

Now days the tremendous increase in environmental pollution is seen. Environmental pollution is due to the introduction of pollutants into natural environment that causes instability disorders, harms or discomfort the living organism in the ecosystem. Pollution can take the form of toxic chemical substance and contaminated water, soil, air which showed effect on the targeted and non- targeted organisms. It has been studied that acute toxicity of endosulfan on *Bufo bufo* gills & stream near sprayed agriculture field, after 24, 48, and 96 hours of exposure showed 50% mortality (LC 50) (Ilaria Bernabo *et.al.*, 2008). It has been found that exposure to natural and synthetic estrogenic chemicals may adversely affect wildlife and human health (Colborn

et al., 1993). There are various pesticides such as, Organochlorine (*Endosulfan, Endrin*) are used in fields and gardens. The environmental toxicological studies on vertebrates is rapidly expanding, fishes have become valuable indicator for the evaluation of the effects of toxic compounds (Khidr and Mekawy, 2008). Histology and histopathology can be used as biomonitoring tools for health in toxicity studies (Meyers and Hendricks, 1985). Histopathological alterations are biomarkers of effect exposure to environmental stressors, revealing alterations in physiological and biochemical function (Hinton *et.al.*, 1992). Histopathology, the study of lesions or abnormalities on cellular and tissue levels is useful tool for assessing the degree of pollution, particularly for sublethal and chronic effects (Bernet *et. al.*, 1999). There are various ways of spread of pesticidal

pollution in the environment such as rain water drained off from the pesticides spread field through which pesticides residues reach to environment and cause toxic effect on the aquatic and other organisms. In past several decades, decline in amphibian population has been occurring all over the world, for unexplained reasons which are thought to be varied but of which pesticides may be a part. Mixtures of multiple pesticides appear to have a cumulative toxic effect on frogs. Tadpoles from ponds with multiple pesticides present in the water lake longer period to metamorphosis into adult frogs, decreasing their ability to catch prey and to avoid predators. (Benoit *et.al.*, 2003) studied that the effect of cadmium , Endosulfan and atrazine on African frog (*Xenopus leavis*) and Bull frog *Rana catesbeiana* showed adverse impact on secretory capacity of adrenal cells of amphibians. Amphibians itself acts as a pest controlling organism plays a vital role in food web and are commonly found in agriculture fields, near ponds and rivers. When they came in contact with pesticides they absorb orally, cutaneously or by inhalation and get affected even though when they feed on the affected insects they get affected indirectly.

II. Material and Method

Adult frog (*Rana tigrina*) of both sexes were collected by net or hand from their spawning ponds at unpolluted and non-agriculture site. The collected frogs were transported to laboratory in covered baskets. Adult frogs of the same size and same weight (35-40 gm) were acclimatized in glass aquarium tank for the time period of 10 to 15 days in laboratory condition; frogs were feed twice a day alternatively by insects. Stock solutions of experimental dose were prepared by using Endosulfan and ethanol as a vehicle. From 0.073 ml stock solution is used as dose in per liter water after acclimatization of 10-15 days frogs has become divided into two groups:

GROUP I - Control

GROUP II – Experimental

Group I and Group II consist of six adult frogs respectively. Group I were placed in plane water glass aquarium. While Group II is treated with 0.073 ml/lit of dose of Endosulfan for 7 days. On eighth day frogs of both groups were sacrifice for further experimentation.

For Histopathological observation, after 7 days, frog of each group were removed and dissected. Small pieces of the intestine and pancreas were taken and immediately fixed in alcoholic bouin's fluid. Fixed tissues were processed routinely for paraffin embedding technique. Embedded tissues were sectioned at 5-7 μ in thickness and then stained with double staining method. Finally the sections were proceed for microscopic studies for observation and collect their respective photographs for observation.

III. Observations

Histopathological changes in intestine exposed to Endosulfan:

Frog *Rana tigrina* exposed to sub lethal concentration of Endosulfan showed the morphological changes. In the present investigation cells of intestine of treated frog showed flaccid and degenerative stage. Cells of intestine showed shrinkage, damaged and ruptured membrane of the cells, and space between the intestinal goblet cells was clearly observed.

Histopathological changes in pancreas exposed to Endosulfan:

Frog *Rana tigrina* exposed to sub lethal concentration of Endosulfan showed the morphological changes in the cells of pancreas. It showed the disruptor of the secretory acini cells and damage of the islets of langerhans cells of pancreas, as well as connective tissue was loosely bounded also observed.

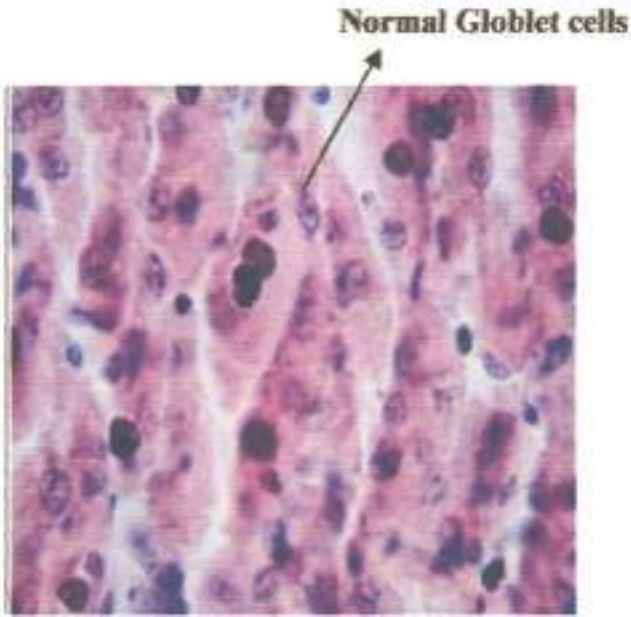


Fig no.1) Section of control frog intestine-Goblet cells of normal size are observed before treatment to endosulphan.



Fig. 1.1 Section of endosulfan treated intestine Goblet cells become swollen and enlarged after 7 days treatment of endosulfan

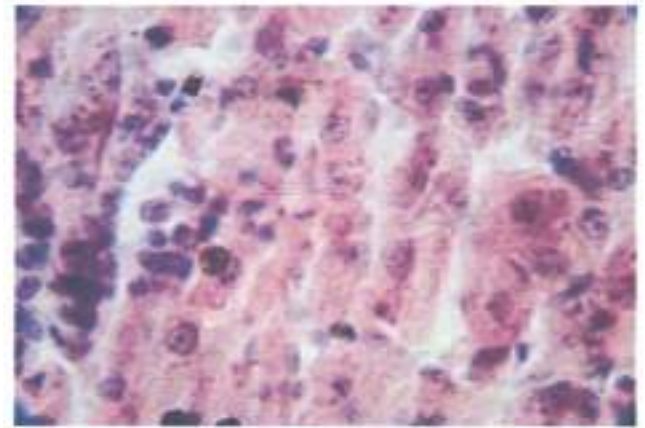


Fig.2) Section of Control frog intestine

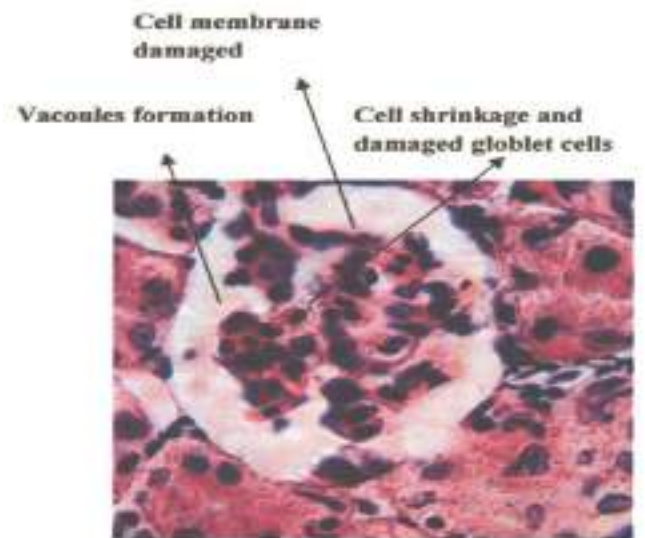


Fig.2.1) Section of Endosulfan treated frog intestine- Shows the endosulfan effects, goblet cells shrinkage, damage of cell membrane and vacoules formation.



Fig. 3) Section of control frog Pancreas- Connective tissue and exocrine secretory normal of frog Rana tigrina

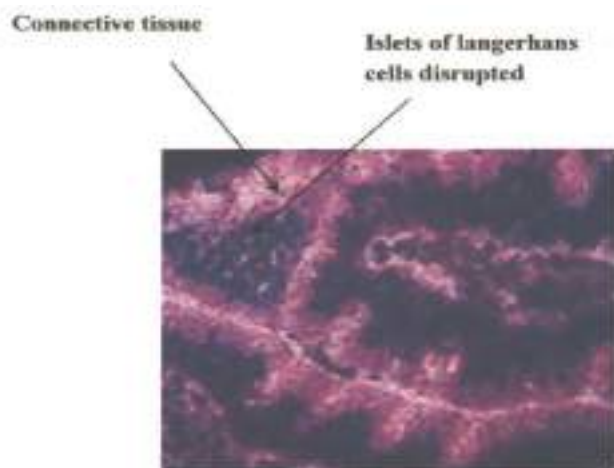


Fig. 3.1) Section of Endosulfan treated Pancreas- shows exocrine secretory acini cells suspended and islets of langerhans cells become disrupted.

IV. RESULTS AND DISCUSSION

Pesticides indirectly and directly affect on non-targeted organism and may interfere the physiological process of the living organism. A pesticide not only affects the physiochemical properties of the ecosystem but also affects the flora and fauna. The proper and safe use of pesticides will be ensuring only a good knowledge of their toxicological characteristics and behavior in biological media to be reviewed. Here we study the effect of Endosulphan on intestine and pancreas of frog *Rana tigrina* which shows morphological changes.

Histological changes in the Pancreas:

In the present investigation the effect of *endosulfan* shows the morphological changes in the intestinal cells the cell become swollen, cell membrane of intestinal cell was highly affected and disturbed, vascularization and degeneration of intestinal cell were also observed. Similar finding by **Chayya Roy Kundu et.al.,(2011)** observed the effect of malathion at sublethal concentration (0.006) on the intestine of cricket frog (*Fegarvarya limnocharis*) was observed for 24 hour to 240 hour of exposure and remarkable histopathological alteration were

observed it showed acute pathological condition in intestinal wall, due to toxicity the cytoplasm of cells disintegrated become empty and vacuolated, cell membrane was ruptured and degenerative villi of intestine also observed.

Similar finding also observed by **T. Braunbeck and S. Appelhum (1999)** on the exposure of endosulfan to carp *Cyprinus carpio* for 5 weeks it showed the liver alteration and enlargement of nucleolus, Golgi complex and rough endoplasmic reticulum and ultra structure of intestine shows complete lack of chylomicrons in epithelial lining which indicates disturbance of intestinal absorption.

Histological changes in the Pancreas:

In the present investigation toxicity of endosulfan showed impact on morphological changes in pancreatic cells .The cells of islets of langerhans secretory acini cells becomes ruptured and degeneration of pancreatic cells were observed. Similar Study were observed by **Ozlem Oznen et.al.,(2010)** studied the exposure of endosulfan and vitamin C on rabbit pancreatic cells. It was showed remarkable degenerative changes and decrease in proinsulin-insulin and amylin secreting cells also slight decrease in glucagon secreting cells. From the present investigation it was suggested that endosulphan may show impact on the endocrine organs of frog, endosulfan is toxic to the amphibians and other organism so there is need to more research to develop a less toxic pesticides which will not harmful for targeted and non-targeted organism.

V. REFERENCES

- [1]. Bernabò, I., Brunelli, E., Berg, C., Bonacci, A., Tripepi, S., (2008): Endosulfan acutotoxicity in *Bufo bufo* gills: ultrastructural changes and nitric oxide synthase localization. *Aquat. Toxicol.* 86, 447–456.

- [2]. Bernet, D., Schmidt, H., Meier, W., Brkhardt-Holm, P. & Wahli, T. (1999): Histopathology in fish: Proposal for a protocol to assess aquatic pollution. *Journal of Fish Diseases*. Vol. 22: 25–34.
- [3]. Benoit N. Goulet., Alice Hontela (2003): Toxicity of cadmium, endosulphan, and atrazine in Adrenal striodogenic cells of two amphibian species, *Xenopus laevis* and *Rana catesbeiana*. *Environmental Toxicology and Chemistry*, Vol. 22, No. 9, pp. 2106–2113,
- [4]. Braunbeck T, Appelbaum S (1999): Ultra structural alterations in the liver and intestine of carp *Cyprinus carpio* induced orally by ultra-low doses of endosulfan. *Dis. Aquat. Org.* 36: 183-200.
- [5]. Colborn, T., Vom Saal, F.S. & Soto, A.M. (1993): Developmental effects of endocrine disrupting chemicals in wildlife and humans. *Environmental Health Perspectives*. Vol. 101: 378–384..
- [6]. Chayya R Kundu., S. Roychoudhury, M. Capcarova (2011): Malathion-induced sublethal toxicity on the intestine of cricket frog (*Fejervarya limnocharis*) *Journal of Enviro Sci and Health Part B Pesticides Food Contaminants and Agricultural Wastes* 46(8):691-6
- [7]. Hinton, D.E., Baumann, P.C., Gardner, G.R., Hawkins, W.E., Hendricks, J.D., Murchelano, R.A. & Okihiro, M.S. (1992): Histopathological biomarkers, In: Huggett, R.J.,
- [8]. Khidr, M. B. & Mekkawy, I.A.A. (2008): Effect of separate and combined lead and selenium on the liver of the cichlid fish *Oreochromis niloticus*: ultrastructural study. *Egyptian Journal of Zoology*. Vol. 50: 89-119.
- [9]. Ozlem Ozmen., M. Firdevs (2010) : Effects of vitamin C on pathology and caspase-3 activity of kidneys with subacute endosulfan toxicity *Biotechnic and Histochemistry* 90(1):1-6 . August 2010
- [10]. Meyers, T.R. & Hendricks, J.D. (1985): Histopathology. In: Loux, D.B., Dorfman, M., (Eds.), *Fundamentals of Aquatic Toxicology: Methods and Applications*, Hemisphere USA, pp. 283–330.

Cite this article as :

Rathod S. H, N. R. Thorat, "Effect of Endosulfan on Intestine and Pancreas of frog *Rana tigrina*", *International Journal of Scientific Research in Science and Technology (IJSRST)*, Online ISSN : 2395-602X, Print ISSN : 2395-6011, Volume 9 Issue 1, pp. 256-260, January-February 2022. Available at doi : <https://doi.org/10.32628/IJSRST229145> Journal URL : <https://ijsrst.com/IJSRST229145>



Effects of water extract of *Parthenium hysterophorus* on Haematological contents of freshwater fish *Labeo rohita*.

(Vikhar)Khedkar, A.P.

Assistant Professor, Department of Zoology,

Vidya Bharati Mahavidyalaya, Camp, Amravati (M.S.)

Email: alkavikhar@gmail.com

ABSTRACT: The present study was conducted and assessed the toxicity of *Parthenium hysterophorus* is on locally available fresh water fish. The present study deals with to determine haematological effects on fresh water fish, *Labeo rohita*. The fishes were collected from nearby water body and brought to the laboratory for determination of haematological parameters up to 96hr. period. The blood parameters are important in diagnosing the structural and functional status of the animal. Hematological and biochemical indices provide extensive information about fish oxygen transport capacity, immune potential, level of stress, disease, intoxication, nutritional status etc.

(KEYWORDS: *Parthenium hysterophorus*, WBC, Platelets count, *Labeo rohita*.)

INTRODUCTION:

Fresh water can be defined as water with less than 500 parts per million (ppm) of dissolved salts (Walter, C. and Maguire, J. 1996). Water is an essential issue for the survival of all living organisms especially aquatic animals like fishes. Water is a critical issue for the survival of all living organisms. To encourage the fishery is necessary for ever increasing demands for protein rich food to earn valuable foreign exchange (Varadharajan, 2012). Water of good quality is required for living organisms. Water quality characteristics influence histopathological appearances of poisonous effects (Bhavan and Geraldine, 2000).

Rohu is the natural inhabitant of freshwater sections of the rivers. Rohu is a bottom feeder. It is diurnal and generally solitary. It is commonly eaten in Bangladesh and the Indian states of Bihar, West Bengal. The higher concentration of toxicants bring the adverse effects on aquatic organisms, at cellular level or molecular level and ultimately lead to disorder in biochemical composition which is useful in determining different toxicants and

protective mechanism of the body to resist the toxic effects of the substances (Jain and Kulshreshta, 2000). *Parthenium hysterophorus* L.(Congress grass) is an exotic weed comes under Asteraceae family, accidentally introduced in India, 1955 in Pune through the imported food grains(Dhawan and Dhawan, 1996). It is harmful to all the living beings; it has nearly destroyed all the useful crops and plants, even though growing near to it. It is known to cause asthma, bronchitis, dermatitis, and hay fever in man and livestock (Narasimhan et al. 1977). Root extracts are useful in dysentery (Singh et al. 1996). It is used as folk remedies in West Indies and Central America (Navie et al. 1996). Sharma and Bhutani, (1988) also reported as *parthenium* is promising remedy against hepatic amoebiasis. It contains several important chemical constituents, mainly histamine, saponin, glucosides and triterpene (sesquiterpene).

Blood is most important and abundant body fluid. Its composition often reflects the total physiological condition Blood of living organisms are very sensitive to changes and are widely used in Ichthyology research, aquaculture research as well as toxicology and biological monitoring (Svoboda et al., 2001; Adedeji et al., 2009; Adeyemo, 2008).WBCs and platelets being the essential cellular components of fish blood Significant reduction in rat WBC count after oral treatment of Parthenium extract signifies its immune system weakening ability (Yadav et al. 2010). Significant decreases in the number of circulating thrombocytes in a teleost leads to increases in clotting time, loss of haemoglobin content and immature erythrocytes.The methanol extract of the flowers showed significant antitumour activity and parthenin exhibited cytotoxic properties against T cell leukaemia, HL-60 and Hela cancer cell lines (Das et al. 2007). Panigrahi and Mishra (1978) have reported a increase in WBC in *Colis fasciatus* exposed to metals. The leucocytosis may also be attributed for the removal of cellular debris of necrosed tissue in the rats under the toxic stress (Mcleay and Brown, 1974). An increase in WBC count after chemical stress recorded in the present study is in accordance with various workers (Sharma et al., 1984; Tyagi, 1984; Goel and Maya, 1986). White blood corpuscles (WBC) play a major role in defence mechanism (Adedeji et al., 2009)The response of the leucocytes to the changes in water quality and chemicals is variable (Nussey et al., 1995). Several other authors also reported an increase in leucocyte count in various fresh water fishes exposed to different heavy metal salts which supports findings in the present study (Ali et al., 2000). Platelets are first line of defense against accidental blood loss and platelets play an important role in the inflammatory response (Smith et al., 1976).

MATERIAL AND METHODS:

Labeo rohita fish were collected from Nal Damayanti Dam,local fish market Amravati washed with 10% solution of Potassium Permagnate to free any fungal infections. Then acclimatized to the laboratory condition for fifteen days in large aquarium. The fish size 15 to 20 cm in length and weight 150 to 200 gm. Fishes maintained in well water and its physico-chemical characteristics analyzed following the method given in APHA (2005). Fishes fed with add libitum food, oil cake and rice bran to keep them more or less in the same state of metabolic requirement.

A group contain ten fishes were taken in both container experimental and control respectively. The dose starting from 10 ml in 10 lit. Well water. The dose increased daily by 10 ml. Their behavioral changes recorded daily and throughout the exposure period. Everyday water change to maintain the concentration of *Parthenium hysterophorus*

extract and histological changes were recorded.

RESULT AND DISCUSSION:

The exposure of *Labeo rohita* to lethal levels of extract resulted in time-and concentrations dependent significant.

A) Total leukocyte count or White Blood Cell

In the present investigation the control values for haematological parameter WBCs for lethal concentration is (9.80±0.017). At different concentration 0.50 ml/l to 6.00 ml/l of extract of *Parthenium hysterophorus* at 24hrs (10.40±0.005 to 14.89±0.005), 48hrs (10.42±0.025 to 14.37±0.025), 72hrs (10.43±0.061 to 12.76±0.38) and 96hrs (10.55±0.032 to 11.37±0.30). It is significantly increase to increasing the concentration but decrease at the end.

Table 1: Total leucocyte count of *Labeo rohita* exposed to *Parthenium hysterophorus* extract at different time intervals.

Total Leucocyte Count ($\times 10\text{mm}^3$)							
Exposure Period (hrs.)	Lethal Concentration						
	0.50	1.00	1.50	3.00	4.00	5.00	6.00
Control	9.80 ±0.017	9.80 ±0.017	9.80 ±0.017	9.80 ±0.017	9.80 ±0.017	9.80 ±0.017	9.80 ±0.017
24hrs.	10.40 ±0.005	10.51 ±0.023	10.86 ±0.030	11.19 ±0.037	11.83 ±0.055	12.43 ±0.035	14.59 ±0.005
48hrs.	10.42 ±0.025	10.57 ±0.045	11.26 ±0.10	11.53 ±0.037	12.22 ±0.045	12.65 ±0.015	14.37 ±0.025
72hrs.	10.43 ±0.061	10.76 ±0.011	11.33 ±0.020	11.89 ±0.15	12.53 ±0.026	13.62 ±0.020	12.76 ±0.38
96hrs.	10.55 ±0.032	11.19 ±0.01	11.82 ±0.060	12.48 ±0.020	12.93 ±0.085	12.22 ±0.015	11.37 ±0.30

Each value is the mean \pm SD of three observations; p value represents significant of differences between control and experimental animal. Significant at $p < 0.05$.

B) Total thrombocyte count or Platelets

While the control values for hematological parameter Thrombocytes or Platelets for lethal concentration is (19.52±0.20). At different concentration 0.50 ml/l to 6.00 ml/l of extract of *Parthenium hysterophorus* at 24hrs (19.68±0.15 to 27.60±0.02), 48hrs (20.2±0.3 to 28.36±0.03), 72hrs (21.30±0.37 to 29.3±0.26) and 96hrs (22.34±0.25

to 30.46 ± 0.20). It is significantly increased to increasing the concentration.

Table2: Total Thrombocyte count of *Labeo rohita* exposed to *Parthenium hysterophorus* extract at different time intervals.

Total Thrombocyte Count ($\times 10^3 \text{ mm}^3$)							
Exposure Period (hrs.)	Lethal Concentration						
	0.50	1.00	1.50	3.00	4.00	5.00	6.00
Control	19.52 ± 0.20	19.52 ± 0.20	19.52 ± 0.20	19.52 ± 0.20	19.52 ± 0.20	19.52 ± 0.20	19.52 ± 0.20
24hrs.	19.68 ± 0.15	20.44 ± 0.15	21.63 ± 0.15	22.32 ± 0.30	24.50 ± 0.26	25.40 ± 0.1	27.60 ± 0.2
48hrs.	20.20 ± 0.3	21.20 ± 0.05	22.14 ± 0.25	23.44 ± 0.35	25.55 ± 0.20	26.17 ± 0.15	28.36 ± 0.30
72hrs.	21.30 ± 0.37	21.86 ± 0.20	23.35 ± 0.40	24.27 ± 0.20	26.47 ± 0.45	26.15 ± 0.35	29.30 ± 0.26
96hrs.	22.34 ± 0.25	23.30 ± 0.1	24.47 ± 0.40	25.10 ± 0.36	27.60 ± 0.26	28.56 ± 0.25	30.46 ± 0.20

Each value is the mean \pm SD of three observations; p value represents significant of differences between control and experimental animal. Significant at $p < 0.05$.

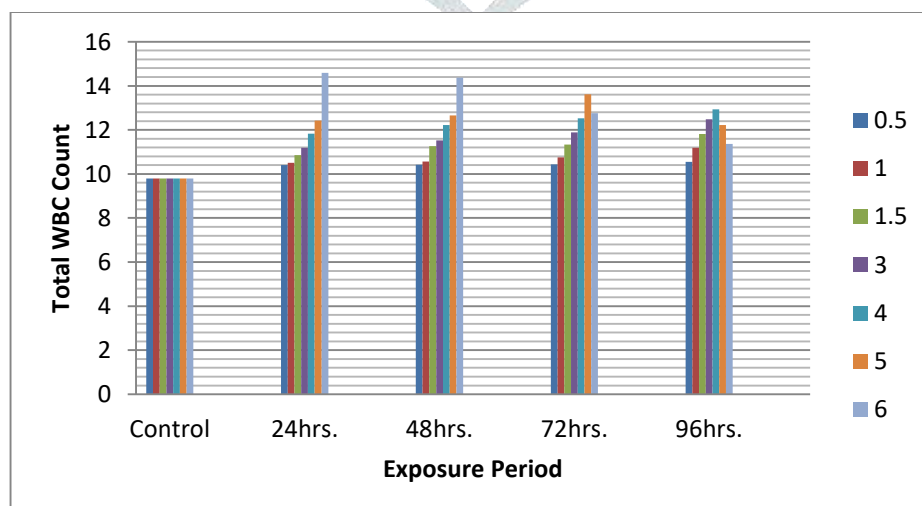


Fig.1: Total leucocytes count for *Labeo rohita* exposed to *Parthenium hysterophorus* water extract of lethal concentration at different time intervals.

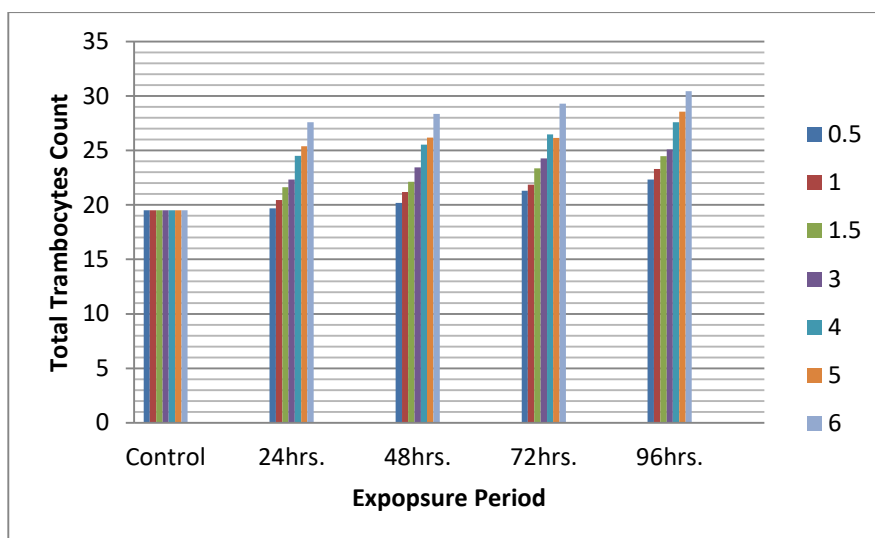


Fig2.: Total Thrombocytes count for *Labeo rohita* exposed to *Parthenium hysterophorus* water extract of lethal concentration at different time intervals.

In the present work the TLC or WBC increase continued in lethal effect at different exposure periods in fish exposed to the two lower concentrations (Table1). But, it reduced in higher concentration at 96 h exposure period in lethal effect it decreases during long exposure period (Table1). White Blood Cells (WBC) count decrease with increasing level of the toxicant and prolonged exposure to extract that affected leucopoiesis. This assumption is also in line with, (Das *et al.*, 2006). Decrease in total WBC counts reflects a state of stress in fish and points to the role of compound present in to extract as a potential environmental stressor. The findings of the present study is in coincide with the work of (Subeena Begum and Navaraj, 2012) who reported that WBC counts are higher in *Labeo rohita* fingerlings treated with *Mangifera indica* when compared to control. Thrombocytes are involved in blood clotting and vital to the haemostatic plug after vascular injury. The increase in the level of thrombocytes in the present study during lethal effect at different exposure period up to 96 h (Table2) suggests the toxicant may have enhanced thrombocytopoiesis through increased rate of conversion of arachidonic acid to thromboxane B2 (Craig *et al.*, 2002). In the present work, the thrombocytes count of fish exposed to the extract steadily increased up to end of the experiment. These findings are supported by (Vijayamohan, 2000). There is steady increase in thrombocyte count as the exposure period increased, in the present study must be due to entry of toxicant root extract in blood, which causes haemolysis and tissue damage that librates more platelets in the circulation.

CONCLUSION:

A) Total leukocyte count or White Blood Cell

In the present investigation the control values for hematological parameters for lethal concentration is (9.80 ± 0.017) . At different concentration 0.50 ml/l to 6.00 ml/l of extract of *Parthenium hysterophorus* at 24hrs (10.40 ± 0.005 to 14.89 ± 0.005), 48hrs (10.42 ± 0.025 to 14.37 ± 0.025), 72hrs (10.43 ± 0.061 to 12.76 ± 0.38) and 96hrs (10.55 ± 0.032 to 11.37 ± 0.30). It is significantly increase to increasing the concentration but decrease at the end.

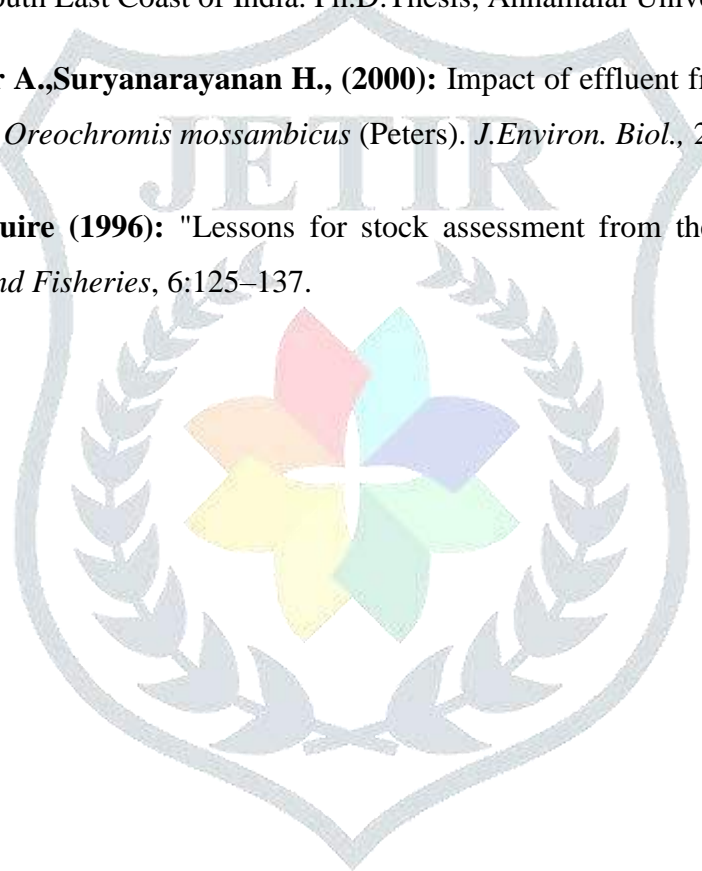
B) Total thrombocyte count or Platelets

In the present investigation the control values for hematological parameters for lethal concentration is (19.52±0.20). At different concentration 0.50 ml/l to 6.00 ml/l of extract of *Parthenium hysterophorus* at 24hrs (19.68±0.15 to 27.60±0.02), 48hrs (20.2±0.3 to 28.36±0.03), 72hrs (21.30±0.37 to 29.3±0.26) and 96hrs (22.34±0.25 to 30.46±0.20). It is significantly increase to increasing the concentration

REFERENCES:

- Adedeji, O.B., Adeyemo, O.K. and Agbede, S.A. (2009):** Effects of diazinon on blood parameters in the catfish (*Clarias gariepinus*). *African J. Biotechnol.*, 8: 3940-3946.
- Adeyemo, O.K. (2008):** Histological alterations observed in the gills and ovaries of *Clarias gariepinus* exposed to environmentally relevant lead concentrations. *J. Environ. Health.* 70 (9): 48-51.
- Bhavan, P.S. and P. Geraldine, (2000):** Histopathology of the hepatopancreas and gills of the prawn *Macrobrachium malcolmsonii* exposed to endosulfan. *Aqua.Toxicol.* 50: 331-339.
- Craig, J.I.O., Haynes, A.P., McClell D.B.L. and Ludlam C.A. (2002):** Avidson's principle and practice of medicine, 19th edition. (*Christopher, H., Edwin R.C., Nicholas A.B. and Nicki, R.C. (eds.)*) *Publisher.City.*
- Das B, Reddy VS, Krishnaiah M, Sharma AVS, Ravi Kumar K, Rao JV, Sridhar V.(2007):** Acetylated pseudoguaianolides from *Parthenium hysterophorus* and their cytotoxic activity. *Phytochemistry.* 2007; 68:2029–2034.
- Das, P.C., Ayyappan, S. and Jena, J. K. (2006):** Haematological changes in the three Indian major carps, *Catla catla* (Ham.), *Labeo rohita* (Ham.) and *Cirrhinus mrigala* (Ham.) exposed to acidic and alkaline pH. *Aquaculture*, 256: 80-87.
- Dhawan, S. R. and Dhawan, P. (1996):** Regeneration in *Parthenium hysterophorus* L. *World Weeds*, 2: 244-249.
- Goel, K. A. and Maya. (1986):** Haematological anomalies in *Clarias batrachus* under the stress of rogor. *Ad.Bios.* 5(II): 187-192.
- Jain, M. and K. Kulshreshta (2000):** Effect of pesticides on fishes. A review of recent studies in India. *J. Nation.* 7 (2): 14-18.
- Mc Leay, D. J. and Brown, A. D. (1974):** Growth stimulation and juvenile coho salmon (*Oncorhynchus kistuch*) exposed to bleached kraft pulpmill effluent for 200 days. *J. Fish. Res. Bd. Can.* 31: 1043- 1049.
- Narasimhan, T. R., Ananth, M., Narayana Swami, Rajendra Babu, Mangla, A. and Subba Rao P. V. (1977):** Toxicity of *Parthenium hysterophorus* L. *Curr. Sci.* 46(1): 15.
- Navie, S.C., Mcfadyen, R.E., Panetta, F.D. and Adkins, S.W. (1996):** The Biology of Australian Weeds 27. *Parthenium hysterophorus* L. *Plant Protection Quarterly* 11 (2) : 76-88.
- Sharma, G.L. and Bhutani, K.K. (1988):** Plant based antiamoebic drugs. Part II. Amoebicidal activity of parthenin isolated from *Parthenium hysterophorus*. *Planta Medica.* 54 : 20-22.

- Singh, U., Wadhvani, A.M. and Johri, B.M. (1996):** Dictionary of economic plants in India. Indian Council of Agricultural Research, New Delhi.
- Smith M J H.; Walker, J. R. Ford Hutchinson, A. W.; Penington, D. G. (1976):** Prostaglandin's inflammation Agents Action., 6 : 701-704.
- Subeena Begum, S. and P.S. Navaraj, (2012):** Synergistic effect of Plant Extracts Supplemented Diets on Immunity and Resistance to *Aeromonas hydrophila* in *Mystus keletius*. *J. Pharm. Bio. Sci.*, 2(4): 30-36.
- Svoboda, M., Luskova, V., Drastichova, J. and Ilabek, V. (2001):** The effect of diazinon on haematological indices of common carp (*Cyprinus carpio*). *Acta Vet.Brno.*, 70: 457-465.
- Varadharajan D. (2012):** Biodiversity and antimicrobial activities of Crabs from Arukkattuthurai to Pasipattinam, South East Coast of India. Ph.D.Thesis, Annamalai University, India.
- Vijayamohan, G., Nair A., Suryanarayanan H., (2000):** Impact of effluent from a TiO₂ factory on the peripheral haematology of *Oreochromis mossambicus* (Peters). *J. Environ. Biol.*, 21(4): 293-296.
- Walters, C. and J. Maguire (1996):** "Lessons for stock assessment from the northern cod collapse", *Reviews in Fish Biology and Fisheries*, 6:125-137.





Effect of *Terminalia Arjuna* Bark Extract on Streptozotocin-Induced Diabetic Rats

Akhare Y. D.¹, Shraddha Sharma²

^{1,2}Dept. of Zoology, VidyaBharati Mahavidyalaya, C. K. Naidu Marg Camp., Amravati. Maharashtra, India 444 602

ABSTRACT: The present study was carried out to evaluate the antihyperglycemic property of *Terminalia arjuna* in streptozotocin-induced type 2 diabetic model rats. To evaluate the antihyperglycemic and antioxidant role of ethanol extract of *Terminalia arjuna* leaf in rats. Hyperglycemia was induced in rats by single intraperitoneal injection of streptozotocin (STZ, 120mg/kg body weight). Three days after STZ induction, the hyperglycemic rats were treated with a dilution of bark extract of *T. arjuna* orally at the dose of 1 ml and 2 ml daily for 15 days. The level of blood glucose before meals were measured on every fifth day during the 15-day treatment. Ethanol extract of *T. arjuna* dose-dependently reduced and normalized blood glucose levels. *T. arjuna* treatment also significantly increases protein levels. The present study shows that *T. arjuna* leaf shows remarkable reduction in blood glucose level in STZ-induced diabetic rats. The *T. arjuna* has an effect in physiological and biochemical processes in rats.

KEYWORDS: Antihyperglycemic, streptozotocin, *Terminalia arjuna* (Combretaceae).

INTRODUCTION

Diabetes mellitus is a condition associated with high blood glucose levels and it is mainly linked with a low level of insulin in the blood or the inability of target organs to use insulin. It is the most prevalent disease in the world affecting 25% of the population and afflicts 150 million people and is predicted to rise to 300 million by 2025 (Vats et al. 2000). It causes numerous complications such as retinopathy, neuropathy, and peripheral vascular insufficiencies (Chehade & Mooradian, 2000). Diabetes is still not completely curable by the present antidiabetic therapy. Insulin therapy is the only satisfactory approach in diabetic Mellitus, even though it has several drawbacks like insulin resistance (Piedrola et al., 2001), anorexia, brain atrophy, and fatty liver in chronic treatment (Weidmann et al., 1993). There are many medications available To reduce blood glucose levels but certain adverse effects and weak effectiveness of them have led to the search for more effective agents.

Terminalia arjuna (Combretaceae), commonly known as Arjuna, is a large tree growing on the banks of rivers, streams, and dry watercourses and it is considered a versatile medicinal plant. In India, the plant has been traditionally used for several medicinal purposes. The fruits of the plant are used as a tonic. Externally, its leaf paste is used as a cover on sores and ulcers. The bark is antidiarrhetic, antipyretic, astringent, cardiotoxic, lithotripsy, and tonic; a powder of the bark acts as a diuretic in cirrhosis of the liver and gives relief in symptomatic hypertension (Chatterjee & Pakrashi, 1994). *T. arjuna* is a potent anti-diabetic and beneficial in the control of diabetes-related abnormalities in serum lipid profile renal markers and oxidative damage in liver and pancreas of streptozotocin-induced rat model (Parveen et al. 2011). The powder of the bark is also given with honey in fractures and contusions with ecchymosis. Besides this, the extract of the bark as an astringent is used for cleaning sores, ulcers, cancers, and so on (Dhiman, 2006). The stem bark of *Terminalia arjuna* is also used for the treatment of various cardiovascular diseases; (Parveen et al., 2012). No pharmacological investigation is still reported on *T. arjuna* leaf. Therefore the present study was performed to investigate the antidiabetic effect of extract of *T. arjuna* leaf against streptozotocin (STZ)-induced diabetic rat. Streptozotocin is a naturally occurring chemical that is particularly toxic to the insulin-producing beta cells of the pancreas in mammals. It is used in medicine for treating certain cancer cells of islets of Langerhans and used in medical research to induce an animal model with Type 1 diabetes in large doses as well as Type 2 diabetes with multiple low doses. (Brentjens R, Saltz, 2001).

Traditional healers claim that the stem bark of the plant possesses anti-diabetic properties. Some scientific studies also show the hypoglycemic property of this plant. However, no published report supports both the acute and chronic hypoglycemic effect of *T. arjuna* on streptozotocin-induced diabetic model rats. As the majority of the diabetic population suffers from diabetes, we undertook the present study to evaluate the anti-diabetic effect of *T. Arjuna* on diabetic model rats and to analyze the possible hypoglycemic /

antihyperglycemic and activity of the extract as well as to investigate the possible chemical constituents responsible for the activity and the target tissue involved in this action.



MATERIALS & METHODS

Plant Extract:-

The experimental animals were administered intramuscularly with active principles of plants extract *T. arjuna*. Dilution of drugs. *T. Arjuna* was made with 100 ml distilled water. From the above stock, 1 ml and 2 ml quantities were administered intramuscularly to the high region of the experimental rat using a 2 ml disposable syringe and 24G disposable needle daily once for up to 15 days.

Experimental Setup:-

Experiments were carried out by dividing the rat of both sexes into various groups.

Group I:- Served as a normal healthy control rat.,

Group II:- Served as experimental rat (EXPT. A. Diabetic induce control rat given 1 ml. bark extract 500 mg/kg PO.)

Group III; - served as experimental rat (EXPT. B. Diabetic induced control rat given 2 ml. *T. arjuna* extract 250 mg/kg PO.)

Group IV: - Served as untreated diabetic induced control rat. (EXPT. C) Induction of diabetes in rat: -120 mg/kg of STZ in sterile saline was induced by a single IP injection. The duration of treatment was 15 days.

Collection of blood, kidney, and liver from the rat:-

After the experimental regimen, the animals were sacrificed by cervical dislocation under mild chloroform anesthesia. Blood was collected on decapitation and serum was separated by centrifugation (for 20 min at 2000 rpm). The liver and kidney were excised immediately and thoroughly washed in ice-cold saline. The serum and tissues were collected and used for biochemical experiments.

Estimation of serum glucose was estimated by GOD/POD method (Trinder, P. 1969).

Determination of total protein in tissue:-

Liver protein analysis can diagnose the liver disorder, nutritional deficiencies of protein, renal failure, and lymphoproliferative disorder. 100 mg(0.10g) of tissue was taken and dried by folding between the folding of filter paper and drop in a beaker kept on ice. It was transferred to a homogenizer tube and homogenized by taking 10 ml of distilled water. It was centrifuged and the supernatants were used for protein estimation. Protein is estimated by different methods as described by the de Lowry method. No method is 100% sensitive. Hydrolyzing the protein and estimating the amino acid alone will give the exact quantification. Specimens used are Liver, Kidney & Pancreas.

Determination of glycogen in tissue

“Glycogen, the branched-chain homopolysaccharide, is present both in liver & muscle. Glycogen can be effectively extracted from the liver by treatment with alkali followed by separation of glycogen by alcoholic precipitation.

Liver tissue was collected from a well-fed rat .10 g of liver tissue was weighed or transferred into a 100ml conical flask & added 20ml of 30% KOH and kept in a boiling water bath for 30 minutes Content was centrifuged at 4000 rpm for 30 min. The supernatant was discarded and the ppt was collected. The glycogen ppt is recovered by centrifugation of the above content with d/w and ethanol at 4000 rpm for 15 min. Discarded supernatant was dried for a glycogen sample in a desiccator containing calcium chloride as a desiccant and the % yield of glycogen was calculated.



RESULT & DISCUSSION

Table 1. Effect of *T. arjuna* stem bark on serum glucose, of control and experimental rats.

Parameters	Group 1	Group II	Group III	Group IV
	Control	Expt. A	Expt. B	Expt. C
Serum Glucose (mg/dl)	98.33 ± 02.66	* 302.67±22.35	* 125.60 ± 24.73	* 82.50 ± 04.72

Values are mean ± SD of five observations, *p<0.05, **p<0.01, and NS =Not significant.

Fig.:- 1

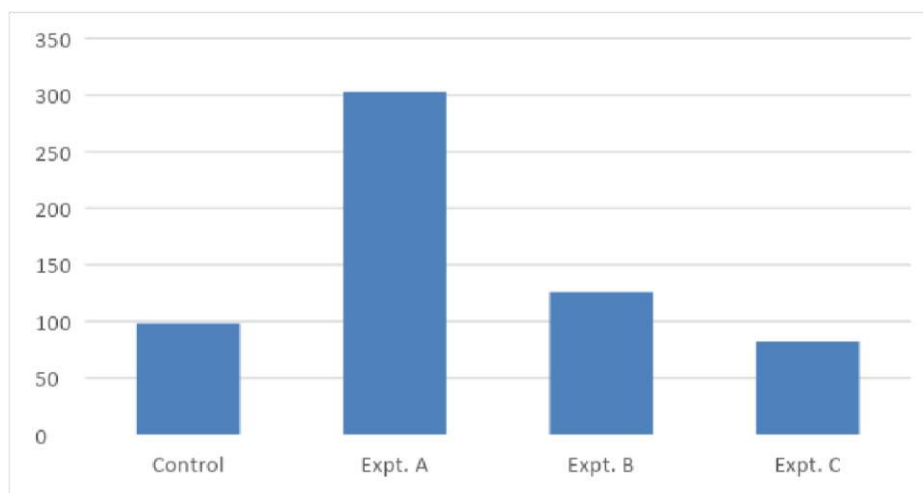


Table 2. Effect of *T. arjuna* plant extract on glycogen% in some organs of the rat.

Tissue\Group	Control	Expt. A	Expt. B	Expt. C
Liver	3.47 ± 0.34	** 1.02 ± 1.44	NS 0.7 ± 1.84	** 0.5 ± 0.37
Kidney	3.58 ± 0.39	** 0.78 ± 0.26	** 0.98 ± 0.57	** 7.52 ± 0.40
Pancreas	2.07 ± 0.27	** 1.71 ± 0.15	** 0.57 ± 0.36	* 1.19 ± 0.63
Muscles	2.25 ± 1.57	NS 0.37 ± 0.3	** 0.28 ± 1.65	NS 13.4 ± 11.14

Values are mean ± SD of five observations, *p<0.05, **p<0.01, and NS =Not significant.



Fig:- 2

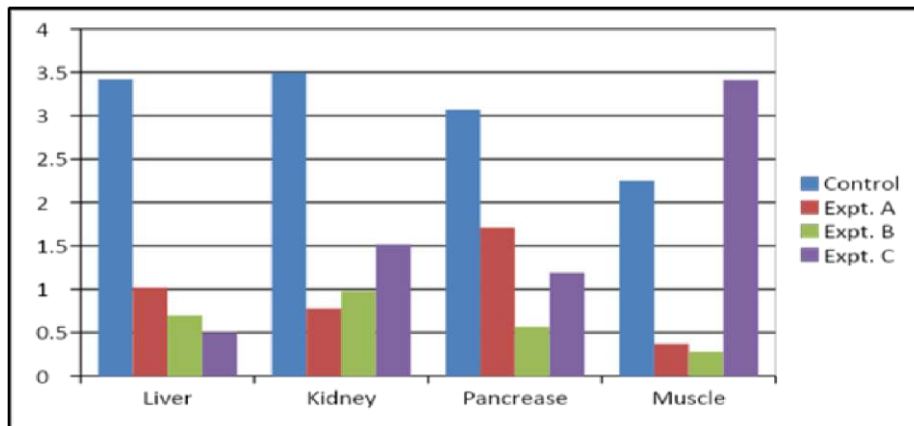
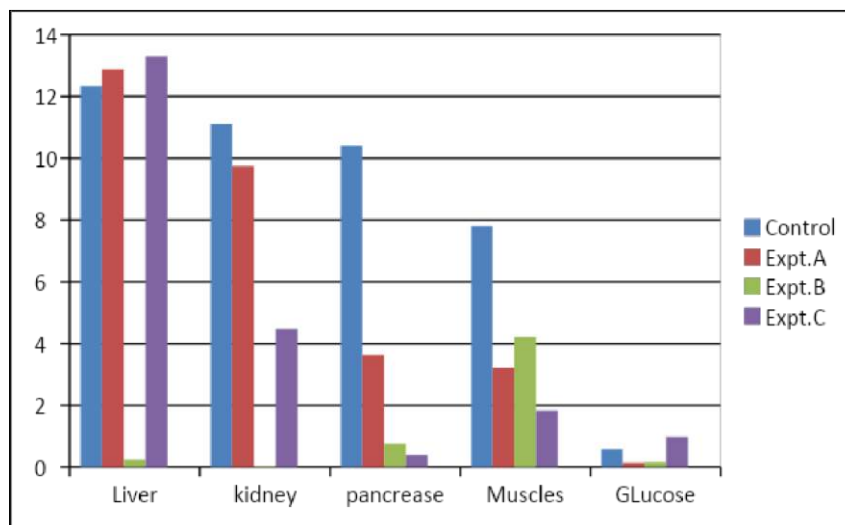


Table No 3. Effect of *T. arjuna* plant extract on protein concentration mg /g in some organs of the wild rat.

Tissues/ groups	Control	Expt. A	Expt. B	Expt. C
Liver	12.33 ± 1.0	* 12.87 ± 0.3	** 0.24 ± 1.97	NS 13.3 ± 0.26
Kidney	11.12 ± 0.39	NS 9.74 ± 2.8	** 0.032 ± 0.08	NS 4.47 ± 2.8
Pancreas	10.4 ± 0.13	** 13.62 ± 0.26	** 0.75 ± 2.28	** 0.4 ± 0.15
Muscles	7.80 ± 0.86	* 3.22 ± 0.72	* 4.21 ± 0.31	* 1.83 ± 0.27

Values are mean ± SD of five observations, *p<0.05, **p<0.01, and NS =Not significant.

Fig:- 3



In our study, administration of *T.arjuna* bark extract resulted in a significant reduction in blood glucose level, when compared with diabetic control animals. The extract containing 500 mg/kg body weight showed a better glucose level reduction than 250 mg/kg body weight. The mechanism may be through the stimulation of b cells for elevated secretion of insulin, thereby increasing the utilization of glucose in various tissues (Prakasam et al., 2002.).



In the present study, the STZ induced diabetic rats were chosen as an experimental model. An experimental rat that was administered with 2 ml Terminalia arjuna bark extract for 15 days showed a significant fall in tissue protein whereas control rats were shown there were small changes in tissue protein i.e the protein level in tissue is moderately decreasing. It might be due to the antihyperglycemic property of *T. arjuna*.

Experimental wild rats administered *T. arjuna* (2ml) showed that the liver protein gradually decreased as compared with the control rat as well as bark extract of 1 ml showed that the liver protein moderately increased as compared to control and again those diabetic induced rats showed liver protein significantly increase. The decrease in total protein content and other diabetic complications such as increased gluconeogenesis and ketogenesis in STZ induced mice may be due to elevated transaminase activities (Ghosh & Suryawanshi, 2001).

The glycogen of liver and glycogen of muscles is decreased given 2 ml dose experimental rat as compared to another rat. Diabetic induced rats showed that the glycogen of muscles moderately increases.

Mythili P. et. al., 2012 have found phytosterols, flavonoids (arjunone, arjunolone, and luteolin, arjunic acid, arjunolic acid, arjungenin, arjun glycosides) in *Terminalia arjuna* bark triterpenoids extract. Flavonoids (Arjunic acid, arjunolic acid, arjungenin acid, and arjunglycosides) being powerful antioxidants are reported to play a role in antihyperglycemic and analgesic activity by targeting pancreas and prostaglandins since the stem bark of *T. arjuna* plant contain flavonoids and glycosides.

CONCLUSION

It can be concluded that the medicinal plant *T. arjuna* has a promising antihyperglycemic effect in streptozotocin-induced model rats and it can be considered a potent source of the antidiabetic agent. The antihyperglycemic activity can be possible due to its antioxidant properties. It also shows that *T. arjuna* decreases the protein level in the tissue of rats.

REFERENCES

1. Bharani A, et al Salutory effect of T.Arjuna involuntary with severe refractor heart failure. Int J Cardio l. (1995).
2. Chatterjee A, Pakrashi SC. (1994). *The Treatise on Indian Medicinal Plants*. New Delhi: Publication and Information Directorate, Council of Scientific and Industrial Research.
3. Chehade JM, Mooradian AD. (2000). A rational approach to drugs 60:95–113. therapy of type 2 diabetes mellitus.
4. Moulis Biswas, Biswakanth Kar, Sanjib Bhattacharya, R.B. Suresh Kumar, Ashoke Kumar Ghosh & Pallab Kanti Haldar Antihyperglycemic activity and antioxidant role of Terminalia Arjuna leaf in streptozotocin-induced diabetic rats. (2011).
5. Mythili P, Parameswari CS and J Dayana. Phytochemical analysis of the bark extract of Terminalia arjuna and its cardioprotective effect. 2nd National level students conference nascent technologies in biomedical, Electrical engineering, and communications (NT-BEECOM 12). 6th August, 2012; 40. Vol. 1.
6. Parveen, A., Babbar, R., Agarwal, S., Kotwani, A., Fahim, M. 2012. Terminalia arjuna enhances baroreflex sensitivity and myocardial function in Isoproterenol-induced chronic heart failure rats. J. Cardiovasc. Pharmacol. Ther. 17: 199-207.
7. Piedrola G, Novo E, Escobar F, García-Robles R. (2001). White blood cell count and insulin resistance in patients with coronary artery disease. Ann Endocrinol (Paris) 62:7–1
8. Prakasam, S., Sethupathy and Pugalendi, K.V. (2002) Antihyperglycaemic effect of
9. Caseariaesculenta root extracts in streptozotocin-induced diabetic rats, Pharmazie. 57,11.
10. Ghosh S, Suryawanshi SA. (2001). Effect of *Vinca rosea* extracts in the treatment of alloxan diabetes in male albino rats. *Indian J Exp Biol* 39:748–759.
11. Trinder, P. (1969) Glucose oxidase method, Ann. Clin. Biochem. 6, 24.
12. Vats R K, Kumar V, Kothari A, Mital A, Ramachandran U. Emerging targets for diabetes. Curr Sci 88:241–247(2000).
13. Weidmann P, Boehlen LM, de Courten M. (1993). Pathogenesis and treatment of hypertension associated with diabetes mellitus. *Am Heart J* 125:1498–1513.

Cite this Article: Akhare Y. D., Shraddha Sharma (2022). Effect of Terminalia Arjuna Bark Extract on Streptozotocin-Induced Diabetic Rats. International Journal of Current Science Research and Review, 5(4), 1044-1048

Effect of Extract of *Emblica officinalis* on Developmental Stages of *Drosophila melanogaster*

Dr. Y. D. Akhare¹, Shrutika S. Kurhekar²

Department of Zoology, Vidya Bharati Mahavidhyalaya, C.K. Naidu Road, Camp, Amravati, Maharashtra, India, 444602

Abstract: *Drosophila melanogaster* is an excellent organism to test on ayurvedic medicines. It also remains one of the most studied organisms in biological research, particularly in genetic analysis. The objective of our study was to explore the potential of *Emblica officinalis* on the life span of *Drosophila melanogaster* using the adult feeding method. *Emblica officinalis* commonly known as the Indian gooseberry is an important medicinal plant in the traditional Indian system of Ayurvedic medicine. In the present study, we noted the phenotypic effect of *Emblica officinalis* extracts on the different developmental stages of *Drosophila melanogaster*. The fruit flies were grown on 10-gram culture media supplemented with a concentration of *E. officinalis* (0.1ml). Further, the size and growth of different life stages of *D. melanogaster* were observed and total protein estimated from it. The increase in the size and protein concentration in different life stages of controlled *D. melanogaster* were recorded.

Keywords: *Drosophila melanogaster*, *Emblica officinalis*, Protein, Lifespan

1. Introduction

India has been known to be a rich depot of medicinal plants. More than 35% of the entire plant species, at one time or other, were used for medicinal purposes. Plants have always been an important source of drugs (Yakob et al., 2012). A large number of the world's populations, especially in developing countries, depend upon medicinal plants as an alternative and complementary drugs therapy for various diseases.

Medicinal plants such as Aloe (*Aloe vera*), Tulsi (*Ocimum tenuiflorum*), Neem (*Azadirachta indica*), Amla (*Emblica officinalis*), Turmeric (*Cucurma longa*), and Ginger (*Zingiber officinale*) cure several common maladies. These are considered home remedies. The extractor plant is identified to be curative against diversified ailments and several pathophysiological or antioxidant property which prevents the damage of cellular compartments from oxidative stress. In Ayurveda "Swaasthasya Rakshanam" is one of the goals of treating disorders, maintaining or preserving good health (Charaka Samhita Sutrasthana, 2000a). However, serious efforts are required in systemic research to identify, isolate and evaluate the chemical constituents for nutritional and therapeutic potentials.

Emblica officinalis is commonly known as Indian gooseberry, Amla in Marathi, and Amalaki in Sanskrit is a deciduous tree of the family Phyllanthaceae. In Ayurveda, Amla is considered to be a potent rejuvenator and immune modulator effective in stalling degenerative processes and senescence and promoting longevity, enhancing digestion, treating constipation, reducing fever and cough, alleviating asthma, strengthening the heart, benefiting the eyes, stimulating hair growth and enliven the body. *E. officinalis* contain high amounts of ascorbic acid (vitamin C) and have a slightly bitter taste that may derive from a high density of ellagitannins, such as emblicanin A (37%), emblicanin B (33%), punigluconin (12%), and pedunculagin (14%). Amla also contains punicalofol and phyllanemblinin A, phyllanemblin other polyphenols, such as flavonoids,

kaempferol, ellagic acid, and gallic acid (Rawal S, et.al., 2014).

Drosophila melanogaster is a species of fly in the family Drosophilidae. The species is known generally as the common fruit fly or vinegar fly. *D. melanogaster* as the test system proved to be an excellent organism to test the effect of many drugs and other chemicals. This is cosmopolitan mostly found in temperate regions. In *Drosophila* and other insects body size is positively linked with mating success, longevity, fecundity, and other fitness characteristics also the best phenotypic heritable characteristics (Santos et al., 1992). All these conclusions demonstrated the advantage of size in mating success and fitness. Various researchers demonstrated the fitness studies in *Drosophila* such as fecundity, fertility (Long et al., 1980), and longevity (Cordts & Partridge, 1996, Partridge Tower, 2008). The species shares large numbers of homologous genes with mammals, 13,601 with humans. These have been studied to identify sequences related to those causing human diseases (Reiter et al. 2001). The present study was carried out with an objective to explore the potential of *Emblica officinalis* drug, on the fitness of *D. melanogaster* using the adult feeding method and found an increase in the life span of *D. melanogaster* flies.

2. Materials and Methods

Collection and culture of *Drosophila melanogaster* fly: *D. melanogaster* has a cosmopolitan distribution. These flies were collected on ripe banana fruit and fed on them then transferred to a conical flask by covering the open mouth of the flask with a muslin cloth so that flies were easily collected with a brush.

For studying *Drosophila* etherization process is used and for culturing those etherized and re-etherized bottles are used. Flies become inactive and we studied them under a microscope. The flies were cultured on potato, dextrose, and agar culture media.

Volume 11 Issue 4, April 2022

www.ijsr.net

[Licensed Under Creative Commons Attribution CC BY](https://creativecommons.org/licenses/by/4.0/)

Preparation of *Embllica officinalis* (Amla) Extract: For making ethanolic *E. officinalis* extract about 415.76gm of powdered material of *E. officinalis* was taken in a clean, flat-bottomed glass container and soaked in 1300ml of 80% ethanol. This container was sealed and kept for 2days accompanied by occasional shaking and stirring. Then the whole mixture is filtered with filter paper and evaporated the filtrate at room temperature. The gummy concentrate was designed as an ethanolic extract.

Experimental setup: Take five culture vials having 0.1 ml of *E. officinalis* extract and 10 gm of normal culture media. *Drosophila* flies are anesthetized with diethyl ether and transferred into culture vials and each vial contains 10 flies (95 females and 5 males). The vials were covered with a cotton plug and placed on the working table. The flies were laid eggs in cultures, hatched, and completed the life cycle within 10 to 15 days. Collecting the larval stages for measurement of total length and protein estimation was done by De Lowry's (1951) method.

The measurements of the size of the various stages of *Drosophila melanogaster* were carried out by the oculometer. The photographs were taken in Carl Zeiss binocular microscope and stereo zoom microscope. All results are presented a mean \pm SD. Test of significance (t-test) was used to analyze the data collected.

3. Observation and Result

It has been observed that *E. officinalis* plant has potent

antioxidant activity. Thus to determine whether supplementation of Amla powder with food, provides any nutritional support, survivorship assay of control (without *E. officinalis* and with *E. officinalis*) fed flies was performed. For this, an equal number of male and female flies were fed on control, and *E. officinalis* mixed food.

In the present study, the 1st, 2nd, 3rd, pupal stage and adult stage flies showed a normal size along with normal growth and development. The concentration of 0.1ml showed extremely significant change by making an increase in size at each stage of development. The percent change of 36.1%, 5%, 95.6%, 87.29%, and 88% was observed in 1st instar, 2nd instar, 3rd instar, pupa, and adult respectively over the control (Table).

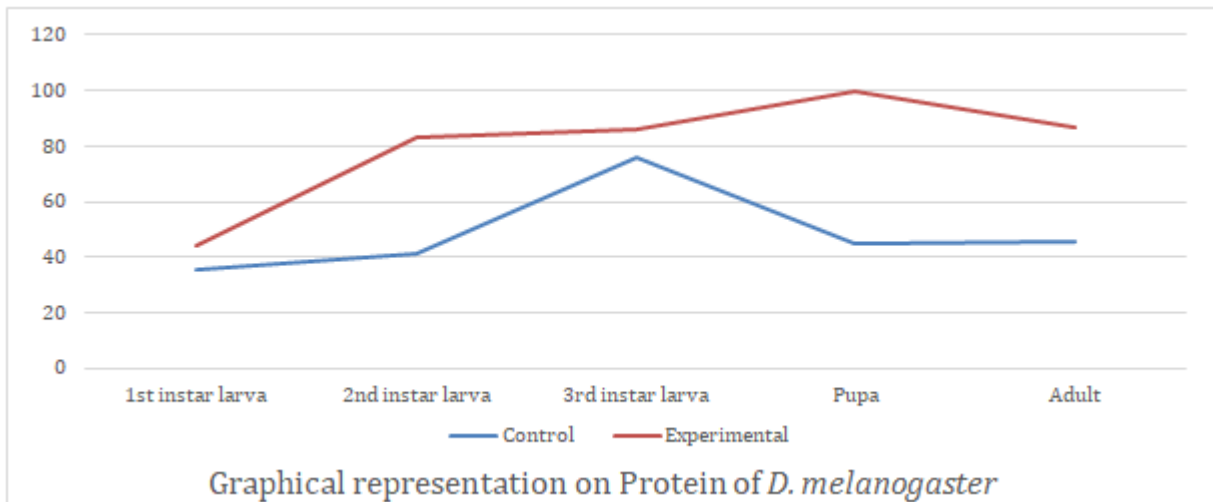
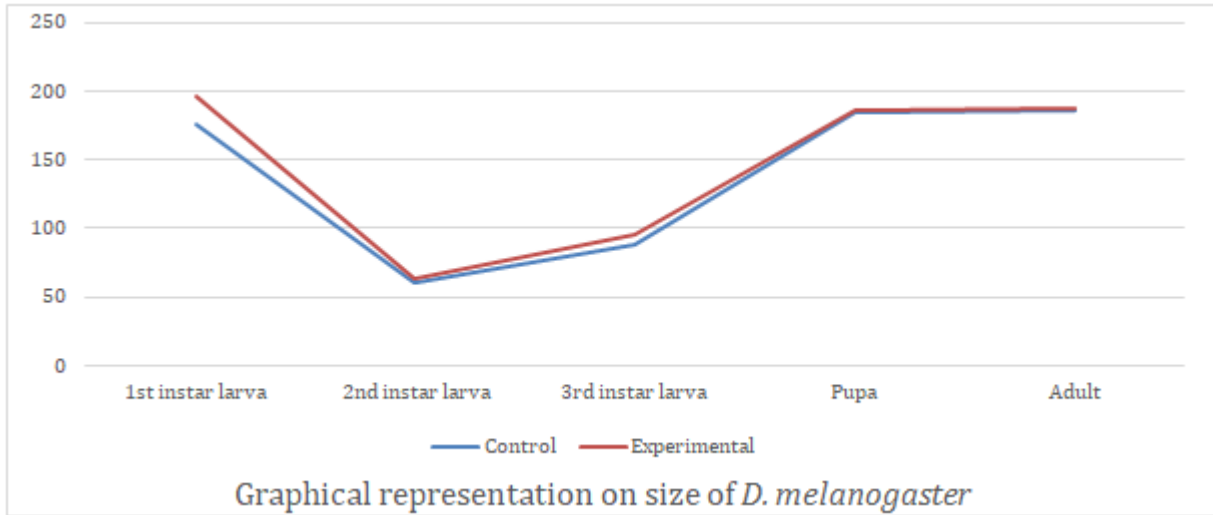
Whereas in protein concentration, significant increases were found in an experimental setup. In 1st instar larvae 55.63%, in 2nd instar larvae 17.21%, in 3rd instar larvae 14.07%, in pupal stage 0.374% and in adult stage 13.648% changes were noted over the control (Table).

The phenotypic characteristics were altered in the developmental stages of the experimental setup. In 1st, 2nd and 3rd larvae increased in size, faded transparency, yellowish to brownish color, and high segmentation was noticed in the experimental setup as compared to normal fruit flies. In pupal and adult stages, increased size, brownish color, no transparency, and high pigmentation were found over the control.

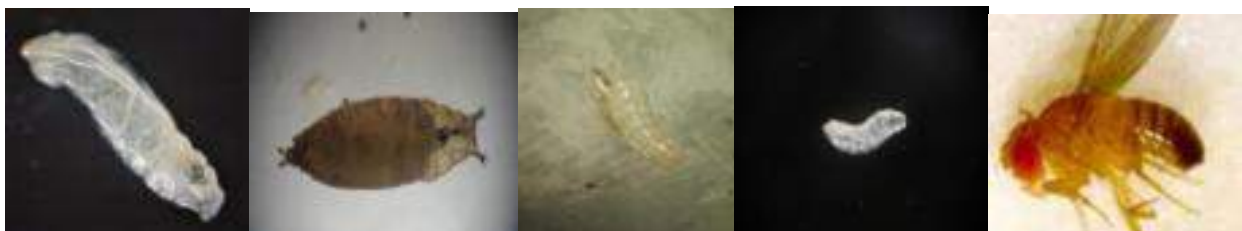
Table: Measurement of size (μm) & Protein (μg) from various larval stages of *Drosophila melanogaster*.

Developmental stages		Size	Protein
1 st instar larva	Control	175.6 \pm 10.65	35.256 \pm 1.57
	Experimental	NS 195.6 \pm 9.85 (36.1%)	NS 44.37 \pm 2.46 (55.63%)
2 nd instar larva	Control	60.45 \pm 8.20	41.488 \pm 0.98
	Experimental	NS 63.9 \pm 5.19 (5%)	NS 82.788 \pm 2.78 (17.21%)
3 rd instar larva	Control	87.7 \pm 13.49	75.6 \pm 10.65
	Experimental	NS 95 \pm 12.27 (95.6%)	NS 85.928 \pm 2.26 (14.072%)
Pupa	Control	184.5 \pm 15.34	45.054 \pm 0.26
	Experimental	NS 186 \pm 12.44 (87.29%)	NS 99.626 \pm 1.88 (0.374%)
Adult	Control	186.5 \pm 15.64	45.458 \pm 0.69
	Experimental	NS 188 \pm 11.15 (88%)	NS 86.352 \pm 2.62 (13.648%)

Mean \pm S.D. of five flies <0.05 (*), P<0.01(**), and P<0.001(***) and NS= not significant. The value mentioned in parenthesis is the % change over the control.



Normal development of *D. melanogaster*:



Experimental development of *D. melanogaster*:



4. Discussion

Amla (*Emblica officinalis*) has a hallowed position in Ayurveda- an Indian indigenous system of medicine. In the present study, we reveal that there was significant increase

in size of experimental set up as compared to the normal flies. It may be due to the beneficial effect of *Emblica officinalis* extract.

Shilpa Rawal (2014) reported the intake of food and

nutrition plays a major role in affecting aging process and longevity. However, the precise mechanisms underlying the aging process are still unclear. To this respect, diet has been considered to be a determinant of aging process. In order to better illustrate this, they used *Drosophila melanogaster* as a model and fed them orally with different concentrations of two commonly used Indian medicinal plant products, *Cucurma longa* (rhizome) and *Emblca officinalis* (fruit). Similar type of results is found in the present study.

The lifespan of *D. melanogaster* was increased in the study after the treatment of *Emblca officinalis*.

Lee K. S (2010) reported that curcumin extended the life span of two different strains of *Drosophila melanogaster*, an effect that was accompanied by protection against oxidative stress, improvement in locomotion, and chemopreventive effects. Life span extension was gender and genotype-specific. Curcumin also modulated the expression of several aging-related genes, including mth, thor, InR and JNK. The observed positive effects of curcumin on the life span of two different *D. melanogaster* strains demonstrate a potential applicability of curcumin treatment in mammals. Whereas in the present study the protein concentration increases were recorded in an experimental setup. (Akhare et al., 2021) reported the protein concentration in control *D. melanogaster* was normal but in experimental protein, the concentration was changed after the effect of neem extract.

In the present study the 1st, 2nd, 3rd, pupa, and adult stage flies showed extremely significant change by making an increase in size at each stage of development. Furthermore, the study was get concluded on the phenotypic character change in the development if *Drosophila melanogaster* in an experimental setup. Like as the phenotypic character involved the change in the transparency of the *Drosophila* body. Along with that, the segmentation on the body surface occurred in each of the stages. The segmentation showed the overall impact of body features on a fly. The effect of *Emblca officinalis* extract showed no pigmentation in an experimental setup but in the control setup, it does. The intake of 0.1ml concentration showed a color change too. All changes in the phenotypic character may be due to the activity of *Emblca officinalis* extract.

Many researchers have made studies on the effect of different chemicals on fertility in *D. melanogaster* (Vasudeva and Krishnamurthy1983). The rate of development is another parameter, which is used to examine some chemicals clinically. In that studies, the genetic constitution, amount of the food, temperature, and space were kept constant. The differences in the development must have been determined by the chemical used and not by the other factors. This type of result on the developmental time by different chemicals in *Drosophila melanogaster* has been shown by Luning (1966).

5. Conclusion

The conclusion of our experiment suggests that extract of *Emblca Officinalis* enhances the development of *D. melanogaster*. This noticeably increases longevity, fertility, fecundity, number along with developmental time. The

Influence of the above fruit extract is found in the parents and F1 generation and not in the F2 generations. The impact of the *E. officinalis* is not carried in further generations after the F1 generation. Lastly, here concludes there is a linear interrelationship between sexual activities and fitness parameters in experimental culture. So, along with life span, other reproductive fitness characteristics of flies like size, pigmentation, segmentation, and protein concentration were undertaken for this study so as to explore the unknown principles of *Emblca Officinalis* extract therapy which improves the quality of the life.

References

- [1] Akhare Y. D., H. A. Patharikar, P. R. Kumare, 2020. Effect of Neem Extract on the developmental Stages of *Drosophila melanogaster*, International Journal of All Research Education and Scientific Methods, 8, 12: 1523-1526.
- [2] Charaka Samhita Sutrasthana. 2000a. Bhagvan Dash vol 1, 6th ed. Chowkamba series office, Varanasi, India; Chowkamba Orientalia: 25: 17.
- [3] Cordts, R. & Partridge, L. 1996. Courtship reduces longevity in male *Drosophila melanogaster*. Animal Behaviour, 52: 269-278.
- [4] Lee, K.S., Lee, B. S., Semnani, S., Avanesian, A., Um, C. Y., Jeon, H.J., & Jafari, M. Curcumin extends life span, improves health span, and modulates the expression of age associated aging genes in *Drosophila melanogaster*. *Rejuvenation Research*, 13(5), 561- 570., 2010.
- [5] Long, C. E., Markow, T. A. & Yeager, P. 1980. Relative male age, fertility, and competitive mating success in *Drosophila melanogaster*. Behaviour genetics, 10: 163.
- [6] Lowry OH, Rosebrough NJ, Farr AL, Randall R. 1951. J.Biol.Chem.(193):265-275.
- [7] Luning, K. G. 1966. *Drosophila* tests in Pharmacology. Nature, 209: 84-86. Raven leaves (Onagraceae) in BALB/c mice. J Ethnopharmacol. 2012; vol.142 (3): 663-81.
- [8] Rawal, S., Singh, P., Gupta, A., & Mohanty, S. Dietary intake of *Cucurma longa* and *Emblca officinalis* increases life span in *Drosophila melanogaster*. *Biomed research international*, 2014.
- [9] Reiter, L. T., Potocki, L., Chein, S., Ghribkov, M. & Bier, E. 2001. A systematic analysis of human disease-associated gene sequences in *Drosophila melanogaster*. Genome Resonance, 11: 1114-25.
- [10] Santos, M. A., Ruiz, J. E., Quezada-Diaz, A., Barbadilla, A. & Fontdevila. 1992. The evolutionary history of *Drosophila buzzatii*. XX. Positive phenotypic covariance between field adult fitness components and body size. Journal Evolutionary biology, 5: 403-422.
- [11] Vasudeva, V. & Krishnamurthy, N. B. 1983. Effect of Dithane M-45 on the rate of development viability, morphology, and fecundity in *Drosophila melanogaster*. J.Mys. University, 29: 79-86.
- [12] Yakob K. H, Manaf Uyub A, Fariza Sulaiman S. 2012 Toxicological evaluation of 80% methanol extract of *Ludwigia octovalvis* (Jacq.) P.H.



Effect of Neem Extract on the Developmental Stages of *Drosophila Melanogaster*

Y. D. Akhare¹, H. A. Patharikar², P. R. Kumare³

^{1,2,3}Dept. of Zoology, VidyaBharati Mahavidyalaya, C. K. Naidu Road, Camp. Amravati, Maharashtra, India, 444602.

ABSTRACT

The common fruit fly, *Drosophila melanogaster* is used as a model organism. *Drosophila melanogaster* share several basic biological, biochemical, neurological and physiological similarities with mammals. In the present study, we noted the toxicological effect of extract of *Azadirachta indica* leaf on different developmental stages of *Drosophila melanogaster*. The fruit flies were grown on 10 gm culture media supplemented with leaf extract of *Azadirachta indica*. Neem is the most promising and medicinal plant, having widely value in biological activity and it is a well known for its insecticidal properties. Further, the size and growth of different life stages of *Drosophila melanogaster* were observed and total protein estimated from it. The reduction in the size of different stages and reduction in protein concentration in controlled *Drosophila melanogaster* were recorded.

Keywords : *Drosophilamelanogaster*, *Azadirachta indica*, aqueous extract, larval stages, protein.

INTRODUCTION

Herbal medicine has long been recognised as one of the oldest forms of remedies used by humans. Many people in developing countries still rely on traditional healing practice and medicinal plants for daily health care needs, in spite of the advancement in modern medicine. *Drosophila melanogaster*, also known as the fruit fly, is an excellent model organism widely used in biological research that has made significant contributions to the greater scientific community over the last century. It is because fruit flies are inexpensive to maintain in the laboratory, have simplified genetics, and short generation times allow for quick experiments with high sample numbers.

Neem is a key ingredient in non-pesticidal management, providing a natural alternative to synthetic pesticides. Because neem contains an antifungal and antibacterial property that effect on bacteria as well as fungal growth. Neem seed grind into a powder that is soaked overnight in water and sprayed onto the crop, to be effective, it must be applied repeatedly at least every ten days. Neem does not directly kill the insect on the crop. It acts as an antifeedant, repellent and egg-laying deterrent protecting the crop from damage. The insect starves and dies within a few days. Neem also suppresses the hatching of pest insect from their eggs. Neem based fertilizer has been effective against the pest Southern armyworm. Neem cake is often used as a fertilizer.

However, *Azadirachta indica* has been found to be toxic to non-target organisms where it induces marked alterations in experimental animals (Siriwattananurongsee, S., Sukontason, K. L., Kuntalue, B., & Sukontason, K. 2008). Biologically *Azadirachta indica* has various bioactive ingredients with diverse applications. These bioactive ingredients are known to have an anti-inflammatory, anti-allergic, insecticidal, nematicidal, spermicidal, antiviral, antifungal, anti-AIDS, anticancer, anti-antimalarial, antipyretic, antiarthritic, hypoglycemic, anti gastric ulcer, immunomodulatory, antihistamine, antitubercular, antiprotozoal, insect repellent, antifeedant, antihormonal and some other biological activities. Each and every part of *Azadirachta indica* tree known to possess biological activity. The biological activity depends upon the ingredients present in that particular part. More than 135 compounds have been isolated from different parts of *Azadirachta indica*. commonly-known as neem.

Neem leaves are dried and placed in the cupboard to prevent insect eating the clothes and in tins where rice in Neem leaves are dried and tropical region to keep away mosquitoes. Neem reduces fecundity and longevity as well as increased the development time of immature insect. On the insect has been shown to act as growth retardant and cause moulting disorder, change in the behaviour and morphogenetic defect. The present work to be undertaken for better understanding the effect of neem extract on development of *Drosophila melanogaster*

MATERIAL AND METHODS

Fruit flies were collected on ripe banana fruit. After that flies were cultured on potato, dextrose and agar culture media. Neem is well known for its marvelous medicinal properties. These plants are one of the richest sources of health for human beings coming from nature.

Preparation of aqueous neem leaf extract: *Azadirachta indica* leaves were obtained from the surrounding area of Vidya Bharati college laboratory, dried and finely chopped, then dissolved in acetone at a concentration of 200gm of dried leaves per 100ml of acetone for 24 hours at room temperature (as described by Khan and Ahmad et al. 1998). The material was filtered and the residue was shaken twice with 100ml acetone. All the extracts were combined and evaporated to almost dryness and redissolved in 1 Molar acetone. The prepared extract was used in the experiment in different dilutions.

Experimental set up: Take five culture vials having 0.1 ml of neem leaf extract and 10gm of normal culture media mixed. *Drosophila* flies anaesthetized with diethyl ether and transferred into culture vial and each vial contains 10 flies (5 female and 5 male). The vial covers with cotton plug and place on the working table. The flies were laid eggs in culture, hatch and completed the life cycle within 10 to 15 days. And collect the larval stages for measurement of total length and protein estimation by De Lowry's method. The measurement of size of the various stages of *Drosophila melanogaster* was carried out with help of coulometer. The total estimation of protein was done by the De Lowry method. Photographs were taken in the Carl Zeiss trinocular microscope and stereo microscope. All results are presented as Mean \pm SD. Test of significance (t-test) was used to analyse the data collected.

RESULT AND DISCUSSION

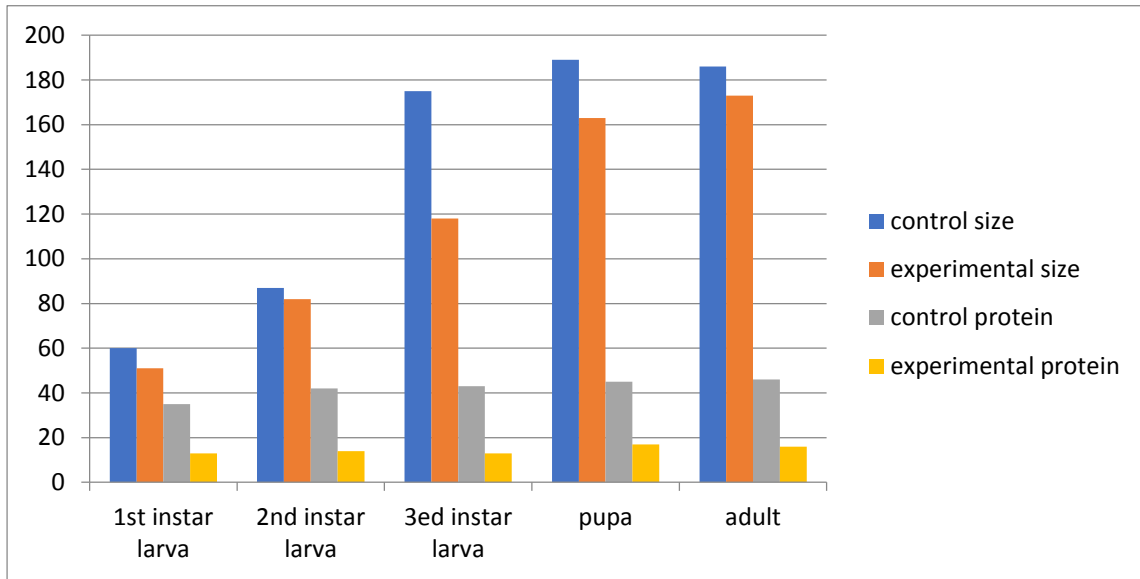
The size and growth in the control *D. melanogaster* is normal but in experimental *D. melanogaster* size and growth rate was changed. In the percentage, the change was 48.5%, 17.7%, 18.7%, 63.7%, 73.3% of 1st, 2nd, 3rd instar larvae, pupa and adult respectively as compared to controlled *D. melanogaster*.

The protein concentration in control *D. melanogaster* is normal but in experimental *D. melanogaster* level of protein, concentration was changed. In percentage, the protein concentration was 62.34%, 66.49%, 69.34%, 61.98%, 19.41% of 1st, 2nd, 3rd instar larvae, pupa and adult respectively as compared to control *D. melanogaster*. Recent studies have also demonstrated neem-induced effect on vitellogenesis and severe degeneration of follicle cell during oogenesis in mosquitoes (Lucantoni L. et al., 2006). When toxicological evaluation of neem extract against third instar larva of *Drosophila melanogaster* was applied the larvae were found to be melanised and failed to pupate (Anjum et al., 2010).

Table No. 1: Measurement of size and protein from various larval stages of *Drosophila melanogaster*

Developmental stages	Groups	Size(μ m)	Protein conc.(μ g)
1 st instar larva	Control	60.45 \pm 8.20	35.256 \pm 1.57
	Experimental	NS 51.5 \pm 6.25 (48.5)	NS 13.276 \pm 0.11 (62.34%)
2 nd instar larva	Control	87.7 \pm 13.49	41.488 \pm 0.98
	Experimental	NS 82.3 \pm 2.45 (17.7%)	NS 13.902 \pm 0.035 (66.49%)
3 rd instar larva	Control	175.6 \pm 10.65	43.216 \pm 0.50
	Experimental	NS 118.7 \pm 1.88 (18.7%)	NS 13.246 \pm 0.040 (69.34%)
Pupa	Control	189.5 \pm 31.13	45.054 \pm 0.26
	Experimental	NS 163.7 \pm 5.94 (63.7%)	NS 17.126 \pm 0.14 (61.98%)
Adult	Control	186.5 \pm 15.64	45.458 \pm 0.69
	Experimental	NS 173.3 \pm 2.98 (73.3%)	NS 16.582 \pm 0.24 (3.05%)

Mean \pm S.D. of five flies P<0.05(*), **P<0.01(***) and P<0.001 and NS = not significant. The value mentioned in parenthesis is the % change over the control.



Control Experimental Control Experimental



1st instar larvae 2nd instar larvae



3rd instar larvae

pupa



Adult female



Adult male

Graph No.1 : Size and protein estimation of control and experimental larvae, pupa adults of *Drosophila melanogaster*

We found a reduction in the size of each development stage as compare to control It may be the inhibitory effect of neem leaf extract which contains inhibitory growth compound so they were induced to reduce the normal growth of flies. Neem elaborates a vast array of biologically active compound that is chemically diverse and structurally complex. More than 140 compounds have been isolated from different part of neem. Neem leaf its constituent has been demonstrated to exhibit immunomodulatory, anti-inflammatory, antihyperglycemic, antibiotic properties (subapriya R,et.al., 2005).

We observed the decreased the size and also noticed that transparency colour, pigmentation and life span were showed variation. It may be due to the effect of the insecticidal activity of neem .R.Gnanamani et al. (2017) reported the efficacy of *Azadirachta indica* on the biochemical estimation of a Lepidopteran pest *pericalliaricini*. The investigation report on the impact of the leaf extract of *Aadirachta indica* on the total protein, carbohydrate and lipid level of fifth instar larvae of *Pericalliaricini*. After treatment with A. Indica leaf extract. Total protein level was reduced to 0.22mg from 0.72mg/g the carbohydrate level was decreased gradually from control. Lipid content in the p.ricini larvae also slowly decreased.

In our study, it is concluded that the impact of neem leaf extract on the stages of *Drosophila melanogaster* was found on the size, protein concentration and phenotypic character. They interrupt protein concentration with the decline of protein synthesis hormones. In contrast, the inhibitory effect of neem was showed by reducing the size it may be due to the azadirachtin. It shows an effect on normal development and growth. It was also found variation over the experimental flies Because neem exert effect on the pest, insect and the flies to control their population rate. So, They were show variation in the phenotypic character as compared to normal. It may be due to the insecticidal of neem.

REFERENCES

- [1]. Anjum, S. I., Yousf, M. J., Ayaz, S., & Siddiqui, B. S. (2010). Toxicological evaluation of chlorpyrifos and Neem extract (Biosal B) against 3RD instars larvae of *Drosophila melanogaster*. *Journal of Animal & Plant Sciences*, 20(1), 9-12.
- [2]. Lucantoni, L., Giusti, F., Cristofaro, M., Pasqualini, L., Esposito, F., Lupetti, P., & Habluetzel, A. (2006). Effects of a neem extract on blood feeding, oviposition and oocyte ultrastructure in *Anopheles stephensi* Liston (Diptera: Culicidae). *Tissue and Cell*, 38(6), 361-371.
- [3]. Mohammad f. khan, S.M. ahomad (1998) Toxicity of crude neem leaf extract against *Musca domestica* L. adult as compared with DDVP, dichlorvos.
- [4]. R. Gnanamani and S. Dhanasekaran (2017). Efficacy of *Azadirachta Indica* Leaf Extract on the Biochemical Estimation of a Lepidopteran Pest *Pericallia ricini* (Lepidoptera: Arctiidae). *World Journal of Agricultural Sciences* 13 (2): 63-67, 2017.
- [5]. Siritattananurongsee, S., Sukontason, K. L., Kuntalue, B., & Sukontason, K. (2008). Ultrastructural alteration of larvae and puparia of blow fly *Chrysomya megacephala* (F.)(Diptera: Calliphoridae) and house fly *Musca domestica* L.(Diptera: Muscidae) exposed to neem extract. *Trop Biomed*, 25(2), 107-116.
- [6]. Subapriya, R., & Nagini, S. (2005). Medicinal properties of neem leaves: a review. *Current Medicinal Chemistry- Anti-Cancer Agents*, 5(2), 149-156. nce, 7(4), 181-188.



Identification of concealed structural alerts using QSTR modeling for *Pseudokirchneriella subcapitata*

Vijay H. Masand^{a,*}, Magdi E.A. Zaki^b, Sami A. Al-Hussain^b, Anis Ben Ghorbal^c,
Siddhartha Akasapu^d, Israa Lewaa^e, Arabinda Ghosh^f, Rahul D. Jawarkar^g

^a Department of Chemistry, Vidya Bharati Mahavidyalaya, Amravati, Maharashtra, 444 602, India

^b Department of Chemistry, Faculty of Science, College of Sciences, Imam Mohammad Ibn Saud Islamic University (IMSIU), Riyadh 13318, Saudi Arabia

^c Department of Mathematics and Statistics, Faculty of Science, College of Sciences, Imam Mohammad Ibn Saud Islamic University (IMSIU), Riyadh 13318, Saudi Arabia

^d Corden Pharma, Colorado, USA

^e Assistant Lecturer of Statistics, Faculty of Business Administration, Department of Business Administration, Economics and Political Science, The British University in Egypt, Cairo, Egypt

^f Microbiology Division, Department of Botany, Gauhati University, Guwahati, Assam, 781014, India

^g Department of Medicinal Chemistry, Dr. Rajendra Gode Institute of Pharmacy, Amravati, Maharashtra, India

ARTICLE INFO

Keywords:

QSTR

Pseudokirchneriella subcapitata

Raphidocelis subcapitata

Toxicity

ABSTRACT

In the present work, QSTR modeling was conducted for microalga *Pseudokirchneriella subcapitata* using a data set of 271 molecules belonging to different types of chemical classes for the prediction of EC₅₀ for 72 hr based assays. The balanced QSTR model encompasses seven easily interpretable molecular descriptors and possesses statistical robustness with high predictive ability. This Genetic Algorithm Multi-linear regression (GA-MLR) model was subjected to internal validation, Y-randomization test, applicability domain analysis, and external validation as per the recommended OECD guidelines. The newly developed model fulfilled the threshold values for more than 20 recommended validation parameters including $R^2 = 0.72$, $Q^2_{LOO} = 0.70$, etc. The developed QSTR model was successful in identifying the type of hybridization or specific type of atoms of previously reported and newer structural alerts. Thus, the model could be useful for data gap filling and expanding mechanistic interpretation of toxicity for different chemicals.

1. Introduction

The 2006 Organization for Economic Cooperation and Development (OECD) Guidelines for the Testing of Chemicals (Alga, Growth Inhibition Test; OECD, 2006) demands reliable toxicity data for growth inhibition on aquatic plants, preferably on algae, because of often greater sensitivity to chemical exposures compared to fishes (Aruoja et al., 2011) and being first acceptors of several toxic chemicals such as herbicides, pharmaceuticals, etc. (Machado and Soares, 2021). Hence, the availability of algal toxicity data of high quality and subsequent analysis could help to minimize the requirements for the testing of large numbers of new chemicals under current regulatory programs. Therefore, several algal species including *Pseudokirchneriella subcapitata*, *Pediastrum duplex*,

Monoraphidium arcuatum, *Nannochloropsis-like sp.*, and *Chlorella sp. 12*, etc. have been used successfully for toxicity testing (Machado and Soares, 2021; Stone et al., 2019).

P. subcapitata, now known as *Raphidocelis subcapitata*, is a microalga with a high sensitivity to the exposure with toxic substances in fresh water (Machado and Soares, 2021). In addition, its ubiquitous distribution in aquatic ecosystems renders it a useful species for the assessment of nutrient levels (Marques et al., 2007). Therefore, it is commonly used for eco-toxicological studies as a bio-indicator species (Yamagishi et al., 2017). However, toxicological *in vivo* studies are expensive and time consuming (Yang et al., 2020b). Consequently, alternative approaches based on computational techniques like QSAR (Quantitative Structural-Activity Relationship) are methods of choice to minimize

Abbreviations: OECD, Organization for economic cooperation and development; FIFRA, U.S. federal insecticide, Fungicide and rodenticide Act; EPA, U.S. Environment protection agency; QSAR, Quantitative structural-activity relationship; QSTR, Quantitative structural-toxicity relationship; GA-MLR, Genetic algorithm multiple linear regression.

* Corresponding author.

E-mail addresses: Mezaki@imamu.edu.sa (M.E.A. Zaki), sahussain@imamu.edu.sa (S.A. Al-Hussain), assghorbal@imamu.edu.sa (A.B. Ghorbal), asidhu09@gmail.com (S. Akasapu), Israa.lewaa@bue.edu.eg (I. Lewaa).

<https://doi.org/10.1016/j.aquatox.2021.105962>

Received 12 March 2021; Received in revised form 10 August 2021; Accepted 1 September 2021

Available online 8 September 2021

0166-445X/© 2021 Elsevier B.V. All rights reserved.

these shortcomings (Yang et al., 2020b). Due to their ability to identify structural alerts for toxicity, deciphering plausible toxic mechanisms and predicting the toxicity of a compound, QSARs have received recognition and approval by different regulatory agencies such as OECD, the European REACH framework, Agency for Toxic Substances and Disease Registry (ATSDR), Atlanta (USA), etc. (Khan et al., 2019) (Fu et al., 2015; Khan and Roy, 2019; Toropov et al., 2017).

In fact, Biocidal Products Regulation (BPR) (BPR, Regulation (EU) 528/2012) promotes the use of QSAR for the adaptation of the data requirements, if the developed model satisfies all the requirements mentioned in annex IV on the general rules (Khan et al., 2019; Khan and Roy, 2019). These requirements include use of adequate and reliable documentation of the applied method for QSAR model development followed by its appropriate validation, and the chemicals must fall within the applicability domain of the model (Eu, 2012). Thus, for chemicals lacking experimental toxicity data or yet to be synthesized, the QSAR models are expected to reliably predict their toxicity.

This prompted a large number of researchers to develop QSTR (Quantitative Structural-Toxicity Relationship) models for *P. subcapitata*. However, most of these studies either used small data sets or data sets lacking diversity in chemical classes (see row 1-3 in Table 1). Though, in 2019, Khan and Roy reported a six-parametric QSTR model for pEC₅₀ (-log₁₀ EC₅₀) (24 hr) using a larger data set of 334 chemicals comprising a variety of compounds (Khan and Roy, 2019). The developed model had a coefficient of determination of R² = 0.72. Later, the same data set was used by Yu (Yu, 2020) for developing a Support Vector Machine (SVM) based eight parameter QSTR model, which has an acceptable R² = 0.75 and MAE = 0.60. However, their models used pEC₅₀ determined for 24 hr assays. Whereas, OECD guidelines (OECD, 2006) prefer 72 hr (or longer) assays for determination of toxicity, which is also recommended by algal toxicologists (Brill et al., 2021). Until to date, no QSTR model has been reported for a larger data set of diverse chemicals based on pEC₅₀s determined for 72 hr assays. Therefore, there is a need to develop a properly validated and balanced (Masand et al., 2021) QSTR model for toxicity against *P. subcapitata* using a larger data set comprising pEC₅₀ determined for 72 hr assays.

The present work is an attempt to develop a QSTR model comprising a balance of descriptive and predictive aspects for pEC₅₀s for an assay of 72 hr for *P. subcapitata* using a larger data set. The work further aimed to identify novel, important and concealed structural features, which are difficult to be captured by visual inspection.

2. Materials and methods

In the present work, the standard procedure recommended by OECD and other regulatory agencies for successful QSTR analysis has been followed (OECD, 2006). The procedure involved data set formation and curation, calculation of molecular descriptors and their pruning, model generation and its thorough validation, and mechanistic interpretation of the model (Cherkasov et al., 2014; Dearden et al., 2009; Fujita and Winkler, 2016; Gramatica, 2007, 2014, 2020; Johnson, 2008).

The present work comprises QSTR analysis using a data set for *P. subcapitata*, which involves organic chemicals belonging to diverse scaffolds with high variation in substitution patterns and toxicity profile. A few representative molecules have been depicted in Fig. 1 as examples to show diversity of molecules in the data set.

The data set was collected from the literature and publicly available databases (Connors et al., 2019). Then, it was subjected to curation before actual QSTR analysis (Fourches et al., 2010). The curation involves removal of duplicate entries, metal derivatives, complexes, salts, and molecules without EC₅₀ values. In case of duplicate entries with two different values of EC₅₀, the entry with a lower value was retained to consider a possible worst case situation (Yang et al., 2020a). Also, EC₅₀ values available as mg/L were converted to mol/L before initiating the analysis. Consequently, the data set consisted of 271 molecules tested for *P. subcapitata* for an assay of 72 hr. The five least and most toxic compounds from the selected data set were depicted in Table 2. The CAS number, SMILES notation and pEC₅₀ for all the molecules are presented in the supplementary materials.

In the next step, molecular descriptor calculation was accomplished using PaDEL (Yap, 2011) and PyDescriptor (Masand et al., 2017a; Masand and Rastija, 2017). This led to a pool of more than 29,000 molecular descriptors. Then, QSARINS 2.2.4 (Gramatica et al., 2013) was used to eliminate highly correlated (|R| > 0.95) and nearly constant (99%) molecular descriptors. This noticeably minimized the number of molecular descriptors for the data set to 2862. However, the reduced pool of molecular descriptors still comprised a variety of descriptors such as atom-pairs, presence/absence of certain functional groups/atoms, molecular properties, etc. Thus, it covers a broad molecular descriptor space to allow to identify prominent structural features.

Then, the data set was randomly split into a training set and an external (or prediction) set in an 80:20 (217:54 molecules) proportion. The training set was used for selection of appropriate number of molecular descriptors, i.e. for model building using Genetic Algorithm (GA) multi-linear regression (MLR) analysis. The molecules in the prediction set were not involved in the selection of molecular descriptors, that is,

Table 1
QSTR models reported for *P. subcapitata* and their comparison with model-A.

S. N.	Equation	Chemical Class	h (Test duration in hr)	Number of molecules in data set	R ²	Reference
1	Log(1/EC ₅₀) = 0.6 logKow - 0.9	propargylic alcohols (primary)	48	15	0.76	Chen et al., (2012)
2	Log(1/EC ₅₀) = 0.0014 logKow - 2.10 Elumo - 0.024 logKow × Elumo + 4.00	propargylic alcohols (secondary)	48	7	0.85	Chen et al., (2012)
3	Log(1/EC ₅₀) = -1.799 + 0.8763LogKow-mix	non-polar narcotic chemicals	72*	50	0.9469	Aruoja et al., (2014)
4	pEC ₁₀ = 1.93569 + 0.02315 CrippenMR + 0.00496 MW + 0.3396 MLOGP + 0.64439 LogKow - 1.71498 nSO2OH + 1.48957 B02[N-S]-2:24679 × 3A - 0.17172 H-051	organic compounds	48	251	0.70	Khan and Roy (2019)
5	pEC ₅₀ = 3.67917 + 0.02888 CrippenMR + 0.39725 MLOGP + 0.60323 B06 [C -N] + 0.78557 B05 [C- Cl] + 1.16506 B02[N-S]- 1.46683 nSO2OH	organic compounds	24	251	0.72	Khan and Roy (2019)
6	pEC ₁₀ = 1.579 + 0.027 CrippenMR - 0.370 MHYD + 0.927 F02[N-S] + 0.768 B05[C-Cl] + 0.497 CCCN + 1.017 R3m	organic compounds	24	334	0.75	Yu (2020)
7	pEC ₅₀ = 3.298 + 0.186 naroCpplus + 0.479 fnotringNsp3C1B + 1.565 fClaroC5B + 0.104 nTRing + 0.326 XLogP + 1.029 SubFP96 -2.439 fsp2Csp3C8B	organic compounds	72*	271	0.72	Present work

* Recommended exposure duration by OECD

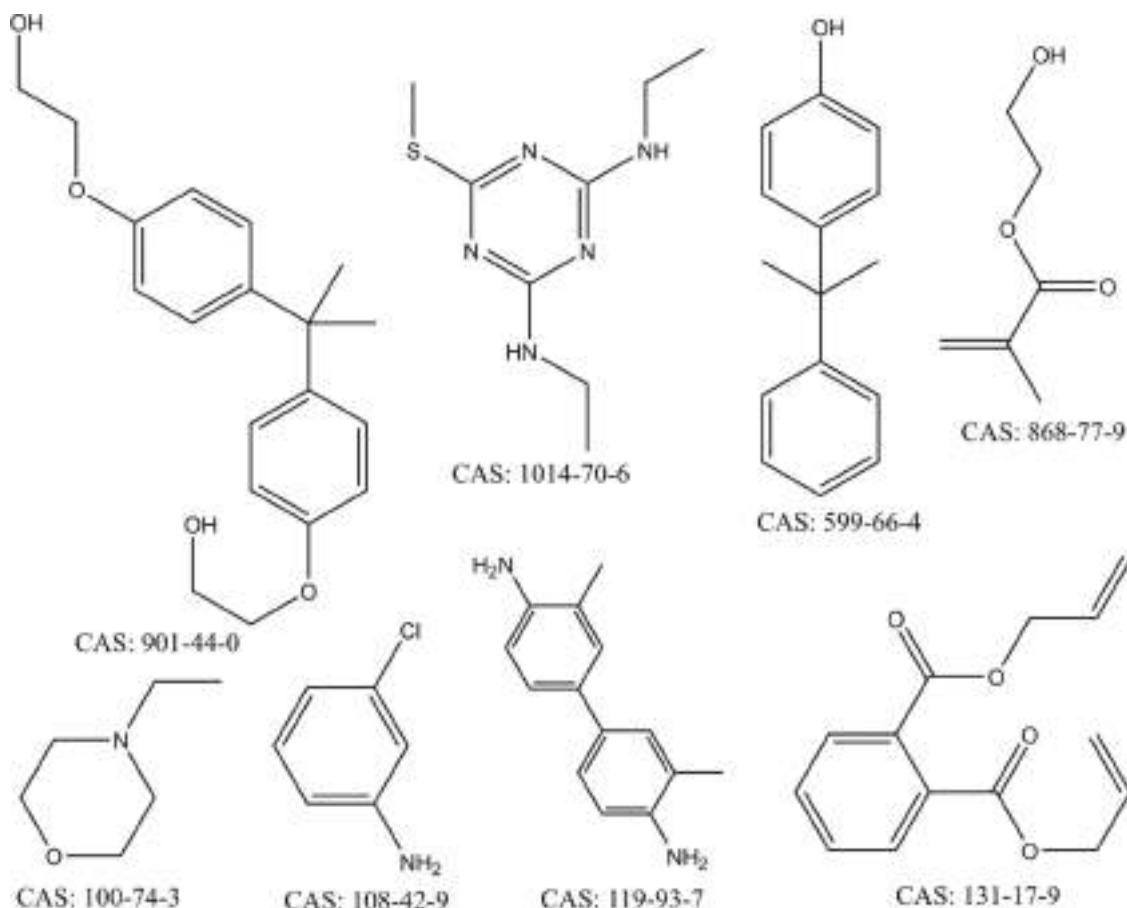


Fig. 1. Different types of scaffolds used for the present QSTR analysis (as representative examples).

Table 2

Five least and most toxic compounds of the present data set.

CAS number	SMILES	pEC ₅₀ (mol/L)
143226	CCCCOCCOCCO	2.422
868779	CC(=C)C(=O)OCCO	2.57
110190	CC(COC(=O)C)C	2.667
96231	ClCC(CCl)O	2.73
108952	Oc1cccc1	2.77
205823	c1ccc2c(c1)ccc1c2ccc3c2c1ccc3	8.022
124287	CCCCCCCCCCCCCN(C)C	8.218
23184669	CCCOCN(c1c(CC)ccc1CC)C(=O)CCl	8.264
51218496	CCCOCN(c1c(CC)ccc1CC)C(=O)CCl	8.318
193395	c1ccc2c(c1)c1cc3ccc4c3c1c2ccc3cc4	9.14

they were used for external validation of the developed model only. The Genetic Algorithm (GA) module available in QSARINS 2.2.4 (Gramatica, 2013; Gramatica, 2020; Gramatica et al., 2012; Gramatica et al., 2013) was used for searching the molecular descriptors using Q^2_{LOO} as a fitness function and number of generations was set to 10,000. The GA-MLR analysis resulted in the generation of a multitude of seven parametric models for the data set. Then, the reported model-A was selected on the basis of its superior statistical performance, which involved fulfilling of recommended threshold values for different internal and external validation parameters, thus ensuring adequate validation of the model-A. The suggested values for different internal and external validation parameters were as follows (Masand et al., 2019a; Masand et al., 2018; Masand et al., 2017b; Masand et al., 2019b):

R^2_{tr} (coefficient of determination) ≥ 0.6 , Q^2_{loo} (Leave-one-out coefficient of determination) ≥ 0.5 , Q^2_{LMO} (Leave-many-out coefficient of determination) ≥ 0.6 , $R^2 > Q^2$, R^2_{ex} (coefficient of determination for

external set) ≥ 0.6 , $RMSE_{tr}$ (Root mean squared error for training set) $< RMSE_{cv}$ (Root mean squared error for cross validation), CCC (Concordance Correlation Coefficient) ≥ 0.80 , Q^2-F^d (Leave-one-out coefficient of determination for test set) ≥ 0.60 with $RMSE$ and MAE (Mean absolute error) close to zero.

All the models not satisfying the threshold values were rejected. In addition, Y-scrambling was performed to eliminate the models developed by chance (Dearden et al., 2009; Gramatica, 2013; Gramatica, 2020; Masand et al., 2014). The applicability domain of the developed model was assessed using Williams plot. Further details of experimental methodology and description of various validation parameters are available in the supplementary materials.

The model-A is a seven parametric model. The description and other details are available in Table 3.

3. Results and discussion

The very first principle and expediency of a QSAR analysis is to gain a deep understanding and knowledge about structural features associated with a bio-activity (descriptive QSAR) and to predict the desired activity of an unknown molecule before its synthesis and bio-screening (predictive QSAR) (Fujita and Winkler, 2016). Therefore, in the present QSTR analysis, molecular descriptors with ease of interpretation in terms of structural features were used during model building. The developed QSTR model for the data set was as follows:

Model-A: Training Set: 80% (217 molecules) and Prediction set: 20% (54 molecules) pEC₅₀ = 3.298 (± 0.154) + 0.186 (± 0.049) * *nar-oCpcplus* + 0.479 (± 0.087) * *fnotringNsp3C1B* + 1.565 (± 0.389) * *fClaroC5B* + 0.104 (± 0.026) * *nTRing* + 0.326 (± 0.048) * *XLogP* + 1.029 (± 0.348) * *SubFP96* -2.439 (± 0.765) * *fsp2Csp3C8B*

Table 3
Molecular descriptors present in model-A.

S. N.	Molecular descriptor	Description	Software used for calculation
1	<i>naroCpplus</i>	Number of positively charged aromatic Carbon atoms	PyDescriptor
2	<i>fnotringNsp3C1B</i>	Frequency of occurrence of sp ³ -hybridized Carbon atom present exactly at one bond from non-ring Nitrogen atoms	PyDescriptor
3	<i>fClaroC5B</i>	Frequency of occurrence of aromatic Carbon atoms exactly at five bonds from Chlorine atom	PyDescriptor
4	<i>nTRing</i>	Number of rings (includes counts from fused rings)	PaDEL
5	<i>SubFP96</i>	Number of C-S, dithiophosphates or Carbodithioic ester moiety in a molecule	PaDEL
6	<i>XLogP</i>	Lipophilicity of a molecule	PaDEL
7	<i>fsp2Csp3C8B</i>	Frequency of occurrence of sp ³ -hybridized Carbon atoms exactly at eight bonds from sp ² -hybridized Carbon atom	PyDescriptor

3.1. Validation parameters for model-A

$R^2_{tr} = 0.718$, $R^2_{adj} = 0.708$, $LOF = 0.416$, $RMSE_{tr} = 0.604$, $MAE_{tr} = 0.506$, $CCC_{tr} = 0.836$, $s = 0.615$, $F = 75.98$, $R^2_{cv} (Q^2_{loo}) = 0.700$, $RMSE_{cv} = 0.623$, $MAE_{cv} = 0.524$, $CCC_{cv} = 0.827$, $Q^2_{LMO} = 0.68$, $R^2_{Yscr} = 0.033$, $RMSE_{ex} = 0.565$, $MAE_{ex} = 0.432$, $R^2_{ex} = 0.693$, $Q^2-F^1 = 0.691$, $Q^2-F^2 = 0.691$, $Q^2-F^3 = 0.752$, $CCC_{ex} = 0.826$

The close values of R^2_{tr} , R^2_{adj} (adjusted coefficient of determination) and $R^2_{cv} (Q^2_{loo})$ indicate that the model has been built using adequate number of variables (Kiralj and Ferreira, 2009). The high value of Q^2_{LMO} , which is essential to confirm internal validation of a model along with low value of Y-scrambling related parameters viz. R^2_{Yscr} (coefficient of determination for Y-scrambling) and Q^2_{Yscr} (Leave-one-out coefficient of determination for Y-scrambling) indicates that the model possesses

statistical robustness and free from chance correlations (Cherkasov et al., 2014; Fujita and Winkler, 2016; Gramatica, 2014, 2020; Masand et al., 2015). The model-A has high external predictive ability, which is reflected by the high value of R^2_{ex} , Q^2-F^1 and CCC_{ex} (Chirico and Gramatica, 2011, 2012; Consonni et al., 2009; Consonni et al., 2019; Gramatica, 2014). The Fischer value F indicates the high statistical significance of the model-A. Thus, a comparison of validation parameters for model-A with the recommended threshold values for the same parameter indicates that the model-A satisfies the suggested criteria. The predicted pEC₅₀, residuals, etc. are available in supplementary material.

The model-A is a seven parametric model. Out of seven parameters, six molecular descriptors *naroCpplus*, *fnotringNsp3C1B*, *fClaroC5B*, *nTRing*, *SubFP96* and *XLogP* (see Table 3 for their description) have positive coefficients in model-A. Therefore, any increase in their values will result in a higher toxicity. The reverse is true for the seventh molecular descriptor *fsp2Csp3C8B* (see Table 3 for its description), as it has a negative coefficient in the model-A.

The molecular descriptors present in the model-A have been depicted in Fig. 2 and their description is available in Table 3. The values of all molecular descriptors for all data set molecules are present in the supplementary materials.

The molecular descriptor *naroCpplus* stands for number of positively charged aromatic Carbon atoms. This descriptor further indicates the importance of aromatic rings and the bonding of a hetero atom with higher electronegativity than Carbon such as F, Cl, Br, O, N, etc. as a substituent on the aromatic ring in deciding the toxicity. From this it appears that positively charged aromatic carbon atoms signify the enhanced electrophilicity, which in turn, is responsible for augmented toxicity. A similar observation has been earlier reported by Rasulev et al. (2010) and He et al., (2012). Thus, the higher the number of aromatic rings along with more electronegative hetero atoms as a substituent on the aromatic ring atoms, the higher the toxicity.

Another similar molecular descriptor which highlights the importance of aromatic carbon atoms as well as chlorine atoms is *fClaroC5B*. The descriptor *fClaroC5B* indicates the frequency of occurrence of aromatic carbon atoms exactly at five bonds from chlorine atom. If the same

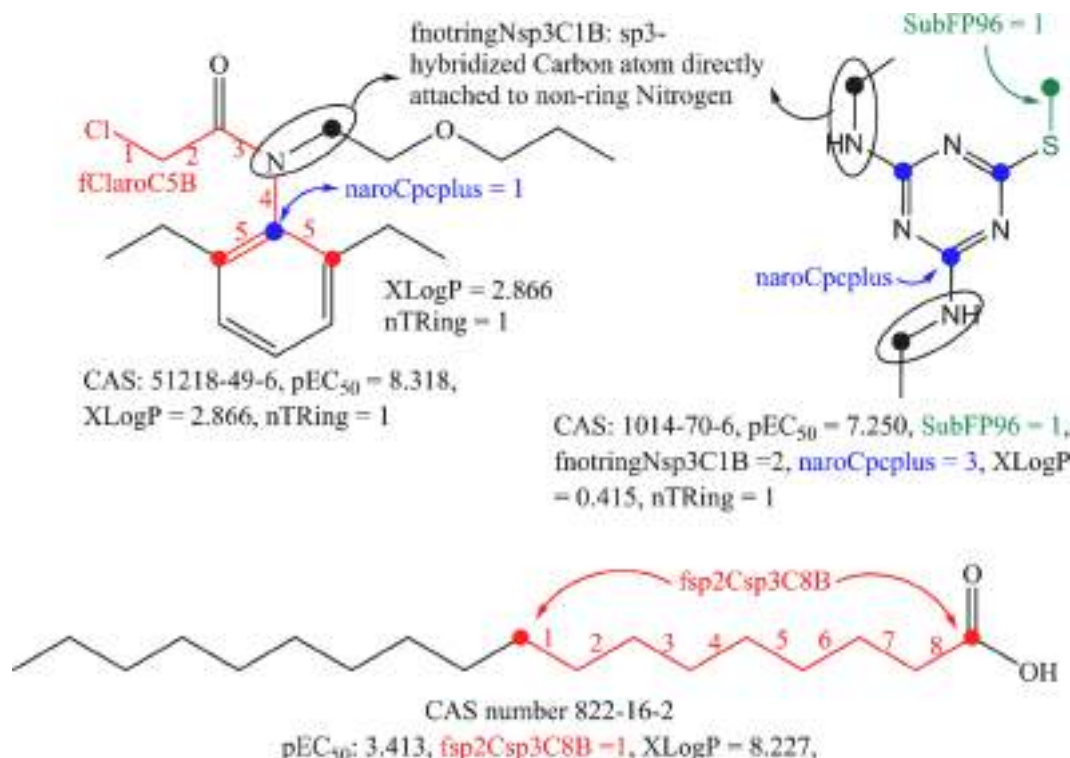


Fig. 2. Depiction of molecular descriptors present in the model-A.

aromatic carbon atom is also present at four or less bonds from any other Chlorine atom than it was rejected during the calculation of this descriptor. Khan and Roy (2019), and later Yu (2020), also identified B05[C-Cl], which stands for the presence/absence of C-Cl at topological distance 5, as an important feature. In the present work, we have further and successfully identified the specific type of carbon atom (i.e. aromatic carbon), which is responsible for augmentation of toxicity in combination with chlorine atom. The increased toxicity could be attributed to either nucleophilic substitution of Cl atom or lipophilic contribution of Cl atoms as in case of Pretilachlor (CAS: 51218-49-6) and Butachlor (CAS: 23184-66-9) (Artemenko et al., 2011; Han and Hatzios, 1991).

A molecular descriptor which is present in model-A and signifies the importance of rings in deciding the toxicity is $nTRing$, a 2D- descriptor which means number of rings (includes counts from fused rings). As the number of rings in the molecule increases, toxicity also increases (Ghavami and Sepehri, 2015). Likewise, another structural feature associated with augmentation of toxicity is $SubFP96$, which counts the number of C-S, dithiophosphates or Carbodithioic ester moiety in a molecule. Giuseppina et al., (2019) have also pointed out that dithiophosphates are fragments associated with aquatic toxicity.

$fnotringNsp3C1B$ is a molecular descriptor which represents the frequency of occurrence of sp^3 -hybridized carbon atom present exactly at one bond from non-ring nitrogen atoms. This molecular descriptor indicates that increasing the number of sp^3 -hybridized carbon atom directly attached to non-ring nitrogen atoms could lead to enhanced toxicity.

$XLogP$ represents the lipophilicity of a molecule. Due to its positive coefficient in model-A, it appears that increasing the lipophilicity, increases toxicity, too. This observation is supported by the fact that there is an overall increase in the toxicity of molecules with an increase in $XLogP$ values and vice-versa. However, several molecules like 822-16-2 (pEC₅₀: 3.413, $XLogP$: 8.227), 143-19-1 (pEC₅₀: 3.652, $XLogP$: 7.711), 111-03-5 (pEC₅₀: 4.172, $XLogP$: 7.132) are exceptions to this observation, which indicates that other molecular descriptors and confounding factors also play crucial role in deciding the overall toxicity of a molecule.

It is a common observation that increased lipophilicity could lead to higher toxicity. Lipophilicity is usually said to increase with number of carbon atoms. Interestingly, in the present work, we found a carbon-based molecular descriptor $fsp2Csp3C8B$ (frequency of occurrence of sp^3 -hybridized carbon atoms exactly at eight bonds from sp^2 -hybridized carbon atom) having a negative coefficient in the developed model. Therefore, increasing its value could lead to lower toxicity. It should be noted that if the same sp^3 -hybridized carbon atom is also present at seven or less bonds from any other sp^2 -hybridized carbon atom, then it was not considered during the calculation of $fsp2Csp3C8B$. There are six molecules which possess such a combination of carbon atoms have relatively lower toxicity (pEC₅₀ = 3.413 to 4.396 M). This is further indicated by molecules like CAS number 822-16-2 (pEC₅₀: 3.413), 143-19-1 (pEC₅₀: 3.652), 111-03-5 (pEC₅₀: 4.172) having high values of $XLogP$ but lower toxicity, which could be attributed to the fact that they have $fsp2Csp3C8B$ = 1, along with the absence of other toxicity enhancing features ($naroCpcplus$ = 0, $fnotringNsp3C1B$ = 0, $fClaroC5B$ = 0, $SubFP96$ = 0, and $nTRing$ = 0). The profound effect of $fsp2Csp3C8B$ is reflected from the fact that the molecule CAS number 683-10-3 had lower toxicity despite having $fnotringNsp3C1B$ = 4. Therefore, $fsp2Csp3C8B$ appears as a novel structural feature associated with the toxicity profile a molecule against *P. subcapitata*; however, careful consideration and further confirmation is required, as only six molecules from the present data set contain $fsp2Csp3C8B$ = 1. In addition, molecule numbers 30, 39, 41, 224, and 227, which possess $fsp2Csp3C8B$ = 1, have appeared as outliers for model-A (see Fig. 3b).

Based on the findings of this study, model-A has acceptable statistical performance and the majority of molecules are present within applicability domain (see Fig. 3 and supplementary materials).

It appears that compound number 206 (CAS 193-39-5) and 41 (CAS 334-48-5) were outliers (see Fig. 3). On excluding these two compounds, the statistical performance of model-A slightly diminished with respect to fitting and cross-validation (R^2_{tr} = 0.696, Q^2_{LOO} = 0.673) but improved significantly for external (prediction or test) set with R^2_{ex} = 0.760, Q^2-F^1 = 0.758, Q^2-F^2 = 0.758, Q^2-F^3 = 0.788, and CCC_{ex} = 0.857. This indicates that these two molecules are able to influence the model performance. Structural analysis of these compounds revealed that

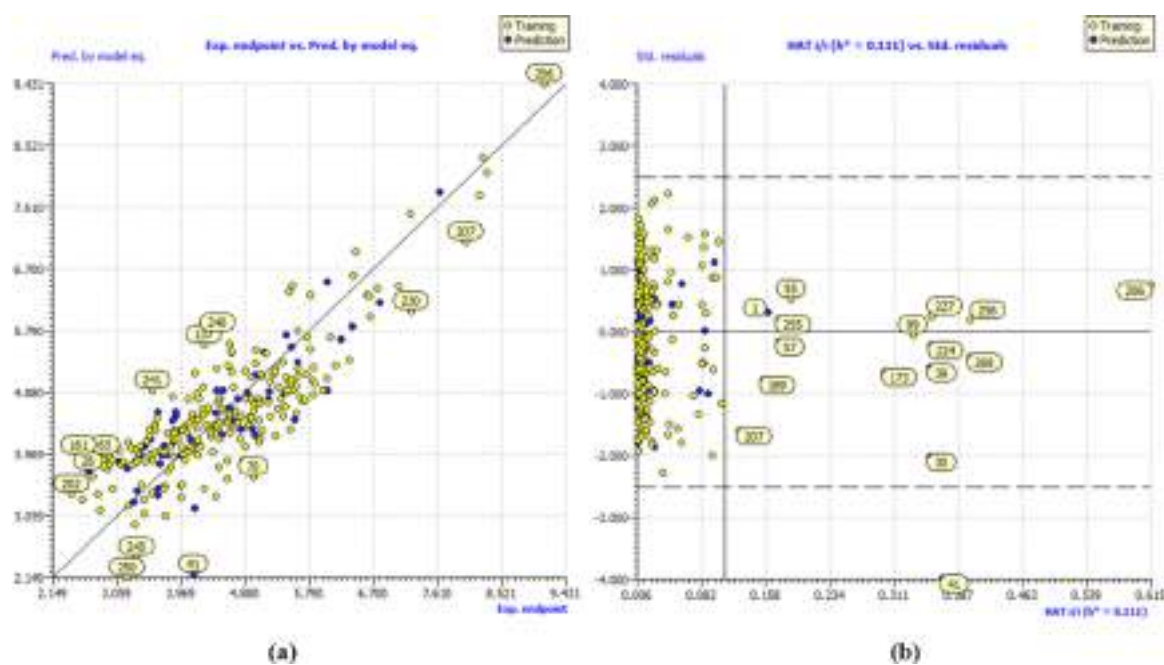


Fig. 3. Different graphs associated with model-A. (a) experimental (x-axis) vs. predicted (y-axis) pEC₅₀ (the solid line represents the regression line and molecules with high residuals have been labelled); (b) Williams plot for applicability domain HAT i/i (x-axis) vs. Std. residuals (y-axis) (the vertical solid line represents $h^* = 0.111$ and horizontal dashed lines represent the upper and lower boundaries for standard residuals and all outliers have been labelled).

molecule **206** (CAS 193-39-5) was Indeno[1,2,3-cd]pyrene, a polycyclic aromatic hydrocarbon (PAH), and **41** (CAS 334-48-5) was decanoic acid (also known as cupric acid), a free fatty acid (FFA). The compound **206** is highly toxic (pEC_{50} : 9.140), which could be attributed to its abnormally high lipophilicity ($XLogP = 6.404$) and $nTRing = 39$, thus due to the presence of six fused rings, it could be an outlier. Likewise, molecules **1**, **30**, **39**, **55**, **57**, etc. are outliers due to abnormally high or low lipophilicity. Molecules **30**, **39**, **41**, **224**, and **227** were structural outliers due to the presence of $fsp2Csp3C8B = 1$.

Conclusion: The identification of novel structural alerts for toxicity are important for policy makers and regulatory agencies. The QSTR model developed using diverse chemicals in the present work has not only successfully identified novel structural alerts but also extended and deepened the information and understanding about known structural alerts. It is statistically robust and validated using a good number of validation parameters showing excellent performance for prediction of external compounds within the applicability domain of the model. The model identified the importance of aromatic carbon atoms, chlorine atoms, lipophilicity, number of rings, etc. as prominent features that govern the toxicity of a molecule. The model should be useful for data gap filling, prediction of toxicity values of new or untested biocides, and minimization of cost involved in toxicity assessment.

Author statement

Vijay H. Masand: Conceptualization, Methodology, Software

Magdi E.A. Zaki: Writing- Original draft preparation, Revising Manuscript

Sami A. Al-Hussain: Funding acquisition, Resources

Anis Ben Ghorbal: Statistical analysis, Validation

Siddhartha Akasapu: Investigation, Writing - Review & Editing

Israa Lewaa: Statistical analysis, Validation

Arabinda Ghosh: Methodology, Data Curation

Rahul D. Jawarkar: Data Curation, Investigation

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors acknowledge the Deanship of Scientific Research at Imam Mohammad Ibn Saud Islamic University, Riyadh, KSA, for its support of this research through research group number RG-21-09-77. We are also thankful to Dr. Paola Gramatica and her team for providing QSARINS 2.2.4.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.aquatox.2021.105962.

References

- Artemenko, A.G., Muratov, E.N., Kuz'min, V.E., Muratov, N.N., Varlamova, E.V., Kuz'mina, A.V., Gorb, L.G., Golius, A., Hill, F.C., Leszczynski, J., Tropsha, A., 2011. QSAR analysis of the toxicity of nitroaromatics in Tetrahymena pyriformis: structural factors and possible modes of action. SAR QSAR Environ. Res. 22, 575–601. <https://doi.org/10.1080/1062936x.2011.569950>.
- Aruoja, V., Sihtmäe, M., Dubourguier, H.-C., Kahru, A., 2011. Toxicity of 58 substituted anilines and phenols to algae Pseudokirchneriella subcapitata and bacteria Vibrio fischeri: comparison with published data and QSARs. Chemosphere 84, 1310–1320. <https://doi.org/10.1016/j.chemosphere.2011.05.023>.
- Brill, J.L., Belanger, S.E., Barron, M.G., Beasley, A., Connors, K.A., Embry, M., Carr, G.J., 2021. Derivation of algal acute to chronic ratios for use in chemical toxicity extrapolations. Chemosphere 263, 127804. <https://doi.org/10.1016/j.chemosphere.2020.127804>.

- Cherkasov, A., Muratov, E.N., Fourches, D., Varnek, A., Baskin, I.I., Cronin, M., Dearden, J., Gramatica, P., Martin, Y.C., Todeschini, R., Consonni, V., Kuz'min, V.E., Cramer, R., Benigni, R., Yang, C., Rathman, J., Terfloth, L., Gasteiger, J., Richard, A., Tropsha, A., 2014. QSAR modeling: where have you been? Where are you going to? J. Med. Chem. 57, 4977–5010. <https://doi.org/10.1021/jm4004285>.
- Chirico, N., Gramatica, P., 2011. Real external predictivity of QSAR models: how to evaluate it? Comparison of different validation criteria and proposal of using the concordance correlation coefficient. J. Chem. Inf. Model. 51, 2320–2335. <https://doi.org/10.1021/ci200211n>.
- Chirico, N., Gramatica, P., 2012. Real external predictivity of QSAR models. Part 2. New intercomparable thresholds for different validation criteria and the need for scatter plot inspection. J. Chem. Inf. Model. 52, 2044–2058. <https://doi.org/10.1021/ci300084j>.
- Connors, K.A., Beasley, A., Barron, M.G., Belanger, S.E., Bonnell, M., Brill, J.L., de Zwart, D., Kienzler, A., Krailler, J., Otter, R., Phillips, J.L., Embry, M.R., 2019. Creation of a Curated Aquatic Toxicology Database: EnviroTox. Environ. Toxicol. Chem. /SETAC 38, 1062–1073. <https://doi.org/10.1002/etc.4382>.
- Consonni, V., Ballabio, D., Todeschini, R., 2009. Comments on the definition of the Q2 parameter for QSAR validation. J. Chem. Inf. Model. 49, 1669–1678. <https://doi.org/10.1021/ci900115y>.
- Consonni, V., Todeschini, R., Ballabio, D., Grisoni, F., 2019. On the Misleading Use of Q2F3 for QSAR Model Comparison. Mol. Inform. 38, e1800029. <https://doi.org/10.1002/minf.201800029>.
- Dearden, J.C., Cronin, M.T., Kaiser, K.L., 2009. How not to develop a quantitative structure-activity or structure-property relationship (QSAR/QSPR). SAR QSAR Environ. Res. 20, 241–266. <https://doi.org/10.1080/10629360902949567>.
- Eu, E.P.A., 2012. Regulation (EU) No 528/2012 of the European Parliament and of the Council of 22 May 2012 Concerning the Making Available on the Market and use of Biocidal Products, p. 2985.
- Fourches, D., Muratov, E., Tropsha, A., 2010. Trust, but verify: on the importance of chemical structure curation in cheminformatics and QSAR modeling research. J. Chem. Inf. Model. 50, 1189–1204. <https://doi.org/10.1021/ci100176x>.
- Fu, L., Li, J.J., Wang, Y., Wang, X.H., Wen, Y., Qun, W.C., Su, L.M., Zhao, Y.H., 2015. Evaluation of toxicity data to green algae and relationship with hydrophobicity. Chemosphere 120, 16–22. <https://doi.org/10.1016/j.chemosphere.2014.05.040>.
- Fujita, T., Winkler, D.A., 2016. Understanding the Roles of the “Two QSARs”. J. Chem. Inf. Model. 56, 269–274. <https://doi.org/10.1021/acs.jcim.5b00229>.
- Ghavami, R., Sepehri, B., 2015. QSPR/QSAR solely based on molecular surface electrostatic potentials for benzenoid hydrocarbons. J. Iran. Chem. Soc. 13, 519–529. <https://doi.org/10.1007/s13738-015-0761-2>.
- Giuseppina, G., Thomas, F., Anna, L., Antonio, C., Emilio, B., 2019. A new QSAR model for acute fish toxicity based on mined structural alerts. J. Toxicol. Risk Assess. 5, 1–8. <https://doi.org/10.23937/2572-4061.1510016>.
- Gramatica, P., 2007. Principles of QSAR models validation internal and external. QSAR Comb. Sci. 26 (5). QSAR & Combinatorial Science 26, 694–701.
- Gramatica, P., 2013. On the development and validation of QSAR models. Methods Mol. Biol. 930, 499–526. https://doi.org/10.1007/978-1-62703-059-5_21.
- Gramatica, P., 2014. External Evaluation of QSAR Models, in Addition to Cross-Validation Verification of Predictive Capability on Totally New Chemicals. Mol. Inf. 33, 311–314.
- Gramatica, P., 2020. Principles of QSAR Modeling. Int. J. Quant. Struct.-Property Relationships 5, 61–97. <https://doi.org/10.4018/IJQSPR.20200701.0a1>.
- Gramatica, P., Cassani, S., Roy, P.P., Kovarich, S., Yap, C.W., Papa, E., 2012. QSAR Modeling is not Push a Button and Find a Correlation: a Case Study of Toxicity of (Benzo-)triazoles on Algae. Mol. Inf. 817–835.
- Gramatica, P., Chirico, N., Papa, E., Cassani, S., Kovarich, S., 2013. QSARINS: a new software for the development, analysis, and validation of QSAR MLR models. J. Comput. Chem. 34, 2121–2132. <https://doi.org/10.1002/jcc.23361>.
- Han, S., Hatzios, K.K., 1991. Uptake, translocation, and metabolism of [¹⁴C]pretilachlor in fenclorim-safened and unsafened rice seedlings. Pestic. Biochem. Physiol. 39, 281–290. [https://doi.org/10.1016/0048-3575\(91\)90123-4](https://doi.org/10.1016/0048-3575(91)90123-4).
- He, G., Feng, L., Chen, H., 2012. A QSAR study of the acute toxicity of halogenated phenols. Procedia Eng. 43, 204–209. <https://doi.org/10.1016/j.proeng.2012.08.035>.
- Johnson, S.R., 2008. The Trouble with QSAR (or How I Learned To Stop Worrying and Embrace Fallacy). J. Chem. Inf. Model. 48, 25–26. <https://doi.org/10.1021/ci700332k>.
- Khan, K., Khan, P.M., Lavado, G., Valsecchi, C., Pasqualini, J., Baderna, D., Marzo, M., Lombardo, A., Roy, K., Benfenati, E., 2019. QSAR modeling of Daphnia magna and fish toxicities of biocides using 2D descriptors. Chemosphere 229, 8–17. <https://doi.org/10.1016/j.chemosphere.2019.04.204>.
- Khan, K., Roy, K., 2019. Ecotoxicological QSAR modelling of organic chemicals against Pseudokirchneriella subcapitata using consensus predictions approach. SAR QSAR Environ. Res. 30, 665–681. <https://doi.org/10.1080/1062936x.2019.1648315>.
- Kiralj, R., Ferreira, M.M.C., 2009. Basic validation procedures for regression models in QSAR and QSPR studies: theory and application. J. Braz. Chem. Soc. 20, 770–787. <https://doi.org/10.1590/s0135-50532009000400021>.
- Machado, M.D., Soares, E.V., 2021. Exposure of the algae Pseudokirchneriella subcapitata to environmentally relevant concentrations of the herbicide metolachlor: Impact on the redox homeostasis. Ecotoxicol. Environ. Saf. 207, 111264. <https://doi.org/10.1016/j.ecoenv.2020.111264>.
- Marques, C.R., Abrantes, N., de Figueiredo, D.R., Pereira, M.J., Gonçalves, F., 2007. Are Pseudokirchneriella subcapitata and Chlorella vulgaris Affected by Environmental Samples from a Rice Field? Water Air Soil Pollut. 189, 49–59. <https://doi.org/10.1007/s11270-007-9554-2>.

- Masand, V.H., El-Sayed, N.N., Mahajan, D.T., Rastija, V., 2017a. QSAR analysis for 6-arylpyrazine-2-carboxamides as Trypanosoma brucei inhibitors. SAR QSAR Environ. Res. 28, 165–177. <https://doi.org/10.1080/1062936X.2017.1292407>.
- Masand, V.H., El-Sayed, N.N.E., Bambole, M.U., Patil, V.R., Thakur, S.D., 2019a. Multiple quantitative structure-activity relationships (QSARs) analysis for orally active trypanocidal N-myristoyltransferase inhibitors. J. Mol. Struct. 1175, 481–487. <https://doi.org/10.1016/j.molstruc.2018.07.080>.
- Masand, V.H., El-Sayed, N.N.E., Bambole, M.U., Quazi, S.A., 2018. Multiple QSAR models, pharmacophore pattern and molecular docking analysis for anticancer activity of α , β -unsaturated carbonyl-based compounds, oxime and oxime ether analogues. J. Mol. Struct. 1157, 89–96. <https://doi.org/10.1016/j.molstruc.2017.12.045>.
- Masand, V.H., El-Sayed, N.N.E., Mahajan, D.T., Mercader, A.G., Alafeefy, A.M., Shibi, I. G., 2017b. QSAR modeling for anti-human African trypanosomiasis activity of substituted 2-Phenylimidazopyridines. J. Mol. Struct. 1130, 711–718. <https://doi.org/10.1016/j.molstruc.2016.11.012>.
- Masand, V.H., Elsayed, N.N., Thakur, S.D., Gawhale, N., Rathore, M.M., 2019b. Quinoxalinones Based Aldose Reductase Inhibitors: 2D and 3D-QSAR Analysis. Mol. Inform. 38, e1800149 <https://doi.org/10.1002/minf.201800149>.
- Masand, V.H., Mahajan, D.T., Nazeruddin, G.M., Hadda, T.B., Rastija, V., Alfeefy, A.M., 2014. Effect of information leakage and method of splitting (rational and random) on external predictive ability and behavior of different statistical parameters of QSAR model. Med. Chem. Res. 24, 1241–1264. <https://doi.org/10.1007/s00044-014-1193-8>.
- Masand, V.H., Mahajan, D.T., Nazeruddin, G.M., Hadda, T.B., Rastija, V., Alfeefy, A.M., 2015. Effect of information leakage and method of splitting (rational and random) on external predictive ability and behavior of different statistical parameters of QSAR model. Med. Chem. Res. 24, 1241–1264.
- Masand, V.H., Patil, M.K., El-Sayed, N.N.E., Zaki, M.E.A., Almarhoon, Z., Al-Hussain, S. A., 2021. Balanced QSAR analysis to identify the structural requirements of ABBV-075 (Mivebresib) analogues as bromodomain and extraterminal domain (BET) family bromodomain inhibitor. J. Mol. Struct. 1229. <https://doi.org/10.1016/j.molstruc.2020.129597>.
- Masand, V.H., Rastija, V., 2017. PyDescriptor: A new PyMOL plugin for calculating thousands of easily understandable molecular descriptors. Chemometrics and Intelligent Laboratory Systems 169, 12–18. doi: 10.1016/j.chemolab.2017.08.003.
- OECD, 2006. OECD Guideline 201 for the Testing of Chemicals, Alga, Growth Inhibition Test. Organisation for Economic Cooperation and Development, Paris, France.
- Rasulev, B., Kušić, H., Leszczynska, D., Leszczynski, J., Koprivnac, N., 2010. QSAR modeling of acute toxicity on mammals caused by aromatic compounds: the case study using oral LD50 for rats. J. Environ. Monit. 12 <https://doi.org/10.1039/b919489d>.
- Stone, S., Adams, M.S., Stauber, J.L., Jolley, D.F., Warne, M.S.J., 2019. Development and application of a multispecies toxicity test with tropical freshwater microalgae. Environ. Pollut. 250, 97–106. <https://doi.org/10.1016/j.envpol.2019.03.058>.
- Toropov, A.A., Toropova, A.P., Marzo, M., Dorne, J.L., Georgiadis, N., Benfenati, E., 2017. QSAR models for predicting acute toxicity of pesticides in rainbow trout using the CORAL software and EFSA's OpenFoodTox database. Environ. Toxicol. Pharmacol. 53, 158–163. <https://doi.org/10.1016/j.etap.2017.05.011>.
- Yamagishi, T., Yamaguchi, H., Suzuki, S., Horie, Y., Tatarazako, N., 2017. Cell reproductive patterns in the green alga Pseudokirchneriella subcapitata (=Selenastrum capricornutum) and their variations under exposure to the typical toxicants potassium dichromate and 3,5-DCP. PLoS One 12, e0171259. <https://doi.org/10.1371/journal.pone.0171259>.
- Yang, L., Wang, Y., Chang, J., Pan, Y., Wei, R., Li, J., Wang, H., 2020a. QSAR modeling the toxicity of pesticides against Americamysis bahia. Chemosphere 258, 127217. <https://doi.org/10.1016/j.chemosphere.2020.127217>.
- Yang, L., Wang, Y., Hao, W., Chang, J., Pan, Y., Li, J., Wang, H., 2020b. Modeling pesticides toxicity to Sheepshead minnow using QSAR. Ecotoxicol. Environ. Saf. 193, 110352 <https://doi.org/10.1016/j.ecoenv.2020.110352>.
- Yap, C.W., 2011. PaDEL-descriptor: an open source software to calculate molecular descriptors and fingerprints. J. Comput. Chem. 32, 1466–1474. <https://doi.org/10.1002/jcc.21707>.
- Yu, X., 2020. Quantitative structure-toxicity relationships of organic chemicals against Pseudokirchneriella subcapitata. Aquatic. Toxicol. 224, 105496 <https://doi.org/10.1016/j.aquatox.2020.105496>.



ORIGINAL ARTICLE

QSAR based virtual screening derived identification of a novel hit as a SARS CoV-229E 3CL^{pro} Inhibitor: GA-MLR QSAR modeling supported by molecular Docking, molecular dynamics simulation and MMGBSA calculation approaches



R.D. Jawarkar ^{a,*}, Ravindrakumar L. Bakal ^a, Magdi E.A. Zaki ^b,
Sami Al-Hussain ^{b,*}, Arabinda Ghosh ^c, Ajaykumar Gandhi ^d, Nobendu Mukerjee ^e,
Abdul Samad ^f, Vijay H. Masand ^g, Israa Lewaa ^h

^a Department of Medicinal Chemistry, Dr. Rajendra Gode Institute of Pharmacy, University-Mardi Road, Amravati, Maharashtra, 444603, India

^b Department of Chemistry, Faculty of Science, Al-Imam Mohammad Ibn Saud Islamic university, Riyadh 13318, Saudi Arabia

^c Microbiology Division, Department of Botany, Gauhati University, Guwahati, Assam 781014, India

^d Department of Chemistry, Government College of Arts and Science, Aurangabad, Maharashtra 431 004, India

^e Department of Microbiology; Ramakrishna Mission Vivekananda Centenary College, Akhil Mukherjee Rd, Chowdhary Para, Rahara, Khardaha, Kolkata, West Bengal 700118, India

^f Department of Pharmaceutical Chemistry, Faculty of Pharmacy, Tishk International University, Erbil, Kurdistan Region, Iraq

^g Department of Chemistry, Vidyabharti Mahavidyalaya, Camp Road, Amravati Maharashtra, India

^h Department of Business Administration, Faculty of Business Administration, Economics & Political Science, The British University in Egypt (BUE), Cairo, Egypt

Received 31 July ; accepted 10 October 2021

Available online 19 October 2021

* Corresponding authors at: Department of Medicinal Chemistry, Dr. Rajendra Gode College of Pharmacy, Mardi Road, Amravati, Maharashtra, India and (Rahul D. Jawarkar). Department of Chemistry, Faculty of Science, Al-Imam Mohammad Ibn Saud Islamic university, Riyadh 13318, Saudi Arabia (Magdi E.A. Zaki).

E-mail addresses: rahuljawarkar@gmail.com (R.D. Jawarkar), rlbakal@gmail.com (R.L. Bakal), Mezaki@imamu.edu.sa (M.E.A. Zaki), sahussain@imamu.edu.sa (Sami Al-Hussain), dra.ghosh@gauhati.ac.in (A. Ghosh), gascjay18@gmail.com (A. Gandhi), nobendumukerjee00@gmail.com (N. Mukerjee), abdul.samad@tiu.edu.iq (A. Samad), vijaymasand@gmail.com (V.H. Masand), Israa.lewaa@bue.edu.eg (I. Lewaa).

Peer review under responsibility of King Saud University.



KEYWORDS

HCoV SARS 3CL^{pro};
GA-MLR;
QSAR based virtual
screening;
Molecular docking and MD
simulation;
Lead;
MMGBSA calculations

Abstract Congruous coronavirus drug targets and analogous lead molecules must be identified as quickly as possible to produce antiviral therapeutics against human coronavirus (HCoV SARS 3CL^{pro}) infections. In the present communication, we bear recognized a HIT candidate for HCoV SARS 3CL^{pro} inhibition. Four Parametric GA-MLR primarily based QSAR model (R²:0.84, R²_{adj}:0.82, Q²_{loo}: 0.78) was once promoted using a dataset over 37 structurally diverse molecules along QSAR based virtual screening (QSAR-VS), molecular docking (MD) then molecular dynamic simulation (MDS) analysis and MMGBSA calculations. The QSAR-based virtual screening was utilized to find novel lead molecules from an in-house database of 100 molecules. The QSAR-vS successfully offered a hit molecule with an improved pEC₅₀ value from 5.88 to 6.08. The benzene ring, phenyl ring, amide oxygen and nitrogen, and other important pharmacophoric sites are revealed via MD and MDS studies. Ile164, Pro188, Leu190, Thr25, His41, Asn46, Thr47, Ser49, Asn189, Gln191, Thr47, and Asn141 are among the key amino acid residues in the S1 and S2 pocket. A stable complex of a lead molecule with the HCoV SARS 3CL^{pro} was discovered using MDS. MM-GBSA calculations resulted from MD simulation results well supported with the binding energies calculated from the docking results. The results of this study can be exploited to develop a novel antiviral target, such as an HCoV SARS 3CL^{pro} Inhibitor.

© 2021 The Author(s). Published by Elsevier B.V. on behalf of King Saud University. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Introduction

Coronaviruses are classified as RNA viruses. To date, seven human coronaviruses (HCoVs) viz. SARS-CoV, Middle East Respiratory Syndrome (MERS) -CoV, and SARS-CoV-2, 229E, Human coronavirus OC43, Human coronavirus NL63 (HCoV-NL63), and Human coronavirus HKU1 (HCoV-HKU1) have been discovered. First, three of the seven coronaviruses, particularly SARS-CoV, MERS-CoV, and SARS-CoV-2 are pathogenic species. Whereas the ultimate four, namely 229E, OC43, NL63, and HKU1 cause mild diseases. Coronaviruses belongs to the order Nidovirales, household Coronaviridae, and subfamily Orthocoronavirinae. Amongst the four coronavirus genera (Alphacoronavirus, Betacoronavirus, Gammacoronavirus, Deltacoronavirus), HCoVs are categorized below Alphacoronaviruses; HCoV-229E and NL63 and Betacoronaviruses; MERS-CoV, SARS-CoV, HCoV-OC43 and HCoV-HKU1. SARS-CoV-2 from the Betacoronavirus genus has fairly close relatedness with two bat-derived CoV-like coronaviruses, viz. bat-SL-CoVZC45 and bat-SL-CoVZXC21 (Malik, 2020). Coronaviruses are of sphere-shape with a diameter of a hundred twenty-five nm with the club-shaped projections on the floor that resemble a photo voltaic corona. Coronavirus has fairly the greatest genome amongst each and every positive-strand RNA viruses. (Lai et al., 2006)

An incredibly transmissible coronavirus to that amount causes lethal respiratory harm was once in the beginning determined of China. The severity of the symptoms is characterized by the increased nasal mucosal plasma exudation and interferon γ (IFN γ) levels in nasal lavage specimens (Linden et al., 1995). The advance peak of respiratory tract viral loads seems within the preceding three days then infection or drops off dramatically within a week, correlating including development and raise within signs and symptoms stability.

Dramatically, there is a consequential considerable vibrancy into the quantity of corona cases. To date, corona infection has reached more than 29 lot humans international

with a mortality dimensions as high as 3.15 % (according in accordance with World Health Organization's (WHO's) report, September 2020). Despite the fact, potent hit in opposition to SARS-CoV-2 is still a dream durability (Konwar and Sarma, 2020).

A 3C-like protease (3CL^{pro}) additionally appear among CoV-229E, the causative agent for the severe acute respiratory sign (SARS) into human. CoV-229E and SARS-CoV exist in a complex with the inhibitors were investigated in the several crystal structures of 3C^{pro} from CVB3 and 3CL^{pro} protein. (Lee et al., 2009) In it concern, numerous investigators hold utilized MD, MDS, quantitative structure-activity relationship (QSAR) studies for virtual screening to to identify a new hit for HCoV SARS 3CL^{pro} inhibition.

QSAR techniques have been effectively implemented not only In the development of a reliable statistics-based mathematical correlation between physicochemical properties of chemical substances and their desired biological activities but also to forecast the biological activity of *de novo* molecules.

In the last couple of decades together with the advances in the computational field, wet-lab chemical experimentation has been substituted by molecular modeling and virtual experimentation that deploy fundamentals of basic sciences such as, mathematics, chemistry, physics, and algorithms (Gini, 2016).

Enriching the utility about QSAR methodologies among the drug search yet development endeavor, especially into the improvement on the doubtlessly potent fresh chemical entities then hit/lead together with diverse bioactivity is a captivating scientific research community. (Garro Martinez et al., 2015) With the advances in computational sciences, QSAR technologies are evolving rapidly and gaining potential makes use of in regulatory science. Food and Drug Administration (FDA) had invested a lot of efforts to facilitate the development of reliable QSAR models in setting up chemical databases using superb and protected experimental statistics accompanied with the aid of the development of computational algorithms (Hong et al., 2016).

Successful application of high throughput screening (HTS) to molecules' libraries to find out the new lead for a particular biological property is one of the core traits in drug discovery. To set up the correlation of the undertaking of a molecule with molecular descriptors, QSAR analysis is frequently used which includes digital molecular filtering and screening based on a mathematical model. This strategy reduces the cost in the failure of a drug candidate in superior (clinical) degrees by filtering combinatorial libraries, rejecting these molecules with an expected toxic effect, and disadvantaged pharmacokinetic profiles, thereby decreasing the number of experiments.

Molecular docking (MD) is one of the widely used, well-established in-silico structure-based drug discovery methods. Docking describes and/or predicts ligand-target interactions at the molecular level, set up structure-activity relationships (SAR) and enable the identification of new lead candidate of therapeutic interest a priori information on the chemical structure of other goal modulators. MD techniques are largely used to discover conformation adopted by means of ligands inside the binding pocket(s) of the macromolecular targets. MD additionally evaluates the ligand-receptor binding free power by way of assessing critical phenomena via a complicated intermolecular recognition system (Ferreira et al., 2015).

Hit identification and lead optimization are abundantly tangled with computational modeling. In drug discovery, structure-based virtual screening (VS) has been indispensable for more than a decade with its drastically studied, underlying computational technique, docking. The parameters for vS may range with the objective, however the usual protocol is very straightforward. Principally in vS a library of small molecules are docked into the binding pocket of a macromolecule (target receptor, protein, etc.). The system ends up by way of returning various solutions per molecule, ranked in the order of acceptance for similarly screening and the identification of the fine possible hit(s) (Kontoyianni, 2017). vS is a time, cost, resources, and labor saving approach and this has marked vS as one of the effective computational techniques to display libraries of small molecules for new hits to be experimentally examined for desired property/activity. Among the vS approaches, QSAR analysis is the most powerful method due to its excessive and speedy throughput and desirable hit rate. A QSAR model once developed and fully validated for robustness and productiveness, can be utilized to the reliable prediction of the biological property of novel compounds. Although the experimental trying out of computational hits is now not an inherent section of QSAR methodology, it is exceptionally preferred and need to be carried out as an ultimate validation of developed models, advisably.

In the present scientific contribution, QSAR primarily based virtual screening strategy is expected for the rapid and less expensive development of medicines to deal with SARS-CoV-2. This tactic is primarily based on discovering the anti-HCoV SARS 3CL^{pro} attainable of leverage molecules beforehand testified to have powerful inhibitory recreation for the same. Critical evaluations of present information on HCoV SARS 3CL^{pro} inhibitors the use of QSAR based vS supported and enriched by way of MD and MDS procedures have been carried out to perceive novel HCoV-229E inhibitors with preferred properties.

2. Experimental

2.1. Preparation of data sets

To begin, we classified the complete ChEMBL information set into two classes, with assays of each the wild kind and mutant form of the target. As documented with the aid of ChEMBL, no in-vitro assay was developed to evaluate Human Coronavirus 229E inhibitory activity towards mutant targets.

After removing structural duplicates, we used the median EC₅₀ value to create the QSAR models. The log-transformed EC₅₀ values were used for the QSAR models. All the compounds test in vitro against Human coronavirus 229E were used in this collection of inhibitor structures from the ChEMBL database. In the end, 39 compounds representing 37 unique compounds were identified as having been tested against Human Coronavirus 229E (see Table 1).

2.1.1. Modeling set preparation from ChEMBL data

ChEMBL's trustworthily determining criteria were used to prefilter the compounds and data in the database: (1) the confidence score (a quantitative indicator of data quality in ChEMBL) is greater than 8; (2) expert-based curation; (3) data source (PubMed); is indicated; (4) EC₅₀ is a parameter of activity measurement; (5) EC₅₀ is precisely define (there is no ">" or "*" signal before EC₅₀); (6) We have not included the structure because it is not a multi-component complex or salt. Therefore, only the compounds tested against the Human Coronavirus 229E inhibition assay were extracted from ChEMBL.

2.2. Structure optimization and molecular descriptor calculation

To create the structures, ACD Labs' chemical sketch program (www.acdlabs.com) was used. The structures were converted into 3D structures using Open-Babel 2.4 and then optimized with MMFF94 force fields. The 3D constructions were optimized using TINKER default settings, and then they were aligned using Open3DAlign.

2.3. Molecular descriptor pruning

Over 30,000 molecular descriptors have been created by way of PyDescriptor and PaDEL for each molecule in all sets. This led to molecular descriptor thinning as it eliminated vain molecular descriptors. As a way to keep away from multi-collinear and counterfeit variables in the GAMLR (Genetic Algorithm-Multi-linear Regression) model, molecular descriptors with excessive co-linearity ($|R|$ greater than 0.90) and tightly constant (greater than 95 %) had been excluded the usage of objective feature selection in QSARINS ver. 2.2.4. Despite the limited variety of molecular descriptors, the condensed pool was once massive adequate to embody the area of 1D to 3D descriptors (Davies et al., 2015; Masand et al., 2018; Masand et al., 2017).

Table 1 showing Experimental end point (experimental pEC₅₀ value in nm), Predicted fitting (Predicted pEC₅₀ value) and Predicted fit residual value (residual).

sn	Status	Exp. endpoint	Pred. fitting	Pred.Fit.Res.	Pred. LOO	Pred. LOO Res.
1	Training	6.699	6.4234	-0.2756	6.3412	-0.3578
2	Training	6.301	6.4887	0.1877	6.5641	0.2631
3	Training	6.222	5.8991	-0.3229	5.8629	-0.3591
4	Training	5.886	5.6832	-0.2028	5.6284	-0.2576
5	Training	5.824	5.5829	-0.2411	5.5531	-0.2709
6	Prediction	5.745	5.475	-0.27	PRED	-0.27
7	Training	5.745	5.8673	0.1223	5.9219	0.1769
8	Prediction	5.745	5.8991	0.1541	PRED	0.1541
9	Training	5.602	5.8991	0.2971	5.9324	0.3304
10	Training	5.284	5.0159	-0.2681	4.9946	-0.2894
11	Training	5.268	5.1013	-0.1667	5.0908	-0.1772
12	Training	5.268	5.1013	-0.1667	5.0908	-0.1772
13	Training	5.268	5.2093	-0.0587	5.2047	-0.0633
14	Training	5.268	5.1013	-0.1667	5.0908	-0.1772
15	Training	5.268	5.6909	0.4229	5.746	0.478
16	Training	5.26	5.1239	-0.1361	5.1067	-0.1533
17	Training	5.102	5.1293	0.0273	5.1346	0.0326
18	Training	5.102	5.1293	0.0273	5.1346	0.0326
19	Training	5.102	5.0213	-0.0807	5.0119	-0.0901
20	Prediction	5.102	5.0866	-0.0154	PRED	-0.0154
21	Training	5.102	5.3523	0.2503	5.3728	0.2708
22	Training	5.073	4.9786	-0.0944	4.9689	-0.1041
23	Training	5.051	4.9934	-0.0576	4.985	-0.066
24	Training	5.051	4.765	-0.286	4.7031	-0.3479
25	Prediction	5.051	4.9079	-0.1431	PRED	-0.1431
26	Training	5.051	5.2093	0.1583	5.2217	0.1707
27	Training	5.051	4.9786	-0.0724	4.9712	-0.0798
28	Prediction	5.048	4.7402	-0.3078	PRED	-0.3078
29	Training	5.025	4.7627	-0.2623	4.6986	-0.3264
30	Prediction	5.025	4.8481	-0.1769	PRED	-0.1769
31	Training	4.928	5.3321	0.4041	5.3673	0.4393
32	Training	4.923	5.0415	0.1185	5.0826	0.1596
33	Training	4.609	4.5289	-0.0801	4.5015	-0.1075
34	Training	4.403	4.4822	0.0792	4.5994	0.1964
35	Training	4.357	4.7999	0.4429	4.9401	0.5831
36	Training	4.347	4.5491	0.2021	4.5866	0.2396
37	Training	4.222	4.4209	0.1989	4.4861	0.2641

2.4. Splitting the data set into training and external sets and subjective feature selection (SFS):

A precise way to avoid statistics leakage is to split the data set into training, prediction, and external/test sets with the appropriate composition and parts prior to exhaustive subjective feature selection (Masand et al., 2017). For bias-free analysis, the dataset was randomly split up into training (80 % = 30 molecules) and prediction (20 % = 7 molecules) sets. To choose a set of molecular descriptors, a training set was used alone, and a prediction/external set was used solely to perform external validation of the model (Predictive QSAR) (See Fig. 1).

We employed QSARINS-2.2.4's GA-MLR method to pick out relevant descriptors for subjective feature decision the usage of Q_{LOO}^2 as a fitness parameter. The variety of molecular descriptors in the model is an essential factor in growing a profitable QSAR model without excessive over-fitting. Using breaking point values drawn from R_{tr}^2 and Q_{LOO}^2 values, a design (see Fig. 2) was plotted between the wide variety of molecular descriptors involved in the QSAR model and the

number of molecular descriptors involved in the model. The breaking point used to be consequently viewed to be the optimal number of the molecular descriptors. According to Fig. 2, there are four variables that determine the breaking point. As a result, we excluded QSAR models with more than 4 descriptors.

2.5. QSAR model building and their validation:

The set of information used to be arbitrarily split using random splitting in QSARINS into a training set and a prediction set (80 % training and 20 % prediction, respectively). After creating the model, the training set was used for external validation, that is, to reveal the model's potential to predict fresh chemical entities (Masand et al., 2016; Masand et al., 2015; Masand et al., 2015; Gramatica, 2020).

With default settings, QSARINS used to be used to create GA-MLR primarily based QSAR models. In GA, the selected fitness purpose to maximize used to be Q^2 , which also covered the double cross-validation. During the improvement of the

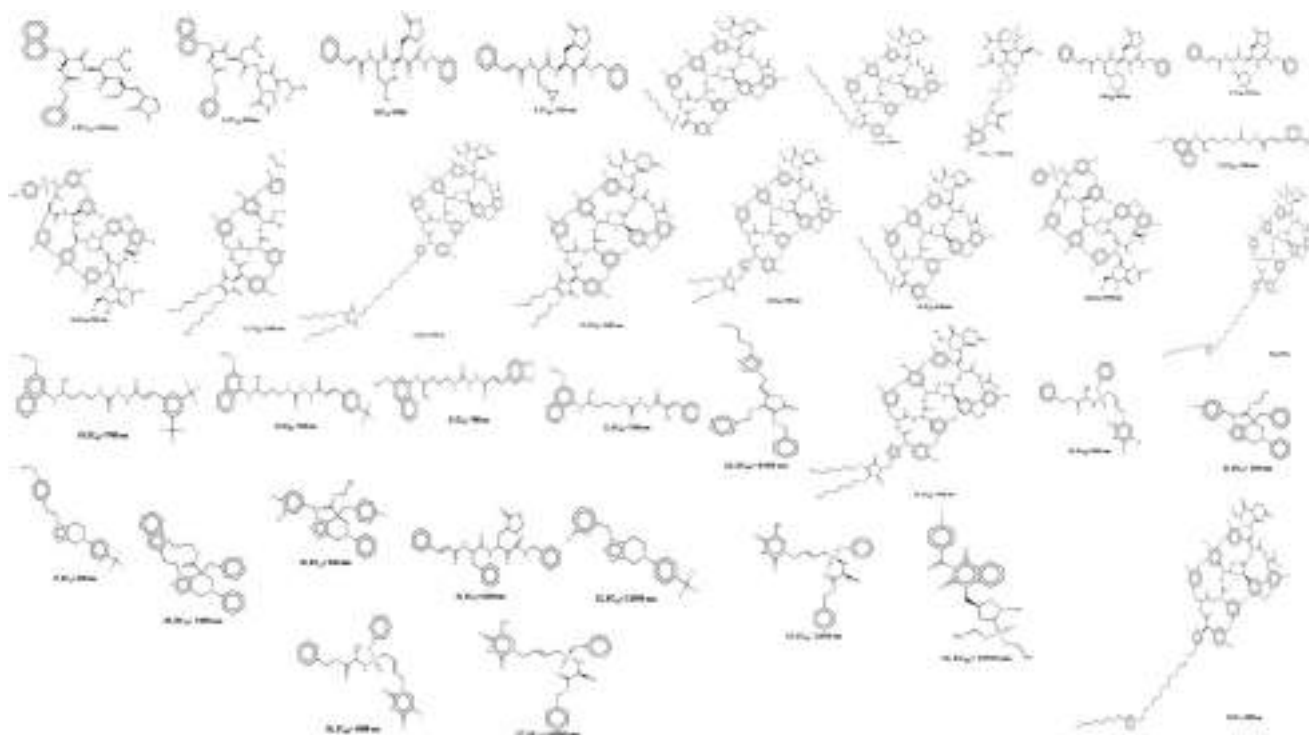


Fig. 1 Depiction of 37 dataset molecules used in QSAR study.

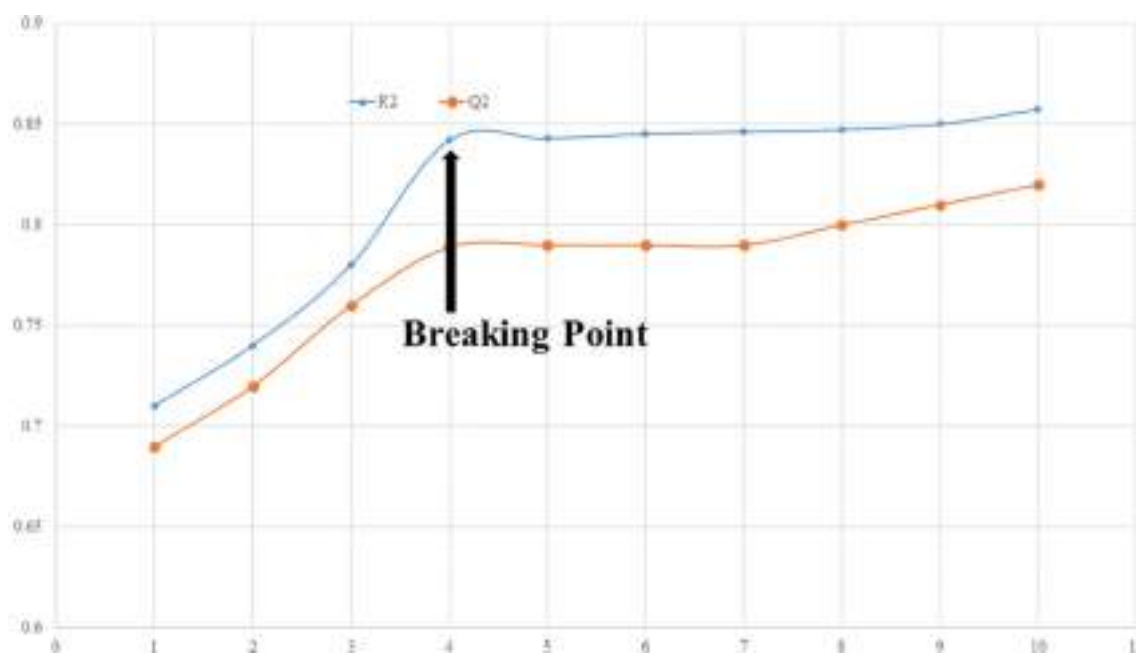


Fig. 2 Plot of number of descriptors against Coefficient of Determination R^2 and Leave-One out Coefficient of Determination Q^2 to identify the optimum number of descriptors.

model, it was once found that the value of Q^2 extended up to 4 variables, but then dropped significantly. To avoid overfitting and construct simple and informative QSAR models, the molecular descriptor vary was once confined to a set of 4 descriptors (Fujita and Winkler, 2016; Gramatica, 2014).

Values for molecular descriptors used in QSAR models can be determined in the extra information for every molecule. Because one of the OECD guidelines advises methodically validating a QSAR model, all of the fashions had been subjected to inside and external validation, Y-scrambling, and

QSARINS model applicability domain (AD) analysis. A GA-MLR based QSAR model's statistical agreeable and strength have been assessed the use of the following criteria: (a) Internal validation primarily based on leave-one-out (LOO) and leave-many-out (LMO) system (i.e. cross-validation (CV)); (b) the usage of External validation; (c) Y-randomization (or Y-scrambling) and (d) fulfilling of precise threshold value for the statistical limits (Consonni et al., 2019; Huang and Fan, 2011) : $R_{tr}^2 \geq 0.6$, $Q_{loo}^2 \geq 0.5$, $Q_{LMO}^2 \geq 0.6$, $R^2 > Q^2$, $R_{ex}^2 \geq 0.6$, $RMSE_{tr} < RMSE_{cv}$, $\Delta K \geq 0.05$, $CCC \geq 0.80$, $Q^2 - Fn \geq 0.60$, $r^2_m \geq 0.6$, $(1-r^2/ro2) < 0.1$, $0.9 \geq ok \geq 1.1$ or $(1-r^2/ro2) < 0.1, 0.9 \geq k' \geq 1.1, |ro2 \geq r'o2| < 0.3$ with RMSE and MAE shut to zero. As a result, any QSAR model that did no longer meet the above-mentioned criteria was once eliminated. The formulae for calculating these statistical parameters are accessible in the [supplementary material](#).

2.6. Molecular docking analysis

The protein data bank provided the pdb file for SARS-CoV 229e 3CLpro (pdb id-2zu2). The pdb 2zu2 was carefully chosen for its X-ray resolution and sequence completion. For docking analysis, the optimised protein is suitable (see Fig. 7). The native ligand (zinc-coordinating and peptidomimetic chemicals) was eliminated before docking study. The binding site for native ligand has been considered as the active site in the present work. Consequently, all the compounds were docked into the active site, where native ligand was bound with SARS-CoV 229e 3CLpr, the docking pose for the most active molecule as a representative is presented here for convenience.

The software NRGSuite was utilized to perform the molecular docking study. (Gramatica, 2007) This is a free and open source software that may be used as a PyMOL plugin. With the help of FlexAID, it can detect surface holes in a protein and use them as target binding sites for docking simulations. It uses a genetic algorithm to operate conformational search, model ligand and side-chain flexibility, and allows for covalent docking simulation. To acquire the great performance using NRGSuite, the flexible-rigid docking method was used with the following default settings: input method for binding sites—spherical shape (diameter: 18 Å); spacing of three dimensional grid—0.375 Å; side chain flexibility-no; ligand flexibility-yes; ligand pose as reference-no; constraints-no; Hetero groups-included water molecules; van der Waals permeability-0.1; solvent types-no type; number of chromosomes-1000; number of generations-1000; fitness model-share; reproduction model-population boom; and number of top complexes-5. For validation of molecular docking, the molecule TG-0204998, a recognized peptidomimetic inhibitor of SARS-CoV 229e 3CLpro, was used to validate the docking protocol.

2.7. MD simulation analysis

The virtual screening results are used to analyze the Hit Molecule 97 with a docking score of -8.043 kcal/mol and Molecule 4 the usage of the Schrodinger Desmond module in molecular dynamics and simulation (MD simulation). Using Hit Molecule 97 and Molecule 4 docking complexes, the SPC (Simple point charge) model used to be employed to bind protein ligands. The OPLS-2005 pressure subject (Gaudreault et al.,

2015) and explicit solvent model with the SPC water molecules have been used in this system (Lee et al., 2009). Na⁺ ions had been introduced to neutralize the charge. 0.15 M, NaCl options delivered to the machine to simulate the physiological environment. The NPT ensemble was set up by the use of the Nose-Hoover chain coupling scheme (Jorgensen et al., 1996) with temperature 300 K, leisure time of 1.0 ps and pressure 1 bar was once maintained in all the simulations. A time step of 2 fs used to be used. The Martyna-Tuckerman-Klein chain coupling scheme barostat technique was used for pressure control with a leisure time of 2 ps.

Long-range electrostatic interactions were calculated the use of the particle mesh Ewald technique (Martyna et al., 1992) with a radius of 9 Å for Coulomb interactions. The non-bonded forces have been calculated using the RESPA integrator. To have a look at the balance of the complex in MD simulations, the root mean square deviation (RMSD), root mean square fluctuation (RMSF), radius of gyration (Rg), and protein ligand interactions had been measured.—

2.8. Molecular mechanics generalized Born and surface area (MMGBSA) calculations

During MD simulations of 2zu2 complexed with dataset compound 4, most active hit molecule 97 and least active hit molecule 70, the binding free energy (Gbind) of docked complexes was calculated using the premier molecular mechanics generalized Born surface area (MM-GBSA) module (Schrodinger suite, LLC, New York, NY, 2017–4). The binding free energy was calculated using the OPLS 2005 force field, VSGB solvent model, and rotamer search methods (Lai et al., 2006). After the MD run, 10 ns intervals were used to choose the MD trajectories frames. The total free energy binding was calculated using Eq. (1):

$$\Delta G_{\text{bind}} = G_{\text{complex}} - (G_{\text{protein}} + G_{\text{ligand}}) \quad (1)$$

Where,

ΔG_{bind} = binding free energy,

G_{complex} = free energy of the complex,

G_{protein} = free energy of the target protein, and

G_{ligand} = free energy of the ligand.

3. Result

In this paper, QSAR and Molecular docking studies were employed to uncover hidden structural information responsible for SARS-CoV 229e 3CLpro inhibition. The QSAR model is created using PyDescriptor, which is simple to understand and link with biological activity. With the availability of easily accessible chemical descriptors and interpretation in terms of structural properties, the four-parameter GA-MLR model shows strong external prediction ability. Even though the current analysis used a straight evaluation of EC₅₀ values of the molecules in the dataset to describe the influence of a precise descriptor, it is important to note that the combined or opposite effect of unknown factors or other molecular descriptors could have a significant impact on the molecule's EC₅₀ value. (See Table 1)

The present QSAR analysis is performed using a data set comprising structurally assorted 37 compounds with experimentally determined EC₅₀ value ranging from 200 to

60,000 nM. Therefore, it encompasses acceptable as well as comprehensive chemical space and data range. This will be utilized for the development of properly validated genetic algorithm combined multilinear regression (GA-MLR) model to assemble or outspread exhaustive evidence about the pharmacophoric features that govern desired bio-activity (Descriptive QSAR) and having adequate external predictive capability (Predictive QSAR). The four variable based GA-MLR QSAR model along with the selected internal and external validation parameters (see [supplementary material](#) for additional parameters) is as follow:

3.1. QSAR model

To accomplish a better knowledge about structural features determining the SARS-CoV 229e 3CL^{pro} inhibitory activity, we have used interpretable molecular descriptors (as structural features) for model development. The GA-MLR QSAR four parametric model with its selected internal and external validation parameters (see [supplementary material](#) for additional parameters), is as follow:

QSAR Model: Training set: 30 (80 %) and Prediction set: 7 (20 %)

$$pEC_{50} = 5.343 (\pm 0.516) + 0.567 (\pm 0.562) * f_{\text{notringNsp3C3B}} + -0.043 (\pm 0.198) * f_{\text{acch4B}} + -0.108 (\pm 0.352) * \text{com_lipohyd_3A} + -0.208 (\pm 0.555) * \text{ringC_sp3N_2B}$$

Validation parameters for QSAR model: R^2 : 0.8425, R_{adj}^2 : 0.8183, Q_{loo}^2 : 0.7762, $R^2-Q_{\text{loo}}^2$: 0.0663, $R^2-R_{\text{adj}}^2$: 0.0242, K_{xx} : 0.2178, Delta K: 0.1545, RMSE_{tr}: 0.2203, RMSE_{cv}: 0.2626, RMSE_{ex}: 0.2014, Sy: 0.2405, F: 34.7785, Q^2-F_1 : 0.6342, Q^2-F_2 : 0.6168, Q^2-F_3 : 0.8683, CCC_{tr}: 0.9145, CCC_{cv}: 0.8776, CCC_{ex}: 0.8584, $r^2_{\text{m av}}$: 0.6534, $r^2_{\text{m de}}$: 0.1706, MAE_{tr}: 0.1896, MAE_{cv}: 0.2272, MAE_{ex}: 0.1779, RSS_{tr}: 1.5043, PRESS_{cv}: 2.1379, PRESS_{ex}: 0.2434, R_{LMO}^2 : 84.6802, Q_{LMO}^2 : 73.6421, R_{Yscr}^2 : 13.6730, Q_{Yscr}^2 : -25.4536

In the present QSAR modeling work, various statistical validation parameters were suggested to justify the internal and external robustness and have typical meaning (see [supplementary material](#) for detailed descriptions and formulae). The high value of unlike statistical parameters like R_{tr}^2 (coefficient of determination), R_{adj}^2 (adjusted coefficient of determination), and R_{cv}^2 (Q^2_{loo}) (cross-validated coefficient of determination for leave-one-out), R_{ex}^2 (external coefficient of determination), $Q^2 - F''$ and CCC_{ex} (Concordance Correlation Coefficient) etc. and low value of LOF (lack-of-fit), RMSE_{tr} (Root mean square error), MAE_{tr} (Mean absolute error), R_{Yscr}^2 (R^2 for Y-scrambling), etc. alongside different graphs obtained in the developed QSAR model explain the statistical robustness as well as excellent internal and external predictive ability with no chancy correlation. Furthermore, the Williams plot specifies that the model is statistically satisfactory (see [Fig. 5](#)). Thus, the developed QSAR model satisfies all the Organisation for Economic Co-operation and Development (OECD) suggested guidelines. (See [supplementary material section 1.3.1 for explanation and calculation method of various statistical parameters](#))

3.2. Discussion

A properly developed and validated QSAR model successfully established a correlation between a salient pharmacophoric

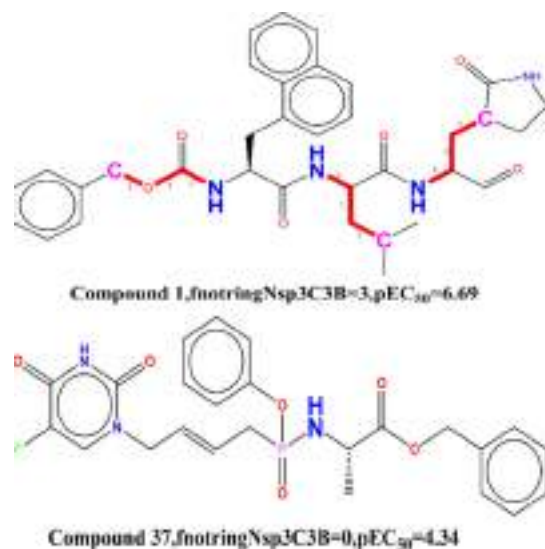


Fig. 3 Display of Descriptor $f_{\text{notringNsp3C3B}}$ exclusively for the molecule 1 and 37.

traits presented by molecular descriptors and their biological activity, that extend hidden information about mechanistic features of molecule, specificity of particular substituents and even presence or absence of various pharmacophoric aspects critical for SARS-CoV 229e 3CL^{pro} inhibition. Although, in the developed QSAR model, we have compared the EC₅₀ value of diverse dataset molecules in correlation and as an effect of certain molecular descriptor, however a similar or opposite effect of other molecular descriptors or unknown features having a prevailing influence in determining the general EC₅₀ value of a molecule cannot be ignored. Moreover, In other words, a single molecular descriptor is not sufficiently expert of fully clarifying the experimental EC₅₀ value for such a diverse set of molecules. That is, the successful application of the established QSAR model depend on the simultaneous usage of constituent molecular descriptors.

3.2.1. Mechanistic interpretation of descriptor

3.2.1.1. $f_{\text{notringNsp3C3B}}$. The descriptor $f_{\text{notringNsp3C3B}}$ point out closer to the frequency of prevalence of sp³ hybridized carbon atom precisely at three bonds from non-ring nitrogen atom. The descriptor $f_{\text{notringNsp3C3B}}$ has nice core-relationship with the pEC₅₀, therefore increase in the wide variety of such combination in molecule, and may similarly enhances the SARS-CoV 229e 3CL^{pro} inhibition.

The calculation of the $f_{\text{notringNsp3C3B}}$ descriptor was dis-allowed if the same sp³ hybridized carbon atom was once simultaneously present at one or two bonds from any different non-ring nitrogen atom. This statement supports, when we have compared the structures of molecule 1 ($pEC_{50} = 6.69$, $f_{\text{notringNsp3C3B}} = 3$) and 37 ($pEC_{50} = 4.22$, $f_{\text{notringNsp3C3B}} = 0$). If, we amplify the value of the descriptor $f_{\text{notringNsp3C3B}}$ from 1 for the molecule 37 to 3 resulted into increase in the pEC₅₀ by means of about 2.47 unit (about 20-fold expand in the SARS-CoV 229e 3CL^{pro} inhibition). Furthermore, the presence of a sp³ hybridized carbon atom at 3 bonds from a non-ring nitrogen atom performs a necessary role in SARs covid viral inhibition when we consider that

it increases hydrophobicity and offers an electrostatic function to the molecule 1. Molecular 37, on the different hand, lacks the same property, which ought to be the reason of the discrepancy in pEC_{50} of these molecules. (See Fig. 3), Similar observation is revealed when we have in contrast molecule 2 ($pEC_{50} = 6.30$ nm, $fnotringNsp3C3B = 3$) with molecule 34 ($pEC_{50} = 4.43$ nm, $fnotringNsp3C3B = 0$).

3.2.1.2. *facch4B*. The descriptor *facch4B* highlights the frequency of hydrogen atoms precisely at four bonds from the acceptor atoms. Because the descriptor has a negative correlation with pEC_{50} , adding greater nitrogen atoms at four bonds from the acceptor atom might also decrease the pEC_{50} value of these molecules. If the identical Hydrogen atom is simultaneously existing at two to three bonds from any acceptor atom, then it was once excluded at some stage in the calculation of *facch4B*.

The poor pEC_{50} for the molecule 33 ($pEC_{50} = 4.60$) and 37 ($pEC_{50} = 4.22$) could be attributed to the frequency of occurrence of such hydrogen atoms exactly at 4 bonds from the acceptor atoms. (*facch4B* = 14). This ought to be the possible reason for the variation in the biological activity of the molecule 1 ($pEC_{50} = 6.69$, *facch4B* = 12), 37 ($pEC_{50} = 4.22$, *facch4B* = 14) resp (See Fig. 4). If, we limit the value of the descriptor *facch4B* from 14 for the molecule 37 to the 12 resulted into increase in the pEC_{50} by about 2.47 unit (about 20-fold amplification in the SARS-CoV 229e 3CLpro inhibition) (see Fig. 5).

Because this descriptor has a negative coefficient in the generated models, the number of hydrogen atoms close to the 4 bonds acceptor atom is a proper combination to hire for SARS-CoV 229e 3CLpro lead/drug optimization. Because hydrogen is the smallest element, it suggests that the bulk in the vicinity of ring Nitrogen atoms be stored to a minimum. To enhance SARS-CoV 229e 3CLpro inhibition, steric bulk close to acceptor atom inside four bonds atoms be decreased or averted in future changes.

3.2.1.3. *Com_lipohyd_3A*. This descriptor signify the prevalence of hydrophobic atoms inside $3A^0$ from the center of mass

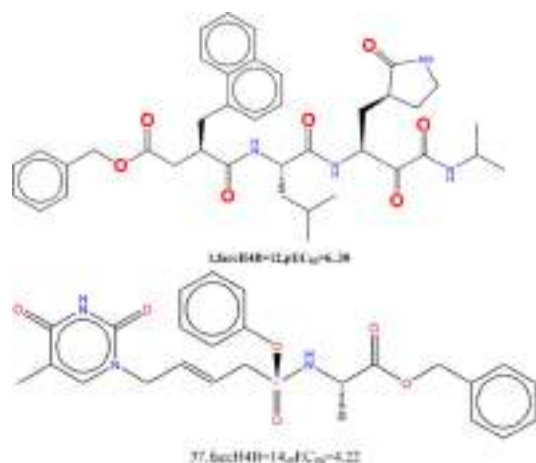


Fig. 4 Presentation of the descriptor *facch4B* for the molecules 1 and 37 only.

of the molecule. It has a negative coefficient in the developed QSAR model; therefore, growing the number of such hydrophobic atoms may want to end result in the reduced EC_{50} value for a molecule for SARs covid virus. The underprivileged pEC_{50} values for the molecules 35 ($pEC_{50} = 4.35$, $com_lipohyd_3A = 4$), 36 ($pEC_{50} = 4.34$, $com_lipohyd_3A = 3$) and 37 ($pEC_{50} = 4.22$, $com_lipohyd_3A = 3$) may want to be associated with the occurrence of hydrophobic atoms inside $3A^0$ from the center of mass of the molecule. This observation supported, when we have compared another pair of molecules, 25 ($pEC_{50} = 5.05$, $com_lipohyd_3A = 3$) with 3 ($pEC_{50} = 6.22$, $com_lipohyd_3A = 1$). If, we limit the value of the descriptor $com_lipohyd_3A$ from 3 for the molecule 25 to 1 will give upward thrust to the 2 expand in the pEC_{50} via about 1.17 unit (about eleven fold increase in the SARS-CoV 229e 3CLpro inhibition). Furthermore, in some molecules namely; 1 ($pEC_{50} = 5.10$, $com_lipohyd_3A = 0$), 18 ($pEC_{50} = 5.10$, $com_lipohyd_3A = 0$) and 20 ($pEC_{50} = 5.10$, $com_lipohyd_3A = 0$), lipo-hydrophobic atom is absent therefore, it exhibits that there is a requirement of only one lipo-hydrophobic atom inside $3A^0$ from the center of mass of the molecule. This statement is supported, when we have in contrast the activity of molecule 1 ($pEC_{50} = 6.69$, $com_lipohyd_3A = 1$) with the molecule 17, 18 and 20.

3.2.1.4. *RingC_sp3N_2B*. The presence of a sp^3 hybridized nitrogen atom within two bonds from ring carbon atoms is represented by this description. In the mounted QSAR model, this descriptor has a negative coefficient; thus, an amplification in the number of such sp^3 hybridized nitrogen atoms should result in a decrease in the EC_{50} value for the molecule for SARS-CoV 229e 3CLpro. The poor EC_{50} for the molecules 34 ($pEC_{50} = 4.40$, $ringC_sp3N_2B = 3$), 35 ($pEC_{50} = 4.35$, $ringC_sp3N_2B = 4$), 23 ($pEC_{50} = 5.05$, $ringC_sp3N_2B = 3$), 25 ($pEC_{50} = 5.05$, $ringC_sp3N_2B = 3$) and 26 ($pEC_{50} = 5.051$, $ringC_sp3N_2B = 1$) may additionally attributed to the high frequency of occurrence of such sp^3 hybridized nitrogen atoms. In the existing dataset, there are around 14 molecules which have 3 to 1 such sp^3 hybridized nitrogen within 2 bonds from the ring carbon atoms. Based on this analysis, it is rationale to say that, close combination of such nitrogen atom and ring carbon atom should be eluded in future to have higher EC_{50} for SARS-CoV 229e 3CLpro inhibition. Alongside, molecule 1 ($pEC_{50} = 6.69$, $ringC_sp3N_2B = 0$), 2 ($pEC_{50} = 6.3$, $ringC_sp3N_2B = 0$), 3 ($pEC_{50} = 6.2$, $ringC_sp3N_2B = 0$), 7 ($pEC_{50} = 5.74$, $ringC_sp3N_2B = 0$), 8 ($pEC_{50} = 5.74$, $ringC_sp3N_2B = 0$), and 9 ($pEC_{50} = 5.60$, $ringC_sp3N_2B = 0$) show absence of such Sp^3 hybridized nitrogen atoms, that ought to be the possible cause for the decrease in the activity of these molecules.

The constituent molecular descriptors obtained in the GA-MLR QSAR model have presented visible and hidden records about the structure landscapes linked to a various set of molecules investigated for their activity against SARS-CoV 229e 3CLpro inhibition in the current QSAR study. It is essential to recognize that no single molecular description can totally explain the supported EC_{50} distribution for such a numerous set of molecules. That is, the performance of the built QSAR model is dependent on the employment of constituent molecular descriptors at the equal time.

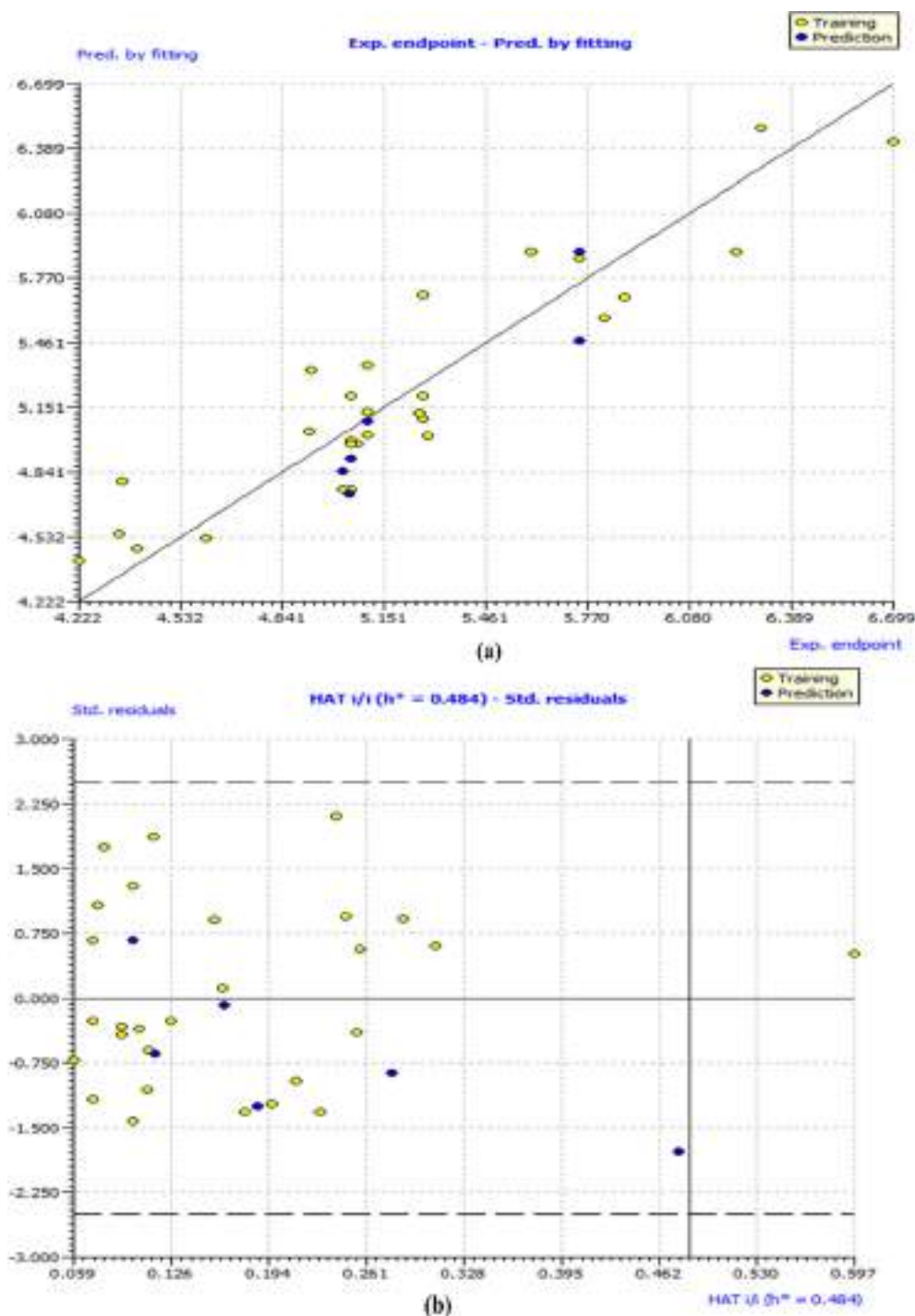


Fig. 5 Different graphs associated with the developed Quantitative Structure – Activity Relationship (QSAR) model: (a) experimental vs predicted pEC₅₀ and (b) Williams plot to assess applicability domain of model, and (c) Insubria Plot.

3.3. QSAR-Based virtual screening

Supplementary Materials include SMILES notations, calculated molecular descriptor values, pEC₅₀, and EC₅₀ for a 100-compound in-house library utilized for virtual screening. We've included the five most active and five least active hit molecules from the in-house library, as predicted through the created QSAR model, for the sake of convenience. (See Fig. 2)

3.3.1. Docking analysis

SARS-CoV 3CL^{pro} is a dimeric protein with three domains in each subunit. 3CL^{pro} has a massive loop between -strands C1 and D1, in accordance to structure-based sequence alignment. SARS-CoV 3CL^{pro}'s C1–D1 loop keeps the P2 facet chain in the S2 hydrophobic pocket. The C1–D1 loop of SARS-CoV 3CL^{pro} secures the S2 hydrophobic pocket for the P2 side chain. Gln as the P1 residue, a hydrophobic residue at the

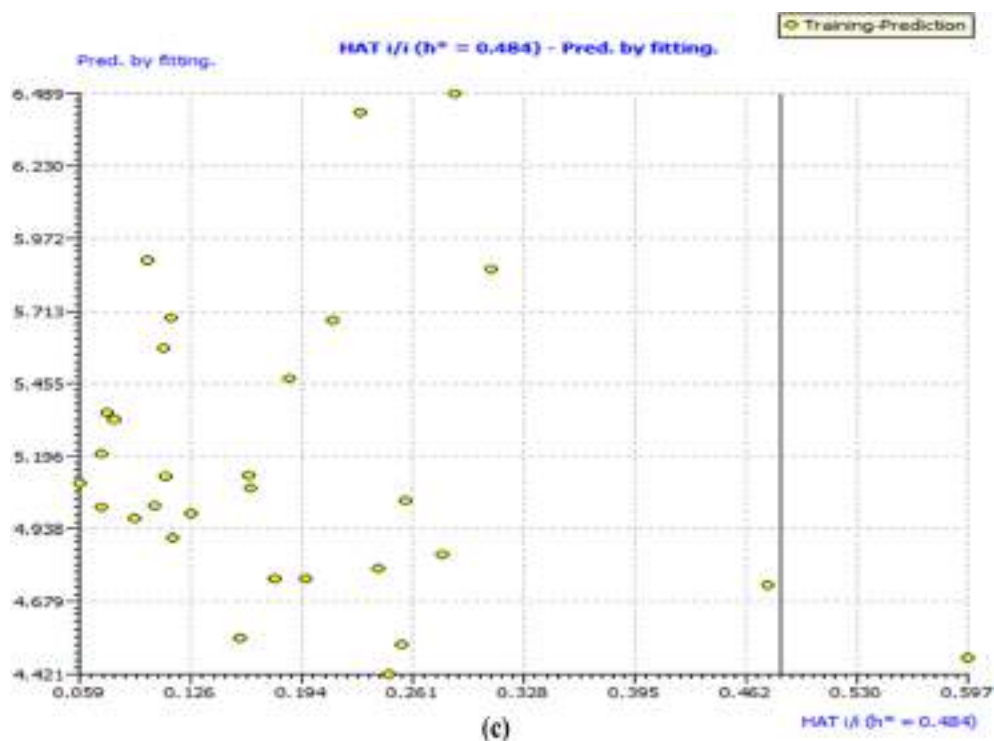


Fig. 5 (continued)

P2 position, and a brief amino acid residue at the *P1* position are all identified by means of 3CLpro with similar substrate specificity. (see Fig. 6).

To unfold binding mode and interactions, the dataset molecule 4 and a known inhibitor such as TG-0204998 were docked into the active binding pocket of SARs Cov 229e 3CL^{pro} in this study.

The substrate binding subsites are chosen as S1, S1', S2, S3, and S4, with preserved water molecules. The catalytic dyad of His-Cys is located in the active site in the cleft between

domains I and II, whereas domain III participates in the protease dimerization. TG-0204998, the unsaturated ethyl ester occupies the S1 site, which is in close proximity to the catalytic center. Therefore, we have selected native binding site of known inhibitors, TG-0204998 as an active site in the docking protocol.

The TG-0204998 is the peptidomimetic inhibitors of SARs Cov 229e 3CLpro, whose X-ray resolution shape is used to validate the docking protocol. The alignment of SARs Cov 229e 3CLpro with the TG-0204998 and the molecule four is

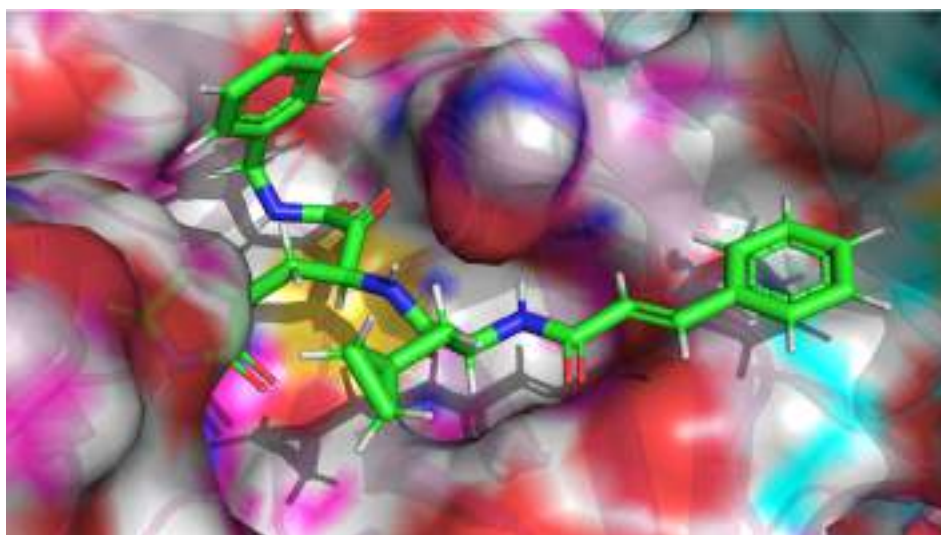


Fig. 6 Depiction of Molecule 4 orientation within the binding pocket of SARS-CoV 229e 3CL^{pro} (pdb id-2zu2).

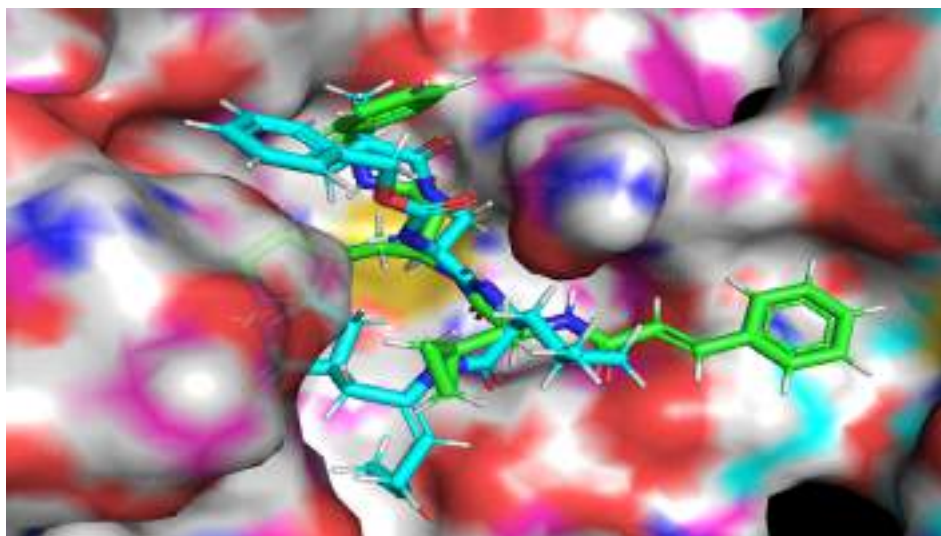


Fig. 7 Display of Superimposed structures of Molecule 4 (Green colored) with Molecule TG-0204988 (Cyan Colored) within the binding pocket of SARS-CoV 229e 3CL^{pro} (pdb id-2zu2).

depicted in Fig. 7, which mean that the docking protocol is agreeable (see Fig. 7). Table 2 divulges the docking scores for the 5 most active and 5 least active dataset molecules.

In this paper, we have identified a novel class of SARs Cov 229e 3CL^{pro} inhibitor by performing a computer aided drug-design protocol. Our experiment contain selection of the dataset containing 37 structurally diverse compounds whose activity was predicted by using QSAR modelling. Further, the developed QSAR model was once used to predict the biological activity of in-house library undertaking of 100 numerous compounds. Successively, we have docked all the hundred compounds into the SARs Cov 229e 3CL^{pro}. On the foundation of docking simulation outcomes (docking score), we have chosen 5 most active and 5 least active hit compounds accompanied through molecular dynamic simulation and binding free energy calculations.

Amongst the five most active molecules, molecule 4 (-8.473 kcal/mol) and molecule 5 (-9.605 kcal/mol) achieved good docking score but, RMSD value become less for the molecule 5; therefore. Molecule four is chosen for the docking analysis. The molecule four with the RMSD value of 1.609 display better fit into the binding pocket of SARS-CoV 229e 3CL^{pro}. The 2D and 3D interactions of molecule 4 with SARS-CoV 229e 3CL^{pro} is presented in Figs. 2 and 3 Even though, some inactive molecule namely, molecule 35 collect a docking score of -10.147 kcal/mol with the RMSD value of 2.53, that should be attributed to the large size of the molecule 10, as in contrast to molecule 4.

3.3.2. Docking pose analysis of the most active molecule 4

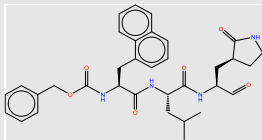
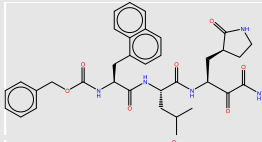
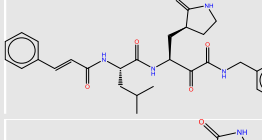
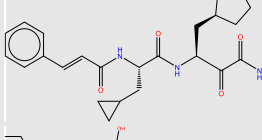
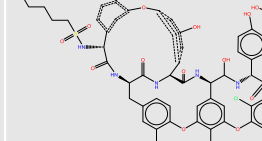
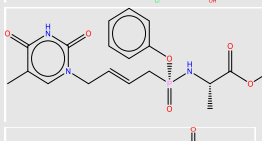
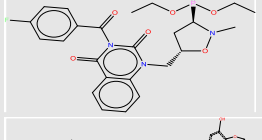
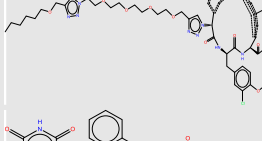
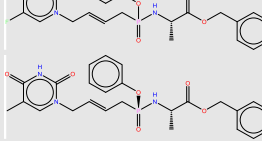
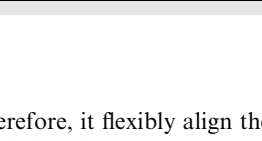
The molecule 4 bind to the active site of SARS-CoV 229e 3CL^{pro} as that of the TG-0204988 in the similar manner. Chemically, molecule 4 is a (S)-N-benzyl-3-((S)-2-cinnamamido-3-cyclopropylpropanamido)-2-oxo-4-((S)-2-oxopyrrolidin-3-yl) butanamide. The terminal benzyl substituent anchored with hydrophobic residue ALA: 1 thru pi-cation interactions. Likewise, 2-cinnamamido moiety subsequent to the benzyl moiety structure a hydrogen bond with the negatively charged residue GLU165 through the involvement of a

water molecule, while cinnamamido nitrogen form a hydrogen bonding contact with polar residue ASN B:14. Here, Glu165 is a necessary residue for keeping the enzyme in proper conformation. (See Fig. 8)

Further, the cyclopropyl substituent show exposure to the solvent whilst adjoining carbonyl oxygen bind with water molecule by hydrogen bonding interactions. Concurrently, oxygen atom of 2-oxopyrrolidin bind with negatively charged residue GLU165 through means of hydrogen bonding interaction whereas, terminal 2-oxo oxygen attached with Gly143 residue the usage of hydrogen bond. At the same time, adjacent benzamide oxygen anchored hydrogen bond with the hydrophobic residue Cys144, which form the S2 pocket of the SARS-CoV 229e 3CL^{pro}. Moreover, pyrrolidine ring form a close contact with the catalytic polar residue His B: 41 Here, it show off Pi-alkyl contact with the His B: 41 which signify the inhibitory impact of SARS-CoV 229e 3CL^{pro}.

The identical observation is indicated by using the descriptor fnotringNsp3C3B, highlight the significance of sp³ hybridized carbon atom exactly at three bonds from non-ring nitrogen atom. In the molecule 4, cinnamamido nitrogen structure hydrogen bonding with the polar residue Asn141 while Sp³ hybridized C1 carbon atom of cyclopropyl flexibly exposed the molecule 4 to the solvent accessible surface area of the receptor. Thus, QSAR and molecular docking results are analogous and are complimentary. Further, the descriptor faccH4B signify that, the presence of least quantity of hydrogen atom precisely at four bonds from acceptor atom in the ligand molecule is recommended for biological activity. The much less density of hydrogen atoms in the close proximity of the acceptor oxygen atom in the molecule 4 leads to the less steric bulk, which facilitate the desirable binding and alignment of molecule into the active pocket. This justify that, QSAR analysis showed right correlation with the docking results. Subsequently, the descriptor com_lipohyd_3A mean the prominence of the lipo-hydrophobic atoms within 3A0 from the center of mass of the molecule. Herein, center of mass of the molecule 4 is located near cyclopropylamido nitrogen atom, where no bulky hydrophobic substituent is present

Table 2 Portrayal of Structures, Docking Score (kcal/mol) and RMSD values for the five most active and five least active dataset molecules.

Molecule	Structures	Docking Score	RMSD
1		-7.1447477	2.7707791
2		-7.8803358	1.7312964
3		-7.3945093	2.4114711
4		-8.4731464	1.6090333
5		-9.605979	2.440057
33		-7.4059458	1.5316099
34		-6.7014847	2.1337159
35		-10.147323	2.534488
36		-6.8096747	2.0087693
37		-7.909008	1.5497004

within the vicinity of 3A0; therefore, it flexibly align the molecule four in the S2 pocket of the SARS-CoV 229e 3CLpro. This implies that, docking outcomes are in good agreement with QSAR analysis. Finally, the descriptor ringC_sp3N_2B gives an idea about the prevalence of sp³ hybridized nitrogen atom within 2 bonds from ring carbon atoms. In this, molecule four don't have such kind of combination which may additionally drop the binding affinity against SARS-CoV 229e 3CLpro. This exhibits that, QSAR analysis correctly identified con-

cealed and hidden structural characteristic decisive for SARS-CoV 229e 3CLpro inhibition.

3.3.3. Docking pose analysis of the most active hit molecule 97

Table 3 shows the molecular docking scores for the six most active and six least lively hit molecules. With a docking score of -8.043 kcal/mol and an RMSD of 1.53257, hit no. molecule 97 emerged as the most active of the 100 hit molecules. It reveals a 6.089 predicted EC₅₀. Although various hit

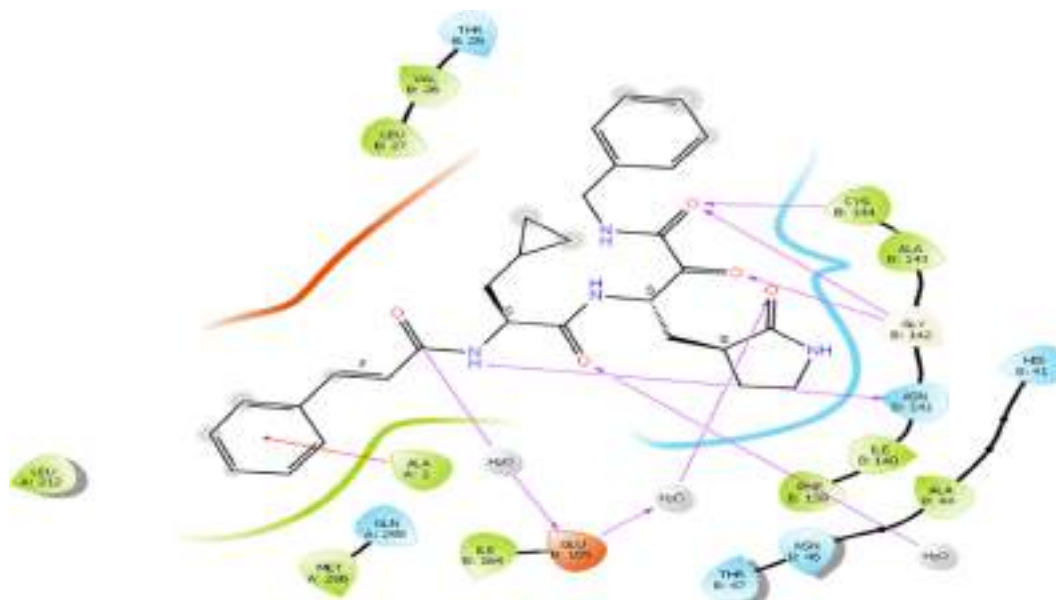


Fig. 8 Presentation of 2D interaction of molecule 4 with SARS-CoV 3CL^{pro} (pdb id-2zu2).

molecules, such as 19, 6, 39, 91, and 38, exhibited robust predictive activity, but they did not have top binding characteristics, consequently hit molecule 97 was chosen as the most outstanding hit for the analysis.

Chemically, hit molecule 97 is (S)-3-(2-((2-(*sec*-butylamino)-2-oxoethyl) thio)-4-oxoquinazolin-3(4H)-yl)-N-(3,4-dimethoxyphenethyl)propanamide. To verify the binding mode of hit molecule 97, we have docked it into the SARS-CoV 3CL^{pro}. The docking pose analysis exhibits that, hit molecule ninety-seven binds with the SARS-CoV 3CL^{pro} via hydrogen bonding and hydrophobic contacts with the involvement of water molecules. The drug receptor interaction gives rise to the docking score of -8.043 with the RMSD of 1.53. This result supports that the molecule 97 has very good fitting and affinity for SARS-CoV 3CL^{pro}. (See [Figs. 9, 10 and 11](#))

In the molecule 97, the 2-oxo-ethylthio oxygen forms a hydrogen bonding interaction with the key amino acid residue His B: 41, which forms a catalytic dyad of SARS-CoV 3CL^{pro} and an important amino acid residue in the S1 pocket (See [Fig. 10](#)). Further, the terminal side chain substituent dimethoxy-substituted phenyl ring structure forms a pi-cation contact with the hydrophobic residue Ala: 1 of the S2 hydrophobic pocket. Next, 4-oxoquinazolin oxygen atoms bind with Thr B: 7 residue through hydrogen bonding interaction with the involvement of water molecules. Moreover, propanamide oxygen in the amide linkage joining quinazolin and dimethoxy phenyl ring, forms a contact of hydrogen bond with the negatively charged Glu B:165 residue, while the amide nitrogen anchors a hydrogen bond with the Asn B:141 residue. Here, binding of hit molecule 97 into the respective S1 and S2 binding pockets of SARS-CoV 3CL^{pro} gives an explanation for its binding specificity. Our docking evaluation outcomes provide a structural basis for

the optimization of the hit molecule 97 and development of potential candidates for antiviral therapies.

3.4. Molecular dynamic simulation based analysis for the molecule 4 and hit molecule 97

Root mean square deviation (RMSD) of the C- α backbone of 229e with ligand coupled complex confirmed a fairly stable structure, with a fluctuation of 80 Å denoted as a pink colored line. However, the ligand Hit Molecule 97, RMSD was once at first slightly distorted until 40 ns, after which it grew to be steady from 70 ns to 100 ns with no additional variations. The RMSD of the C- α spine of HCoV 229e with ligand bound complex, on the other hand, confirmed a relatively stable structure with a fluctuation of 80, as proven by the green colored line. However, the ligand Molecule 4, RMSD exhibited minor distortions until 50 ns, after which it became stable from 60 ns to a hundred ns, with no similar fluctuations. (See [Fig. 12](#))

On the other hand, root mean square fluctuations of respective amino acids of C- α spine of 2 displayed least fluctuations signifying the stable protein structure ([Fig. 13](#)).

Ligand-protein interactions might also be tracked for the duration of the simulation. There are 4 sorts of interactions: additive, multiplicative, functions, and symmetric. Hydrogen bonds, hydrophobic, ionic, and water bridges are classifications of protein-ligand interactions. Ligand interaction of Hit Molecule 97 with the binding site residues of 229e and molecule 4 with the binding site residues of Hcov_229e; displayed the formation of non-bonded interactions such as hydrophobic interaction as shown in [Fig. 14\(A\) & \(B\)](#).

The radius of gyration (Rg) is the indicator of the size and compactness of the protein in the ligand-bound state displayed in [Fig. 12](#). We have observed the Rg plot of C α -backbone of

Table 3 Presentation of Structures, Docking Score (kcal/mol), RMSD and pEC₅₀ values for the five most active and five least active Hits obtained in QSAR Modeling Based Virtual Screening.

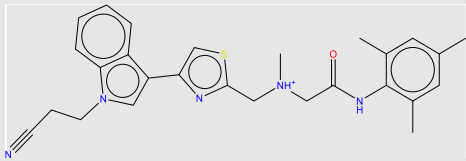
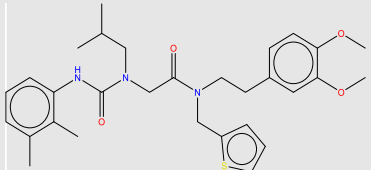
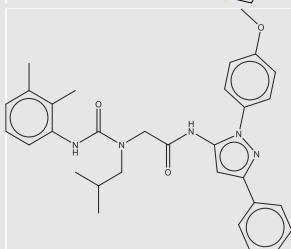
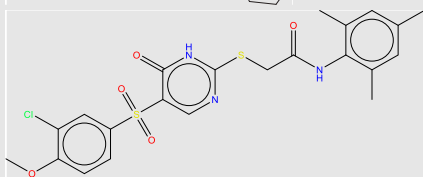
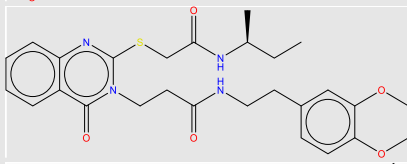
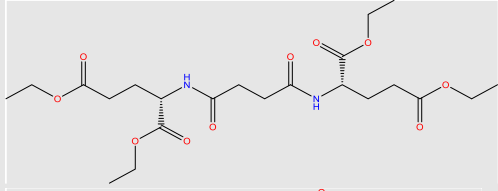
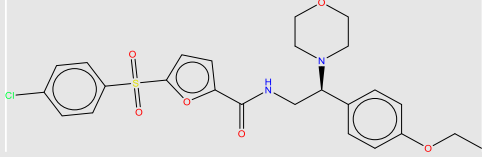
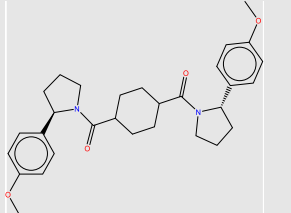
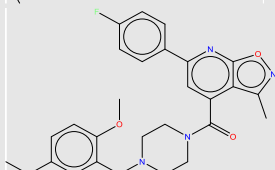
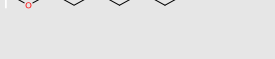
sn	Molecule	Structure	Docking score	RMSD	pEC ₅₀	status
1	19		-6.960	1.368	6.872	Most active
2	6		-7.126	2.978	6.743	Most active
3	39		-7.126	1.598	6.678	Most active
4	91		-6.728	2.3425	6.175	Most active
5	97		-8.043	1.53257	6.089	Most active
6	38		-7.335	1.666	6.025	Most active
7	4		-6.627	2.1068397	3.937	Least active
8	9		-7.485	1.298	3.921	Least active
9	59		-6.756	1.814	3.743	Least active
10	98		-6.948	0.9944	3.657	Least active

Table 3 (continued)						
sn	Molecule	Structure	Docking score	RMSD	pEC ₅₀	status
11	94		-7.533	1.638	3.635	Least active
12	70		-7.042	2.106	3.592	Least active

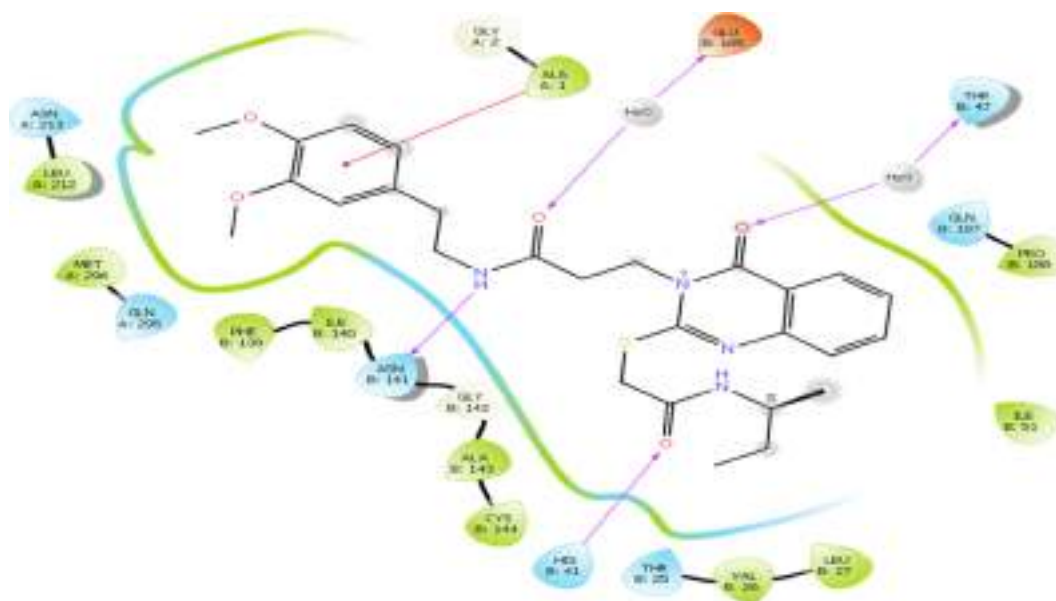


Fig. 9 Presentation of 2D interaction of Hit molecule 97 with SARS-CoV 3CL^{pro} (pdb id-2zu2).

229e- Hit Molecule 97 (red) and Hcov_229e-complex molecule 4(black) bound protein complex in Fig. 12 having significant compactness well after the last 40 ns with an average of 25.8 Å deviation indicating the significant convergence. We observed less Rg score in complex molecule 4 bound 229e complex with 25.6 Å throughout the simulation (Figure S1, black, See Supplementary material). But we observed the high lowering of Rg in Hcov_229e- Hit Molecule 97 bound complex (Figure S1, red, see Supplementary material), which signify less compactness and lesser stability comparatively.

We have also recorded the average hydrogen bonds formed between complexes Hcov_229e- Hit Molecule 97 (red) & 229e-complex molecule 4(black), during the 100 ns simulation in (Figure S2, see Supplementary material). The average hydrogen bond formed for complex Hcov_229e- Hit Molecule 97 (red) is 1 and for complex 229e-complex molecule 4(black) is 2.

A chronology of the interactions and contacts that were listed on the preceding page. The (Figure S3, See Supplementary material) displays the total number of distinct interactions the protein makes with the ligand during the journey. Ligand-

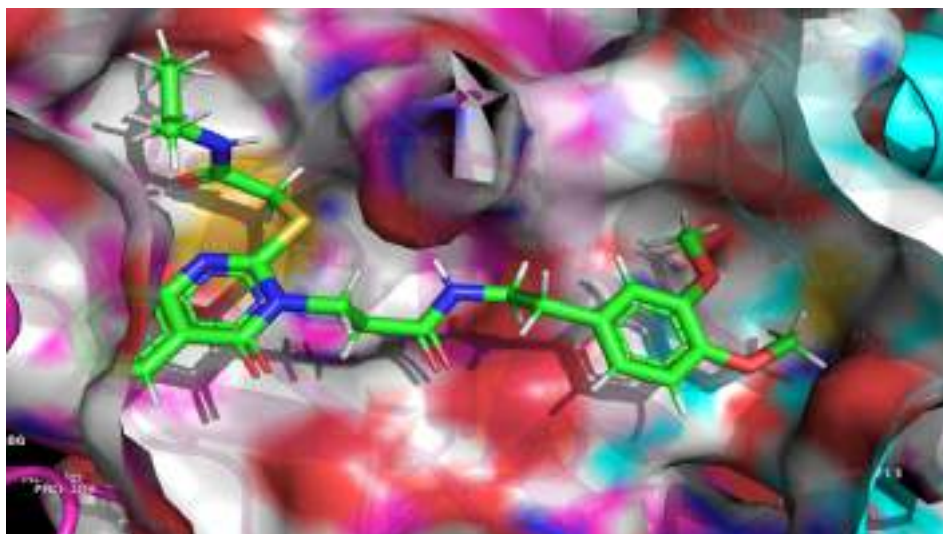


Fig. 10 Depiction of Hit Molecule 97 orientation within the binding pocket of SARS-CoV 229e 3CL^{pro} (pdb id-2zu2).

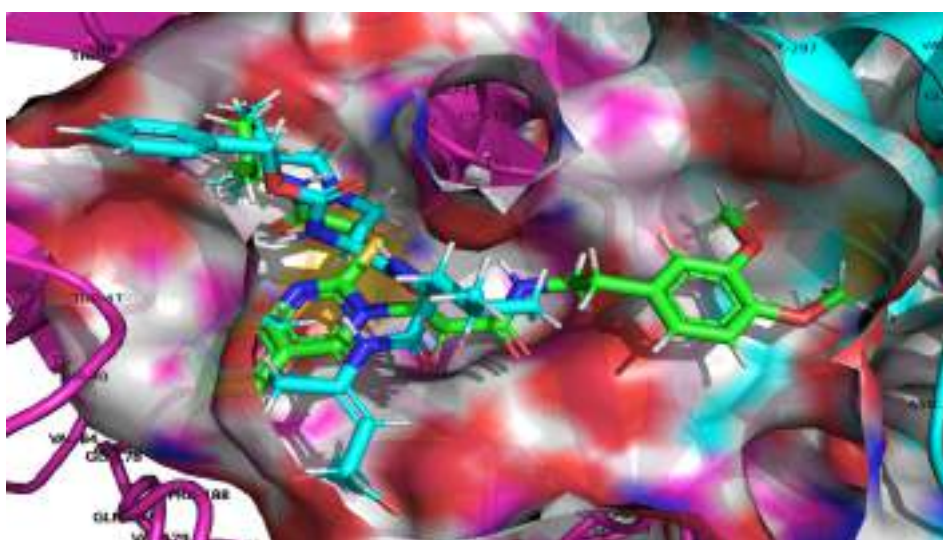


Fig. 11 Display of Superimposed structures of Hit Molecule 97 (Green colored) with Molecule TG-0204988 (cyan Colored) within the binding pocket of SARS-CoV 229e 3CL^{pro} (pdb id-2zu2).

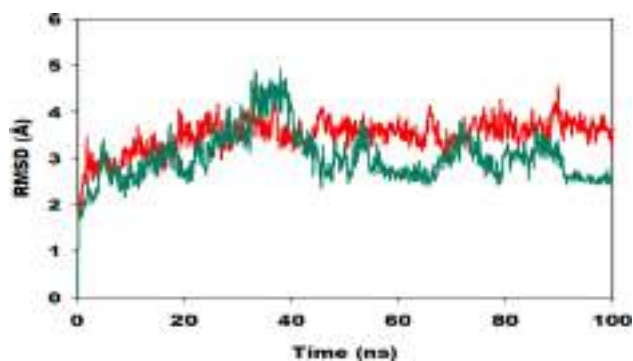


Fig. 12 Root mean square deviation (RMSD) of C- α backbone of 229e (red) with **Hit Molecule 97** and Hcov_229e (green) with ligand compound 4 for 100 ns simulation exhibiting a stable configuration of 229e-hit6 & Hcov_229e-compound4.

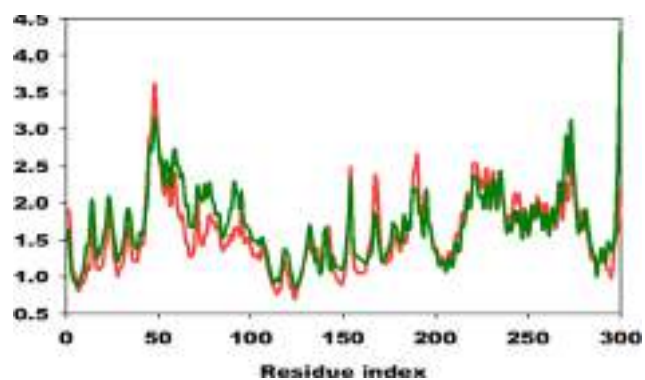


Fig. 13 Root mean square fluctuation of C- α backbone of 229e (red) & Hcov_229e (green) at its respective amino acid residues for 100 ns simulation exhibiting a stable configuration.

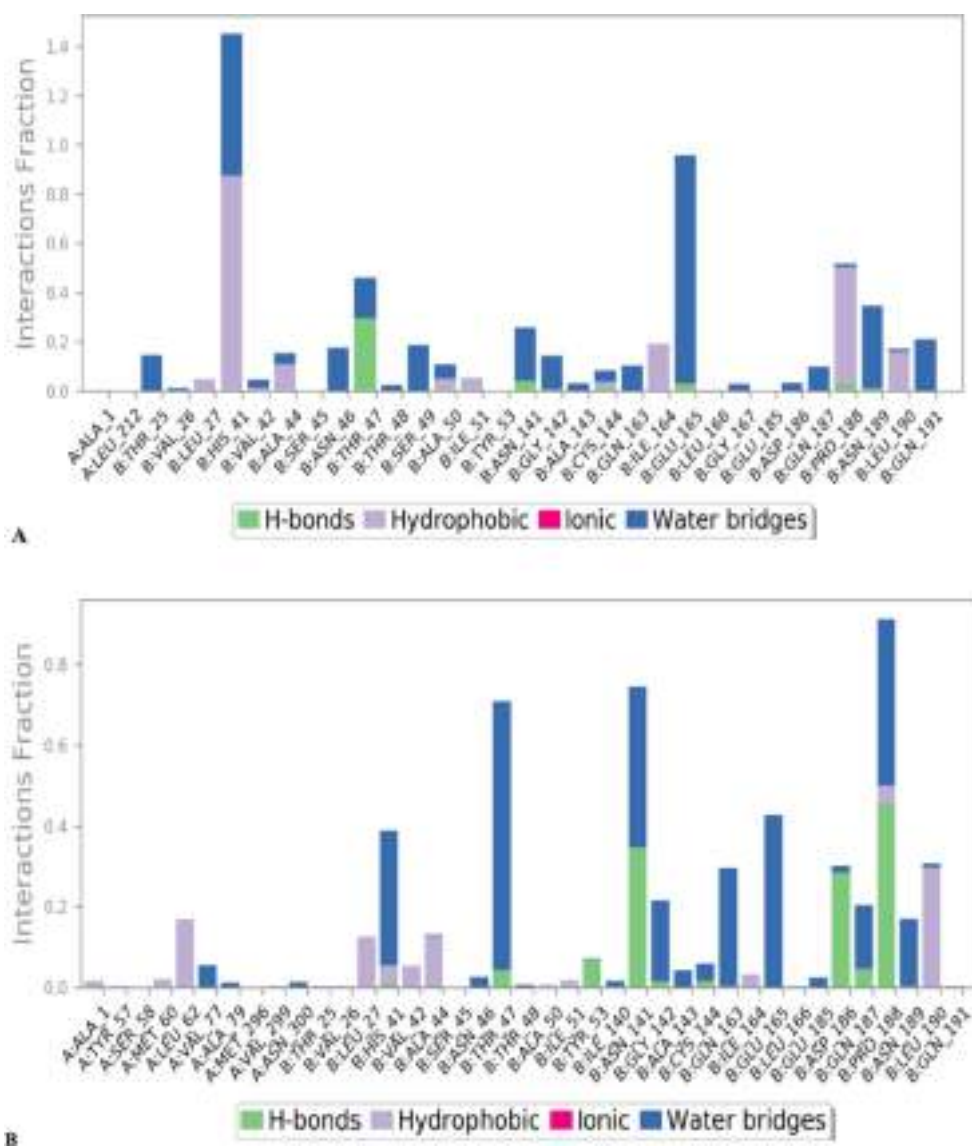


Fig. 14 The types of bonds and the amino acid residues that participated during 100 ns of simulation; (A)229e- **Hit Molecule 97**, (B) Hcov_229e-complex molecule 4.

interacting residues are shown on the bottom panel in [Figure S4 in Supplementary material](#). Some residues make several specific contacts with the ligand, which is shown by a deeper shade of orange on the y-axis. The range of distribution of distinct forms of the molecule is determined by the RMSD of a ligand to the reference conformation. Calculate the radius of gyration using the ligand's "extendedness," which is equal to the ligand's moment of inertia. In a ligand molecule, the number of intramolecular hydrogen bonds (intramolecular HB).

Use a probe radius of 1.4 to estimate the molecular surface area. This is the van der Waals area. Use the formula with the oxygen and nitrogen atoms acting as the entire composition to get the PSA. [Figure S5 \(See Supplementary material\)](#) highlights the ligand characteristics such as RMSD, radius of gyration (rGyr), intramolecular hydrogen bond, molecular surface area

(MolSA), solvent accessible surface area (SASA), and polar surface area (PSA). Both the ligands possess an intramolecular hydrogen bond.

A detailed molecular structure drawing that depicts the ligand molecules' molecular structure as well as specific amino acid residue interactions with protein residues. Interactions that occur 12.0 % or more of the simulation time are reported if the simulation lasts from 0.00 to 100.00 nsec. Some residues are capable of interacting with the same ligand atom via a variety of interactions. From the [Fig. 15\(A\)](#) it can be concluded that Glu165 is charge negatively; Ile164, Pro188, Leu190 are having hydrophobic interactions; while Thr25, His41, Asn46, Thr47, Ser49, Asn189, Gln191 are showing polar interactions with the ligand Hit Molecule 97. While from the [Fig. 15 \(B\)](#) it can be concluded that Pro188 is showing hydrophobic

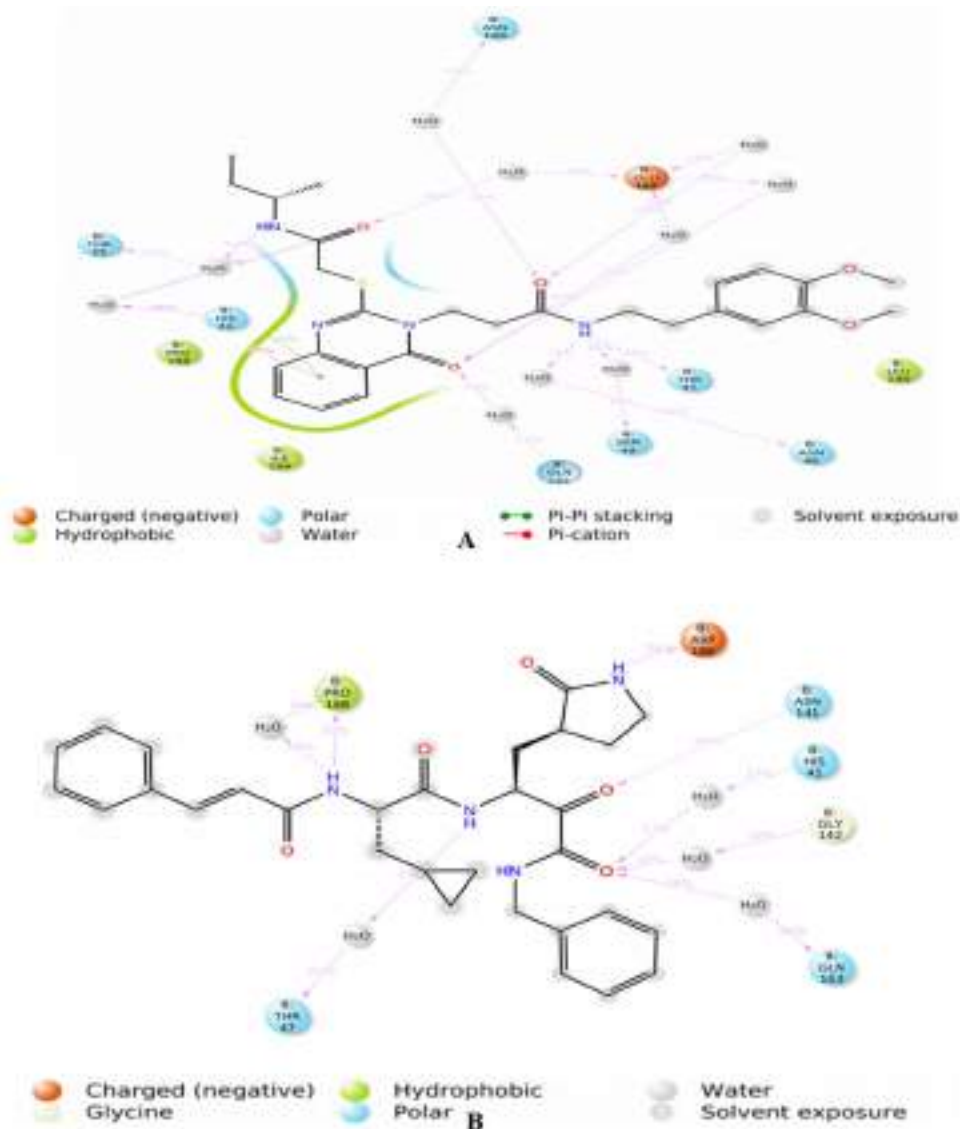


Fig. 15 2D interaction plots showing ligand interactions of 229e with (A) Hit Molecule 97 & (B) complex molecule 4 with the binding cavity residues of SARS-CoV 3CL^{pro}.

interactions; Asp186 is negatively charged; Gly142 is showing a hydrophilic interaction; while His41, Thr47, Asn141, Gln163 are having a polar interaction with the ligand complex molecule 4. (see Figs. 16 and 17 for 2D and 3D depiction of hit molecule 97 and molecule 4)

3.5. Molecular mechanics generalized Born and surface area (MMGBSA) calculations

MMGBSA is a popular method in calculating the binding energy of ligand to protein molecules. The estimation of the binding free energy of each of the protein–ligand complexes, as well as the role of other non-bonded interactions energies were estimated. The average binding energy of the ligands dataset compound 4 (229e-complex4), most active hit molecule 97 (229e-hit6) and least active hit molecule 70 with SARS CoV-229E 3CL^{pro} were found to be -32.2 ± 7.6 , -53.81 ± 6.7 and -7.2 ± 3.4 , respectively (Table 4). The ΔG_{bind} is influ-

enced by of various types of non-bonded interactions, including ΔG_{bind} Coulomb, ΔG_{bind} Covalent, ΔG_{bind} Hbond, ΔG_{bind} Lipo, ΔG_{bind} SolvGB and ΔG_{bind} vdW interactions. Among all the types of interactions ΔG_{bind} vdW, ΔG_{bind} Lipo and ΔG_{bind} Coulomb energies contributed most to achieve the average binding energy. In contrast, ΔG_{bind} SolvGB and ΔG_{bind} covalent energies contributed the lowest to attain the final average binding energies (see Fig. 18).

3.5.1. Results presented in Mean \pm SD

In addition, the values of ΔG_{bind} Hbond interaction of hit molecule 97, dataset compound 4 and 70-inactive protein complexes showed the stable hydrogen bonds with the amino acid residues. In all the complexes ΔG_{bind} SolvGB and ΔG_{bind} Covalent showed unfavorable energy contributions and thus opposed binding. It is observed from Fig. 18, at pre-simulation (0 ns) dataset compound 4, most active hit molecule 97 and least active hit molecule 70 at the binding

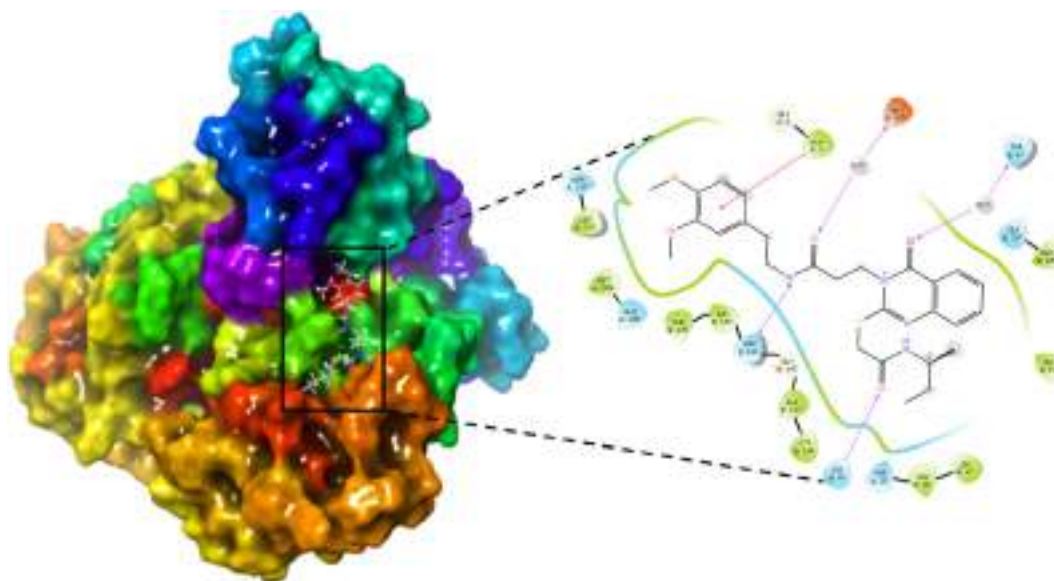


Fig. 16 Presentation of 3D and 2D interaction of Hit Molecule 97 in complex with SARS-CoV 3CL^{pro}.

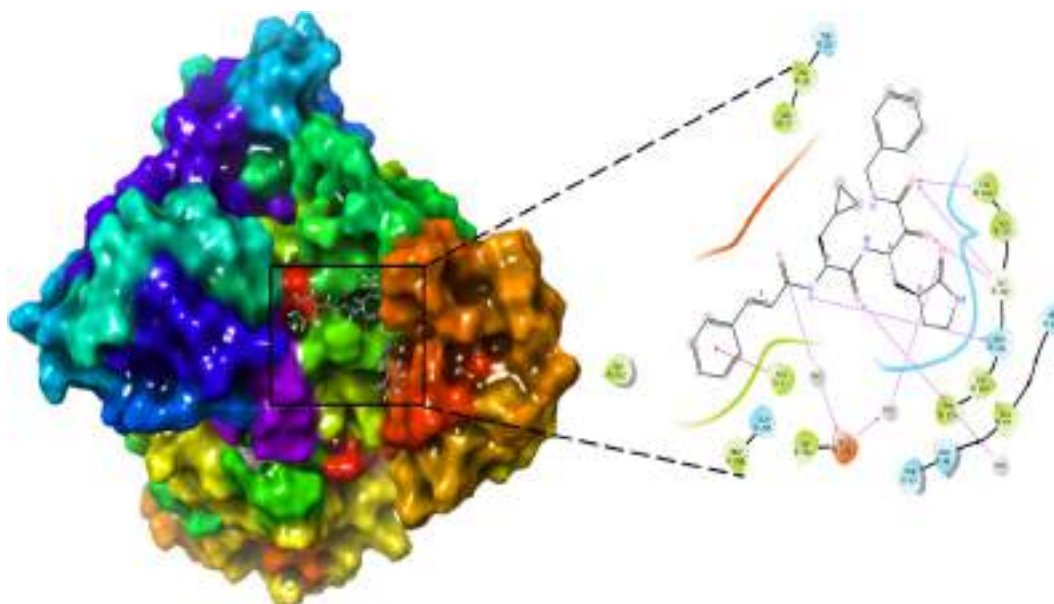


Fig. 17 Presentation of 3D and 2D interaction of Molecule 4 in complex with SARS-CoV 3CL^{pro}.

pocket of SARS CoV-229E 3CL^{pro} undergone substantial angular movement of the pose (curved to straight) after post simulation (100 ns). These conformational changes consequences the better acquisition at the binding pocket as well as the interaction with the residues for higher stability and better binding energy.

Thus MM-GBSA calculations resulted, from MD simulation trajectories well corroborated with the binding energies calculated from the docking results. Therefore, it can be suggested that the dataset compound 4, most active hit molecule 97 has good affinity for the major target SARS CoV-229E 3CL^{pro}. However, least active hit molecule 70 displayed least

Table 4 MMGBSA binding energy contribution by non bonded interactions by 229e-complex4, 229e-hit6 and 70-inactive molecules with the target protein.

Energies (kcal/mol)	229e-complex4	229e-hit97	70-inactive
ΔG_{bind}	-32.2 ± 7.6	-53.81 ± 6.7	-7.2 ± 3.4
$\Delta G_{\text{bindLipo}}$	-13.8 ± 2.9	-19.5 ± 2.4	-5.6 ± 1.1
$\Delta G_{\text{bindVdW}}$	-38.1 ± 7.7	-52.2 ± 7.2	-4.8 ± 6.0
$\Delta G_{\text{bindCoulomb}}$	-8.1 ± 7.9	-14.0 ± 9.1	-2.8 ± 0.9
$\Delta G_{\text{bindH}_{\text{bond}}}$	-0.14 ± 0.2	-0.95 ± 0.1	-0.49 ± 0.3
$\Delta G_{\text{bindSolvGB}}$	23.6 ± 9.1	30.6 ± 5.4	2.2 ± 0.7
$\Delta G_{\text{bindCovalent}}$	4.9 ± 2.3	2.8 ± 1.9	3.1 ± 3.5

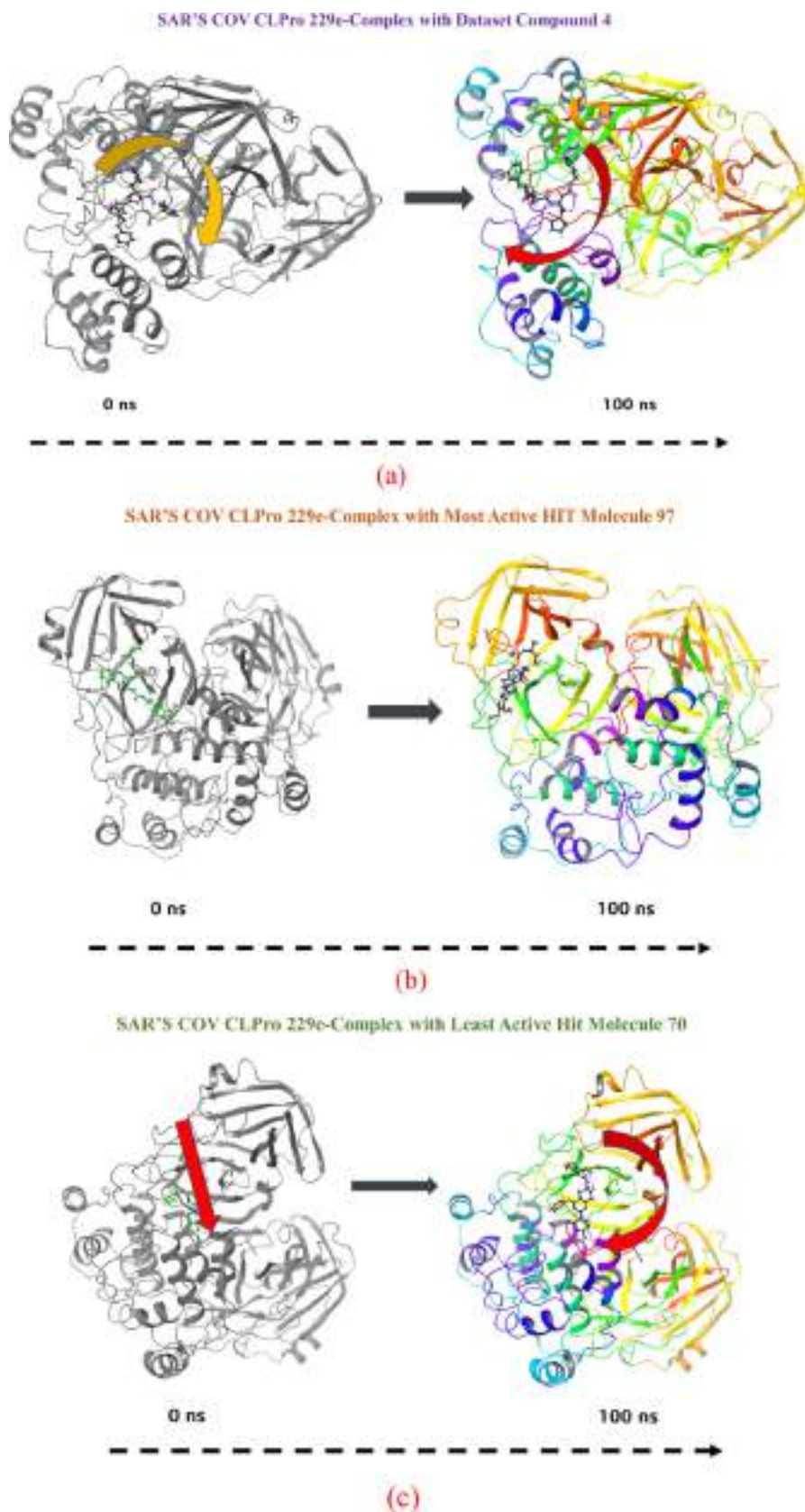


Fig. 18 MMGBSA trajectory (0 ns, before simulation and 100 ns, after simulation) exhibited conformational changes of dataset compound 4(a), most active hit molecule 97(b) and least active hit molecule 70 upon binding with the protein SARS CoV-229E 3CLpro. The arrows indicating the overall positional variation (movement and pose) of dataset compound 4, most active hit molecule 97 and least active hit molecule 70 at the binding site cavity.

binding energy with SARS CoV-229E 3CL^{pro}. The MMGBSA trajectories displayed the conformational changes in the dataset compound 4, most active hit molecule 97 and least active hit molecule 70 to achieve the best fitting in the binding cavity of the protein.

4. Conclusion

Throughout of this paper, QSAR modelling, QSAR-based virtual screening, molecular docking, and MD simulation reality findings are used to uncover the new molecule as a SARS-CoV 229e 3CL^{pro} inhibitor. Expending four descriptors, a GAMLR based QSAR model is invented to understand the essential pharmacophoric prospect accountable for the SARS-CoV 3CL^{pro} inhibition. Ensuing OECD directions, the QSAR model was once appraise for both internal and external validation measures. Pharmacophoric characters counting fno-tringNsp3C3B, faccH4B, com_lipohyd_3A, and ringC_sp3N_2B seems as prominent aspects that deliver SARS-CoV 3CL^{pro} inhibition, concurrent to the cutting-edge investigation. Internal and external validation specification in the derived model have a high value. In addition, QSAR-based virtual screening yielded a compound with a lower pEC_{50} value of 5.88 nm and a higher pEC_{50} value of 6.08. Furthermore, molecular docking investigation of molecule 4 into the SARS-CoV 3CL^{pro} proclaim the key pharmacophoric moieties implicated in the binding interactions that are accountable for the inhibitory potential. The MD simulation and Molecular Docking evaluation divulge the imperative pharmacophoric centers like benzene ring, phenyl ring, amide oxygen and nitrogen etc. plays vital position in executing hydrogen bonding and hydrophobic interactions with the key amino acid residues namely; Ile164, Pro188, Leu190, Thr25, His41, Asn46, Thr47, Ser49, Asn189, Gln191, His41, Thr47, Asn141. In order to produce effective and selective SARS-CoV 3CL^{pro} inhibitors, QSAR and molecular docking yielded a consensus as well as complimentary pharmacophoric features, which should be kept in the future. Finally, the extraordinary high docking score of hit molecule 97 with SARS-CoV 3CL^{pro} explains the higher affinity and opens up new domain for a novel SARS-CoV 3CL^{pro} inhibitor drug.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

The authors extend their appreciation to the Deanship of Scientific Research at Imam Mohammad bin Saud Islamic University for funding this work through Research Group no-21-09-77. Authors Rahul D. Jawarkar and Magdi E.A. Zaki are thankful to Dr. Paola Gramatica for providing free copy of QSARINS-2.2.4's.

Appendix A. Supplementary material

- See supplementary material section 1.3.1 for explanation and calculation method of various statistical parameters.

- Section 1.3.3 Figure S1, S2, S3, S4, S5.
- Excel file contains smiles for dataset compound along with EC_{50} & pEC_{50} . Calculated descriptors for dataset molecule and Smiles notation for 100 in house library of compounds used in QSAR based Virtual Screening and their Predicted pEC_{50} values.

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.arabj.2021.103499>.

References

- Consonni, V., Todeschini, R., Ballabio, D., Grisoni, F., 2019. On the Misleading Use of Q_{2F3} for QSAR Model Comparison. *Molecular informatics* 38, (1–2). <https://doi.org/10.1002/minf.201800029>. In this issue e1800029.
- Davies, M., Nowotka, M., Papadatos, G., Dedman, N., Gaulton, A., Atkinson, F., Bellis, L., Overington, J.P., 2015. ChEMBL web services: streamlining access to drug discovery data and utilities. *Nucleic acids research* 43 (W1), W612–W620. <https://doi.org/10.1093/nar/gkv352>.
- Ferreira, L.G., Dos Santos, R.N., Oliva, G., Andricopulo, A.D., 2015. Molecular docking and structure-based drug design strategies. *Molecules (Basel, Switzerland)* 20 (7), 13384–13421. <https://doi.org/10.3390/molecules200713384>. In this issue.
- Fujita, T., Winkler, D.A., 2016. Understanding the Roles of the “Two QSARs”. *Journal of chemical information and modeling* 56 (2), 269–274. <https://doi.org/10.1021/acs.jcim.5b00229>. In this issue.
- Garro Martinez, J.C., Vega-Hissi, E.G., Andrada, M.F., Estrada, M. R., 2015. QSAR and 3D-QSAR studies applied to compounds with anticonvulsant activity. *Expert opinion on drug discovery* 10 (1), 37–51. <https://doi.org/10.1517/17460441.2015.968123>. In this issue.
- Gaudreault, F., Morency, L.P., Najmanovich, R.J., 2015. NRGsuite: a PyMOL plugin to perform docking simulations in real time using FlexAID. *Bioinformatics* 31 (23), 3856–3858. <https://doi.org/10.1093/bioinformatics/btv458>. In this issue.
- Gini, G., 2016. QSAR Methods. *Methods in molecular biology (Clifton, N.J.)* 1425, 1–20. https://doi.org/10.1007/978-1-4939-3609-0_1. In this issue.
- Gramatica, P., 2007. Principles of QSAR models validation: internal and external. *QSAR & combinatorial science* 26 (5), 694–701. <https://doi.org/10.1002/qsar.200610151>. In this issue.
- Gramatica, P., 2014. External Evaluation of QSAR Models, in Addition to Cross-Validation: Verification of Predictive Capability on Totally New Chemicals. *Molecular informatics* 33 (4), 311–314. <https://doi.org/10.1002/minf.201400030>. In this issue.
- Gramatica, P., 2020. Principles of QSAR modeling: comments and suggestions from personal experience. *International Journal of Quantitative Structure-Property Relationships (IJQSPR)* 5 (3), 61–97. <https://doi.org/10.4018/IJQSPR.20200701.0a1>. In this issue.
- Hong, H., Chen, M., Ng, H.W., Tong, W., 2016. QSAR models at the US fda/nctr. In: *In Silico Methods for Predicting Drug Toxicity*. Humana Press, New York, NY. doi, pp. 431–459. https://doi.org/10.1007/978-1-4939-3609-0_18.
- Huang, J., Fan, X., 2011. Why QSAR fails: an empirical evaluation using conventional computational approach. *Molecular pharmaceuticals* 8 (2), 600–608. <https://doi.org/10.1021/mp100423u>. In this issue.
- Jorgensen, W.L., Maxwell, D.S., Tirado-Rives, J., 1996. Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *Journal of the American Chemical Society* 118 (45), 11225–11236. In this issue.
- Kontoyianni, M., 2017. Docking and Virtual Screening in Drug Discovery. *Methods in molecular biology (Clifton, N.J.)* 1647, 255–266. https://doi.org/10.1007/978-1-4939-7201-2_18. In this issue.

- Konwar, M., Sarma, D., 2020. Advances in developing small molecule SARS 3CLpro inhibitors as potential remedy for corona virus infection. *Tetrahedron* 131761. <https://doi.org/10.1016/j.tet.2020.131761>. In this issue.
- Lai, M.M.C., Perlman, S., Anderson, J.L., 2006. Coronaviridae. In: Knipe, D.M., Howley, P.M. (Eds.), *Fields Virology*. 5th edition. Lippincott Williams & Wilkins, Philadelphia, pp. 1305–1335. In this issue.
- Lee, C.C., Kuo, C.J., Ko, T.P., Hsu, M.F., Tsui, Y.C., Chang, S.C., Yang, S., Chen, S.J., Chen, H.C., Hsu, M.C., Shih, S.R., Liang, P. H., Wang, A.H., 2009. Structural basis of inhibition specificities of 3C and 3C-like proteases by zinc-coordinating and peptidomimetic compounds. *The Journal of biological chemistry* 284 (12), 7646–7655. <https://doi.org/10.1074/jbc.M807947200>. In this issue.
- Linden, M., Greiff, L., Andersson, M., Svensson, C., Akerlund, A., Bende, M., Andersson, E., Persson, C.G., 1995. Nasal cytokines in common cold and allergic rhinitis. *Clinical and experimental allergy: journal of the British Society for Allergy and Clinical Immunology* 25 (2), 166–172. <https://doi.org/10.1111/j.1365-2222.1995.tb01022.x>. In this issue.
- Malik, Y.A., 2020. Properties of coronavirus and SARS-CoV-2. *The Malaysian journal of pathology* 42 (1), 3–11. In this issue.
- Martyna, G.J., Klein, M.L., Tuckerman, M., 1992. Nosé-Hoover chains: The canonical ensemble via continuous dynamics. *The Journal of chemical physics* 97 (4), 2635–2643. <https://doi.org/10.1063/1.463940>. In this issue.
- Masand, V.H., El-Sayed, N.N., Mahajan, D.T., Rastija, V., 2017. QSAR analysis for 6-arylpyrazine-2-carboxamides as Trypanosoma brucei inhibitors. *SAR and QSAR in environmental research* 28 (2), 165–177. <https://doi.org/10.1080/1062936X.2017.1292407>. In this issue.
- Masand, V.H., El-Sayed, N.N., Bambole, M.U., Quazi, S.A., 2018. Multiple QSAR models, pharmacophore pattern and molecular docking analysis for anticancer activity of α , β -unsaturated carbonyl-based compounds, oxime and oxime ether analogues. *Journal of Molecular Structure* 1157, 89–96. <https://doi.org/10.1016/j.molstruc.2017.12.045>. In this issue.
- Masand, V.H., El-Sayed, N.N., Mahajan, D.T., Mercader, A.G., Alafeefy, A.M., Shibi, I.G., 2017. QSAR modeling for anti-human African trypanosomiasis activity of substituted 2-Phenylimidazopyridines. *Journal of Molecular Structure* 1130, pp.711–718. <https://doi.org/10.1016/j.molstruc.2016.11.012>. In this issue.
- Masand, V.H., Mahajan, D.T., Nazeruddin, G.M., Hadda, T.B., Rastija, V., Alafeefy, A.M., 2015. Effect of information leakage and method of splitting (rational and random) on external predictive ability and behavior of different statistical parameters of QSAR model. *Medicinal Chemistry Research* 24 (3), 1241–1264. <https://doi.org/10.1007/s00044-014-1193-8>. In this issue.
- Masand, V.H., Mahajan, D.T., Maldhure, A.K., Rastija, V., 2016. Quantitative structure–activity relationships (QSARs) and pharmacophore modeling for human African trypanosomiasis (HAT) activity of pyridyl benzamides and 3-(oxazolo [4, 5-b] pyridin-2-yl) anilides. *Medicinal chemistry research* 25 (10), 2324–2334. <https://doi.org/10.1007/s00044-016-1664-1>. In this issue.



OPEN

Identification of potential edible mushroom as SARS-CoV-2 main protease inhibitor using rational drug designing approach

Debanjan Sen¹, Bimal Debnath², Pradip Debnath³, Sudhan Debnath⁴✉, Magdi E. A. Zaki⁵✉ & Vijay H. Masand⁶

Severe Acute Respiratory Syndrome Coronavirus-2 (SARS-CoV-2) is highly pathogenic to humans and has created health care threats worldwide. This urgent situation has focused the researchers worldwide towards the development of novel vaccine or small molecule therapeutics for SARS-CoV-2. Although several vaccines have already been discovered and are in use for the masses, no therapeutic medication has yet been approved by FDA for the treatment of COVID-19. Keeping this in view, in the present study, we have identified promising hits against the main protease (M^{pro}) of SARS-CoV-2 from edible mushrooms. Structure-based virtual screening (VS) of 2433 compounds derived from mushrooms was performed with M^{pro} protein (6LU7). Four promising hits, namely, Kynapcin-12 (M_78), Kynapcin-28 (M_82), Kynapcin-24 (M_83), and Neonambiterphenyls-A (M_366) were identified based on the result of docking, Lipinski's rule, 100 ns molecular dynamics (MD) simulation and MM/PBSA binding free energy calculations. Finally, the inhibitory properties of these hits were compared with three known inhibitors, baicalein (1), baicalin (2), and biflavonoid (3). Data indicated that M_78, M_82 and M_83 compounds present in edible mushroom *Polyozellus multiplex* were potent inhibitors of M^{pro}protein (6LU7). It could be concluded that edible mushroom *Polyozellus multiplex* has potential activity against SARS-CoV-2 infection and identified molecules could be further explored as therapeutic inhibitors against SARS-CoV-2.

Pandemic COVID-19 caused by SARS-CoV-2 virus has posed serious challenges to the research community, health workers, and government officials worldwide. Due to the rapid human-to-human contagious nature of SARS-CoV-2, the disease adversely affected 241.2537 million people, with 4.9116 million fatalities in 223 countries and territories around the world (report as of 17th October 2021; <https://www.worldometers.info/coronavirus/>). The disease drastically hit the economic growth worldwide and pushed millions of people towards unemployment. The SARS-CoV-2 was initially detected at the end of December 2019, in Wuhan City, characterized by an atypical pneumonia outbreak¹⁻³. The World Health Organization (WHO) declared COVID-19 as a global public health emergency of international concern and in March 2020, it was declared as a pandemic. Under this emergency, scientists all over the world have been working for the development of novel vaccines and drug molecules to prevent and treat the COVID-19 disease. Recently, several manufacturers such as Pfizer Biotech, AstraZeneca University of Oxford, Serum Institute of India Pvt. Ltd, Moderna Biotech, Sinopharm / BIBP have launched vaccines in the market to combat COVID-19. Unfortunately, no therapeutic medication has yet been approved by FDA for the treatment of this disease.

The human coronaviruses genome has several conserved structural proteins such as Spike (S) glycoprotein, envelope (E) protein, membrane (M) protein, and nucleocapsid (N) protein. It has at least four non-structural proteins (nsPs) such as- 3-chymotrypsin-like protease (3CL^{pro}) also known as Main protease (M^{pro}), papain-like

¹BCDA College of Pharmacy & Technology, Jessore Road South, Hridaypur, Kolkata, West Bengal 700127, India. ²Department of Forestry and Biodiversity, Tripura University, Suryamaninagar, Tripura 799022, India. ³Department of Chemistry, Majaraja Bir Bikram College, Agartala, Tripura 799004, India. ⁴Department of Chemistry, Netaji Subhash Mahavidyalaya, Udaipur, Tripura 799114, India. ⁵Department of Chemistry, Faculty of Science, Imam Mohammad Ibn Saud Islamic University, Riyadh 13318, Saudi Arabia. ⁶Department of Chemistry, Vidya Bharati Mahavidyalaya, Amravati, Maharashtra 444 602, India. ✉email: bcsdebnath@gmail.com; Mezaki@imamu.edu.sa

protease (PL^{pro}), helicase, and RNA-dependent RNA polymerase (RdRp)⁴. Protein sequence alignment analyses of SARS-CoV-2 indicated that catalytic sites of the four SARS-CoV-2 enzymes which could serve as antiviral targets are highly conserved and show a total of 79.9% genomic similarity with SARS-CoV⁵. This attribute could be utilized to understand and inhibit the replication cycle of SARS-CoV-2. The non-structural proteins (nsPs), 3CL^{pro} and PL^{pro}, which are two important proteases, play a crucial role in the viral replication process through the extensive proteolysis of two replicase polyproteins, pp1a and pp1ab into 16 non-structural proteins (nsP1–nsP16)⁶. These nsPs are assembled and form the replication-transcription complex which regulates various functions of virus replication viz. replication of the viral genome, sub-genomic RNA processing, and packaging of the new virion⁷. Interrupting any replication process would become a potential molecular target to develop therapeutics against coronavirus.

The urgent need for drugs to treat COVID-19 has led scientists to focus on protease inhibitors as potential drugs for the treatment of COVID-19 patients. In this regard, M^{pro} has been found to be highly sensitive, therefore, it has been considered as a key therapeutic target for the development of a drug against coronavirus^{8,9}. As a treatment strategy against COVID-19, a combination of anti-HIV protease drugs, lopinavir and ritonavir, was currently employed to treat the COVID-19 patients with mild and moderate infections^{10,11}. However, the patients' outcome treated with this combination suggested that the curative effect of these drugs is minimal with potentially toxic side effects that might be harmful to the patients¹². Some other repurposed drugs are also currently used, taking the advantage of drug safety, to treat the COVID-19 patients as a short-term and non-specific solution^{13–16}. Identifying bioactive compounds from the natural sources, which could inhibit SARS-CoV-2 main protease, has been considered as an alternative approach to combat COVID-19. In silico techniques provide promising preliminary evidence for drug discovery in a shorter span of period. Recently, several researchers have focused on identifying potential biomolecules active against SARS-CoV-2 from natural sources by implication of in silico drugs designing approach^{17–21}. This is because phytochemicals have been used as a good source of antiviral drugs in folk medicine to treat viral infections. Moreover, drug molecules identified from natural resources, especially plants have minimal side effects associated with them. Therefore, the development of more targeted inhibitors from natural sources could be an efficient therapeutic strategy to combat COVID-19.

Mushrooms are rich in low-calorie fibre, protein, health-boosting vitamins, and minerals. It is used as food due to its great taste and amazing health benefits worldwide. Mushrooms raised with exposure to ultraviolet light are a good source of Vitamin D²². There is a common belief that supplementation of mushrooms in dietary meals reduces the health care expenditure and remove the fear of the influenza outbreak²³. Recent pharmacological studies indicated that mushrooms are an exceptional source of several bioactive molecules, possessing antiviral^{24,25}, anti-inflammatory^{26,27}, antioxidants, antifungal, anticancer, antibacterial, and inhibition of platelet aggregation activities²⁸. Mushrooms exhibited strong anti-viral properties when used against influenza-A virus^{29–31}, Dengue virus serotype 2³², HIV-1, HIV-2^{33,34}, type-2 herpes simplex viruses³⁵, pandemic H1N1 and human H3N2³⁶. Many patients have continually suffered from inflammatory complications due to cytokine storms because of the elevated levels of ILs, IFN- γ , tumour necrosis factor α (TNF- α), interferon gamma-induced protein (IP10), and granulocyte colony-stimulating factor (GCSF)³⁷. The major life-threatening event associated with the COVID-19 infection is cytokine storm³⁷. More importantly, studies have shown that several edible mushrooms boost up immune responses by stimulating the immune effect or cells like cytotoxic T lymphocytes (TCL), T-cells, dendritic cells (DCs), natural killer cells, and macrophages, which further induced the expression and secretion of cytokines including interleukins (ILs) and interferon-gamma (INF- γ)^{38,39}. These exciting medicinal properties of mushrooms have led us to investigate their therapeutic potential against the COVID-19⁴⁰. The main aim of the present study was to identify potential edible mushrooms with compounds having a high binding affinity towards SARS-CoV-2 M^{pro}. Recently, Rangsinth et al. have carried out an in-silico study of mushroom compounds against the main protease of SARS-CoV-2⁴¹. However, they have investigated only 36 compounds that have been reported to possess anti-HIV protease properties. In our present study, we carried out a detailed investigation of 2433 mushroom compounds for their potential as SARS-CoV-2 main protease inhibitors using VS, MD simulation, Lipinski's rule, MM/PBSA binding free energy calculation, and comparison with known inhibitors. We observed that several phenolic compounds of mushrooms exhibited strong binding affinity with the main protease of SARS-CoV-2. Based on the results obtained, we believe that further in-vitro and in vivo studies of the reported compounds may provide more scientific information about the inhibitory properties of these mushrooms.

Materials and methods

Data collection and preparation target protein. In the present study, a databank of 2433 compounds was retrieved from different mushrooms available in the literature^{42–47} and mushroom compounds from the food databank (<https://foodb.ca/>). Structures of all compounds retrieved from the literature were drawn using ChemDraw Professional 15.1 and saved in the sdf format. After importing all the ligand files in the Maestro version 12.3 used under academic license, a single file was prepared. Then the prepared single-file was imported into the PyRx software tool. The UFF force field⁴⁸ was used to convert all the ligands in the pdbqt format, followed by energy minimization. The X-ray crystal structure of SARS-CoV-2 M^{pro} (PDB ID: 6LU7, resolution: 2.16 Å) was downloaded from the RCSB protein databank (<http://www.rcsb.org/>)^{49,50}. The previously prepared protein pdbqt file of M^{pro}¹⁷ was used for the docking purpose in both AutoDock Vina (ADV) in PyRx and AutoDock 4.2 (AD), both are open-source software. The top fourteen hits resulted from virtual screening were again re-docked using AutoDock 4.2⁵¹ software considering identical grid parameters.

Receptor grid generation, RMSD calculation, virtual screening, and molecular docking. The AutoDock Vina integrated with PyRx software^{52–54} was used to perform the virtual screening installed in a

Windows 10 Operating System supported by Intel i5 8600 K processor with 8 GB RAM. The grid dimension of the main protease was fixed by selecting active site amino acid residue information (HIS-41, MET-49, PHE-140, LEU-141, GLY-143, CYS-145, HIS-163, GLU-166, GLN-189). The grid centre coordinate of M^{Pro} were -10.88, 13.93, 68.21 along the X, Y and Z axis, respectively and grid size were 58, 68, 70 along X, Y and Z axis, respectively with grid spacing 0.375. The energy range was set at 4 and exhaustiveness was set at 8.0. For docking with protonated target, protonation was done by using H++ server version 3.2 (<http://biophysics.cs.vt.edu/H++>)⁵⁵ at pH = 6.5, internal dielectric = 10.0, external dielectric = 80.0 and salinity 0.15. The docking score of coligand (N3) was considered as the standard reference. Validation of the docking protocol is a crucial step before performing docking-based virtual screening. The docking protocol was validated by measuring the root mean square deviation (RMSD) using PyMOL 2.5. The compounds and co-ligand was prepared using the default parameters of PyRx. For calculation of RMSD, each docking poses of N3 generated during the docking program were superimposed on the native conformation of N3, using the "pair_fit" command in PyMOL software (<http://www.pymol.org>). The output compound and protein in pdbqt format were imported in PyMOL 2.5 for visual inspection of binding poses, followed by the export of the protein–ligand complex in PDB. The protein–ligand complexes imported in ProteinsPlus server (<https://proteins.plus>)^{56,57} and their 2D interactions were analyzed.

Drug-likeness properties prediction. Nearly 40% of the identified candidate drugs fail in the clinical trials due to the poor ADME properties⁵⁸. Therefore, prediction of the five physicochemical parameters such as molecular weight, number of H-bond acceptors, number of H-bond donors, molar refractivity, n-octanol/water partition coefficient, i.e., Lipinski's rule of five⁵⁹ of the selected hits was performed using publicly available online server SwissADME (<http://www.swissadme.ch>)⁶⁰.

Molecular dynamic simulation protocol. All-atom molecular dynamics simulation (MDS) of the selected hits was conducted by Gromacs 2018.1⁶¹ software supported by NVIDIA RTX 2070 GPU and Intel i7 990 k processor running over Linuxmint 19.3 Operating System (OS). The pdb2gmx program of the Gromacs 2018.1 package with Charmm36⁶² force field was used to prepare the protein topology. Topology for each ligand was obtained from the SwissParamTool⁶³, an online server-based parameterization program. After rejoining the protein and ligand topology, each system was solvated using TIP3P⁶⁴ water model into a (10Åx10Åx10Å) cubic box. Adequate numbers of Na⁺ and Cl⁻ ions (0.15 M) were added to neutralize each solvated protein–ligand system. The steepest descent algorithm⁶⁵ was used to minimize each system with a maximum of 50,000 steps, and the force was set to less than 10.0 kJole/mol. In the two-stage equilibration step, the 1st step is the NVT ensemble step in which the volume, temperature, and number of particles were kept constant and maintained for 2 ns. The 2nd step is the NPT ensemble step which has constant pressure along with equilibration of temperature and numbers of particles for 10 ns. For each equilibration step, 100 ns positional restraint of C_α atoms were applied. Free movement of the solvent molecules was allowed to maintain the solvent equilibrium. The linear constraint solver algorithm⁶⁶ was used to constrain the covalent bonds of the system. The particle mesh Ewald (PME)⁶⁷ method was applied for long-range electrostatic interaction setting cutoff of 1.2 nm and Fourier spacing of 1.2 nm. The V-rescale weak coupling method⁶⁸ was used to regulate the temperature (310.15 K) of the system. The Parrinello–Rahman method⁶⁹ was used to regulate 1 atm pressure, density, and total energy of the system. Each equilibrated system with acceptable geometry and solvent orientation was subjected to 100 ns production run without setting any restraint followed by a 2 fs time step. The structural coordinates were recorded in every 2 ps interval. After the successful completion of the MDS, water and ions were stripped out, followed by PBC correction to refine the trajectories. From the refined trajectories, various parameters like root mean square deviation (RMSD)⁷⁰, root mean square fluctuation (RMSF)⁷¹, the radius of gyration (Rg)⁷², and solvent accessible surface area (SASA)⁷³ that occurred in between ligand and protein were calculated considering co-crystal coordinates as a reference structure. The VMD1.9⁷⁴ program was used to visualize the trajectory and render images. Grace 5.1.25 software (<https://plasma-gate.weizmann.ac.il/Grace>) was used as a plotting program. The stability of MD complexes was also evaluated using the centre of mass (CoM) distance.

Molecular mechanics Poisson-Boltzmann surface area (MM/PBSA) calculation protocol. The ligand–protein binding interaction was quantitatively estimated by a widely acceptable Molecular Mechanics Poisson-Boltzmann Surface Area (MM/PBSA) approach⁷⁵. The g_mmpbsa script program⁷⁶, a high throughput MM/PBSA calculation tool in Gromacs software, was used to perform MM/PBSA based binding free energy (ΔG_{bind}) calculation [https://rashmikumari.github.io/g_mmpbsa/Tutorial.html]. The g_mmpbsa script program, along with the APBS 1.4 program⁷⁷ was used to calculate the above terms. To perform the calculations mentioned above, snapshots of the last 10 ns frames were extracted from the total trajectory using the gmx-trjconv command. The total 100 ns trajectory frames were supplied as an input for the g_mmpbsa program to calculate the binding energy.

Results and discussion

Molecular docking and Lipinski's rule analysis. The M^{Pro} binding site interactions were assigned by importing the protein–ligand crystal structure in the ProteinsPlus server. The detailed active site interacting amino acid residues included LEU-4, GLY-143, HIS-163, GLU-166, and GLN-189. Therefore, a receptor grid box was prepared by selecting these active site amino acid residues. The RMSD value between the coligand and docking pose of the same was 1.47Å¹⁷. This value was less than 2.0 Å; therefore, the validation of the docking protocol was regarded to be successful. This indicates that there is little visible difference between the docked pose of coligand to the original crystallographic bound ligand pose. In the first step, databases containing 2433 numbers of compounds derived from mushrooms were docked with a ligand-free M^{Pro} active site using

AutoDock Vina in PyRx. Amongst the top-scored hits extracted from the docking, fourteen hits with docking scores ≤ -7.0 kcal/mol were subjected to AD to eliminate the false positives. Lower the value of docking score (i.e. higher negative value) of a ligand indicates a higher binding affinity towards the target protein. Out of eight docked conformations generated during the docking with ADV of each ligand, their highest binding energy conformation was selected for 2D visualization of interactions (Fig. S1). In the crystal structure of SARS-CoV-2 M^{pro} (PDB ID: 6LU7), the co-ligand (N3) of SARS-CoV-2 M^{pro} was covalently bonded with CYS-145. This co-ligand was isolated from the protein–ligand crystal structure and re-docked into the active site of M^{pro}. The docking score of co-ligand was -7.2 kcal/mol and this value was used as a control value to reduce the chemical space after docking. The hits with the docking score of ≤ -7.2 were selected as SARS-CoV-2 M^{pro} inhibitors, source mushroom of the selected hits and trivial name of the hits have been summarized in Table 1. The docking score predicted by the ADV of top ten hits ≤ -8.0 , are M_78, M_82, M_83, M_88, M_111, M_112, M_201, M_366, M_421 and M_505. The structural insights into the binding interaction of these hits with SARS-CoV-2 M^{pro} were analyzed using ProteinsPlus are shown in Fig. S1. Zhang et al. showed that α -ketoamide inhibitor bound with the active site of M^{pro}, involving, HIS-41, HIS-164 and CYS-145 amino acid residues⁷⁸. Yoshino et al. also performed long-time molecular dynamics (MD) simulation of M^{pro} with three drug-like peptide candidates and identified the crucial active site amino acid residues HIS-41, GLY-143, and GLU-166 involved in the inhibition of M^{pro}⁷⁹. Along with docking score, interaction with crucial amino acid residues may be other important criteria in the selection of potential inhibitors. The ADV docking score of ten hits was less than the ADV docking score of known inhibitors 1 and was comparable to other known inhibitors 2 and 3 (Table S1). The hits M_78, M_82, M_83, and M_201 interacted with the two crucial amino acid residues viz. HIS-41, and GLU-166 along with the other interacting residues. The hit M_111 interacted with the three crucial amino acid residues viz. HIS-41, CYS-145, GLU-164, along with five more interactions. Along with the other interacting amino acid residues, the hits M_88 and M_112 interacted with the GLY-143, GLU-166, and CYS-145, GLU-166, respectively. The other hits M_366, M_421, and M_505 interacted with the crucial amino acid residues, HIS-41 or GLU-166. All the top hits also showed the AD docking score less than AD docking score of two known inhibitors 1 and 2 (Table 1). The docking results of the fourteen selected hits with protonated M^{pro} (Fig. S2) were very close to the docking score of hits with nonprotonated M^{pro}. The docking score with protonated M^{pro} have been shown in Table S2. Further, all the ten top hits were deeply inserted into the active site of M^{pro}. The 3D surface topology pose of hits M_78, M_82, M_83, M_88, M_111, M_112, M_201, M_366, M_421 and M_505 in the active site have been shown in Fig. S3. The binding pose of inhibitors M_78, M_82, M_83, M_88, M_111, M_112, and M_336 superimposed on the binding pose of known inhibitors 1 and 2. It was observed that the pharmacophore of most of the selected top-scored hits matched with the known inhibitors. A visual inspection of selected inhibitors superimposed on known inhibitors is shown in Fig. S4.

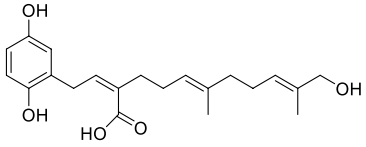
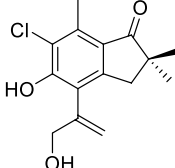
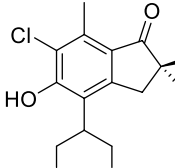
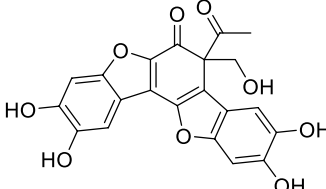
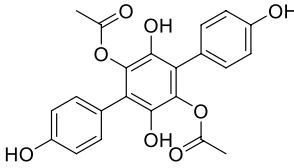
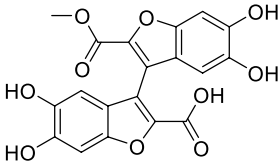
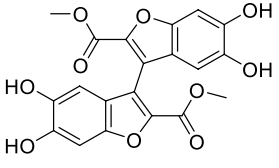
Most of the top ten hit exhibited close (< 3.0 Å) hydrogen bond donor and hydrogen bond acceptor interactions. These short-distance interactions revealed the efficient binding affinity of the hits towards M^{pro}. The detailed view of the 3D interaction often hits with SARS-CoV-2 M^{pro} active site amino acid residues have been depicted in Fig. 1. The molecular weights of all the top hits were in the range of 340.33–540.60. The acceptable range of molecular weight is ≤ 500 . Therefore, the molecular weight of all the hits, except M_421 and M_505 were in the acceptable range. The number of hydrogen bond acceptors of all the hits was also in the range of 6–10. The acceptable range of this is ≤ 10 . Therefore, the numbers of hydrogen bond acceptor values of all the selected hits were in the acceptable range. The range of hydrogen bond donor groups of all the selected hits was 4–5, and the acceptable range is ≤ 5 ; this property was also in the acceptable range.

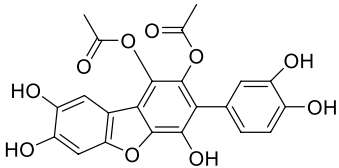
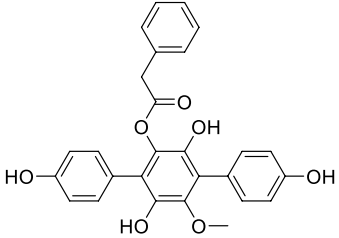
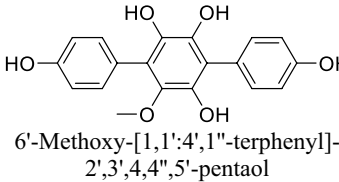
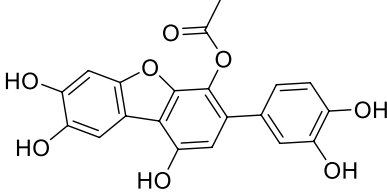
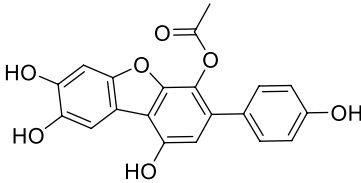
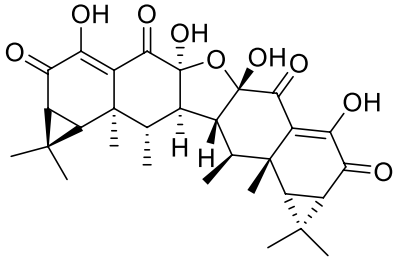
For a drug-like molecule, the molar refractivity should be between 40 and 130, and here all the hits were in this range except M_421 and M_505. The *n*-octanol/water partition coefficient of all the hits should be ≤ 5 ; this value for all the selected hits was in the acceptable range in the present study. Therefore, the Lipinski rule of five dealing with 90% of the orally active drugs that have achieved phase II clinical status was obeyed by the hits M_78, M_82, M_83, M_88, M_111, M_112, M_201, M_366. Among the mushroom compounds based on docking score and interaction with amino acid residues, and drug-like characters M_78, M_82, M_83, M_88, M_111, M_112, M_201, M_366 were selected for the MD simulation study. The physicochemical parameters of the top ten proposed inhibitors have been shown in Table 2.

Molecular dynamics simulation properties analysis. The application of the molecular dynamics simulation is a widely accepted approach for predicting the protein–ligand complex's stability. The 100 ns atomistic MD simulation was performed to explore the dynamics property of each identified protein–ligand complex and was compared with the dynamic behavior of the ligand-free protein (LFP) co-crystalline inhibitor bound protein. The average values of every parameter calculated from molecular dynamics (MD) trajectories has been depicted in Table 3.

The acceptable average RMSD value for globular protein is ≤ 3.0 Å⁸⁰. Nonetheless, lower RMSD values are ideally acceptable. It was observed that the protein backbones average RMSD of the selected M^{pro}–ligand (M_78, M_82, M_83, M_111, M_112, M_201, M_366) were less than 3.0 Å (Table 3). The average RMSD value exhibited by M_82, M_83, M_112, M_201, and M_366–protein systems was lower than the RMSD value of apo-protein. The average RMSD of M_366–protein system was found to be ~ 1.98 Å, which was even lower than the standard inhibitor baicalein (~ 2.15 Å)⁸¹. The RMSD profile of the protein–M_88 system was > 3.0 Å. To compare the RMSD of the protein backbone profile of protein–ligand (M_78, M_82, M_83, M_366) system with apo-protein, the RMSD was plotted against time from the 100 ns molecular dynamics trajectories as shown in Fig. 2. Similarly, the RMSD profiles of the protein–ligand (M_88, M_111, M_112, M_201) system has been depicted in Fig. S5.

To analyze the fluctuation of the individual amino acid residues, the RMSF parameter was calculated for each protein–ligand complex system from the 100 ns molecular dynamics trajectories. Lower the RMSF value

Sl no	Compound source	Structure of the selected hits and trivial name	*ADV score kcal/mol	**BSI	#AD score kcal/mol	Ki
1	M_01 <i>Ganoderma pfeifferi</i> Bres Family: Ganodermataceae	 Ganomycin A	- 7.5	H-bonding: ASN-142, HIS-163, THR-190; Hydrophobic: MET-165, GLN-189, ARG-188	- 6.47	18.02 μM
2	M_60 <i>Lepista sordida</i> (Schumacher), Edible mushroom, Family: Tricholomataceae	 Lepistatins A	- 7.0	H-bonding: GLU-166, ARG-188 Hydrophobic: MET-165, GLN-189	- 7.38;	3.87 μM
3	M_62 <i>Lepista sordida</i> (Schumacher) Edible mushroom, Family: Tricholomataceae	 Lepistatins C	- 7.0	H-bonding: GLU-166, ARG-188 Hydrophobic: MET-165, GLN-189	- 7.37	3.97 μM
4	M_77 <i>Polyozellus multiplex</i> (Underw.) Murrill Korean edible mushroom, Family: Thelephoraceae	 Benzofurans - kynapcin-9	- 7.4	H-bonding: GLU-166, ASP-187; Hydrophobic: MET-49, MET-165	- 8.03	1.29 μM
5	M_78 <i>Polyozellus multiplex</i> (Underw.) Murrill, Korean edible mushroom, Family: Thelephoraceae	 Benzofurans - kynapcin-12	- 8.1	H-bonding: PHE-140, GLU-166, ASP-187; Hydrophobic: HIS-41, GLU-166	- 7.77	2.02 μM
6	M_82 <i>Polyozellus multiplex</i> (Underw.) Murrill Korean edible mushroom, Family: Thelephoraceae	 Benzofurans-kynapcin-28	- 8.5	H-bonding: LEU-141, SER-144, GLU-166, ASP-187; Hydrophobic: HIS-41, MET-165	- 7.54	3.72 μM
7	M_83 <i>Polyozellus multiplex</i> (Underw.) Murrill Korean edible mushroom, Family: Thelephoraceae	 Benzofurans-kynapcin-24	- 8.3	H-bonding: LEU-141, SER-144, GLU-166, ASP-87; Hydrophobic: HIS-41, MET-165	- 8.58	512.25 nM
Continued						

Sl no	Compound source	Structure of the selected hits and trivial name	*ADV score kcal/mol	**BSI	#AD score kcal/mol	Ki
8	M_88 <i>Sarcodonium bricatus</i> (L.) P. Karst, Edible mushroom, Family: Bankeraceae	 p-hydroxybenzoic acid	- 8.0	H-bonding: GLU-166, ASP-187, GLY-143	- 9.08	219.43 nM
9	M_111 <i>Thelephora ganbajun</i> M. Zang, Edible mushroom, Family: Thelephoraceae	 Ganbajunin G	- 8.8	H-bonding: PHE-140, CYS-145, HIS-164, ASP-187; Hydrophobic: HIS-41, MET-49, MET-165, GLN-189	- 10.2	33.64 nM
10	M_112 <i>Thelephora ganbajun</i> M. Zang, Edible mushroom, Family: Thelephoraceae	 6'-Methoxy-[1,1':4',1''-terphenyl]-2',3',4,4'',5'-pentaol	- 8.0	H-bonding: PHE-140, CYS-145, GLU-166, ASP-187;	- 7.52	3.85 μM
11	M_201 <i>Neonothopanus nambi</i> (bioluminescent Mushroom)	 Neonambiterphenyls B	-8.5	H-bonding: CYS-44, GLU-166 Hydrophobic: HIS-41, MET-165, ARG-188	-8.00	1.21 μM
12	M_366 <i>Neonothopanus nambi</i> (bioluminescent Mushroom)	 Neonambiterphenyls A	- 8.3	H-bonding: CYS-44 Hydrophobic: HIS-41, MET-165, ARG-188	- 8.73	401.89 nM
13	M_421 <i>Neonothopanus nambi</i> (bioluminescent Mushroom)	 aurisin A	- 9.3	H-bonding: GLU-166	- 8.45	638.15 nM
Continued						

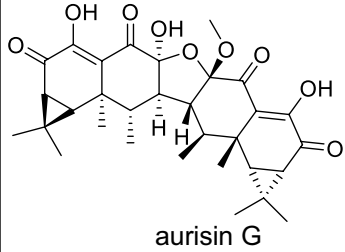
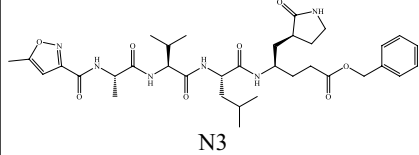
Sl no	Compound source	Structure of the selected hits and trivial name	*ADV score kcal/mol	**BSI	#AD score kcal/mol	Ki
14	M_505 <i>Neonothopanus nambi</i> (bioluminescent Mushroom)	 aurisin G	- 8.8	H-bonding: HIS-41	- 8.92	290.54 nM
15	Co-ligand N3	 N3	- 7.2	H-bonding: ALA-2, GLY-143, HIS-163, GLU-166; Hydrophobic: LEU-4, MET-49, GLN-189	- 6.63	13.71 μM

Table 1. Source mushroom of the compounds, the structure of the top fourteen hits, trivial name of the compounds, docking score (*ADV = Auto Dock Vina, **BSI = Binding Site Interaction, #AD = Auto Dock 4.2), interacting active site residues.

generalizes that after binding with the ligand, the fluctuation of the amino acid residues under consideration is reduced. These facts infer stable protein ligands binding, i.e., after binding with ligand, the fluctuation of residues is minimized. The amino acid residues THR-24 to GLN-192 lies in the binding site region of this protein. Amongst the residues, the HIS-41 and CYS-145 are the important residues that regulate the functionality of this protein⁷⁸. The average RMSF value of each protein–ligand system has been mentioned in Table 3. The plots of protein–ligand (M_78, M_82, M_83, M_366) system amino acid residues and RMSF has been shown in Fig. 3. Similarly, the plots of protein–ligand (M_88, M_111, M_112, M_201) system residue vs RMSF have been depicted in Fig. S6.

It can be ascertained from Table 3 that each system other than M_88 and M_201 bound system, exhibited lower RMSF values in comparison to the apo-protein RMSF. In contrast, M_111 system depicted a slightly higher RMSF (~1.38 Å) value. Each residue of M_88 and M_201 system showed significantly higher RMSF profiles (Fig. S6). Visual inspection of the trajectory concluded unacceptable changes taking place in the M_201 bound system during the 100 ns simulation time. The radius of gyration parameter further confirmed that event.

The parameter radius of gyration (Rg) furnishes information about the compactness of the protein. The higher value of Rg indicates that the protein changes its conformation of distortion during the simulation. The average Rg value of the protein–ligand (M_78, M_82, M_83, M_112, M_201, M_366) system was lower than the Rg value of apo-protein (Table 3). The average Rg value of the protein–ligand (M_78, M_82, M_83, M_201) system was lower than the Rg value of the standard inhibitor (22.44 Å). The Rg profile of the hits M_78, M_82, M_83, and M_366 have been depicted in Fig. 4 and Rg profile of M_88, M_111, M_112, M_201 has been shown in Fig. S7.

The protein–ligand system's solvent-accessible surface area parameter (SASA) was calculated for each protein–ligand system and plotted against the time in nanosecond. Considering that ligand binding is a solvent replacement process, the lower values of the SASA parameter indicate that the binding pocket is less solvent-exposed, and the ligand retains inside the binding pocket during the simulation. The average SASA value of the protein–ligand (M_78, M_82, M_83, M_201, M_366) system was lower than the SASA value of apo-protein (Table 3). The average SASA value of the protein–ligand (M_82, M_201) system was lower than the SASA value of the standard inhibitor (1472 Å²). The changes in SASA of M^{Pro}–ligand (M_78, M_82, M_83, M_366) systems and apo-protein during 100 ns simulation time have been shown in Fig. 5. Similarly, the changes of SASA value of M_88, M_111, M_112, M_201 protein system have been shown in Fig. S8. Each protein–ligand system, other than a protein–ligand system of M_111, and M_112, showed a lower average SASA value in comparison to the apo-proteins average SASA value (Table 3). Nevertheless, the M_201 system exhibited a SASA value of ~1456.0 Å² (lower than the apo-protein SASA value) which indicated that the ligand M_201 resided deeply into the binding pocket. However, this binding caused significant changes in the protein structure.

It was observed from Fig. S9 that the center of mass (CoM) distance for all the ligands resided within ~5.0 Å distance from the main protease binding site. The hits M_78 and M_82 consistently depicted ~2.0 Å distance from the protein binding pocket throughout the simulation. Ligand M_83 after ~85th ns showed slightly higher displacement (~1.0 Å). After ~95 ns, the protein binding pocket M_83 distance reduced to ~2.2 Å. The M_366 after ~36th ns showed a higher distance (~4.1 Å) from the protein binding pocket. However, after ~90 ns the distance reduced and stabilized at ~2.8 Å.

Binding free energy (ΔG_{bind}). The MM/PBSA based binding free energy (ΔG_{bind}) was calculated from the total 100 ns of the molecular dynamics trajectories. The ΔG_{bind} of the standard inhibitor was -180.50 kJ/mol, and the ΔG_{bind} of M_78, M_201, and M_366 were -193.55, -193.50 -190.10 kJ/mol, respectively, greater than the ΔG_{bind} of the standard inhibitor (Table 3). The ΔG_{bind} of M_82 was -180.10 kJ/mol, which was nearly the

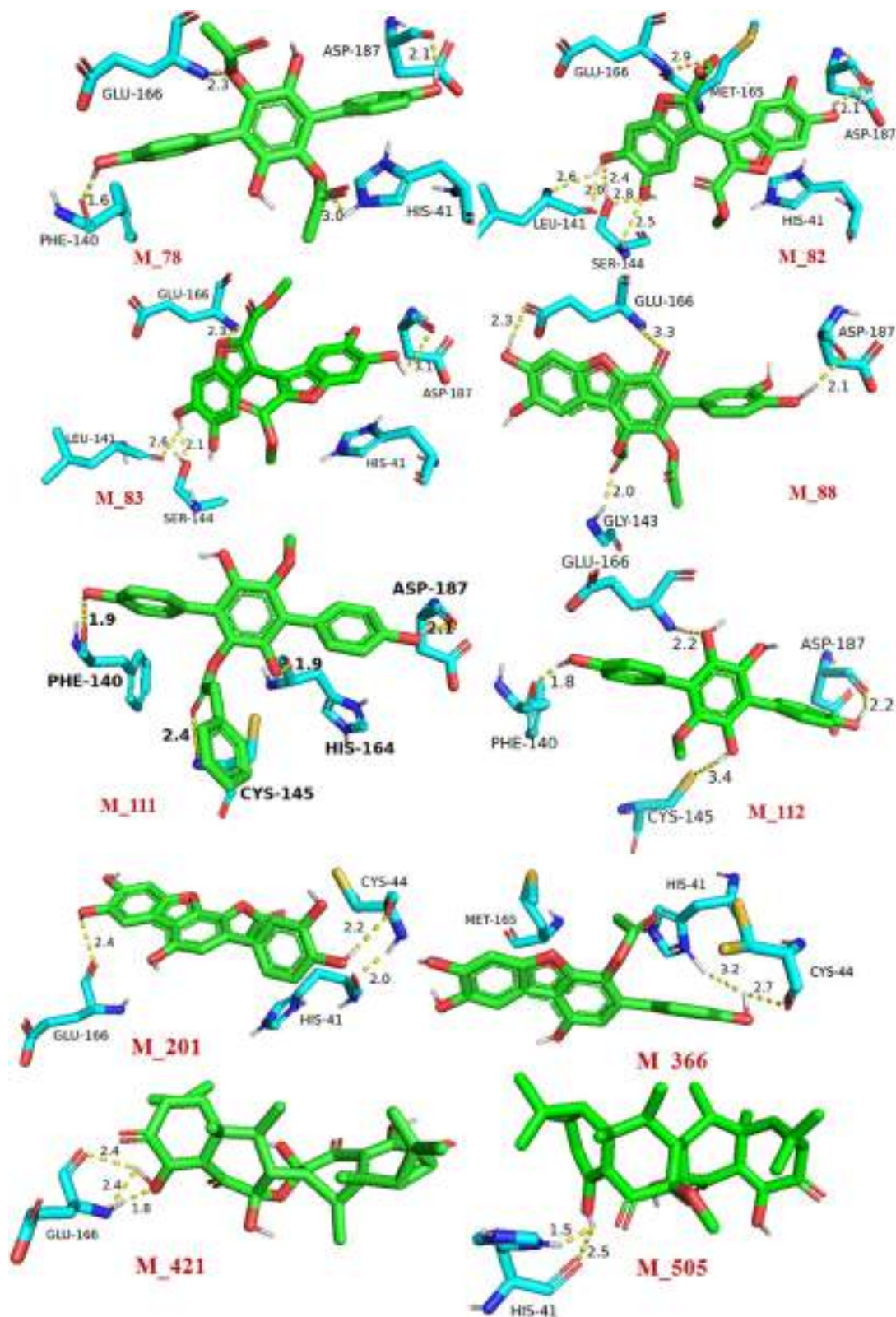


Figure 1. Detailed view of 3D interactions of selected hits (green stick) with SARS-CoV-2 M^{pro} active site amino acid residues (cyan stick) and their interacting distances. The hydrogen-bonding interactions were depicted in a yellow dotted line.

Parameters	M_78	M_82	M_83	M_88	M_111	M_112	M_201	M_366	M_421	M_505
MW	410.37	399.28	414.32	439.35	458.46	340.33	382.32	366.32	540.60	554.63
Acceptable Range	≤ 500	≤ 500	≤ 500	≤ 500	≤ 500	≤ 500	≤ 500	≤ 500	≤ 500	≤ 500
NHBA	8	10	10	10	7	6	8	7	9	9
Acceptable Range	≤ 10	≤ 10	≤ 10	≤ 10	≤ 10	≤ 10	≤ 10	≤ 10	≤ 10	≤ 10
NHBD	4	4	4	4	4	5	5	4	4	3
Acceptable Range	≤ 5	≤ 5	≤ 5	≤ 5	≤ 5	≤ 5	≤ 5	≤ 5	≤ 5	≤ 5
MR	108.40	95.81	102.07	110.38	127.88	93.92	100.77	98.75	136.97	141.70
Acceptable Range	40–130	40–130	40–130	40–130	40–130	40–130	40–130	40–130	40–130	40–130
iLOGp	2.39	1.48	2.37	2.26	2.72	2.15	1.74	1.83	3.811	2.93
Acceptable Range	≤ 5	≤ 5	≤ 5	≤ 5	≤ 5	≤ 5	≤ 5	≤ 5	≤ 5	≤ 5

Table 2. Physicochemical parameters of selected SARS-CoV-2 M^{Pro} proposed inhibitors from mushrooms. The parameters MW molecular weight, NHBA number of H-bond acceptors, NHBD number of H-bond donors, MR molar refractivity, iLOGp n-Octanol/Water partition coefficient).

Sl no	Compound ID	Average RMSD (Å)	Average RMSF (Å)	Average Rg (Å)	Average SASA (Å ²)	Binding energy (kJ/mol)
1	Apo protein	2.38	1.34	22.51	1508.6	–
2	Baicalein*	–	–	–	–	–
3	M_78	2.42	1.05	22.43	1483.26	– 193.55 ± 4.8
4	M_82	2.20	1.06	22.42	1460.12	– 180.10 ± 2.6
5	M_83	2.25	1.14	22.16	1493.40	– 174.73 ± 4.4
6	M_88	3.18	1.46	22.75	1520.00	– 177.73 ± 6.2
7	M_111	2.60	1.38	22.56	1518.49	– 147.71 ± 89
8	M_112	1.50	1.07	22.45	1517.50	– 146.60 ± 2.9
9	M_201	2.52	3.13	22.40	1456.00	– 153.50 ± 7.69
10	M_366	1.98	1.18	22.50	1486.14	– 190.46 ± 0.18

Table 3. Average values of Root Mean Square Deviation (RMSD), Root Mean Square Fluctuation (RMSF), Radius of Gyration (Rg), Solvent Accessible Surface Area (SASA), MM/PBSA based binding free energy calculated from 100 ns molecular dynamics trajectories. *Data available in Supplementary⁸¹.

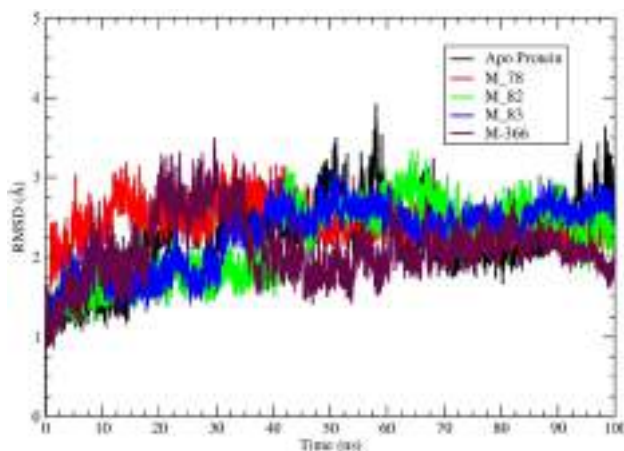


Figure 2. A visual representation of M^{Pro} backbone RMSD (Å) of M^{Pro}–ligand (M_78, M_82, M_83, M_366) complexes and apo-protein obtained from 100 ns MD simulation trajectories. Different ligands represented by different colours.

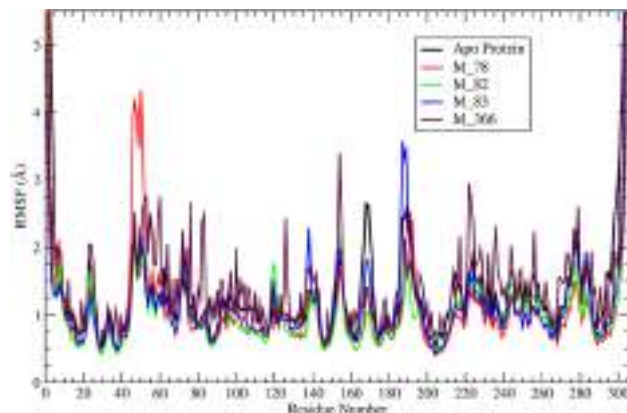


Figure 3. A visual representation of M^{Pro} backbone RMSF vs residue number of M^{Pro} –ligand (M_78, M_82, M_83, M_366) systems and apo-protein during 100 ns simulation.

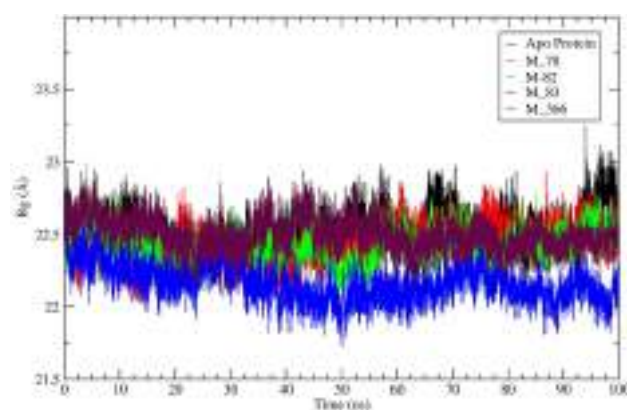


Figure 4. A visual representation of radius of gyration (Rg) vs time of M^{Pro} –ligand (M_78, M_82, M_83, M_366) systems and apo-protein during 100 ns MD simulation.

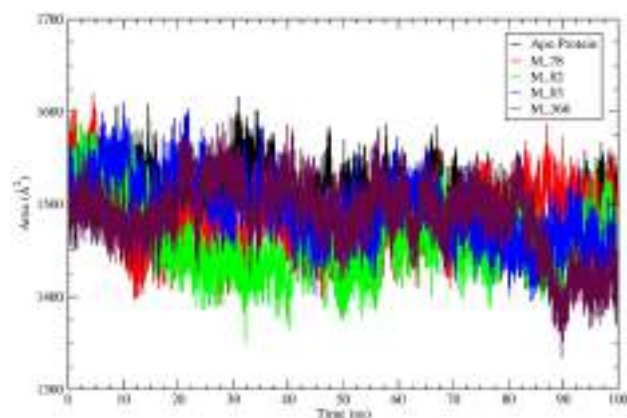


Figure 5. A visual representation of solvent accessible surface area (SASA) of M^{Pro} –ligand (M_78, M_82, M_83, M_366) systems and apo-protein during 100 ns simulation time.

same as the ΔG_{bind} value of the standard inhibitor. The ΔG_{bind} value of M_112 and M_111 was low in comparison to the ΔG_{bind} value of the standard inhibitor. Finally, the RMSD value of the hit M_88 was greater than the acceptable range and therefore was not considered as a promising hit. Due to the low ΔG_{bind} value of the hits, M_111 and M_112 which were -147.71 and -146.60 kJ/mol and greater SASA values in comparison to the apo-protein, the hits M_111 and M_112 were excluded from the list of promising hits. The per-frame binding

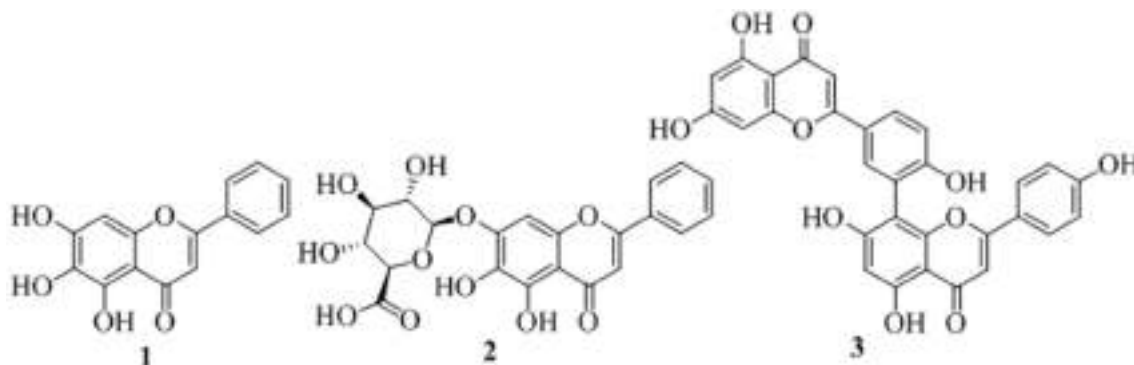


Figure 6. Structures of three known SARS-CoV-2 flavone inhibitors **1** (baicalein), **2** (baicalin) and **3** (biflavonoids).

energy over the simulated time of M_78, M_82, M_83, and M_366 are depicted in Fig. S10. The RMSD value of hit M_201 was comparatively high (2.52 Å), and its ΔG_{bind} was also comparatively low and therefore excluded from the list of promising hits. From the analysis of various parameters like RMSD, RMSE, Rg, SASA, and MM/PBSA calculated from the MD trajectories, it can be stated that the ligand M_78, M_82, M_83, and M_366 might have the potential to form a stable complex with SARS-CoV-2M^{Pro}.

There have been a large report of antiviral activities of mushroom compounds (Table S3) and therefore identification of SARS-CoV-2 M^{Pro} inhibitors from reported mushroom compounds is of great interest. All the proposed compounds are novel, structurally diverse, and contain phenolic functional groups. Polyphenolic compounds perform a series of defensive activities in the human body. One of the important classes of phenolic compounds are flavonoids, which showed blocking potential against different viral proteins like M^{Pro}, PL^{Pro}, Spike against SARS-CoV and MARS-CoV⁸². Two flavones found in different plants, baicalein (**1**) and baicalin (**2**) (Fig. 6), exhibited in vitro inhibition against SARS-CoV-2 M^{Pro} and IC₅₀ values, of which were 6.41 and 0.94 μM , respectively⁸³. Another biflavonoids (**3**) found in *Torreyanucifera* showed inhibitory activity against SARS-CoVM^{Pro} and the IC₅₀ value of which was 8.3 μM ⁸⁴.

The ADV docking scores of the proposed hits were similar to the ADV docking score of the known inhibitors **1**, **2**, and **3** and the docking scores were -7.7 , -8.6 , and -8.7 kcal/mol, respectively (Table S1). The best docking poses of the proposed hits M_78, M_82, M_83, and M_366 in the active site superimposed on the two SARS-CoV-2 known inhibitors **1** and **2** have been shown in Fig. S3. It was observed that the pyran ring, chromen-4-one-oxygen atom, pyran-1-oxygen atom of inhibitors **1** matched with the middle ring, an acetylated-phenolic-oxygen atom of the middle ring, and phenolic-OH of the middle ring, respectively, of hit M_78. Two rings of one of the benzofuran moiety; 5,6-dihydroxy group of one of the benzofuran moiety, furan oxygen of benzofuran moiety of hit M_82 matched with two rings of inhibitor **1**; a 6,7-dihydroxy group of **1**, and pyran-1-oxygen of **1**, respectively. The matching of the pharmacophore of hit M_83 was very close to M_82. The pyran ring, 6-hydroxyl group, and pyran-1-oxygen of compound **1** matched with one of the aromatic rings of dibenzofuran, 1-hydroxy group of dibenzofuran, and furan oxygen, respectively with M_366. The matching of the pharmacophores of the proposed hits in the most stable binding pose with standard inhibitor baicalein was again a piece of evidence in the favour of the potential of the selected hits, against SARS-CoV-2 M^{Pro}.

Conclusion

Based on the detailed in-silico studies, it can be concluded that compounds M_78, M_82, M_83, and M_366 showed docking scores greater than co-ligand and comparable docking scores with the two known SARS-CoV-2 inhibitors. The binding affinity of the hits M_78, M_82, M_83, M_366 against M^{Pro} protein of SARS-CoV-2 was also supported by all the MD parameters such as RMSD, RMSE, Rg, SASA, MM-PBSA binding energy (ΔG_{bind}). Moreover, all the proposed hits obeyed the Lipinski rule of five and matches pharmacophores with known inhibitors in the active site. The promising hits Kynapcin-12 (M_78), Kynapcin-28 (M_82), Kynapcin-24 (M_83) are available in edible mushroom *Polyozellus multiplex* and another promising hit Neonambiterphenyls-A (M_366) is available in the poisonous mushroom *Neonothopanus nimbi*. The mushroom *P. multiplex* contains three potential hits which might be used as a remedy against COVID-19 after the appropriate biological screening. These novel phenolic scaffolds may be further developed as more potential SARS-CoV-2 inhibitors.

Received: 30 July 2021; Accepted: 3 January 2022

© The Author(s) 2022

References

- Huang, C. *et al.* Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet* **395**, 497–506. [https://doi.org/10.1016/S0140-6736\(20\)30183-5](https://doi.org/10.1016/S0140-6736(20)30183-5) (2020).
- Wu, F. *et al.* A new coronavirus associated with human respiratory disease in China. *Nature* **579**, 265–269. <https://doi.org/10.1038/s41586-020-2008-3> (2020).
- Israeli, E. Novel coronavirus that recently emerged in China. *Harefuah* **159**, 70–71 (2020).

4. Pillaiyar, T., Meenakshisundaram, S. & Manickam, M. Recent discovery and development of inhibitors targeting coronaviruses. *Drug Discov. Today* **25**, 668–688. <https://doi.org/10.1016/j.drudis.2020.01.015> (2020).
5. Zhou, P. *et al.* A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* **579**, 270–273. <https://doi.org/10.1038/s41586-020-2012-7> (2020).
6. Báez-Santos, Y. M., St John, S. E. & Mesecar, A. D. The SARS-coronavirus papain-like protease: Structure, function and inhibition by designed antiviral compounds. *Antiviral Res.* **115**, 21–38. <https://doi.org/10.1016/j.antiviral.2014.12.015> (2015).
7. Fehr, A. R., Jankevicius, G., Ahel, I. & Perlman, S. Viral macrodomains: Unique mediators of viral replication and pathogenesis. *Trends Microbiol.* **26**, 598–610. <https://doi.org/10.1016/j.tim.2017.11.011> (2018).
8. McKee, D. L. *et al.* Candidate drugs against SARS-CoV-2 and COVID-19. *Pharmacol Res.* **157**, 104859. <https://doi.org/10.1016/j.phrs.2020.104859> (2020).
9. Zhang, L. *et al.* Crystal structure of SARS-CoV-2 main protease provides a basis for design of improved α -ketoamide inhibitors. *Science* **368**, 409–412. <https://doi.org/10.1126/science.abb3405> (2020).
10. Ye, X. T. *et al.* Clinical efficacy of lopinavir/ritonavir in the treatment of Coronavirus disease 2019. *Eur. Rev. Med. Pharmacol. Sci.* **24**, 3390–3396. https://doi.org/10.26355/eurrev_202003_20706 (2020).
11. Nutho, B. *et al.* Why are lopinavir and ritonavir effective against the newly emerged coronavirus 2019? Atomistic insights into the inhibitory mechanisms. *Biochemistry* **59**, 1769–1779. <https://doi.org/10.1021/acs.biochem.0c00160> (2020).
12. Cao, B. *et al.* A trial of lopinavir-ritonavirs in adults hospitalized with severe COVID-19. *N. Engl. J. Med.* **382**, 1787–1799. <https://doi.org/10.1056/NEJMoa2001282> (2020).
13. Wang, M. *et al.* Remdesivir and chloroquine effectively inhibit the recently emerged novel coronavirus (2019-nCoV) in vitro. *Cell Res.* **30**, 269–271. <https://doi.org/10.1038/s41422-020-0282-0> (2020).
14. Li, G. & De Clercq, E. Therapeutic options for the 2019 novel coronavirus (2019-nCoV). *Nat. Rev. Drug Discov.* **19**, 149–150. <https://doi.org/10.1038/d41573-020-00016-0> (2020).
15. Gao, J., Tian, Z. & Yang, X. Breakthrough: Chloroquine phosphate has shown apparent efficacy in treatment of COVID-19 associated pneumonia in clinical studies. *Biosci. Trends.* **14**, 72–73. <https://doi.org/10.5582/bst.2020.01047> (2020).
16. Colson, P. *et al.* Chloroquine and hydroxychloroquine as available weapons to fight COVID-19. *Int. J. Antimicrob. Agents.* **55**, 105932. <https://doi.org/10.1016/j.ijantimicag.2020.105932> (2020).
17. Sen, D. *et al.* Identification of potential inhibitors of SARS-CoV-2 main protease and spike receptor from 10 important spices through structure-based virtual screening and molecular dynamic study. *J. Biomol. Struct. Dyn.* **18**, 1–22. <https://doi.org/10.1080/07391102.2020.1819883> (2020).
18. Hu, X. *et al.* Possible SARS-Coronavirus-2 inhibitor revealed by simulated molecular docking to viral main protease and host toll-like receptor. *Future Virol.* **18**, 1746–1808. <https://doi.org/10.2217/fvl-2020-0099> (2020).
19. Ibrahim, M. A. A. *et al.* In silico drug discovery of major metabolites from spices as SARS-CoV-2 main protease inhibitors. *Comput. Biol. Med.* **126**, 104046. <https://doi.org/10.1016/j.compbiomed> (2020).
20. Ibrahim, M. A. A. *et al.* Rutin and flavone analogs as prospective SARS-CoV-2 main protease inhibitors: In silico drug discovery study. *J. Mol. Graph Model.* **105**, 107904. <https://doi.org/10.1016/j.jmgm.2021.107904> (2021).
21. Ibrahim, M. A. A. *et al.* Natural-like products as potential SARS-CoV-2 M^{pro} inhibitors: In-silico drug discovery. *J. Biomol. Struct. Dyn.* **39**, 5722–5734. <https://doi.org/10.1080/07391102.2020.1790037> (2021).
22. Cardwell, G., Bornman, J. F., James, A. P. & Black, L. J. A review of Mushrooms as a potential source of dietary vitamin D. *Nutrients* **10**, 1498. <https://doi.org/10.3390/nu10101498> (2018).
23. Chunchao, H. & Guo, J. Y. A hypothesis: Supplementation with mushroom-derived active compound modulates immunity and increases survival in response to influenza virus (H1N1) infection. *Evid. Based Complement. Altern. Med.* **2011**, 252501. <https://doi.org/10.1093/ecam/nej037> (2011).
24. Linnakoski, R. *et al.* Antiviral agents from fungi: Diversity, mechanisms and potential applications. *Front. Microbiol.* **9**, 2325. <https://doi.org/10.3389/fmicb.2018.02325> (2018).
25. Suwannarach, N. *et al.* Natural bioactive compounds from fungi as potential candidates for protease inhibitors and immunomodulators to apply for coronaviruses. *Molecules* **25**, 1800. <https://doi.org/10.3390/molecules25081800> (2020).
26. Elsayed, E. A., El Enshasy, H., Wadaan, M. A. & Aziz, R. Mushrooms: A potential natural source of anti-inflammatory compounds for medical applications. *Mediat. Inflamm.* **2014**, 805841. <https://doi.org/10.1155/2014/805841> (2014).
27. Muszynska, B., Grzywacz-Kisielewska, A., Kała, K. & Gdula-Argasinska, J. Anti-inflammatory properties of edible mushrooms: A review. *Food Chem.* **243**, 373–381. <https://doi.org/10.1016/j.foodchem.2017.09.149> (2018).
28. Badalyan, S. M. & Rapior, S. The neurotrophic and neuroprotective potential of macrofungi. In *Medicinal Herbs and Fungi* (eds Agrawal, D. C. & Dhanasekaran, M.) (Springer, 2021).
29. Nishihira, J. *et al.* Maitake mushrooms (*Grifolafrondosa*) enhances antibody production in response to influenza vaccination in healthy adult volunteers concurrent with alleviation of common cold symptoms. *Funct. Foods Health Dis.* **7**, 462–482. <https://doi.org/10.31989/fhd.v7i7.363> (2017).
30. Obi, N. *et al.* Inhibitory effect of TNF- α produced by macrophages stimulated with grifolafrondosa extract (ME) on the growth of influenza A/Aichi/2/68 virus in MDCK cells. *Am. J. Chin. Med.* **36**, 1171–1183. <https://doi.org/10.1142/S0192415X08006508> (2008).
31. Adhikari, M. Some Antiviral Mushrooms of Nepal. *Nepal J. Sci. Technol.* **19**, 54–59. <https://doi.org/10.3126/njst.v19i1.29767> (2020).
32. Ellan, K. *et al.* Anti-viral activity of culinary and medicinal mushroom extracts against dengue virus serotype 2: An in-vitro study. *BMC Complement Altern. Med.* **19**, 260. <https://doi.org/10.1186/s12906-019-2629-y> (2019).
33. Tochikura, T. S., Nakashima, H. & Yamamoto, N. Antiviral agents with activity against human retroviruses. *J. Acquir. Immune Defic. Syndr.* **2**, 441–447 (1988).
34. El Dine, R. S., Halawany, A. M. E., Ma, C. M. & Hattori, M. Anti-HIV-1 protease activity of lanostanetripenes from the Vietnamese mushroom *Ganodermacolossium*. *J. Nat. Prod.* **71**, 1022–1026. <https://doi.org/10.1021/np8001139> (2008).
35. Mothana, R. A. *et al.* Antiviral lanostanoidtriterpenes from the fungus *Ganodermapfeifferi*. *Fitoterapia* **74**, 177–180. [https://doi.org/10.1016/s0367-326x\(02\)00305-2](https://doi.org/10.1016/s0367-326x(02)00305-2) (2003).
36. Lee, S. *et al.* The anti-Influenza virus effect of *Phellinusignarius* extract. *J. Microbiol.* **51**, 676–681. <https://doi.org/10.1007/s12275-013-3384-2> (2013).
37. Lull, C., Wichers, H. J. & Savelkoul, H. Antiinflammatory and immunomodulating properties of fungal metabolites. *Mediators Inflamm.* **2005**, 63–80. <https://doi.org/10.1155/MI.2005.63> (2005).
38. Moradali, M. F., Mostafavi, H., Ghods, S. & Hedjaroude, G. A. Immunomodulating and anticancer agents in the realm of macrofungi (macrofungi). *Int. Immunol. Pharmacol.* **7**, 701–724. <https://doi.org/10.1016/j.intimp.2007.01.008> (2007).
39. Costela-Ruiz, V. J. *et al.* SARS-CoV-2 infection: The role of cytokines in COVID-19 disease. *Cytokine Growth Factor Rev.* **54**, 62–75. <https://doi.org/10.1016/j.cytogfr.2020.06.001> (2020).
40. Suwannarach, N. *et al.* Natural bioactive compounds from fungi as potential candidates for protease inhibitors and immune modulators to apply for coronaviruses. *Molecules* **25**, 1800. <https://doi.org/10.3390/molecules25081800> (2020).
41. Rangsinth, P. *et al.* Mushroom-derived bioactive compounds potentially serve as the inhibitors of SARS-CoV-2 main protease: An in silico approach. *J. Tradit. Complement Med.* **11**, 158–172. <https://doi.org/10.1016/j.jtcme.2020.12.002> (2021).
42. Thu, Z. M. *et al.* Bioactive phytochemical constituents of wild edible mushrooms from southeast Asia. *Molecules* **25**, 1972. <https://doi.org/10.3390/molecules25081972> (2020).

43. Thongbai, B. *et al.* *Hericiumerinaceus*, an amazing medicinal mushroom. *Mycol Progress*. **14**, 1–23. <https://doi.org/10.1007/s11557-015-1105-4> (2015).
44. Lindequist, U., Niedermeyer, T. H. & Jülich, W. D. The pharmacological potential of mushrooms. *Evid. Based Complement Altern. Med.* **2**, 285–299. <https://doi.org/10.1093/ecam/neh107> (2005).
45. Lin, H. & Ji-Kai, L. Two novel phenylacetoxylated p-terphenyls from *Thelephoraganbajun* Zang. *Z. Naturforsch C J. Biosci.* **56**, 983–987. <https://doi.org/10.1515/znc-2001-11-1213> (2001).
46. Öztürk, M. *et al.* Mushrooms: A source of exciting bioactive compounds. *Stud. Nat. Prod. Chem.* **45**, 363–456. <https://doi.org/10.1016/B978-0-444-63473-3.00010-1> (2015).
47. Sangsopha, W. *et al.* New p-terphenyl and benzoquinone metabolites from the bioluminescent mushroom *Neonothopanusnambi*. *Nat. Prod. Res.* **34**, 2186–2193. <https://doi.org/10.1080/14786419.2019.1578763> (2020).
48. Rappé, A. K. *et al.* UFF, a full periodic table force field for molecular mechanics and molecular dynamics simulations. *J. Am. Chem. Soc.* **114**, 10024–10035. <https://doi.org/10.1021/ja00051a040> (1992).
49. Jin, Z. *et al.* Structure of M^{pro} from SARS-CoV-2 and discovery of its inhibitors. *Nature* **582**, 289–293. <https://doi.org/10.1038/s41586-020-2223-y> (2020).
50. Liu, X. *et al.* The crystal structure of COVID-19 main protease in complex with an inhibitor N3. *Protien Data Bank* <https://doi.org/10.2210/pdb6LU7/pdb> (2020).
51. Morris, G. M. *et al.* AutoDock 4 and AutoDockTools4: Automated docking with selective receptor flexibility. *J. Comput. Chem.* **30**, 2785–2791. <https://doi.org/10.1002/jcc.21256> (2009).
52. Kitchen, D. B., Decornez, H., Furr, J. R. & Bajorath, J. Docking and scoring in virtual screening for drug discovery: Methods and applications. *Nat. Rev. Drug Discov.* **3**, 935–949. <https://doi.org/10.1038/nrd1549> (2004).
53. Trott, A. J. & Olson, O. Auto Dock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.* **31**, 455–461. <https://doi.org/10.1002/jcc.21334> (2010).
54. Dallakyan, S. & Olson, A. J. Small-molecule library screening by docking with PyRx. *Methods Mol. Biol.* **1263**, 243–250. https://doi.org/10.1007/978-1-4939-2269-7_19 (2015).
55. Anandakrishnan, R., Aguilar, B. & Onufriev, A. V. H++ 3.0: Automating pK prediction and the preparation of biomolecular structures for atomistic molecular modeling and simulation. *Nucleic Acids Res.* **40**(W1), W537–541. <https://doi.org/10.1093/nar/gks375> (2012).
56. Lippert, T. & Rarey, M. Fast automated placement of polar hydrogen atoms in protein-ligand complexes. *J. Chem. Inform.* **1**, 13. <https://doi.org/10.1186/1758-2946-1-13> (2009).
57. Bietz, S., Urbaczek, S., Schulz, B. & Rarey, M. Protoss: A holistic approach to predict tautomers and protonation states in protein-ligand complexes. *J. Chem. Inform.* **6**, 12. <https://doi.org/10.1186/1758-2946-6-12> (2014).
58. QikProp. *Schrödinger* (LLC, 2021).
59. Lipinski, C. A., Lombardo, F., Dominy, B. W. & Feeney, P. J. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug Deliv. Rev.* **46**, 3–26. [https://doi.org/10.1016/s0169-409x\(00\)00129-0](https://doi.org/10.1016/s0169-409x(00)00129-0) (2001).
60. Daina, A., Michielin, O. & Zoete, V. SwissADME: A free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. *Sci. Rep.* **7**, 42717. <https://doi.org/10.1038/srep42717> (2017).
61. Abraham, M. J. *et al.* Gromacs: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *Software X* **1–2**, 19–25. <https://doi.org/10.1016/j.softx.2015.06.001> (2015).
62. Huang, J. & Mackerell, A. D. CHARMM36 all-atom additive protein force field: Validation based on comparison to NMR data. *J. Comput. Chem.* **34**, 2135–2145. <https://doi.org/10.1002/jcc.23354> (2013).
63. Zoete, V., Cuendet, M. A., Grosdidier, A. & Michielin, O. SwissParam: A fast force field generation tool for small organic molecules. *J. Comput. Chem.* **32**, 2359–2368. <https://doi.org/10.1002/jcc.21816> (2011).
64. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**, 926–935. <https://doi.org/10.1063/1.445869> (1983).
65. Chen, C., Huang, Y., Ji, X. & Xiao, Y. Efficiently finding the minimum free energy path from steepest descent path. *J. Chem. Phys.* **138**, 164122. <https://doi.org/10.1063/1.4799236> (2013).
66. Hess, B. P-LINCS: A parallel linear constraint solver for molecular simulation. *J. Chem. Theory Comput.* **4**, 116–122. <https://doi.org/10.1021/ct700200b> (2008).
67. Huang, Y., Chen, W., Wallace, J. A. & Shen, J. All-atom continuous constant pH molecular dynamics with particle mesh ewald and titratable water. *J Chem Theory Comput.* **12**, 5411–5421. <https://doi.org/10.1021/acs.jctc.6b00552> (2016).
68. Eslami, H., Mojahedi, F. & Moghadasi, J. Molecular dynamics simulation with weak coupling to heat and material baths. *J. Chem. Phys.* **133**, 84105. <https://doi.org/10.1063/1.3474951> (2010).
69. Martoňák, R., Laio, A. & Parrinello, M. Predicting crystal structures: The Parrinello-Rahman method revisited. *Phys. Rev. Lett.* **90**, 5503. <https://doi.org/10.1103/PhysRevLett.90.075503> (2003).
70. Schreiner, W., Karch, R., Knapp, B. & Ilieva, N. Relaxation estimation of RMSD in molecular dynamics immunosimulations. *Comput. Math. Methods. Med.* **2012**, 173521. <https://doi.org/10.1155/2012/173521> (2012).
71. Martínez, L. Automatic identification of mobile and rigid substructures in molecular dynamics simulations and fractional structural fluctuation analysis. *PLoS ONE* **10**, e0119264. <https://doi.org/10.1371/journal.pone.0119264> (2015).
72. Lobanov, M. Y., Bogatyreva, N. S. & Galzitskaya, O. V. Radius of gyration as an indicator of protein structure compactness. *Mol. Biol.* **42**, 701–706. <https://doi.org/10.1134/S0026893308040195> (2008).
73. Huang, H. & Simmerling, C. Fast pairwise approximation of solvent accessible surface area for implicit solvent simulations of proteins on CPUs and GPUs. *J. Chem. Theory Comput.* **14**, 5797–5814. <https://doi.org/10.1021/acs.jctc.8b00413> (2018).
74. Humphrey, W., Dalke, A. & Schulten, K. V. M. D. Visual molecular dynamics. *J. Mol. Graph.* **14**, 33–38. [https://doi.org/10.1016/0263-7855\(96\)00018-5](https://doi.org/10.1016/0263-7855(96)00018-5) (1996).
75. Wang, C. *et al.* Recent developments and applications of the MMPBSA method. *Front. Mol. Biosci.* **4**, 87. <https://doi.org/10.3389/fmolb.2017.00087> (2018).
76. Kumari, R., Kumar, R., Lynn, A. & G-mmpbsa, A. GROMACS tool for high-throughput MM-PBSA calculations. *J. Chem. Inf. Model.* **54**, 1951–1962. <https://doi.org/10.1021/ci500020m> (2014).
77. Jurrus, E. *et al.* Improvements to the APBS biomolecular solvation software suite. *Protein Sci.* **27**, 112–128. <https://doi.org/10.1002/pro.3280> (2018).
78. Zhang, L. *et al.* Crystal structure of SARS-CoV-2 main protease provides a basis for design of improved α -ketoamideinhibitors. *Science* **368**, 409–412. <https://doi.org/10.1126/science.abb3405> (2020).
79. Yoshino, R., Yasuo, N. & Sekijima, M. Identification of key interactions between SARS-CoV-2 main protease and inhibitor drug candidates. *Sci. Rep.* **10**, 12493. <https://doi.org/10.1038/s41598-020-69337-9> (2020).
80. Maiorov, V. N. & Crippen, G. M. Significance of root-mean-square deviation in comparing three-dimensional structures of globular proteins. *J. Mol. Biol.* **235**, 625–634. <https://doi.org/10.1006/jmbi.1994.1017> (2010).
81. Sen, D. *et al.* Potentiality of *Moringa oleifera* against SARS-CoV-2: Identified by a rational computer aided drug design method. *J. Biomol. Struct. Dyn.* **15**, 1–18. <https://doi.org/10.1080/07391102.2021.1898475> (2021).
82. UlQamar, M. T., Alqahtani, S. M., Alamri, M. A. & Chen, L. L. Structural basis of SARS-CoV-2 3CLpro and anti-COVID-19 drug discovery from medicinal plants. *J. Pharm. Anal.* **10**, 313–319. <https://doi.org/10.1016/j.jpba.2020.03.009> (2020).

83. Su, H. X. *et al.* Anti-SARS-CoV-2 activities in vitro of Shuanghuanglian preparations and bioactive ingredients. *Acta Pharmacol. Sin.* **41**, 1167–1177. <https://doi.org/10.1038/s41401-020-0483-6> (2020).
84. Ryu, Y. B. *et al.* Biflavonoids from *Torreya nucifera* displaying SARS-CoV 3CLpro inhibition. *Bioorg. Med. Chem.* **18**, 7940–7947. <https://doi.org/10.1016/j.bmc.2010.09.035> (2010).

Acknowledgements

This research was supported by the Deanship of Scientific Research, Imam Mohammad Ibn Saud Islamic University (IMSIU), Saudi Arabia, Grant No. (21-13-18-070).

Author contributions

D.S.: MD simulation and analysis. B.D.: data collection, P.D.: initial drafting. S.D.*: conceptualization, supervision, review and final editing (*Corresponding author 1). M.E.A.Z*: fund arrangement and supervision (*Corresponding author 2). V.H.M.: drafting.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-022-05349-x>.

Correspondence and requests for materials should be addressed to S.D. or M.E.A.Z.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022

Article

Exploring the Prominent and Concealed Inhibitory Features for Cytoplasmic Isoforms of Hsp90 Using QSAR Analysis

Magdi E. A. Zaki^{1,*}, Sami A. Al-Hussain¹, Syed Nasir Abbas Bukhari², Vijay H. Masand^{3,*}, Mithilesh M. Rathore³, Sumer D. Thakur⁴ and Vaishali M. Patil⁵

¹ Department of Chemistry, Faculty of Science, Imam Mohammad Ibn Saud Islamic University, Riyadh 13318, Saudi Arabia; sahussain@imamu.edu.sa

² Department of Pharmaceutical Chemistry, College of Pharmacy, Jouf University, Al Jouf 72388, Saudi Arabia; sbukhari@ju.edu.sa

³ Department of Chemistry, Vidya Bharati Mahavidyalaya, Amravati 444 602, Maharashtra, India; mithsrathore@gmail.com

⁴ Department of Chemistry, RDIK and NKD College, Badnera-Amravati 444 701, Maharashtra, India; sdthakur11@gmail.com

⁵ Department of Pharmaceutical Chemistry, KIET School of Pharmacy, KIET Group of Institutions, Delhi-NCR, Ghaziabad 201206, Uttar Pradesh, India; vaishuwise@gmail.com

* Correspondence: mezaki@imamu.edu.sa (M.E.A.Z.); vijaymasand@gmail.com (V.H.M.)



Citation: Zaki, M.E.A.; Al-Hussain, S.A.; Bukhari, S.N.A.; Masand, V.H.; Rathore, M.M.; Thakur, S.D.; Patil, V.M. Exploring the Prominent and Concealed Inhibitory Features for Cytoplasmic Isoforms of Hsp90 Using QSAR Analysis. *Pharmaceuticals* **2022**, *15*, 303. <https://doi.org/10.3390/ph15030303>

Academic Editor: Osvaldo Andrade Santos-Filho

Received: 21 January 2022

Accepted: 23 February 2022

Published: 1 March 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Abstract: Cancer is a major life-threatening disease with a high mortality rate in many countries. Even though different therapies and options are available, patients generally prefer chemotherapy. However, serious side effects of anti-cancer drugs compel us to search for a safer drug. To achieve this target, Hsp90 (heat shock protein 90), which is responsible for stabilization of many oncoproteins in cancer cells, is a promising target for developing an anti-cancer drug. The QSAR (Quantitative Structure–Activity Relationship) could be useful to identify crucial pharmacophoric features to develop a Hsp90 inhibitor. Therefore, in the present work, a larger dataset encompassing 1141 diverse compounds was used to develop a multi-linear QSAR model with a balance of acceptable predictive ability (Predictive QSAR) and mechanistic interpretation (Mechanistic QSAR). The new developed six-parameter model satisfies the recommended values for a good number of validation parameters such as $R2_{tr} = 0.78$, $Q2_{LMO} = 0.77$, $R2_{ex} = 0.78$, and $CCC_{ex} = 0.88$. The present analysis reveals that the Hsp90 inhibitory activity is correlated with different types of nitrogen atoms and other hidden structural features such as the presence of hydrophobic ring/aromatic carbon atoms within a specific distance from the center of mass of the molecule, etc. Thus, the model successfully identified a variety of reported as well as novel pharmacophoric features. The results of QSAR analysis are further vindicated by reported crystal structures of compounds with Hsp90.

Keywords: Hsp90; cancer; QSAR; machine learning; pharmacophores

1. Introduction

Cancer kills; therefore, medicinal chemists are continuously trying to develop therapeutic agents that could retard the growth of cancer cells. In cancer cells, a protein Hsp90 (heat shock protein 90, also known as HSPC) is overexpressed [1]. It is a highly conserved, non-fibrous, and chaperone protein with a key role in many cellular processes like proper folding of other proteins, apoptosis, cell cycle control, cell viability, and degradation and signaling events [1–6]. As the name indicates, Hsp (heat shock proteins) shield cells when stressed by higher temperatures. The number “90” comes from the fact that it weighs about 90 kDa. There are two isoforms of Hsp90: Hsp90 α (the inducible form) and Hsp90 β (the constitutive form), which are found in cytoplasm and share 85% sequence identity [1–6]. These two isoforms are like flexible biological catalysts and interact with a good number of newly synthesized proteins, such as Akt2, CDKs, PKC, MAP kinases, steroid receptors,

BCL-6, CAR, p53, Oct4, etc., to avoid their aggregation or mistakes in their folding [6]. Despite a crucial role, in cancer cells, these are responsible for the stabilization of a number of oncoproteins required for tumor growth, leading to their overexpression [1–6]. Consequently, Hsp90 is an attractive target for developing a drug for cancer.

The majority of Hsp90 inhibitors occupy the ATP (adenosine tri-phosphate) pocket in the N-terminal domain of Hsp90, leading to limited ATPase activity [1–6]. At present, several natural and semi-synthetic Hsp90 inhibitors (see Figure 1) are in different stages of clinical trials for a variety of cancers [2,3,7–9].

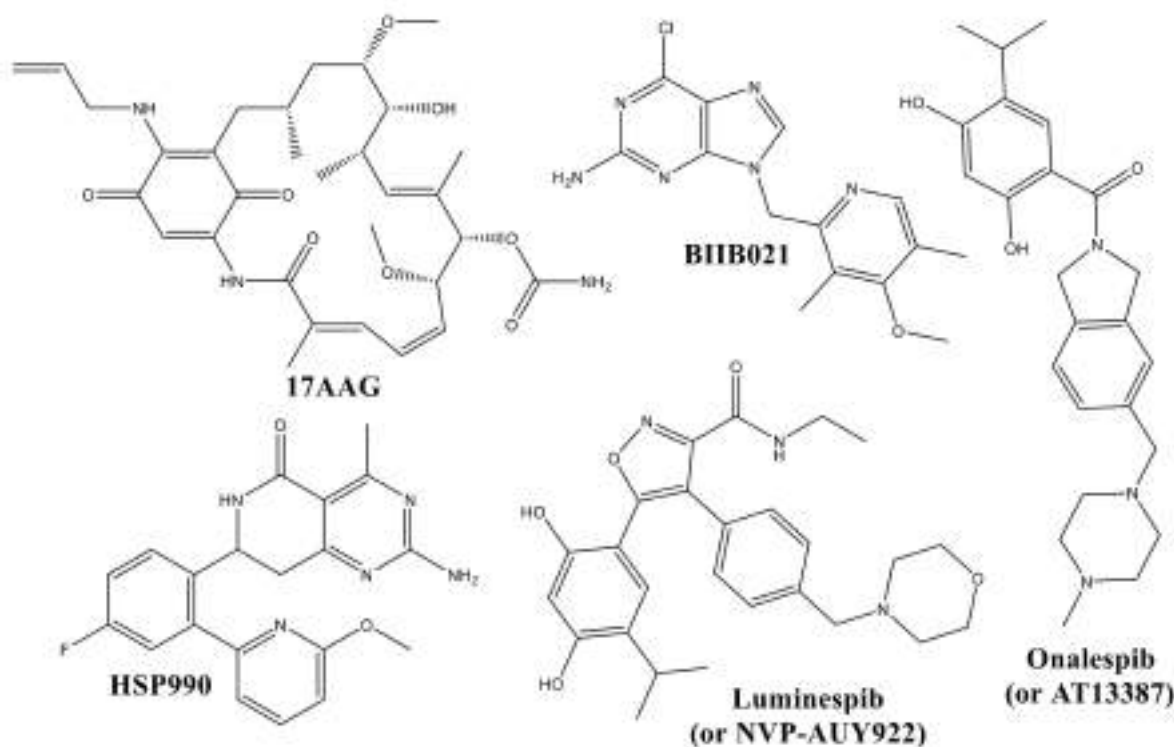


Figure 1. Different clinical trial candidates as inhibitors of Hsp90.

Unfortunately, several inhibitors have shown hepatotoxicity and ocular toxicity [2,10]; consequently, there is a need to modify them with retention of activity against Hsp90, which could be achieved on knowing the structural features responsible for their Hsp90 inhibitory activity. A simple, cost-effective, and faster yet effective strategy to know crucial pharmacophoric features is to use QSAR (Quantitative Structure–Activity Relationships), a successful, contemporary, and widely used branch of computer assisted-drug designing [11–16].

In QSAR analysis, generally, a good number of inhibitors are analyzed using a suitable technique like machine learning, deep learning, etc. There are two main advantages of using the QSAR approach [11,17,18]: (a) the analysis helps to identify the prominent structural features or patterns that influence the bioactivity profile of molecules (Mechanistic interpretation or Qualitative QSAR), and (b) the analysis could be used to predict the desired bioactivity of a molecule prior to its synthesis and lab testing (Predictive ability or Predictive QSAR). Therefore, many researchers prefer QSAR as a method of choice for drug/lead optimization. Nowadays, a QSAR analysis with a balance of mechanistic interpretation with predictive ability is highly preferred.

The literature survey reveals that QSAR analyses have been reported for Hsp90, but they are either based on a small dataset, lack general applicability, have poor predictive ability, are deficient of a mechanistic interpretation, or a combination of these factors, which limit their use [9,19–22]. Therefore, in the present work, we accomplished QSAR analysis for a larger and diverse dataset of Hsp90 inhibitors, and followed the OECD

(Organization for Economic Cooperation and Development) guidelines while developing a QSAR model to have a balance of mechanistic interpretation with predictive ability.

2. Results

The exhaustive and heuristic search resulted in the development of a six-descriptor-based QSAR model (see model-A), which was subjected to thorough statistical validation for internal and external validations.

Model-A: $pIC_{50} (M) = 3.903 (\pm 0.134) + 0.101 (\pm 0.013) \times \text{com_ringChyd_4A} + 0.433 (\pm 0.058) \times \text{faroCN2B} + 0.714 (\pm 0.214) \times \text{aroCminus_sumpc} + 0.065 (\pm 0.005) \times \text{aroC_aroN_5B} + 0.266 (\pm 0.048) \times \text{fringNsp3C5B} + 0.59 (\pm 0.082) \times \text{da_amdN_6B}$

Statistical validation of model-A:

$N_{tr} = 915$, $N_{ext} = 226$, $R^2_{tr} = 0.779$, $R^2_{adj.} = 0.777$, $R^2_{tr} - R^2_{adj.} = 0.002$, $LOF = 0.244$, $K_{XX} = 0.219$, $\Delta K = 0.122$, $RMSE_{tr} = 0.487$, $MAE_{tr} = 0.404$, $RSS_{tr} = 217.321$, $CCC_{tr} = 0.876$, $s = 0.489$, $F = 533.134$, $R^2_{cv} (Q^2_{loo}) = 0.775$, $R^2 - R^2_{cv} = 0.004$, $RMSE_{cv} = 0.491$, $MAE_{cv} = 0.407$, $PRESS_{cv} = 220.839$, $CCC_{cv} = 0.874$, $Q^2_{LMO} = 0.775$, $R^2_{Yscr} = 0.007$, $Q^2_{Yscr} = -0.009$, $RMSE_{ex} = 0.474$, $MAE_{ex} = 0.383$, $PRESS_{ext} = 50.675$, $R^2_{ex} = 0.779$, $Q^2 - F^1 = 0.778$, $Q^2 - F^2 = 0.778$, $Q^2 - F^3 = 0.791$, $CCC_{ex} = 0.876$, $R^2 - ExPy = 0.779$, $R'_o{}^2 = 0.727$, $k' = 0.989$, $1 - (R^2 / R'_o{}^2) = 0.066$, $r'^2m = 0.602$, $R_o{}^2 = 0.779$, $k = 1.005$, $1 - (R^2 - ExPy / R_o{}^2) = 0$, $r^2m = 0.766$

Different researchers have recommended the above statistical parameters to judge the robustness and external predictive ability of a QSAR model [11–16,23–31]. The formula to calculate them is available in the Supplementary Materials. It is clear that model-A fulfils the recommended threshold for many validation parameters and other criteria. A high value of different parameters like R^2_{tr} (coefficient of determination), $R^2_{adj.}$ (adjusted coefficient of determination), and R^2_{cv} (Q^2_{loo} , cross-validated coefficient of determination for leave-one-out), R^2_{ex} (external coefficient of determination), $Q^2 - F^n$, and CCC_{ex} (Concordance Correlation Coefficient), etc., and a low value of LOF (lack-of-fit), $RMSE_{tr}$ (root mean square error), MAE_{tr} (mean absolute error), R^2_{Yscr} (R^2 for Y-scrambling), etc. along with the different graphs (see Figure 2) associated with the model indicate that the model possesses statistical robustness with excellent internal and external predictive ability as well as free from chance correlations. Additionally, the Williams plot specifies that the model is statistically acceptable (see Figure 2d). Therefore, it fulfils all the OECD recommended guidelines for creating a useful QSAR model.

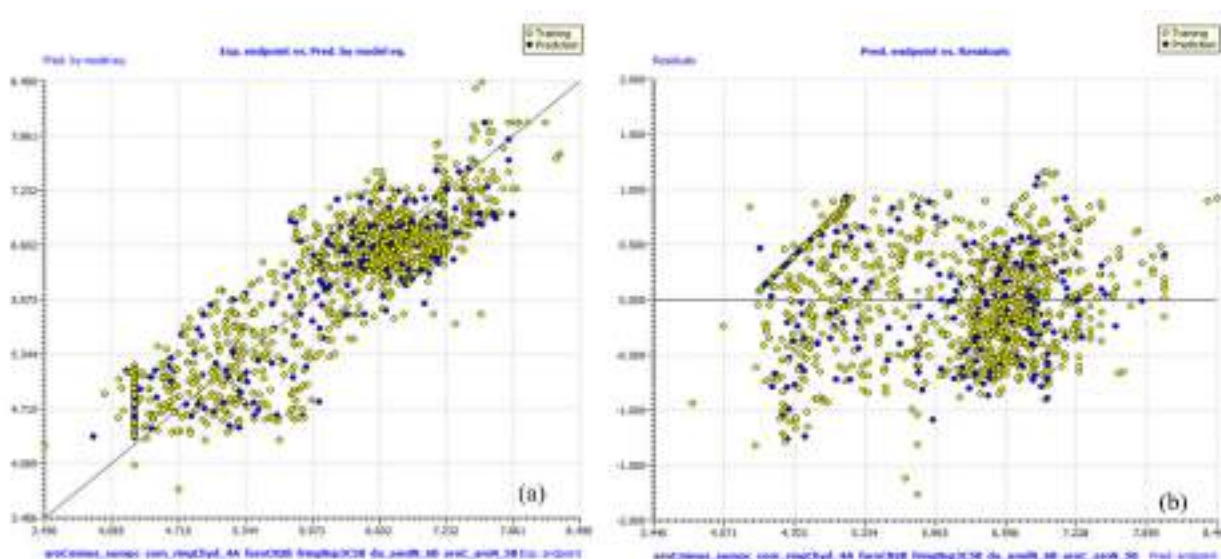


Figure 2. Cont.

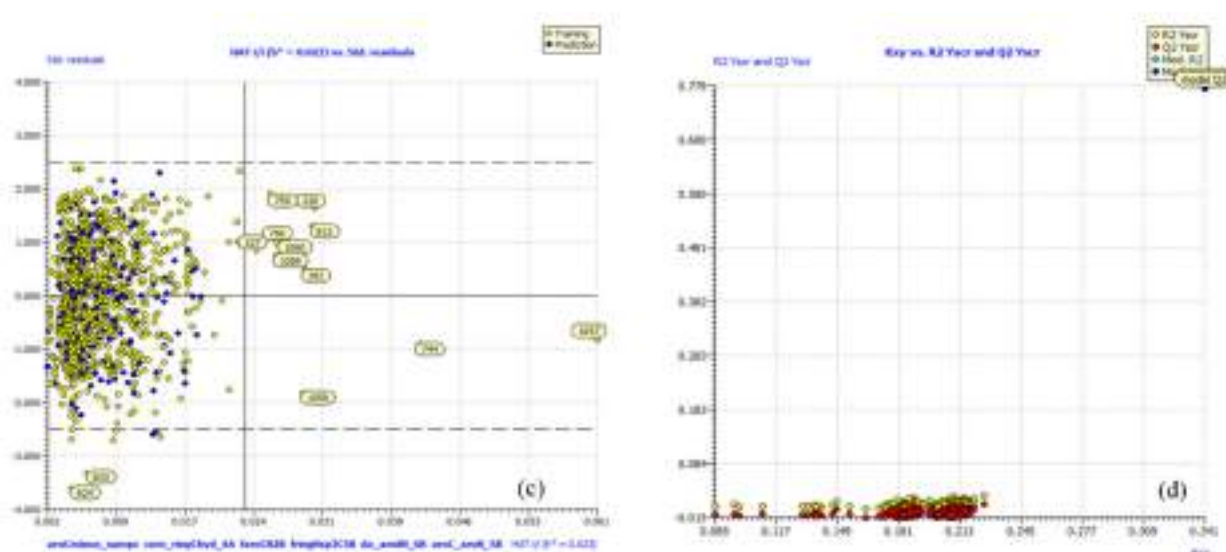


Figure 2. Different graphs associated with model-A: (a) experimental vs. predicted pIC_{50} (the solid line represents the regression line), (b) experimental vs. residuals, (c) Williams plot for applicability domain (the vertical solid line represents $h^* = 0.023$ and horizontal dashed lines represent the upper and lower boundaries for applicability domain), and (d) Y-randomization.

3. Discussion

Mechanistic Interpretation of QSAR Model

A very crucial aspect of a useful QSAR analysis is to gain deep insight into the pharmacophore or structure-oriented linking of molecular descriptors [17,32]. This not only helps throughout the drug discovery process, but also expands the information and understanding of mechanistic aspects of different types of molecules. Though, in the present work, a specific molecular descriptor was used to equate the pIC_{50} values of different molecules, but an extending or reverse influence of unknown factors or other molecular descriptors, having a dominant effect in deciding the final pIC_{50} value of a molecule, cannot be ignored. To simplify, a single molecular descriptor (in turn structure feature) cannot decide the overall experimental pIC_{50} value of a molecule. In other words, the effective use of an appropriately validated QSAR model depends on the synchronous consideration of all constituent molecular descriptors. Interestingly, in model-A, all the molecular descriptors have positive coefficients, which indicates that increasing their value could result in a better Hsp90 inhibitory activity.

The descriptor **com_ringChyd_4A** represents the total number of hydrophobic ring carbons, having partial charge in the range ± 0.2 , within 4\AA from the com (center of mass) of the molecule. From this, it appears that mere total number of ring carbons is very important, but replacing **com_ringChyd_4A** with **nringC** (number of ring carbon atoms) or **naroC** (number of aromatic carbon atoms) significantly reduced the statistical performance of the model ($R^2 = 0.72$). To add further, **com_ringChyd_4A** has a positive correlation of $R = 0.488$ with pIC_{50} , whereas **nringC** and **naroC** have a correlation of $R = 0.461$ and 0.405 , respectively. **com_ringChyd_3A** and **com_ringChyd_5A** represent the total number of ring carbons, having partial charge in the range ± 0.2 , within 3\AA and 5\AA from the com (center of mass) of the molecule, respectively. Replacement of **com_ringChyd_4A** with **com_ringChyd_3A** or **com_ringChyd_5A** resulted in slightly reduced performance of the model with $R^2 = 0.75$ and 0.76 , respectively. This indicates that the optimum distance is 4\AA .

The importance of hydrophobic ring carbon atoms is supported by the X-ray-resolved structure of a good number of Hsp90 inhibitors because the active site of Hsp90 consists of lipophilic side chains of Leu48, Ile91, Val186, Leu315, Ile388, and Val391 [33,34], which favors the presence of hydrophobic moiety in the inhibitors. For example, a comparison of molecule 988 ($pIC_{50} = 6.009$, **com_ringChyd_4A** = 10) with 1007 ($pIC_{50} = 6.481$,

com_ringChyd_4A = 15) highlights the importance of **com_ringChyd_4A**. Another pair of molecules, viz. 794 and 814, also supports this observation. The molecular descriptor **com_ringChyd_4A** is depicted in Figure 3 for different molecules.

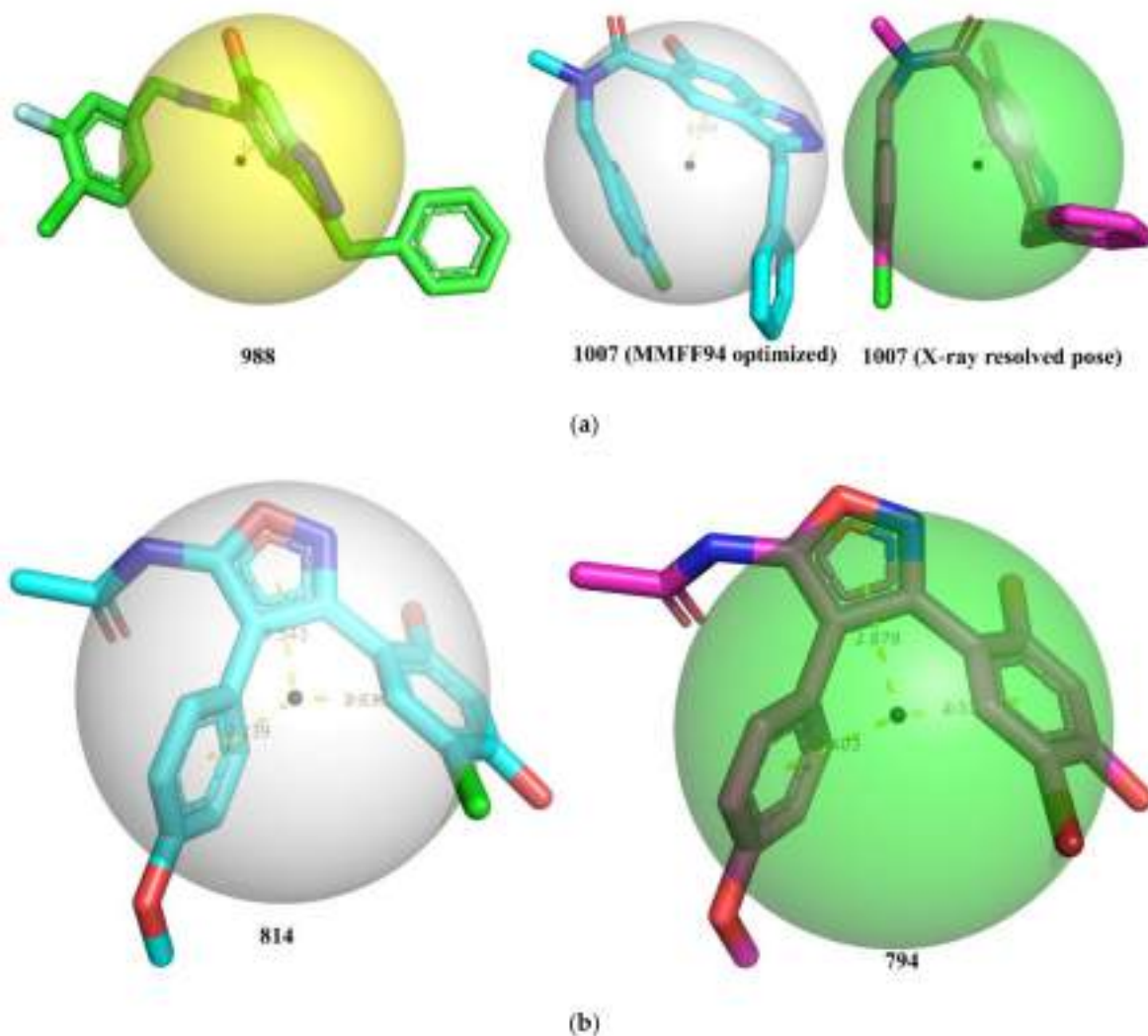


Figure 3. Depiction of **com_ringChyd_4A** using different molecules: (a) molecules 988, 1007 (MMFF94 optimized), and 1007 (X-ray resolved dock pose from pdb 6EY8); (b) molecules 794 and 814 (both X-ray-resolved poses from pdb 5XR9 and 4LWE, respectively). The small black sphere represents the com (center of mass) and the bigger transparent sphere represents the distance of 4Å from the center of mass. The dotted yellow line represents the distance (Å) of com from the centers of the different nearest rings.

From Figure 3, it is clear that the lowest energy conformer of molecule 988 has **com_ringChyd_4A** = 10 due to the closer presence of com (distance 1.206 Å) to the benzene ring of indazole ring. In case of molecule 1007 (MMFF94-optimized and X-ray-resolved pose from pdb 6EY8), the com is located slightly away from the benzene ring of Indazole ring at a distance > 2.40 due to specific conformation, thereby increasing the value of **com_ringChyd_4A** to 15. This could be a plausible reason for the difference in the bioactivity of these two compounds. Similarly, a better Hsp90 inhibitory activity of molecule 794 than 814 could be attributed to difference in their **com_ringChyd_4A** values.

Another molecular descriptor that has a positive effect on Hsp90 activity is **faroCN2B**, which signifies the presence of nitrogen exactly at two bonds from aromatic carbon atoms. If the same nitrogen atom is also present at two or less bonds from any other aromatic carbon atom, then it was excluded while calculating **faroCN2B**. This descriptor highlights the

importance of nitrogen atoms separated from aromatic ring (Benzene, etc.) by two bonds. As the majority of nitrogen atoms act as either an H-bond donor or acceptor; therefore, the presence of nitrogen atoms in the vicinity of aromatic rings could be useful in enhancing interactions with the polar residues of receptor (Hsp90). Additionally, the descriptor further points out the crucial role played by the aromatic rings undoubtedly due to their lipophilic nature. Taken together, the descriptor **faroCN2B** signifies the importance of two important structural features: aromatic rings and their vicinal nitrogen atoms.

This observation is confirmed when we compare the X-ray-resolved structures of molecule 727 ($pIC_{50} = 6.654$, **faroCN2B** = 1, pdb = 4O09) with 725 ($pIC_{50} = 7.137$, **faroCN2B** = 2, pdb = 4O05) depicted in Figure 4. The nitrogen atoms responsible for **faroCN2B** are highlighted by blue dotted circles. From Figure 4, it is clear that the aromatic ring B of both the molecules is responsible for hydrophobic interactions with the residue Met98. The nitrogen atom of ring A present in both the molecules is not only a constituent of **faroCN2B**, but also responsible for H-bonding with the residue Asp93. Thus, such a combination of aromatic carbons and nitrogen is highly beneficial to enhance the interactions with the receptor. In case of molecule 733, an additional nitrogen atom is present in ring E, which is a constituent of **faroCN2B**, and responsible for the H-bond interaction with the nearby water molecule. Thus, the present QSAR analysis revealed an important structural feature, which is also visible in X-ray-resolved structures of the same inhibitors with the same target enzyme Hsp90.

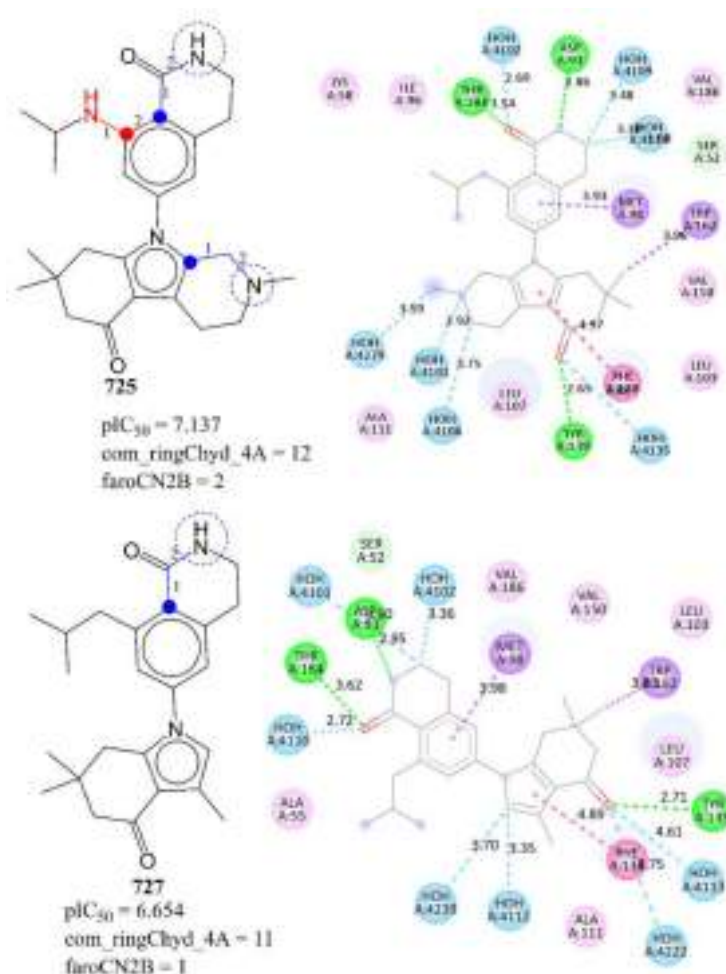


Figure 4. Depiction of **faroCN2B** using representative examples only.

A comparison of the following pairs of molecules further vindicates the importance of **faroCN2B** in determining the bioactivity: 213 ($pIC_{50} = 6.523$, **faroCN2B** = 2) with

212 ($pIC_{50} = 6.469$, **faroCN2B** = 1) and 758 ($pIC_{50} = 7.444$, **faroCN2B** = 2) with 759 ($pIC_{50} = 7.569$, **faroCN2B** = 3).

The importance of aromatic carbon atoms is further emphasized with the presence of **aroCminus_sumpc** as a constituent variable of model-A. The molecular descriptor **aroCminus_sumpc** represents the sum of partial charges on negatively charged aromatic carbon atoms. The positive coefficient for **aroCminus_sumpc** indicates that the higher the value of this descriptor, the better the activity profile. The sum of partial charges on negatively charged aromatic carbon atoms will always be negative; therefore, in reality, this descriptor actually decreases the pIC_{50} value. Further, the replacement of **aroCminus_sumpc** by **aroCplus_sumpc** (sum of partial charges on positively charged aromatic carbon atoms) led to a model with almost identical statistical performance ($R^2_{tr} = 0.772$, $Q^2_{LMO} = 0.767$, $R^2_{ex} = 0.78$, $CCC_{ex} = 0.876$). In fact, **aroCplus_sumpc** has a better correlation ($R = 0.33$) with pIC_{50} than **aroCminus_sumpc** ($R = 0.10$). From this it is clear that, if aromatic carbons are positively charged than the molecule possesses better Hsp90 inhibitory activity. Therefore, the best strategy is to attach atoms or groups that enhance lipophilic and mild polar interactions with the receptor (for example -Cl, etc.) to the aromatic carbon atoms. In short, substituted aromatic rings are preferable for better activity. This observation is supported by comparing following pairs of molecules: 2 with 3, 1054 with 1059, and 214 with 212.

aroC_aroN_5B, which represents the total number of aromatic carbon atoms within five bonds from aromatic nitrogen atoms, again points out the key role played by aromatic carbon atoms in deciding Hsp90 inhibitory activity. It also underlines the usefulness of aromatic nitrogen atoms. This descriptor has a positive correlation with pIC_{50} with $R = 0.63$. Therefore, an increase in number of aromatic carbon atoms within five bonds from aromatic nitrogen atoms leads to better Hsp90 inhibitory activity. The following pairs of the molecules support this observation: 888 ($pIC_{50} = 7.523$, **aroC_aroN_5B** = 22) with 887 ($pIC_{50} = 6.046$, **aroC_aroN_5B** = 20) and 107 ($pIC_{50} = 5.953$, **aroC_aroN_5B** = 13) with 108 ($pIC_{50} = 4.874$, **aroC_aroN_5B** = 10), to mention a few. Further, the 50 most active molecules possess relatively higher value of **aroC_aroN_5B** (range 8–17) than the 50 least active molecules (range 0–8).

fringNsp3C5B stands for the number of sp^3 -hybridized carbon atoms exactly at five bonds from the ring nitrogen atom. If the same sp^3 -hybridized carbon atom is also present at four or less bonds from any other ring nitrogen atom, then it was excluded while calculating **fringNsp3C5B**. It is interesting to note that the 50 most active molecules, except molecule 618, possess at least one or more of such a combination of carbon and ring nitrogen, whereas the 50 least active molecules either lack it or have **fringNsp3C5B** = 1. In the majority of compounds, the sp^3 -hybridized carbon atoms are present either as a linker between two rings or as a substituent, which therefore enhances conformational flexibility of the molecule to adopt a bioactive conformer or lipophilic characters of the molecule. A comparison of 895 ($pIC_{50} = 7.071$, **fringNsp3C5B** = 2) with 896 ($pIC_{50} = 6.777$, **fringNsp3C5B** = 1), 859 ($pIC_{50} = 7.237$, **fringNsp3C5B** = 2) with 896 ($pIC_{50} = 7.071$, **fringNsp3C5B** = 1), 326 ($pIC_{50} = 6.921$, **fringNsp3C5B** = 1) with 328 ($pIC_{50} = 7.046$, **fringNsp3C5B** = 2), and 412 ($pIC_{50} = 7.155$, **fringNsp3C5B** = 1) with 411 ($pIC_{50} = 6.959$, **fringNsp3C5B** = 0) and 410 ($pIC_{50} = 6.854$, **fringNsp3C5B** = 0) confirms the importance of **fringNsp3C5B** in deciding the activity.

A molecular descriptor that identifies the relation of total number amide nitrogen atoms within six bonds from the H-bond donor and acceptor atoms is **da_amdN_6B**. In the majority of compounds in the present dataset, the amide group is present as a substituent on aromatic ring or as a linker between two rings. The descriptor **da_amdN_6B** suggests the significance of amide group and its correlation with the H-bond donor and acceptor atoms. This observation is confirmed on comparing molecule **A** with molecules **B** and **C** (see Figure 5).

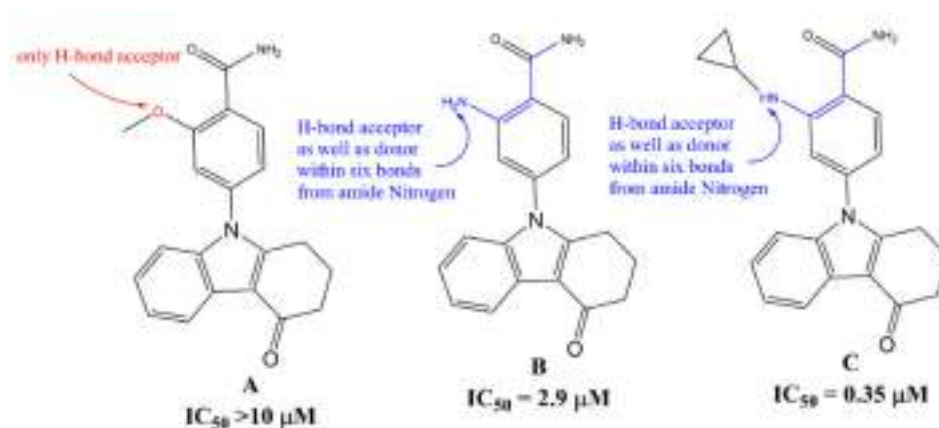


Figure 5. Pictorial representation of **da_amdN_6B** using representative examples only.

A good number of researchers have also pointed out that the amide group is crucial for Hsp90 inhibitors to establish H-bonding with residues of the active site (see pdb 4AWO). For example, Zhao et al. [4] pointed out that the distance between the nitrogen atoms on the piperidine ring and the amide are important for Hsp90 inhibition. Similarly, Baruchello and co-workers [35] studied a library of 3,4-isoxazole diamides for Hsp90 binding and found that a substantial reduction in Hsp90 binding affinity when the amide was replaced with substituted amines. In addition, a H-bond donor at the C-4 position on the isoxazole is vital for retaining the activity. Davies et al. [36] observed that S-acetamide derivatives of compounds have better bioactivity profile than the S-alkylamines. The importance of **da_amdN_6B** was further confirmed by comparing following pair of the molecules: 856 ($pIC_{50} = 6.848$, **da_amdN_6B** = 0) with 861 ($pIC_{50} = 7.114$, **da_amdN_6B** = 1). The earlier work identified the role of amide group, and in the present work, we successfully identified that a combination of amide group with H-bond donor/acceptor within six bonds is a better strategy to have better Hsp90 inhibitory activity. Therefore, such a combination of the amide nitrogen atom and H-bond donor/acceptor should be retained in future optimizations.

In short, three molecular descriptors emphasize the importance of ring carbon atoms, especially aromatic carbon atoms. This could be attributed to the lipophilic character of the active site of Hsp90. Likewise, four molecular descriptors underline the significance of different types of nitrogen atoms, which are responsible for the establishment of the polar or H-bond interactions with polar residues and water molecules present inside the active site of Hsp90. Hence, the present work is successful in identifying reported as well as novel pharmacophoric features of Hsp90 inhibitors.

4. Materials and Methods

The OECD (Organization for Economic Cooperation and Development) guidelines and a standard protocol recommended by different researchers [11–13,16,18,25,26,29,30,37] involve the sequential execution of (1) data collection and its curation, (2) structure generation and calculation of molecular descriptors, (3) objective feature selection (OFS), (4) splitting the dataset into training and external validation sets, (5) subjective feature selection involving building a regression model and validation of the developed model, which have all been followed to build a widely applicable QSAR model for Hsp-90 inhibitory activity. This also ensures thorough validation and successful application of the model.

4.1. Data Collection and Its Curation

The dataset of Hsp-90 inhibitory activity used for building, training, and validating the QSAR model in the present work was downloaded from BindingDB (<https://www.bindingdb.org/bind/index.jsp>, accessed on 24 December 2021), which is a free and publicly accessible database. Initially, the dataset comprised 1839 molecules. Then, as a part of data curation, entries with ambiguous IC_{50} values, duplicates, salts, metal-based inhibitors, etc.

were omitted [11–13,16,18,25,26,29,30,37]. The final dataset comprises 1141 structurally diverse molecules with remarkable variation in structural scaffolds, which were tested experimentally for potency in terms of IC_{50} (nM) (see the MS Excel file ‘SupplementaryMaterial-Final’ in the Supplementary Materials). The dataset includes N-terminal inhibitors of Hsp90. The experimental IC_{50} values have a sufficient variation ranging from 5 to 350,000 nM. After that, IC_{50} values were converted to their negative logarithmic value ($pIC_{50} = -\log_{10}IC_{50}$) so that a comparison of their values became easier. In Table 1 and Figure 6, some of the most and least active molecules are included as examples only.

Table 1. SMILES notation, IC_{50} (nM) and pIC_{50} (M) of the five most and least active molecules of the selected dataset.

S.N.	Ligand SMILES	IC_{50} (nM)	pIC_{50} (M)
308	<chem>COC1cccc(n1)-c1cc(F)ccc1[C@H]1Cc2nc(N)nc(C)c2C(NOC2C[C@H](O)[C@H](O)C2)=N1</chem>	5	8.301
908	<chem>CCNC(=O)c1noc(c1NC(=O)[C@H]1CC[C@H](CNS(=O)(=O)c2cc(F)cc2)CC1)-c1cc(C(C)C)c(O)cc1O</chem>	5.4	8.268
770	<chem>CCNC(=O)c1nnn(c1-c1ccc(CNC2CCCC2)cc1)-c1cc(C(C)C)c(O)cc1O</chem>	6.8	8.167
767	<chem>CCNC(=O)c1nnn(c1-c1ccc(CN2CCCC2CCO)cc1)-c1cc(C(C)C)c(O)cc1O</chem>	10	8
749	<chem>CCNC(=O)c1nnn(c1-c1ccc(CNCCCN(CC)CC)cc1)-c1cc(C(C)C)c(O)cc1O</chem>	12	7.921
775	<chem>Oc1cc(O)c2C[C@@H](OC(=O)[C@H]3CC[C@H](F)CC3)[C@H](Oc2c1)c1ccc(O)c(O)c1</chem>	69,000	4.161
1073	<chem>COC(COCCOc1ccc(Br)cc1)CN1CCN(CC1)c1ccccc1C(C)(C)C</chem>	70,430	4.152
1141	<chem>CO[C@H]1C[C@H](C)C2c(OC)c(O)cc3NC(=O)\C(C)=C\C[C@H](O)C[C@H](OC)[C@H](OC(N=O)\C(C)=C\C[C@H](C)[C@H]1O)c23</chem>	96,000	4.018
778	<chem>Oc1cc(O)c2C[C@H](OC(=O)c3ccccc3)[C@H](Oc2c1)c1ccccc1</chem>	120,000	3.921
207	<chem>CSc1nc(C)nc(N)n1</chem>	350,000	3.456

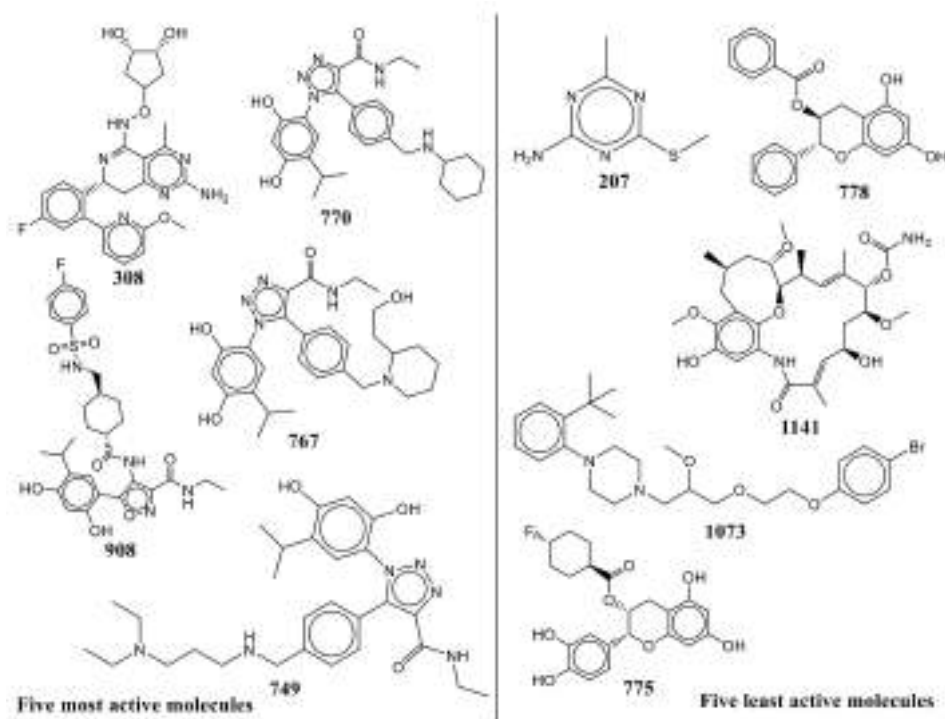


Figure 6. Representative examples from the selected dataset (the five most active and five least active molecules).

4.2. Calculation of Molecular Descriptors and Objective Feature Selection (OFS)

A crucial step before the calculation of molecular descriptors is to convert the SMILES notations to 3D-optimized structures and partial charge assignment, which was accomplished using OpenBabel 3.1 [38] using MMFF94 force field. In the present work, the X-ray-resolved structure of molecule **1007** (pdb 6YE8) was used to identify the parameter tuning in OpenBabel, required to get a better optimized structure, until there was a high similarity between the MMFF94-optimized structure and X-ray-resolved structure. This enhances the chances of getting a bioactive conformer, which in turn is highly beneficial for further optimization of Hsp90 inhibitors in the drug discovery pipeline. A comparison of the X-ray-resolved structures of molecules **1007** and **33** (pdb 2VCJ) and their respective MMFF94-optimized structures are represented in Figure 7.

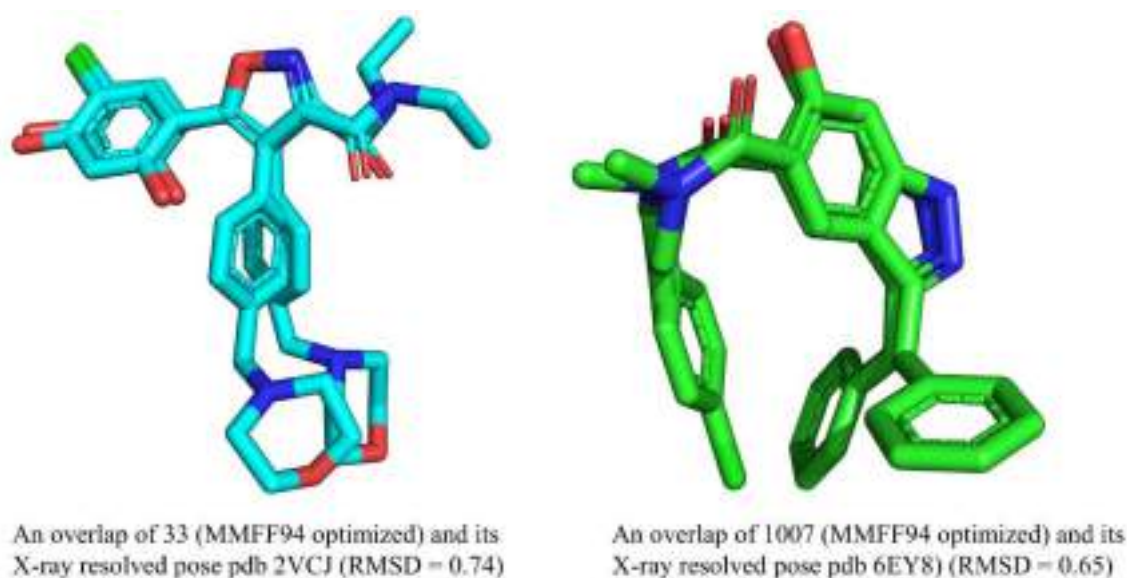


Figure 7. A comparison of X-ray-resolved and MMFF94-optimized structures of molecules **1007** and **33**.

From Figure 7, it is clear that there is a high similarity between the X-ray-resolved and MMFF94-optimized structure of molecules **1007** and **33**, which indicates that appropriate parameter tuning was achieved to optimize the rest of the molecules. That is, the same parameter tuning in OpenBabel was used to optimize the other molecules of the selected dataset. The parameters are as follows: geometry optimization, steepest descent, number of steps: 1500; cut off: 0.01.

In the next step, the 3D-optimized structures of all molecules in the dataset were used to calculate a good number of molecular descriptors. It is important to note that calculation of diverse molecular descriptors enhances the chances of a successful QSAR analysis and significantly helps in mechanistic interpretation. However, descriptor pruning is very useful as it further strengthens the diminished risk of overfitting from noisy redundant descriptors. To fulfil these objectives, more than 40,000 molecular descriptors were generated using *PyDescriptor* [39]. After that, OFS involved elimination of the near constant (90% molecules) and highly intercorrelated ($|R| > 0.90$) molecular descriptors. For this, QSARINS-2.2.4 was used. The final set of molecular descriptors comprises 1228 molecular descriptors, which still comprise manifold descriptors (1D- to 3D-), leading to coverage of a broad descriptor space.

4.3. Splitting the Dataset into Training and External Sets and SFS (Subjective Feature Selection)

Subjective feature selection involves selection of appropriate number and set of molecular descriptors to build a model using suitable algorithm. Prior to SFS, it is essential to divide the dataset into training and test (also known as external or prediction set) sets with a proper composition and proportions to circumvent information leakage and to verify the

predictive ability of a model [11–13,16,18,25,26,29,30,37]. Hence, the dataset was randomly split into training (80% = 915 molecules) and prediction or external (20% = 226 molecules) sets. It is to be noted that the training set was used for the selection of optimum number of molecular descriptors, and the sole purpose of prediction/external set was to validate the external predictive ability of the model (Predictive QSAR). A GA-MLR-based QSAR model is free from over-fitting if it comprises an optimum number of molecular descriptors. Therefore, in the present work, a simple yet effective method of identifying the breaking point was used. Generally, the continuous inclusion of molecular descriptors in the GA-MLR model significantly increases the value of Q^2_{LOO} , but after the breaking point, the value of Q^2_{LOO} does not increase significantly [24]. The number of molecular descriptors corresponding to the breaking point was considered optimum for model building. A graph (see Figure 8) was plotted between the number of molecular descriptors involved in the model and Q^2_{LOO} values, which indicated that the breaking point agreed with the six molecular descriptors. Consequently, QSAR models comprising more than six descriptors were not considered. For SFS, the set of molecular descriptors was selected using the genetic algorithm integrated with multilinear regression (GA-MLR) method available in QSARINS-2.2.4 (generations per size: 10,000; population size: 50; mutation rate: 60; significance level: 0.05; fitness parameter: Q^2_{LOO}).

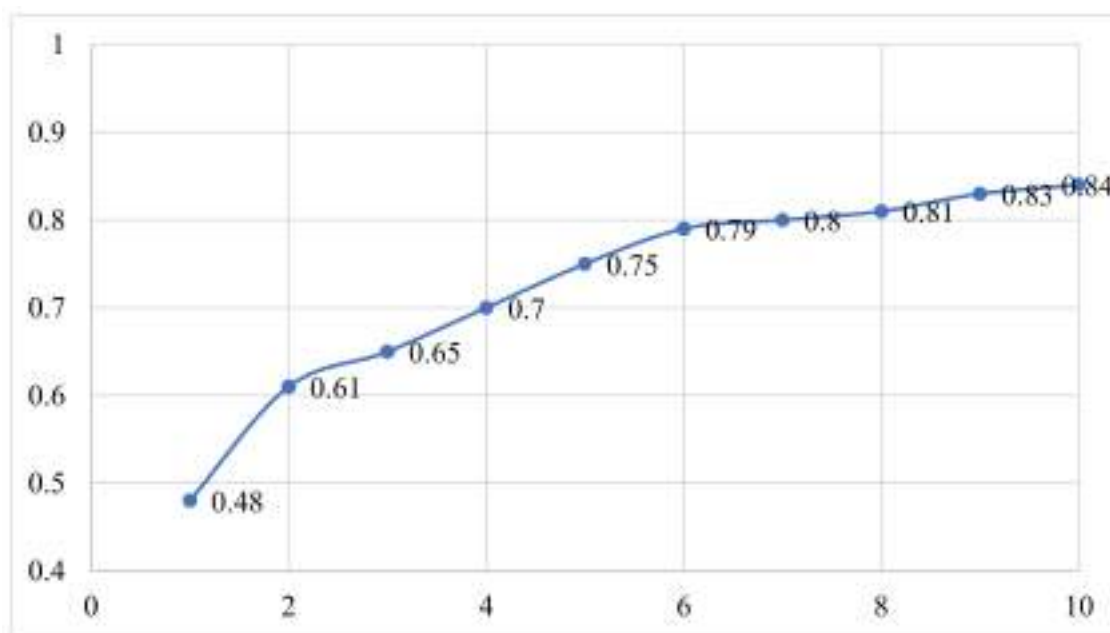


Figure 8. Plot of number of descriptors against leave-one-out coefficient of determination Q^2_{LOO} to identify the optimum number of descriptors.

4.4. Building Regression Model and Its Validation

The GA-MLR approach resulted in the generation of a good number of models having good to excellent statistical performance. Therefore, the following stringent parameters and criteria suggested by different researchers were used to select the best model [11–13,16,18,25,26,29,30,37,40]: $R^2_{tr} \geq 0.6$, $Q^2_{loo} \geq 0.5$, $Q^2_{LMO} \geq 0.6$, $R^2 > Q^2$, $R^2_{ex} \geq 0.6$, $RMSE_{tr} < RMSE_{cv}$, $\Delta K \geq 0.05$, $CCC \geq 0.80$, $Q^2-F^n \geq 0.60$, $r^2_m \geq 0.5$, $(1-r^2/r_o^2) < 0.1$, $0.9 \leq k \leq 1.1$ or $(1-r^2/r_o^2) < 0.1$, $0.9 \leq k' \leq 1.1$, $|r_o^2 - r_o'^2| < 0.3$, $RMSE_{ex}$, MAE_{ex} , R^2_{ex} , Q^2_{F1} , Q^2_{F2} , and Q^2_{F3} , and low R^2_{Yscr} , $RMSE$, and MAE . The details of these statistical parameters are available in the Supplementary Materials. An important aspect of validation of a QSAR model is to identify the applicability domain. In the present work, the William's plot was plotted to assess the applicability domain of the QSAR model [11–13,16,18,25,26,29,30,37,41,42].

5. Conclusions

In the present work, a relatively large and structurally diverse dataset of 1141 Hsp90 inhibitors was used for developing a six-descriptor-based and extensively validated GA–MLR QSAR model with $R^2_{tr} = 0.78$, $Q^2_{LMO} = 0.77$, $R^2_{ex} = 0.78$, and $CCC_{ex} = 0.88$. The inclusion of easily understandable descriptors resulted in identification of important pharmacophoric features that are correlated with Hsp90 inhibitory activity. The present QSAR analysis effectively captured a mixture of reported as well as novel significant structural features. The analysis vindicates that ring and aromatic carbons are important in deciding the activity. In addition, different types of nitrogen atoms in correlation with different types of carbon atoms influence the Hsp90 inhibitory activity. A good balance of external predictive ability and mechanistic interpretations, which are further supported by the reported crystal structures of Hsp90 inhibitors, make the QSAR model useful for the future optimization of molecules in the pipeline as a better Hsp90 inhibitor.

Supplementary Materials: The following are available online at <https://www.mdpi.com/article/10.3390/ph15030303/s1>.

Author Contributions: Conceptualization, V.H.M., M.E.A.Z. and S.A.A.-H.; formal analysis and data curation, V.H.M. and V.M.P.; writing, M.E.A.Z., V.H.M., M.M.R., S.N.A.B. and S.D.T.; Revisions, M.E.A.Z., V.H.M., S.D.T., M.M.R. and S.N.A.B.; editing and proofreading, V.H.M., M.E.A.Z. and S.N.A.B. All authors have read and agreed to the published version of the manuscript.

Funding: The authors acknowledge the Deanship of Scientific Research at Imam Mohammad Ibn Saud Islamic University, Riyadh, Saudi Arabia, for its support of this research through research group number RG-21-09-76.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data is contained within the article and Supplementary Materials.

Acknowledgments: The authors acknowledge the Deanship of Scientific Research at Imam Mohammad Ibn Saud Islamic University, Riyadh, Saudi Arabia, for its support of this research through research group number RG-21-09-76. V. H. Masand is thankful to Paola Gramatica (Italy) and her team for providing the free copy of QSARINS 2.2.4.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

SMILES	Simplified molecular-input line-entry system
GA	Genetic algorithm
MLR	Multiple linear regression
QSAR	Quantitative structure–activity relationship
WHO	World Health Organization
ADMET	Absorption, distribution, metabolism, excretion, and toxicity
OLS	Ordinary least square
QSARINS	QSAR Insubria
OECD	Organization for Economic Cooperation and Development

References

1. Ho, N.; Li, A.; Li, S.; Zhang, H. Heat Shock Protein 90 and Role of Its Chemical Inhibitors in Treatment of Hematologic Malignancies. *Pharmaceuticals* **2012**, *5*, 779–801. [[CrossRef](#)]
2. Li, L.; Wang, L.; You, Q.-D.; Xu, X.-L. Heat Shock Protein 90 Inhibitors: An Update on Achievements, Challenges, and Future Directions. *J. Med. Chem.* **2019**, *63*, 1798–1822. [[CrossRef](#)]
3. Bhat, R.; Tummalapalli, S.R.; Rotella, D.P. Progress in the Discovery and Development of Heat Shock Protein 90 (Hsp90) Inhibitors. *J. Med. Chem.* **2014**, *57*, 8718–8728. [[CrossRef](#)] [[PubMed](#)]

4. Zhao, H.; Moroni, E.; Colombo, G.; Blagg, B.S.J. Identification of a New Scaffold for Hsp90 C-Terminal Inhibition. *ACS Med. Chem. Lett.* **2013**, *5*, 84–88. [[CrossRef](#)]
5. Li, Y.; Zhang, T.; Schwartz, S.J.; Sun, D. New developments in Hsp90 inhibitors as anti-cancer therapeutics: Mechanisms, clinical perspective and more potential. *Drug Resist. Updates* **2009**, *12*, 17–27. [[CrossRef](#)]
6. Hoter, A.; El-Sabban, M.; Naim, H. The HSP90 Family: Structure, Regulation, Function, and Implications in Health and Disease. *Int. J. Mol. Sci.* **2018**, *19*, 2560. [[CrossRef](#)]
7. Zuehlke, A.D.; Moses, M.A.; Neckers, L. Heat shock protein 90: Its inhibition and function. *Philos. Trans. R. Soc. B Biol. Sci.* **2017**, *373*, 20160527. [[CrossRef](#)]
8. Biamonte, M.A.; Van de Water, R.; Arndt, J.W.; Scannevin, R.H.; Perret, D.; Lee, W.-C. Heat Shock Protein 90: Inhibitors in Clinical Trials. *J. Med. Chem.* **2009**, *53*, 3–17. [[CrossRef](#)] [[PubMed](#)]
9. Patil, V.M.; Masand, N.; Gupta, S.P.; Blagg, B.S.J. QSAR Studies to Predict Activity of HSP90 Inhibitors. *Curr. Top. Med. Chem.* **2021**, *21*, 2272–2291. [[CrossRef](#)] [[PubMed](#)]
10. Jhaveri, K.; Taldone, T.; Modi, S.; Chiosis, G. Advances in the clinical development of heat shock protein 90 (Hsp90) inhibitors in cancers. *Biochim. Biophys. Acta BBA Mol. Cell Res.* **2012**, *1823*, 742–755. [[CrossRef](#)] [[PubMed](#)]
11. Gramatica, P. Principles of QSAR Modeling. *Int. J. Quant. Struct.-Prop. Relatsh.* **2020**, *5*, 61–97. [[CrossRef](#)]
12. Cherkasov, A.; Muratov, E.N.; Fourches, D.; Varnek, A.; Baskin, I.I.; Cronin, M.; Dearden, J.; Gramatica, P.; Martin, Y.C.; Todeschini, R.; et al. QSAR modeling: Where have you been? Where are you going to? *J. Med. Chem.* **2014**, *57*, 4977–5010. [[CrossRef](#)]
13. Gramatica, P. On the development and validation of QSAR models. *Methods Mol. Biol.* **2013**, *930*, 499–526. [[CrossRef](#)]
14. Gramatica, P.; Cassani, S.; Roy, P.P.; Kovarich, S.; Yap, C.W.; Papa, E. QSAR Modeling is not Push a Button and Find a Correlation: A Case Study of Toxicity of (Benzo-)triazoles on Algae. *Mol. Inform.* **2012**, *31*, 817–835. [[CrossRef](#)]
15. Li, J.; Gramatica, P. The importance of molecular structures, endpoints' values, and predictivity parameters in QSAR research: QSAR analysis of a series of estrogen receptor binders. *Mol. Divers.* **2010**, *14*, 687–696. [[CrossRef](#)]
16. Muratov, E.N.; Bajorath, J.; Sheridan, R.P.; Tetko, I.V.; Filimonov, D.; Poroikov, V.; Oprea, T.I.; Baskin, I.I.; Varnek, A.; Roitberg, A.; et al. QSAR without borders. *Chem. Soc. Rev.* **2020**, *49*, 3525–3564. [[CrossRef](#)]
17. Fujita, T.; Winkler, D.A. Understanding the Roles of the “Two QSARs”. *J. Chem. Inf. Model.* **2016**, *56*, 269–274. [[CrossRef](#)]
18. Zaki, M.E.A.; Al-Hussain, S.A.; Masand, V.H.; Sabnani, M.K.; Samad, A. Mechanistic and Predictive QSAR Analysis of Diverse Molecules to Capture Salient and Hidden Pharmacophores for Anti-Thrombotic Activity. *Int. J. Mol. Sci.* **2021**, *22*, 8352. [[CrossRef](#)]
19. Zhao, H.; Moroni, E.; Yan, B.; Colombo, G.; Blagg, B.S.J. 3D-QSAR-Assisted Design, Synthesis, and Evaluation of Novobiocin Analogues. *ACS Med. Chem. Lett.* **2012**, *4*, 57–62. [[CrossRef](#)]
20. Barta, T.E.; Veal, J.M.; Rice, J.W.; Partridge, J.M.; Fadden, R.P.; Ma, W.; Jenks, M.; Geng, L.; Hanson, G.J.; Huang, K.H.; et al. Discovery of benzamide tetrahydro-4H-carbazol-4-ones as novel small molecule inhibitors of Hsp90. *Bioorg. Med. Chem. Lett.* **2008**, *18*, 3517–3521. [[CrossRef](#)]
21. Bussenius, J.; Blazey, C.M.; Aay, N.; Anand, N.K.; Arcalás, A.; Baik, T.; Bowles, O.J.; Buhr, C.A.; Costanzo, S.; Curtis, J.K.; et al. Discovery of XL888: A novel tropane-derived small molecule inhibitor of HSP90. *Bioorg. Med. Chem. Lett.* **2012**, *22*, 5396–5404. [[CrossRef](#)]
22. Abbasi, M.; Sadeghi-Aliabadi, H.; Amanlou, M. Prediction of new Hsp90 inhibitors based on 3,4-isoxazolidinamide scaffold using QSAR study, molecular docking and molecular dynamic simulation. *DARU J. Pharm. Sci.* **2017**, *25*, 17. [[CrossRef](#)]
23. Gramatica, P. External Evaluation of QSAR Models, in Addition to Cross-Validation Verification of Predictive Capability on Totally New Chemicals. *Mol. Inform.* **2014**, *33*, 311–314. [[CrossRef](#)]
24. Gramatica, P.; Chirico, N.; Papa, E.; Cassani, S.; Kovarich, S. QSARINS: A new software for the development, analysis, and validation of QSAR MLR models. *J. Comput. Chem.* **2013**, *34*, 2121–2132. [[CrossRef](#)]
25. Chirico, N.; Gramatica, P. Real external predictivity of QSAR models. Part 2. New intercomparable thresholds for different validation criteria and the need for scatter plot inspection. *J. Chem. Inf. Model.* **2012**, *52*, 2044–2058. [[CrossRef](#)]
26. Chirico, N.; Gramatica, P. Real external predictivity of QSAR models: How to evaluate it? Comparison of different validation criteria and proposal of using the concordance correlation coefficient. *J. Chem. Inf. Model.* **2011**, *51*, 2320–2335. [[CrossRef](#)]
27. Gramatica, P.; Pilutti, P.; Papa, E. Approaches for externally validated QSAR modelling of Nitrated Polycyclic Aromatic Hydrocarbon mutagenicity. *SAR QSAR Environ. Res.* **2007**, *18*, 169–178. [[CrossRef](#)]
28. Gramatica, P. Principles of QSAR models validation internal and external. *QSAR Comb. Sci.* **2007**, *26*, 694–701. [[CrossRef](#)]
29. Martin, T.M.; Harten, P.; Young, D.M.; Muratov, E.N.; Golbraikh, A.; Zhu, H.; Tropsha, A. Does rational selection of training and test sets improve the outcome of QSAR modeling? *J. Chem. Inf. Model.* **2012**, *52*, 2570–2578. [[CrossRef](#)]
30. Tropsha, A.; Gramatica, P.; Gombar, V.K. The Importance of Being Earnest Validation is the Absolute Essential for Successful Application and Interpretation of QSPR Models. *QSAR Comb. Sci.* **2003**, *22*, 69–77. [[CrossRef](#)]
31. Golbraikh, A.; Shen, M.; Xiao, Z.; Xiao, Y.D.; Lee, K.H.; Tropsha, A. Rational selection of training and test sets for the development of validated QSAR models. *J. Comput.-Aided Mol. Des.* **2003**, *17*, 241–253. [[CrossRef](#)]
32. Polishchuk, P. Interpretation of Quantitative Structure–Activity Relationship Models: Past, Present, and Future. *J. Chem. Inf. Model.* **2017**, *57*, 2618–2639. [[CrossRef](#)]
33. Jackson, S.E. Hsp90: Structure and Function. In *Molecular Chaperones*; Springer: Berlin/Heidelberg, Germany, 2012; pp. 155–240.

34. Vallée, F.; Carrez, C.; Pilorge, F.; Dupuy, A.; Parent, A.; Bertin, L.; Thompson, F.; Ferrari, P.; Fassy, F.; Lambertson, A.; et al. Tricyclic Series of Heat Shock Protein 90 (Hsp90) Inhibitors Part I: Discovery of Tricyclic Imidazo[4,5-c]pyridines as Potent Inhibitors of the Hsp90 Molecular Chaperone. *J. Med. Chem.* **2011**, *54*, 7206–7219. [[CrossRef](#)]
35. Baruchello, R.; Simoni, D.; Grisolia, G.; Barbato, G.; Marchetti, P.; Rondanin, R.; Mangiola, S.; Giannini, G.; Brunetti, T.; Alloatti, D.; et al. Novel 3,4-Isoxazolidiamides as Potent Inhibitors of Chaperone Heat Shock Protein 90. *J. Med. Chem.* **2011**, *54*, 8592–8604. [[CrossRef](#)]
36. Davies, N.G.M.; Browne, H.; Davis, B.; Drysdale, M.J.; Foloppe, N.; Geoffrey, S.; Gibbons, B.; Hart, T.; Hubbard, R.; Jensen, M.R.; et al. Targeting conserved water molecules: Design of 4-aryl-5-cyanopyrrolo[2,3-d]pyrimidine Hsp90 inhibitors using fragment-based screening and structure-based optimization. *Bioorg. Med. Chem.* **2012**, *20*, 6770–6789. [[CrossRef](#)]
37. Fourches, D.; Muratov, E.; Tropsha, A. Trust, but verify: On the importance of chemical structure curation in cheminformatics and QSAR modeling research. *J. Chem. Inf. Model.* **2010**, *50*, 1189–1204. [[CrossRef](#)]
38. O'Boyle, N.M.; Banck, M.; James, C.A.; Morley, C.; Vandermeersch, T.; Hutchison, G.R. Open Babel: An open chemical toolbox. *J. Cheminform.* **2011**, *3*, 33. [[CrossRef](#)]
39. Masand, V.H.; Rastija, V. PyDescriptor: A new PyMOL plugin for calculating thousands of easily understandable molecular descriptors. *Chemom. Intell. Lab. Syst.* **2017**, *169*, 12–18. [[CrossRef](#)]
40. Zaki, M.E.A.; Al-Hussain, S.A.; Masand, V.H.; Akasapu, S.; Bajaj, S.O.; El-Sayed, N.N.E.; Ghosh, A.; Lewaa, I. Identification of Anti-SARS-CoV-2 Compounds from Food Using QSAR-Based Virtual Screening, Molecular Docking, and Molecular Dynamics Simulation Analysis. *Pharmaceuticals* **2021**, *14*, 357. [[CrossRef](#)]
41. Kar, S.; Roy, K.; Leszczynski, J. Applicability Domain: A Step Toward Confident Predictions and Decidability for QSAR Modeling. In *Computational Toxicology*; Humana Press: New York, NY, USA, 2018; pp. 141–169.
42. Gramatica, P.; Kovarich, S.; Roy, P.P. Reply to the comment of S. Rayne on “QSAR model reproducibility and applicability: A case study of rate constants of hydroxyl radical reaction models applied to polybrominated diphenyl ethers and (benzo-)triazoles”. *J. Comput. Chem.* **2013**, *34*, 1796. [[CrossRef](#)]



Original article

Identification of potent aldose reductase inhibitors as antidiabetic (Anti-hyperglycemic) agents using QSAR based virtual Screening, molecular Docking, MD simulation and MMGBSA approaches

Ravindra L. Bakal^a, Rahul D. Jawarkar^{a,*}, J.V. Manwar^b, Minal S. Jaiswal^b, Arabinda Ghosh^c, Ajaykumar Gandhi^d, Magdi E.A. Zaki^{e,*}, Sami Al-Hussain^{e,*}, Abdul Samad^f, Vijay H. Masand^g, Nobendu Mukerjee^{h,i}, Syed Nasir Abbas Bukhari^j, Praveen Sharma^k, Israa Lewaa^l

^a Department of Medicinal Chemistry, Dr. Rajendra Gode Institute of Pharmacy, University-Mardi Road, Amravati, Maharashtra, India

^b Department of Medicinal Chemistry and Pharmacognosy, Dr. Rajendra Gode College of Pharmacy, University-Mardi Road, Amravati, Maharashtra, India

^c Microbiology Division, Department of Botany, Gauhati University, Guwahati, Assam 781014, India

^d Department of Chemistry, Government College of Arts and Science, Aurangabad, Maharashtra 431 004, India

^e Department of Chemistry, Faculty of Science, Al-Imam Mohammad Ibn Saud Islamic University, Riyadh 13318, Saudi Arabia

^f Department of Pharmaceutical Chemistry, Faculty of Pharmacy, Tishk International University, Erbil, Kurdistan Region, Iraq

^g Department of Chemistry, Vidyabharti Mahavidyalaya, Camp Road, Amravati, Maharashtra, India

^h Department of Microbiology, Ramakrishna Mission Vivekananda Centenary College, West Bengal 700118, Kolkata, India

ⁱ Department of Health Sciences, Novel Global Community Educational Foundation, Australia

^j Department of Pharmaceutical Chemistry, College of Pharmacy, Jouf University, Aljouf, Sakaka, 2014, Saudi Arabia

^k Department of Pharmaceutics, Vinayaka College of Pharmacy, Hathod, Indore, Madhya Pradesh, India

^l Department of Business Administration, Faculty of Business Administration, Economics and Political Science, British University in Egypt, Elshrouk City, Cairo 11837, Egypt

ARTICLE INFO

Article history:

Received 23 October 2021

Accepted 1 April 2022

Available online xxx

Keywords:

QSAR

Antidiabetic

Aldose reductase

GA-MLR

Molecular Docking

ABSTRACT

The aldose reductase (AR) enzyme is an important target enzyme in the development of therapeutics against hyperglycaemia induced health complications such as retinopathy, etc. In the present study, a quantitative structure activity relationship (QSAR) evaluation of a dataset of 226 reported AR inhibitor (ARI) molecules is performed using a genetic algorithm – multi linear regression (GA-MLR) technique. Multi-criteria decision making (MCDM) analysis furnished two five variables based QSAR models with acceptably high performance reflected in various statistical parameters such as, $R^2 = 0.79-0.80$, $Q_{LO}^2 = 0.78-0.79$, $Q_{MO}^2 = 0.78-0.79$. The QSAR model analysis revealed some of the molecular features that play crucial role in deciding inhibitory potency of the molecule against AR such as; hydrophobic Nitrogen within 2 Å of the center of mass of the molecule, non-ring Carbon separated by three and four bonds from hydrogen bond donor atoms, number of sp² hybridized Oxygen separated by four bonds from sp²

Abbreviations: Ari, Aldose Reductase Inhibitors; **QSAR**, Quantitative structure activity Relationship; **MD**, Molecular Dynamic; **MMGBSA**, Molecular mechanics generalized born surface area; **CADD**, Computer Aided Drug Designing; **SMILES**, Simplified Molecular, Input Line-Entry System; **GA**, Genetic Algorithm; **MLR**, Multiple Linear Regression; **QSAR**, Quantitative Structure-Activity Relationship; **OLS**, Ordinary Least Square; **QSARINS**, QSAR Insurbria; **OECD**, Organization for Economic Co-operation and Development; **CCC**, Concordance Correlation Coefficient; **allminus_SASA**, Solvent Accessible surface area of the all negatively charged atoms; **Minus_don_3B**, Occurrence of a donor within three bonds from a negatively charged atom; **don_ringC_6Ac**, Occurrence of the partially charged ring carbon atoms within 6Å of the donor; **don_notringC_4B**, Occurrence of the non-ring carbon atom within four bonds from the donor; **fsp20sp2C4B**, the frequency of occurrence of a sp² hybridized carbon atom exactly four bonds from a sp² hybridized oxygen atom; **com_ringCminus_2A**, Encodes information on the number of negatively charged ring Carbon atoms within 2Å from the center of mass of the molecule; **com_Nhyd_2A**, presence of hydrophobic nitrogen within 2Å of the center of mass; **H_ringN_2B**, H_ringN_2B represents a combination of ring nitrogen and hydrogen separated by within 2 bonds.

* Corresponding authors.

E-mail addresses: rbakal@gmail.com (R.L. Bakal), rahuljawarkar@gmail.com (R.D. Jawarkar), jvmanwar@gmail.com (J.V. Manwar), jaiswalminal77@gmail.com (M.S. Jaiswal), dra.ghosh@gauhati.ac.in (A. Ghosh), gascajay18@gmail.com (A. Gandhi), Mezaki@imamu.edu.sa (M.E.A. Zaki), sahussain@imamu.edu.sa (S. Al-Hussain), abdul.samad@tiu.edu.iq (A. Samad), vijaymasand@gmail.com (V.H. Masand), nabendu21@rkmvcraharaha.org (N. Mukerjee), sbukhari@ju.edu.sa (S. Nasir Abbas Bukhari), praveen140581@gmail.com (P. Sharma), israa.lewaaa@gmail.com (I. Lewaa).

Peer review under responsibility of King Saud University.



Production and hosting by Elsevier

<https://doi.org/10.1016/j.jsps.2022.04.003>

1319-0164/© 2022 The Author(s). Published by Elsevier B.V. on behalf of King Saud University.

This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Please cite this article as: R.L. Bakal, R.D. Jawarkar, J.V. Manwar et al., Identification of potent aldose reductase inhibitors as antidiabetic (Anti-hyperglycemic) agents using QSAR based virtual Screening, molecular Docking, MD simulation and MMGBSA approaches, Saudi Pharmaceutical Journal, <https://doi.org/10.1016/j.jsps.2022.04.003>

MDS
Virtual screening

hybridized Carbon atoms, etc. 14 in silico generated hits, using a compound 18 (a most potent ARi from present dataset with $pIC_{50} = 8.04$ M) as a template, on QSAR based virtual screening (QSAR-VS) furnished a scaffold 5 with better ARi activity ($pIC_{50} = 8.05$ M) than template compound 18. Furthermore, molecular docking of compound 18 (Docking Score = -7.91 kcal/mol) and scaffold 5 (Docking Score = -8.08 kcal/mol) against AR, divulged that they both occupy the specific pocket(s) in AR receptor binding sites through hydrogen bonding and hydrophobic interactions. Molecular dynamic simulation (MDS) and MMGBSA studies right back the docking results by revealing the fact that binding site residues interact with scaffold 5 and compound 18 to produce a stable complex similar to co-crystallized ligand's conformation. The QSAR analysis, molecular docking, and MDS results are all in agreement and complementary. QSAR-VS successfully identified a more potent novel ARi and can be used in the development of therapeutic agents to treat diabetes. © 2022 The Author(s). Published by Elsevier B.V. on behalf of King Saud University. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Introduction

The aldose reductase (AR) belongs to the Aldo-Keto Reductase superfamily (AKRs) composed of 190 enzymes. AR catalyzes the reduction of carbonyl substrates such as sugar aldehydes along with few other biomolecules and marks their role in lipid, carbohydrates, and xenobiotic metabolism (Jez et al., 1997). AR is the first, rate-limiting enzyme in the polyol pathway and it causes sorbitol accumulation in the insulin-independent tissues. Hence, AR has been linked to retinopathy – cataract in particular – and the pathophysiology of diabetes sequelae such as angiopathy, neuropathy, and nephropathy (Moon et al., 2006).

AR inhibitors reduce diabetes-related disorders, particularly in the tissues that exhibit insulin-independent glucose uptake such as; neural tissues, the lens, and glomeruli. Various AR inhibitors such as, alrestatin, benurestat, epalrestat, fidarestat, imirestat, lidorestat have been developed to treat secondary complications in diabetes (Krans 1993, van Gerven and Tjon-A-Tsien 1995, Tsai and Burnakis 2016). ONO Pharmaceuticals developed Epalrestat, the only AR inhibitor approved for the treatment of diabetic neuropathies in Japan, India, and China (Steele et al., 1993, Kucerova-Chlupacova et al., 2020). Many of them have been studied in clinical trials, but eventually discontinued due to no effect or harmful side effects such as fever, nausea, diarrhea, increased liver enzymes, rash including toxic epidermal necrolysis, Stevens-Johnson syndrome, marked thrombocytopenia, lymphadenopathy, splenomegaly and adult dyspnea syndrome (Foppiano and Lombardo 1997). Due to these concerns, the need for the development of the de novo ARi inhibitor has been increased.

In QSAR study, statistical approaches are used to identify a mathematical correlation between structural properties of similar molecules and their bioactivity. Traditional QSAR study employs a variety of scientific disciplines including chemistry, computer science, mathematics, statistics, and biology. Following are the stages in a standard QSAR analysis procedure: (1) to Assemblage molecules with the specified activity/property (referred as a dataset); (2) to draw structures in 2D, convert them into 3D and to optimize them using an appropriate force field; (3) to calculate of as large as possible number of and variety of molecular descriptors and subsequent data pruning using a suitable statistical method; (4) to employ an appropriate feature (molecular descriptor) selection algorithm to build a QSAR model; and (5) adequate validation of the developed QSAR model (Pourbasheer et al., 2015, Masand et al., 2016).

To find out structural and molecular features of the molecule that govern the expected activity of the molecule is the main goal of a QSAR study (i.e. descriptive QSAR). Whereas prediction of desired activity of the molecule prior to its wet lab synthesis and bio-testing is a secondary goal of the QSAR study (i.e. statistical QSAR) (Fujita and Winkler 2016). A good balance of descriptive and statistical elements in a QSAR model not only provides more

information about the structural configurations that have a positive association with an intended activity/property of a drug candidate, but it also improves understanding of the drug's mechanism of action. To find a potential hit as an AR inhibitor, we have used QSAR, QSAR-based virtual screening, molecular docking, MD simulation and MMGBSA analysis in this study.

2. Materials and methods

2.1. Preparation of data sets

A crude dataset of 432 compounds with experimental AR inhibitory potency measured in terms of IC_{50} values is procured from ChEMBL (Gaulton et al., 2017) database. After removing structural duplicates, multi-component compounds or salts, and compounds with imprecise IC_{50} values finally a limited data set of 226 AR inhibitors with accurate IC_{50} values is obtained. IC_{50} values in nanomolar (nM) units were first converted into molar (M) and then into pIC_{50} M ($pIC_{50} = -\log IC_{50}$ M) for ease of data set handling. (See supplementary information Table S1 for SMILES notations for all the 226 compounds with experimental IC_{50} and pIC_{50} M values).

2.2. Model development and validation

We did use ChemSketch to create 2D structures of all the 226 molecules, an Open Babel 3.1.1 programme (O'Boyle et al., 2011) to transfigure them into corresponding 3D structures, and a hyperchem programme (Ivanciuc 1996) to optimize these 3D molecular structures by employing the semi-empirical PM6 method (Bikadi and Hazai 2009). The resultant conformers were used for the calculation of the molecular descriptors using PyDescriptor available as a PyMOL plugin. More than 40,000 descriptors makeup the PyDescriptor, which covers topological, geometric, and constitutional chemical space of molecules (Masand and Rastija 2017).

Furthermore, the pre-filtration of descriptors (excluding semi-constant descriptors (greater than 80%) and highly intercorrelated descriptors (greater than 95%)) give rise to a contracted dataset of 2546 descriptors. The present data set of 226 molecules was divided into training (an 80% i.e. 181) for the QSAR model development and test sets (a 20% i.e. 45) for its validation. The QSAR Models were built by multiple linear regression. The descriptor selection for the training set was carried out by using the entire subset and Genetic function algorithm available in QSARINS 2.2.4 software (Gramatica 2020) (Dearden et al., 2009, Gramatica 2013, Cherkasov et al., 2014, Gramatica 2014, Fujita and Winkler 2016, Harit et al., 2017).

To ascertain the robustness of the develop model various validation criteria are reported in the literature. The internal predictivity and statistical quality of the built model is verified by the parameters like the coefficient of determination (r^2), the

leave-one-out cross-validation (Q_{LOO}^2), and leave-many-out cross-validation (Q_{LMO}^2). Additionally, for every developed model the standard error of estimate (s) is described. To add an extra portion of the accuracy for the reported QSAR models, Root Mean Squared Error (RMSE) for the training ($RMSE_{tr}$) and external prediction ($RMSE_{ext}$) set that account on the whole error of the model are also described. To ensure least possible inter-correlation among descriptors, the QUICK rule was fixed to 0.05. A Y-randomization test at 2000 iterations was performed to ascertain the reliability of the developed QSAR model and to rule out the possibility of any speculative correlation.

The external validation of all the models were verified with the subsequent validation criteria; r_{ext}^2 (external determination coefficient), Q_{F1}^2 , Q_{F2}^2 , Q_{F3}^2 , CCC_{ext} , r_m^2 , and Δr_m^2 . The parameter R^2_m (overall) penalizes a model for large differences between observed and predicted values of the compounds of the whole set (considering both training and test sets). The Δr_m^2 estimates the indulgence between the values of the predicted and the resultant experimental activity data (pIC_{50} M value). It has been suggested that the observed value for the r_m^2 should be greater than 0.5. All of the QSAR models were also tested for validation parameters, such as Golbraikh and Tropsha's criteria to justify model reliability and robustness.

Generally, good predictive ability of the developed QSAR model depends upon the closeness of predicted value against experimental biological activity value. Even, presence of a single outlier diminishes the predictive capability of the developed QSAR model. Subsequently, we have tried to highlight the outlier on the basis of those compounds who showed significantly high residual value in GA-MLR based QSAR models. Moreover, we have identified the outlier compounds by comparing the predicted value with the standardized residual values. Likewise, structural variation in

database compounds was observed by leverage effect in Williams plot. The applicability domain of the developed QSAR model is ascertained by merging the leverage and the standard residuals.

The multi criteria decision making (MCDM) function available in QSARINS v4.1.1 software is used to rank the developed models. It includes some criteria related to the external and internal validation whose values fall between 0 and 1, wherein 0 represents worst validation while 1 represents the best validation performance. The geometric average of all the values obtained in internal and external validation gives rise to MCDM values. The model with the high MCDM agreement among the best selected validation criteria is sorted as the best QSAR model for the analysis. In addition, the best model is evaluated for OECD (organization for economic corporation and development) guidelines (Gramatica et al., 2013, Gramatica et al., 2014, Consonni et al., 2019).

2.3. QSAR based virtual screening

In QSAR based virtual screening, we have carried out scaffold hopping by using the RDKit module. Herein, we have used the most active compound 18 as a template molecule to generate different variants. This has given rise to 14 different scaffolds with enhanced chemical space. Accordingly, 14 variants of compound 18 were used for QSAR-based virtual screening. Erstwhile to molecular descriptor calculations, the 3D- structures of the molecules were arranged in the same way as a modeling set. Then molecular descriptors were calculated and the appropriately validated five parametric QSAR model was used to envisage the biological property of novel compounds (Jawarkar et al., 2021). The structures of some representative scaffolds are given in Fig. 1.

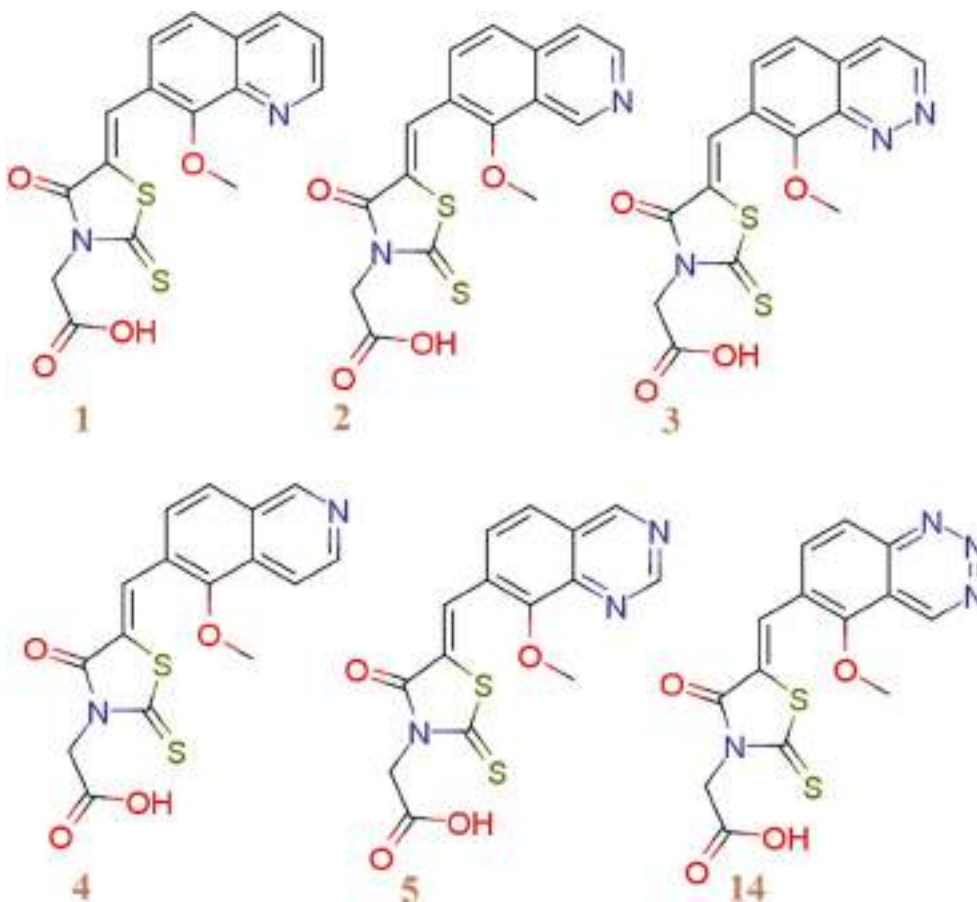


Fig. 1. Showing few representative examples of compounds generated by scaffold hopping.

(Depiction of 14 scaffolds, their smiles notations, calculated descriptors and predicted IC₅₀ Values is available as **Table S2** in [supplementary material](#)).

2.4. Molecular docking study experimental

The human AR in the pdb file was obtained from the Protein Data Bank (<https://www.rcsb.org/structure/1fzd>). The pdb:1fzd was selected on the origin of X-ray resolution and conclusion of the sequence. The optimized protein is suitable for docking analysis. All the active compounds were docked in the active site, but for the sake of ease, herein, the docking pose for most active molecule 18 as a symbol has been depicted.

The software NRGSuite was used to investigate molecular docking. This open source programme is available as a PyMOL plugin (<https://www.pymol.org>). It has the competence to ascertain the surface cavities in a protein and use them as target binding-sites for docking simulations with the aid of FlexAID. It practices genetic algorithms for conformational search, simulates ligand and side-chain flexibility and permits for the simulation of covalent docking. In the current work, flexible-rigid docking procedure was hired with succeeding default settings to get finest performance from NRGSuite: binding sites input method, spherical shape; spacing of three dimensional grid: 0.375 Å; side chain flexibility- no; ligand flexibility- yes; ligand pose as reference- no; constraints- no; HET groups- comprised water molecules; van der Waals permeability- 0.1; solvent types- no type; number of chromosomes- 1000; number of generations- 1000; fitness model- share; reproduction model- population boom; number of TOP complexes-5, Biovia Discovery studio software was used to visualize the docking results ([Gaudreault et al., 2015](#)).

2.5. MD simulations study

Based on the virtual screening results, the scaffold 5 with a docking score of -8.08 kcal/mol and Molecule 18 (-7.91 kcal/mol) is further investigated in molecular dynamics and simulation using the Schrodinger Desmond module (MD simulation) and MMGBSA binding free energy analysis. The protein ligand docking complexes of scaffold 5 and molecule 18 docking complexes were created using the Desmond module's SPC (Simple point charge) configuration. The OPLS-2005 force field ([Bowers et al., 2006](#)) and explicit solvent model with the SPC water molecules were used in this system ([Jorgensen et al., 1996](#), [Shivakumar et al., 2010](#)). Na⁺ ions were added to neutralize the charge. 0.15 M, NaCl solutions added to the system to simulate the physiological environment. The NPT ensemble was set up by using the Nose-Hoover chain coupling scheme ([Martyna et al., 1992](#)) with temperature 300 K, relaxation time of 1.0 ps and pressure 1 bar was maintained in all the simulations. A time step of 2 fs was used. The Martyna-Tuckerman-Klein chain coupling scheme ([Martyna et al., 1994](#)) barostat method was used for pressure control with a relaxation time of 2 ps. The particle mesh Ewald method ([Toukmaji and Board 1996](#)) was used for calculating long-range electrostatic interactions and the radius for the Coulomb interactions was fixed at 9 Å. RESPA integrator was used to calculate the non-bonded forces. The root mean square deviation (RMSD), root mean square fluctuation (RMSF), radius of gyration (Rg), protein ligand interactions were monitored to observe the stability of the complex in MD simulations.

2.6. Molecular mechanics generalized born and surface area (MMGBSA) calculations

During MD simulations of aldose reductase enzyme in complexed with molecule 18 and scaffold 5, the binding free energy

(Gbind) of docked complexes was calculated using the premier molecular mechanics generalized Born surface area (MM-GBSA) module (Schrodinger suite, LLC, New York, NY, 2017-4). The binding free energy was calculated using the OPLS 2005 force field, VSGB solvent model, and rotamer search methods ([Jawarkar et al., 2022](#), [Kumar et al., 2022](#)). After the MD run, 10 ns intervals were used to choose the MD trajectories frames. The total free energy binding was calculated using equation (1):

$$\Delta G_{bind} = G_{complex} - (G_{protein} + G_{ligand}) \quad (1)$$

where, ΔG_{bind} = binding free energy, $G_{complex}$ = free energy of the complex, $G_{protein}$ = free energy of the target protein, and G_{ligand} = free energy of the ligand. The MMGBSA outcome trajectories were analyzed further for post dynamics structure modifications.

2.7. In silico ADMET property prediction

The ADME (Absorption, Distribution, Metabolism, and Excretion) properties of the chosen phytochemical were calculated using the SWISS adme web server (<https://www.swissadme.ch/>). The molecular as well as pharmacokinetic behaviour of molecule 18 and scaffold 5 were assessed for drug like properties, total polar surface area (TPSA), water solubility, blood brain barrier (BBB) permeant, gastro intestinal (GI) absorption, Lipinski, Ghose, Veber, Egan and Muegge violations and synthetic accessibility. If the test molecules e.g. molecule 18 and scaffold 5 pass the drug like filters therefore, it can be predicted to be candidate molecule against aldose reductase. For the analysis of ADME, the SMILES format of molecule 18 and scaffold 5 have been used individually to retrieve the outcomes from the server.

3. Results

In the present investigation, QSAR analysis is performed using a dataset composed of 226 ARI inhibitors with reported IC₅₀ values and molecular docking is accomplished to pinpoint the important structural features. The QSAR model was developed by using easily interpretable molecular descriptors to correlate them with structural features. The five parametric GA-MLR based QSAR model has decent external predictive ability with the incidence of easily comprehensible molecular descriptors alongside the interpretation in terms of structural features. For the aim of model selection, Multi-Criteria Decision Making (MCDM) analysis has been employed in this study. As a result of the MCDM study, a robust QSAR model was chosen for the analysis. The findings are given below.

3.1. Multi-Criteria decision making (MCDM) analysis

In the current QSAR analysis, we have implemented Multi-Criteria Decision Making (MCDM) ([Pavan and Todeschini 2009](#)) technique that involves combining the performances of a certain number of criteria simultaneously, as a single number (score) between 0 and 1.

The MCDM is achieved by associating to each validation criteria a desirability function whose value falls in the range from 0 to 1 (where 0 designates the worst validation criteria value and 1 the best). Using the geometric average of all the values obtained from the desirability functions gives the MCDM value. By default, the MCDM of fitting (maximizing R², R_{adj}² and CCC_{tr}, while minimizing R²-R_{adj}²), cross validation (maximizing Q_{LOO}², Q_{TMO}² and CCC_{cv}, while minimizing R_{scr}²) and external validation (maximizing Q_{F1}², Q_{F2}² and CCC_{EXT}), are automatically calculated using all the above criteria in QSARINS v2.2.4 programme. If any one of the criteria is missing, then the MCDM model will not be obtained. The model with the best

Table 1
Display of five variable models in MCDM.

Model id	Size	Variables/Descriptor
134	5	com_ringCminus_2A H_ringN_2B com_Nhyd_2A notringC_don_3B allminus_SASA
133	5	H_ringN_2B lipo_don_3Bc com_Nhyd_2A notringC_don_5B allminus_SASA
132	5	H_ringN_2B lipo_don_3Bc com_Nhyd_2A notringC_don_3B allminus_SASA
129	5	minus_H_3B H_ringN_2B com_Nhyd_2A notringC_don_3B allminus_SASA
131	5	all_HASA3 H_ringN_2B com_Nhyd_2A notringC_don_3B allminus_SASA
130	5	all_HASA3 H_ringN_2B com_Nhyd_2A notringC_don_5B allminus_SASA
128	5	H_ringN_2B ringN_H_8Ac com_Nhyd_2A notringC_don_5B allminus_SASA
127	5	H_ringN_2B ringN_H_8Ac com_Nhyd_2A notringC_don_3B allminus_SASA
126	5	H_ringN_2B ringN_H_5Ac com_Nhyd_2A notringC_don_5B allminus_SASA

MCDM criteria is selected for analysis (See Table 1). In the present analysis, model no 134 has satisfied the best MCDM criteria (See Table 2 for various MCDM parameters), therefore it is selected for the analyses (see Fig. 2).

In Fig. 3, model no 135 depicts very good fitting performance but it is lacking in external predictive performance. For instance, model no 133 showed very good external predictive performance but not satisfactory in fitting. Therefore, model no 134 in between 133 and 135 showed better compromise between fitting and predictivity, hence selected as best model for analysis and QSAR based virtual screening. Moreover, apart from model no 134, we have developed a full set model.

3.2. QSAR model 1.1 (Model no 134)

$pIC_{50} = 2.603 (\pm 0.447) - 0.193 (\pm 0.085) \text{ com_ringCminus_2A} + 0.249 (\pm 0.074) \text{ H_ringN_2B} - 1.331 (\pm 0.226) \text{ com_Nhyd_2A} - 0.34 (\pm 0.113) \text{ notringC_don_3B} + 0.012 (\pm 0.001) \text{ allminus_SASA}$.

(The experimental and predicted pIC_{50} M value for the divided dataset model is available as Table no S3 in supplementary material).

$R_2: 0.80, R_{adj}^2: 0.79, R^2 - R_{adj}^2: 0.01, LOF: 0.26, K_{xx}: 0.20, \Delta K: 0.09, RMSE_{tr}: 0.49, MATer: 0.40, RSS_{tr}: 42.75, CCC_{tr}: 0.89, s: 0.50, F: 140.68, Q_{loo}^2: 0.79, R^2 - Q_{loo}^2: 0.02, RMSE_{cv}: 0.50, MAE_{cv}: 0.41, PRESS_{cv}: 45.25, CCC_{cv}: 0.88, Q_{LMO}^2: 0.79, R^2 Y_{scr}: 0.03, Q_2 scr: -0.04, RMSE AV Y_{scr}: 1.07, RMSE_{ext}: 0.54, MAE_{ext}: 0.44, PRESS_{ext}: 12.98, Q_{F1}^2: 0.80, Q_{F2}^2: 0.79, Q_{F3}^2: 0.75, CCC_{ext}: 0.88, r^2 m aver.: 0.68, r^2 m delta: 0.18.$

3.3. QSAR model 1.2 (Full set model)

$pIC_{50} = 2.619 (\pm 0.399) - 0.085 (\pm 0.033) \text{ minus_don_3B} + 0.431 (\pm 0.086) \text{ don_ringC_6Ac} - 0.353 (\pm 0.166) \text{ don_notringC_4B} - 0.584 (\pm 0.123) \text{ fsp2Osp2C4B} + 0.012 (\pm 0.001) \text{ allminus_SASA}$ (The experimental and predicted pIC_{50} M value for the full dataset model is available as table no S4 in supplementary material).

$R^2 = 0.79, R_{adj}^2 = 0.79, R^2 - R_{adj}^2 = 0.01, LOF = 0.28, K_{xx} = 0.16, \Delta K = 0.14, RMSE_{tr} = 0.50, MAE_{tr} = 0.40, RSS_{tr} = 57.41, CCC_{tr} = 0.89, s = 0.51, F = 168.60, Q_{loo}^2 = 0.78, R^2 - Q_{loo}^2 = 0.01, RMSE_{cv} = 0.52, MAE_{cv} = 0.41, PRESS_{cv} = 60.39, CCC_{cv} = 0.88, Q_{LMO}^2 = 0.78, R^2 Y_{scr} = 0.02, Q^2 Y_{scr} = -0.03, RMSE AV Y_{scr} = 1.10.$

The statistical validation parameters listed above are recommended for judging internal and external robustness, and they have the same meaning as before (see Supplementary Material Table S5 for detailed formulae and Table S6 for the detailed

descriptions of the descriptors). The high value of R_{tr}^2 (coefficient of determination), R_{adj}^2 (adjusted coefficient of determination), R_{cv}^2 (Q_{loo}^2) (cross-validated coefficient of determination for leave-one-out), R_{ex}^2 (external coefficient of determination), Q_{fn}^2 , and CCC_{ex} (concordance correlation coefficient), and the low value of LOF (lack-of-fit), $RMSE_{tr}$ (root mean square error). The various graphs associated with the model (see Figs. 3 and 5), such as MAE_{tr} (mean absolute error), R_{Yscr}^2 (R^2 for Y-scrambling), and others, show that the model is statistically robust, with high internal and external predictive capacity, and is free of chancy correlation. Furthermore, the Williams plot (see Fig. 4) demonstrates that the model is statistically acceptable. Moreover, insubris As a result, it complies with all of the OECD's suggested standards for developing an effective QSAR model.

4. Discussion

4.1. Allminus_SASA

The solvent accessible surface area, integrates shape and electrical property which additionally can be contributed by the presence of hydrogen bond donor/acceptor atoms (heteroatoms, such as N, O, etc.) in a molecule and is determined by mapping atomic partial charges. Owing to the electronegativity difference between such heteroatoms and carbon/hydrogens there occurs a partial charge separation, subsequent dipole formation, thereby intensifying the drug - water solubility and facilitating drug - biological target interaction, ultimately. The molecular descriptor, allminus_SASA (solvent accessible surface area (\AA^2) of all negatively charged atoms) is in a positive relationship with pIC_{50} M (positive coefficient in both QSAR equation) and hence, increase in its value may possibly lead to better AR inhibitory activity.

This can be well illustrated by comparing compound 190 (allminus_SASA = 131.34, Number hydrogen bond donor/acceptor = 2; pIC_{50} = 4.0 M) and 16 (allminus_SASA = 411.21, Number hydrogen bond donor/acceptor = 7; pIC_{50} = 7.48 M) wherein possibly, for about three fold increase in allminus_SASA value leads to more than 3000 fold increase in AR inhibitory potency. Illustrative Fig. 6 depicts the allminus_SASA.

4.2. Minus_don_3B

The minus_don_3B (presence of a donor within three bonds from a negatively charged atom) has a negative relationship with the inhibitory activity of AR and its value should be kept as low as possible. The importance of this bioactivity governing feature can be supported by comparing compound 20 (minus_don_3B = 0; pIC_{50} = 6.40 M) with the compound 172 (minus_don_3B = 12; pIC_{50} = 4.69 M). Both, compounds 20 and 172 constitute a comparable number of hydrogen bond donors/acceptors which are expected to bring very close AR inhibitory potency to both (as explained for allminus_SASA in preceding section). But contrast observation in case of the compound 20 and 172 pair, highlights the importance of absence of not all but specific donor atoms or negatively charged carbons, defined by minus_don_3B. Furthermore, the lone pair of electrons, associated with the donor NH & OH groups can be deployed to make a covalent link with a biological target. Plausibly, a good compromise between phenyl and moderately long chain ether functionality induced lipophilicity and donor atoms induced hydrophilicity, enhanced the AR inhibitory potency of compound 20 (see Fig. 7).

4.3. Don_ringC_6Ac

The Don_ringC_6Ac (presence of partially charged ring carbon atoms within 6 \AA of the donor) being having positive correlation

Table 2
Presentation of different MCDM parameters.

Model id	R^2	R_{adj}^2	Q_{loo}^2	Q_{LMO}^2	CCC_{tr}	CCC_{cv}	Q_{f1}^2	Q_{f2}^2	Q_{f3}^2
134	0.7999	0.7942	0.7881	0.7888	0.8888	0.8824	0.7975	0.7935	0.7488
133	0.7984	0.7926	0.7848	0.7821	0.8879	0.8805	0.8164	0.8128	0.7724
132	0.7965	0.7908	0.7838	0.7803	0.8867	0.8798	0.8163	0.8127	0.7722
129	0.7953	0.7895	0.7816	0.7782	0.886	0.8785	0.7435	0.7385	0.682
131	0.7951	0.7893	0.7833	0.7814	0.8858	0.8794	0.7945	0.7904	0.7451
130	0.7947	0.7889	0.7817	0.7766	0.8856	0.8785	0.7882	0.784	0.7373
128	0.7942	0.7883	0.7802	0.7773	0.8853	0.8777	0.7859	0.7817	0.7345
127	0.7933	0.7874	0.78	0.7767	0.8847	0.8775	0.7873	0.7831	0.7362
126	0.7927	0.7869	0.7782	0.7759	0.8844	0.8765	0.7814	0.7771	0.729

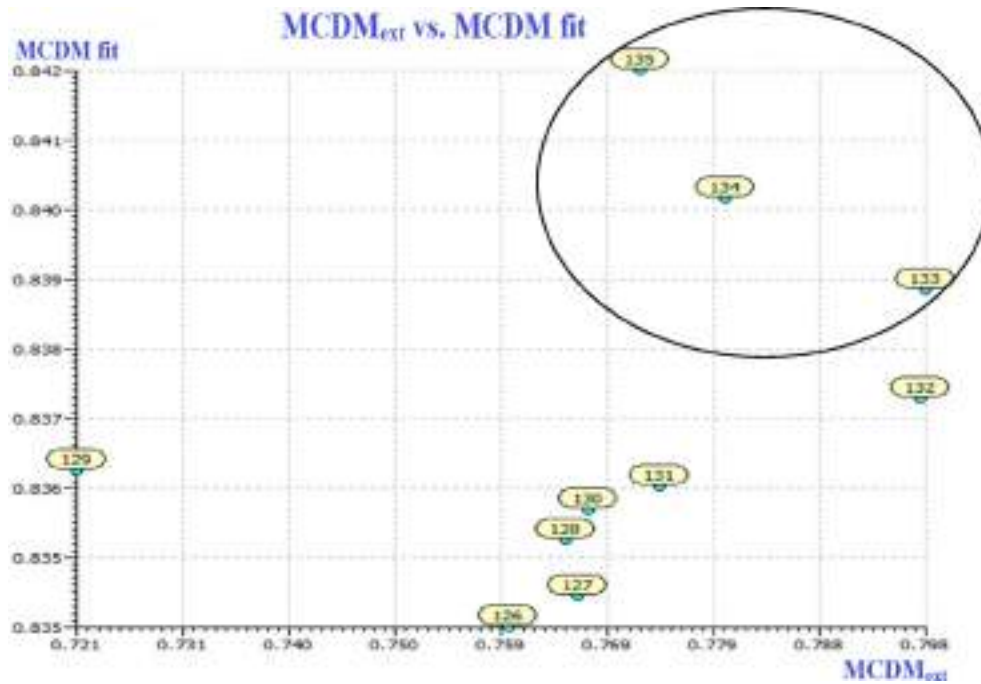


Fig. 2. Depiction of MCDM graph for model 135. The \times axis indicates MCDM fit while y axis shows MCDM ext (Black circle showing three best models).

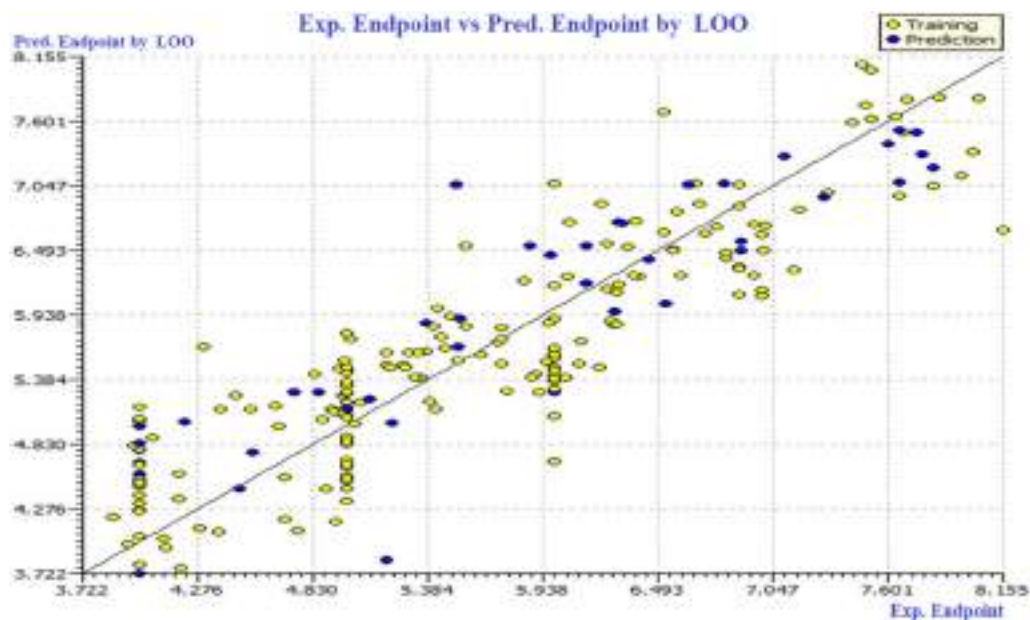


Fig. 3. Depiction of Scatter plot of experimental vs. predicted data by LOO.

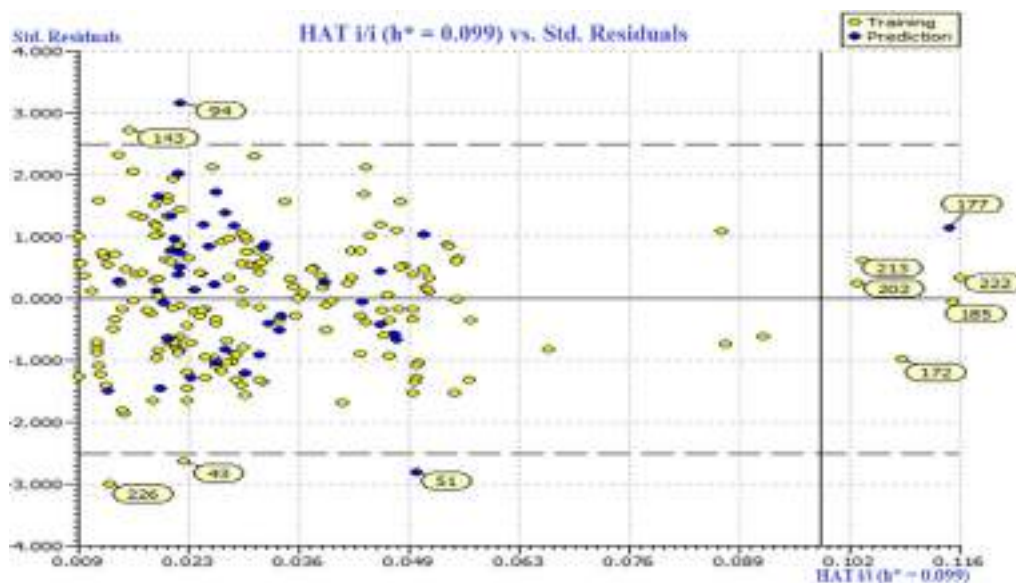


Fig. 4. Display of Williams plot using data predicted by LOO. (Molecules out of applicability domain have been shown with their serial numbers).

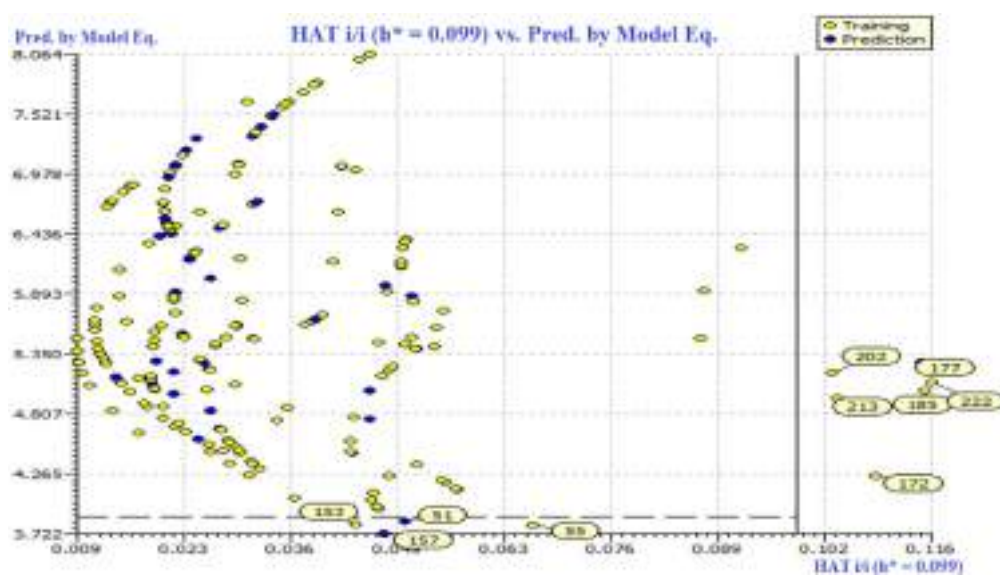


Fig. 5. Insurbia plot for Divided set model 134. (Molecules out of applicability domain have been shown with their serial numbers).



Fig. 6. Depiction of the descriptor allminus_SASA for compound 16 & 190. (The negatively charged atoms are depicted by a bold red color).

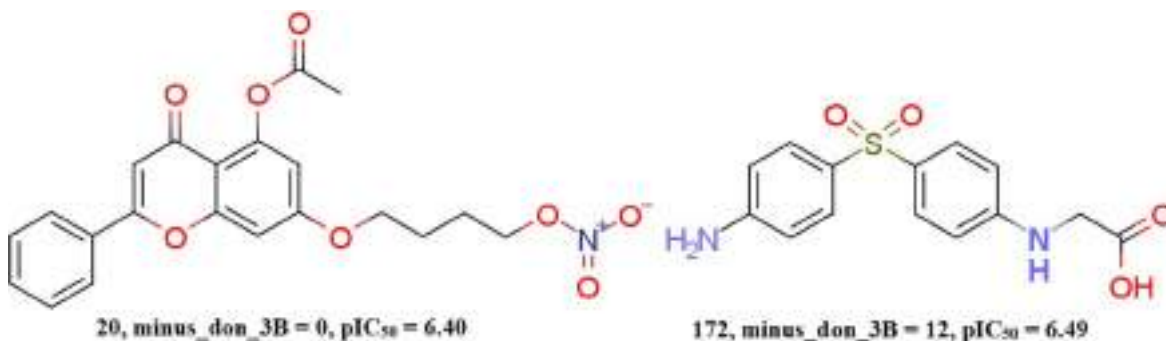


Fig. 7. Depiction of the descriptor minus_don_3B for the compound 20 and 172 only. (Donors are indicated by the dotted circle while negatively charged atoms are shown by red bold color).

with the pIC₅₀ (M), high possible value of don_ringC_6Ac is advisable for better AR inhibitor lead optimization. In Compound 85 (don_ringC_6Ac = 2.49, number of H-bond donor/acceptor = 9; pIC₅₀ = 7.52 M), a blend of hydrogen bond donors capable of forming electrostatic interactions, with equally important lipophilic carbons (in unsaturated region of the isoquinoline and the pyrrolidine ring) enabling lipophilic interactions of molecule, for AR inhibition makes it 2.83 unit more potent than compound 103 (don_ringC_6Ac = 0, number of H-bond donor/acceptor = 7; pIC₅₀ = 4.59 M) (see Fig. 8).

4.4. Don_notringC_4B

The don_notringC_4B (non-ring carbon atom within four bonds from the donor) has negative correlation with biological activity and hence minimum possible value of this molecular descriptor while AR inhibitor lead optimization is recommended. Absence of non-ring carbon atoms within four bonds from the donor atom probably makes compound 46 about 660 times more potent than compound 185 where there are two such non-ring carbons present (Fig. 9). In corollary, a non-ring carbon atom within a molecule must be placed at a distance of five or more bonds from the donor atoms to develop leads with desired potency.

4.5. fsp2Osp2C4B

A molecular descriptor fsp2Osp2C4B (the frequency of occurrence of a sp² hybridized carbon atom exactly four bonds from a sp² hybridized oxygen atom) has a negative correlation with biological activity and keeping its value minimum possible is advisable for better AR inhibitory potency. Absence of this specific

combination of sp²-O a distance of four bonds from the sp²-C atom in compound 37 possibly make it more potent than compound 1 in which such a combination is observed twice. (Fig. 10). Furthermore, overall lipophilicity of the molecule to the large extent is contributed by unsaturated ring carbons. Importance of lipophilicity in molecules for its AR inhibitory potency highlighted in foregone discussion support and supported by the importance of absence of not all but specific combination of sp²C and sp²O.

4.6. com_ringCminus_2A

com_ringCminus_2A encodes information on the number of negatively charged ring Carbon atoms within 2 Å from the center of mass of the molecule. AR inhibitory potency is inversely proportional to the value of this molecular feature and hence, low possible value of com_ringCminus_2A in a molecule is advisable for better AR inhibitory potency. This observation can be supported by Comparison of 203 (com_ringCminus_2A = 4; pIC₅₀ = 4.000 M) with 109 (com_ringCminus_2A = 0; pIC₅₀ = 7.95 M) wherein decrease in value of com_ringCminus_2A from 4 in compound 203 to 0 in compound 109 cause about 10,000 fold increase in AR inhibitory activity (see Fig. 11).

Generally, molecules in the receptor pocket should be more steadily balanced if its center of mass is below the balance point, but unstable if it is above the balance point. The appropriate orientation of the molecule within the active site pocket is determined by the molecule's center of mass. This finding supports the idea that having a larger number of negatively charged carbon atoms within the 2 Å radius can cause problems with molecular alignment within the receptor active site pocket. This could be the cause



Fig. 8. Representation of descriptor Don_ringC_6Ac for the compound 85 and 103 only. (Donors are highlighted by a dotted circle while carbon atoms within 6 Å are indicated by red bold color).

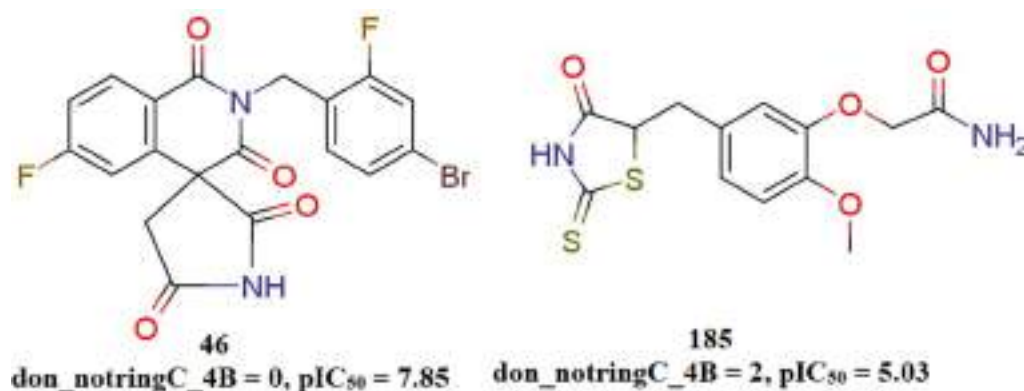


Fig. 9. Depiction of descriptor don_notringC_4B for the compound 46 and 185.

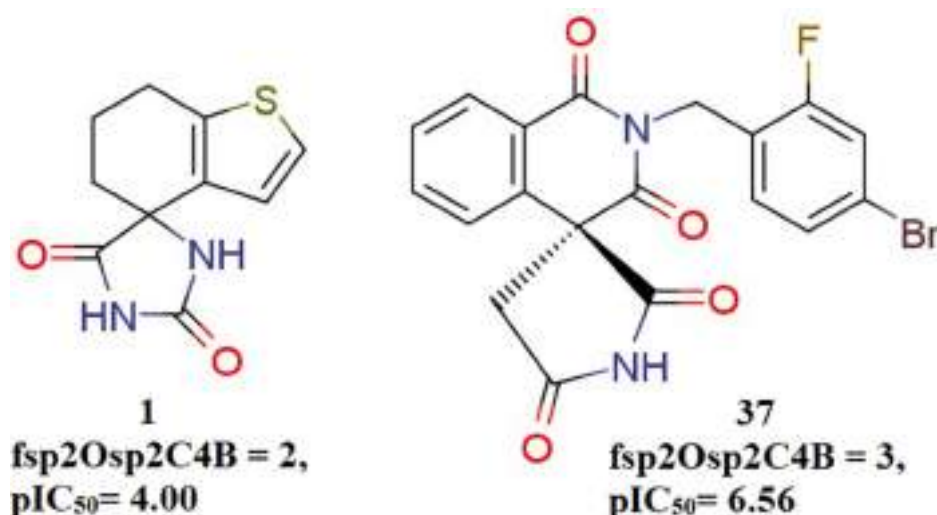


Fig. 10. pictorial depiction of descriptor fsp2Osp2C4B for the compounds 1 and 37.

of the differences in biological activity between molecules 203 and 109 (see Fig. 11).

4.7. com_Nhyd_2A

In the established QSAR model, the molecular descriptor com_Nhyd_2A (presence of hydrophobic nitrogen within 2 Å of the center of mass) is negatively connected with biological activity and hence it is advisable to abstain from introduction of such hydrophobic nitrogen (partial charge in between -0.2 to $+0.2$) to avoid deterioration of biological activity. Here, in compound 57, one hydrophobic nitrogen is present with a charge of -0.09 while another nitrogen has a charge of 0.06 . (see Fig. 12).

The poor activity for the molecule 16, 22, 36, 26, 32, 33 and 51 can be correlated with the high value of such hydrophobic Nitrogens within 2Å⁰ from the center of mass of the molecule (com_Nhyd_2A = 1).

4.8. H_ringN_2B

The descriptor H_ringN_2B represents a combination of ring nitrogen and hydrogen separated by within 2 bonds. This descriptor has a positive coefficient in the QSAR model; therefore the combination of ring nitrogen within 2 bonds from the hydrogen

element is the favorable combination to be used for hit/lead optimization. (see Fig. 13).

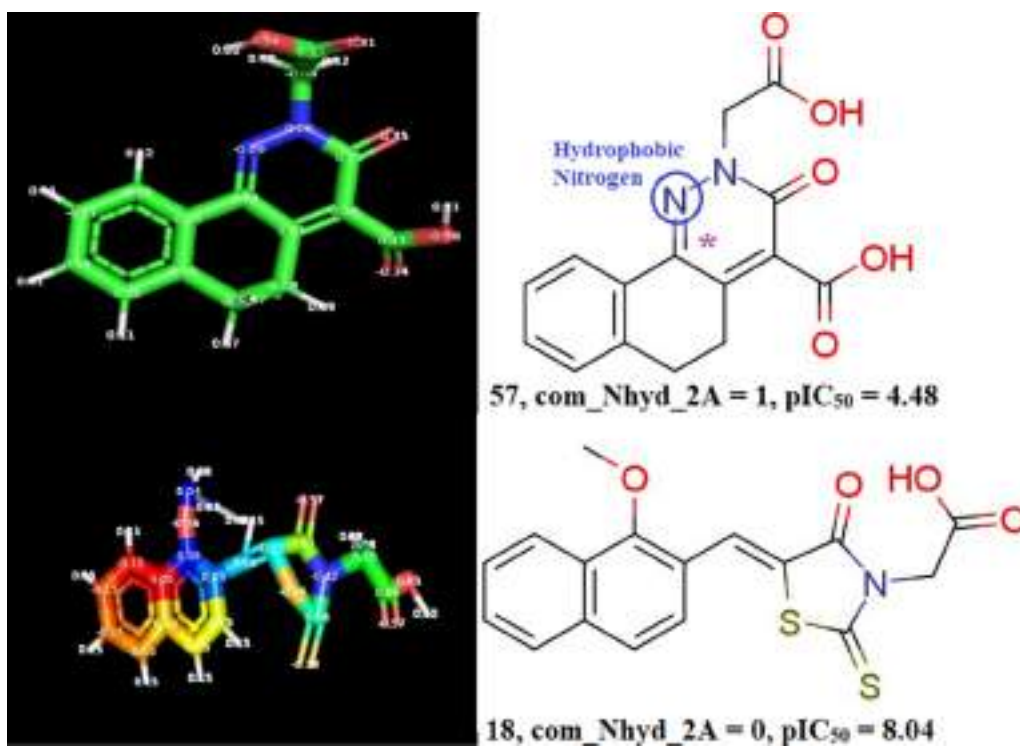
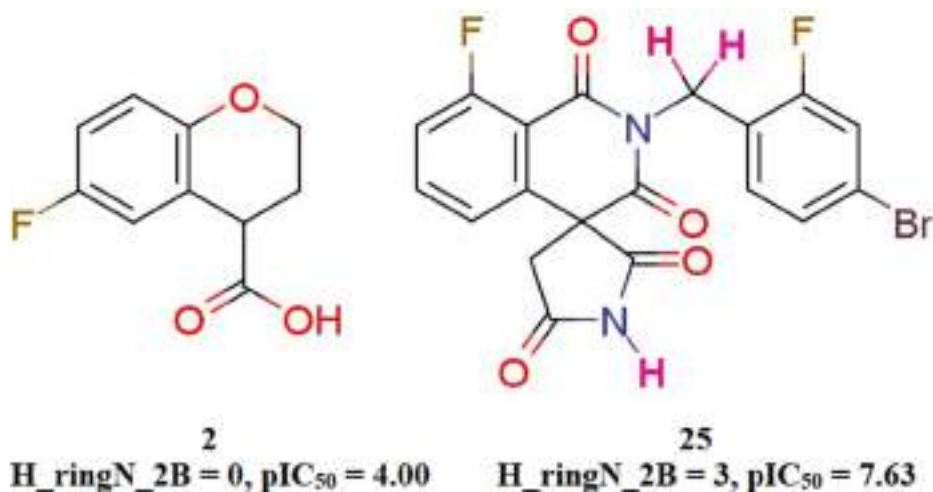
Generally, Hydrogen is the smallest element, it implies that there should be minimum bulk in the vicinity of ring Nitrogen atoms. Therefore, in future modifications, steric bulk nearer to ring Nitrogen atoms should be avoided to have better activity against AR.

A poor activity profile of molecules 112, 113, 100, 102, 118 and 133 is due to absence of such hydrogen and ring nitrogen combinations within 2 bonds ($H_{ringN_2B} = 0$), therefore higher the number of such combinations, greater will be AR inhibitory activity. This may be a plausible reason for the difference in the biological activity of these molecules.

4.9. QSAR based virtual screening

The QSAR based virtual screening was performed by employing a divided set QSAR model.

Among these 14 scaffolds furnished by QSAR-VS, scaffold 5 turned out as the more potent AR inhibitor than a template compound 18. Close observation of the values of molecular features for both scaffold-5 and compound 18 revealed that the large value of two molecular descriptors viz. number of Hydrogen atoms that are two bond away from ring Nitrogen atoms (H_{ringN_2B}) and solvent accessible surface area of all the negatively charged atoms

Fig. 11. Presentation of the descriptor com_ringCminus_2A for the compound 203 and 109.Fig. 12. Representation of the descriptor com_Nhyd_2A for the compound 57 and 18 (Star in the figure indicates center of mass of the molecules).Fig. 13. Depiction of the descriptor H_ringN_2B for the compound 2 and 25.

(allminus_SASA) which are in positive correlation with AR inhibitory potency suffice to enhance the AR inhibitory potency of the scaffold-5 ($H_{ringN_2B} = 4$, allminus_SASA = 353.8; $pIC_{50} = 8.05$ M) over template compound-18 ($H_{ringN_2B} = 2$, allminus_SASA = 337.4; $pIC_{50} = 8.04$ M). In other scaffolds, increase in either H_{ringN_2B} or allminus_SASA or both is not as in case of scaffold-5. Hence, the synchronous effect of H_{ringN_2B} and allminus_SASA is the possible reason for the increase in the potency of AR inhibitor (scaffold 5).

4.10. Molecular docking analysis of molecule 18

Human AR (Pdb-1fzd), a 36 kDa enzyme with a supple and malleable active site, is the target protein under consideration. The enzyme folds into a TIM-barrel and is thought to be involved in diabetic retinopathy and angiopathy. (Docking score, RMSD and predicted pIC_{50}^M values are displayed in Table 3).

As a result, it was created to be a promising drug target. It uses NADPH as a reducing cofactor to convert various aldehydes (including glucose in diabetic circumstances) to their corresponding alcohols. NADPH contributes a hydride ion to the carbonyl carbon of the aldehydes, resulting in a negatively charged intermediate,

although the exact mechanism is still debated. Furthermore, a subsequent proton transfer from one of the nearby acidic active site residues is used to trace this step. The binding site consists of two sub-pockets, one encompassing the residues possibly involved in catalysis (Tyr48, Lys77, and His110) sideways with the nicotinamide moiety of the cofactor, while the second so-called specificity pocket is formed by Trp111, Ala299, Leu300 and Phe122.

With a docking score of -7.912 kcal/mol, molecule 18 orients in the specificity pocket in the same way as the pdb-1fzd ligand (see Fig. 14). Through typical hydrogen bonding, carbon hydrogen bonding, and hydrophobic interactions, molecule 18 forms a complex with human AR. Molecule 18 resides in the second pocket, known as the specificity pocket, in which the 4-oxothiazolidine oxygen acts as an acceptor, forming a conventional hydrogen bond with the NH2 hydrogen atom of specificity pocket residue A: LEU300 (1.90 Å) and so acting as a hydrogen bond donor. Further the same 4-oxo thiazolidine oxygen acceptor atom anchored carbon hydrogen bond with hydrogen atom of ALA299 (2.74 Å) residue therefore, acts as hydrogen bond donor in the drug receptor interaction. (see Fig. 15).

Furthermore, molecule 18's terminal acceptor carboxyl oxygen forms a carbon hydrogen bond with the specificity pocket A:

Table 3

Depiction of docking Results along with pIC_{50} M and IC_{50} by QSAR based virtual screening for the Molecule 18 and the series of 14 Scaffolds.

Sn	DockingScore (kcal/mol)	RMSD	pIC_{50} M	IC_{50} in nM
Molecule 18	-7.91	0.95659	8.04	9.12
Scaffold 1	-7.89	1.02	6.21	616.5
Scaffold 2	-7.69	1.28	7.83	14.79
Scaffold 3	-7.86	1.41	7.61	24.54
Scaffold 4	-7.81	1.65	7.83	14.79
Scaffold 5(Pose 1)	-8.08	1.29	8.05	8.91
Scaffold 5(Pose 2)	-8.00	0.88		
Scaffold 6	-7.91	1.20	7.84	14.4
Scaffold 7 (Pose 1)	-8.06	1.11	7.83	14.7
Scaffold 7(Pose 2)	-8.04	1.25		
Scaffold 8	-7.87	1.82	7.52	30.2
Scaffold 9	-7.91	1.44	7.79	16.2
Scaffold 10	-7.83	1.18	7.97	10.7
Scaffold 11	-7.91	0.81	7.74	18.1
Scaffold 12	-7.96	1.50	7.61	24.5
Scaffold 13	-8.01	1.21	7.75	17.7
Scaffold 14	-8.00	0.99	7.83	14.7

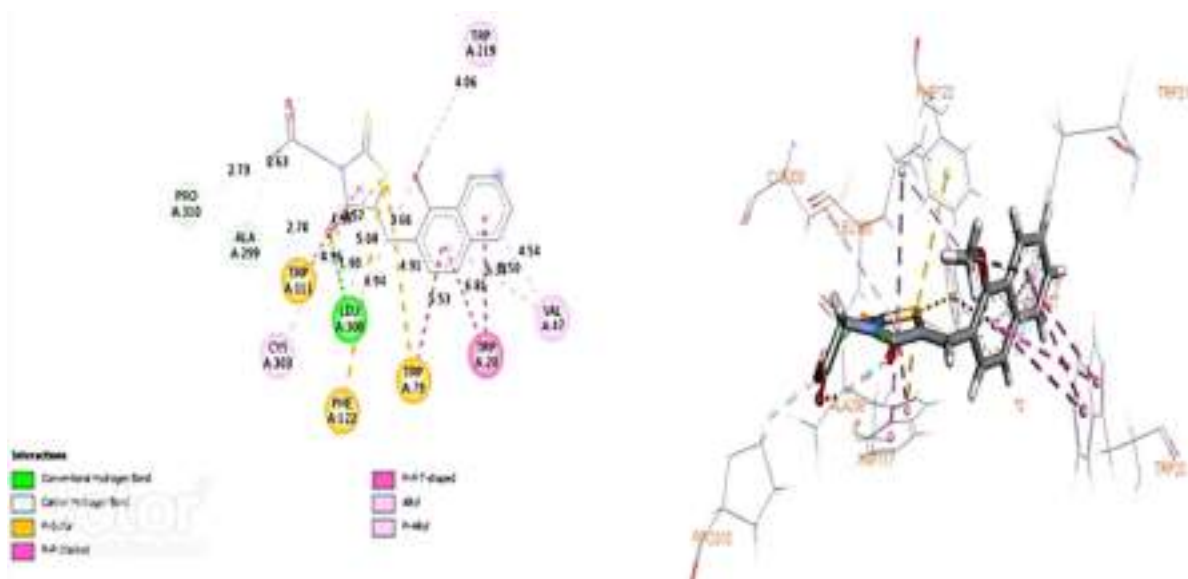


Fig. 14. Showing 3D and 2D interaction of Molecule 18 with Human aldose Reductase.

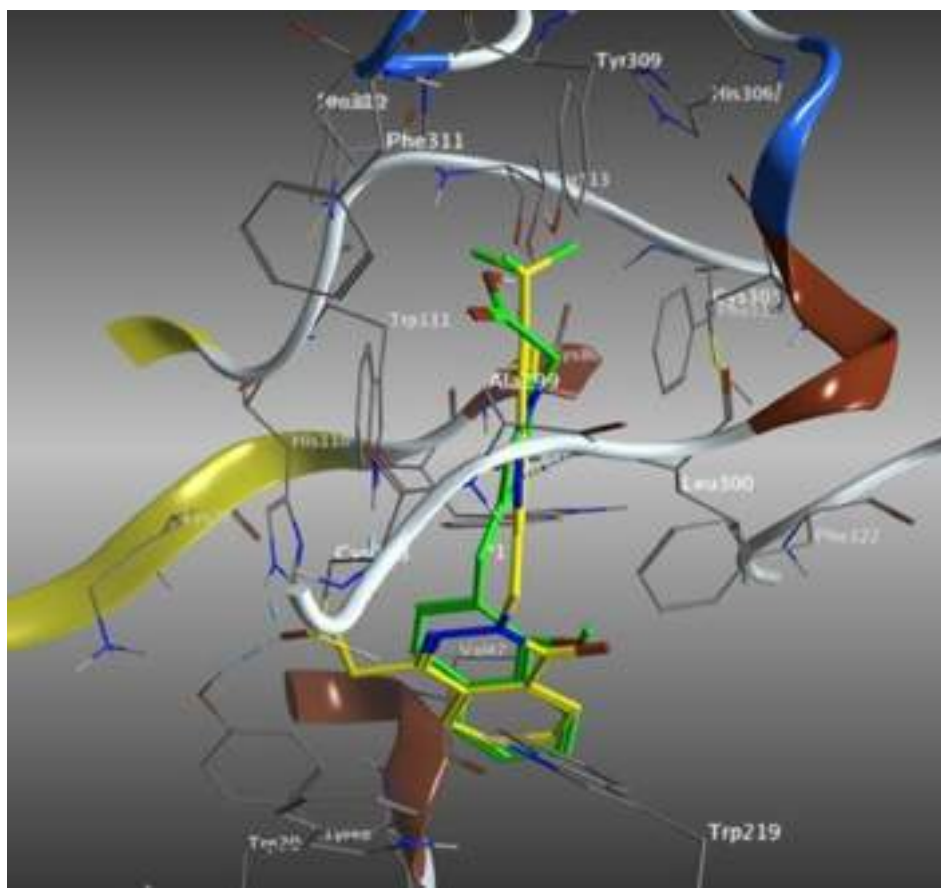


Fig. 15. Depiction of superimposed orientation of Molecule 18 (green) with Pdb-1fzd ligand (yellow) in specificity pocket.

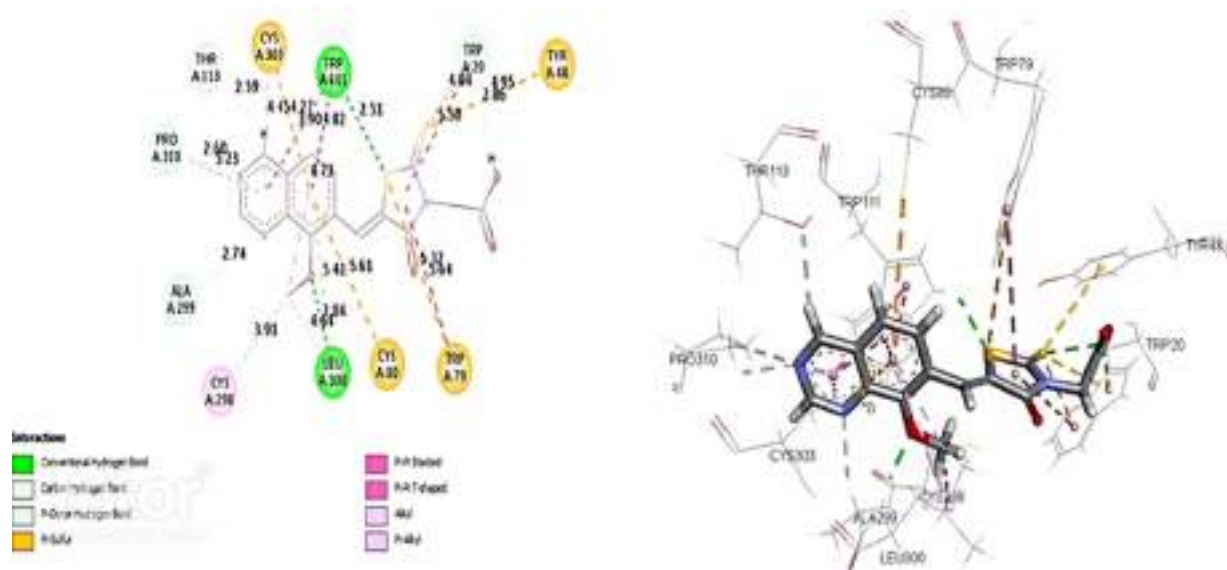


Fig. 16. Showing 3D and 2D interaction of Scaffold 5 with Human aldose Reductase.

ALA299 (2.62Å) residue, where the side chain carbon atom acts as a donor in drug receptor interactions. Furthermore, the acceptor hydroxy group of the terminal carboxyl substituent forms a carbon hydrogen bond with the donor hydrogen atom of the secondary amino group of PRO310 (2.74 Å) residue, producing the human

aldose reductase specificity pocket. Following that, the naphthalene component of Molecule 18 intercalates with the benzene ring of TRP 20(5.21 Å) residue, resulting in a π - π stacking hydrophobic interaction between the naphthalene ring and the benzene ring of TRP20 residue due to the involvement of π orbitals.

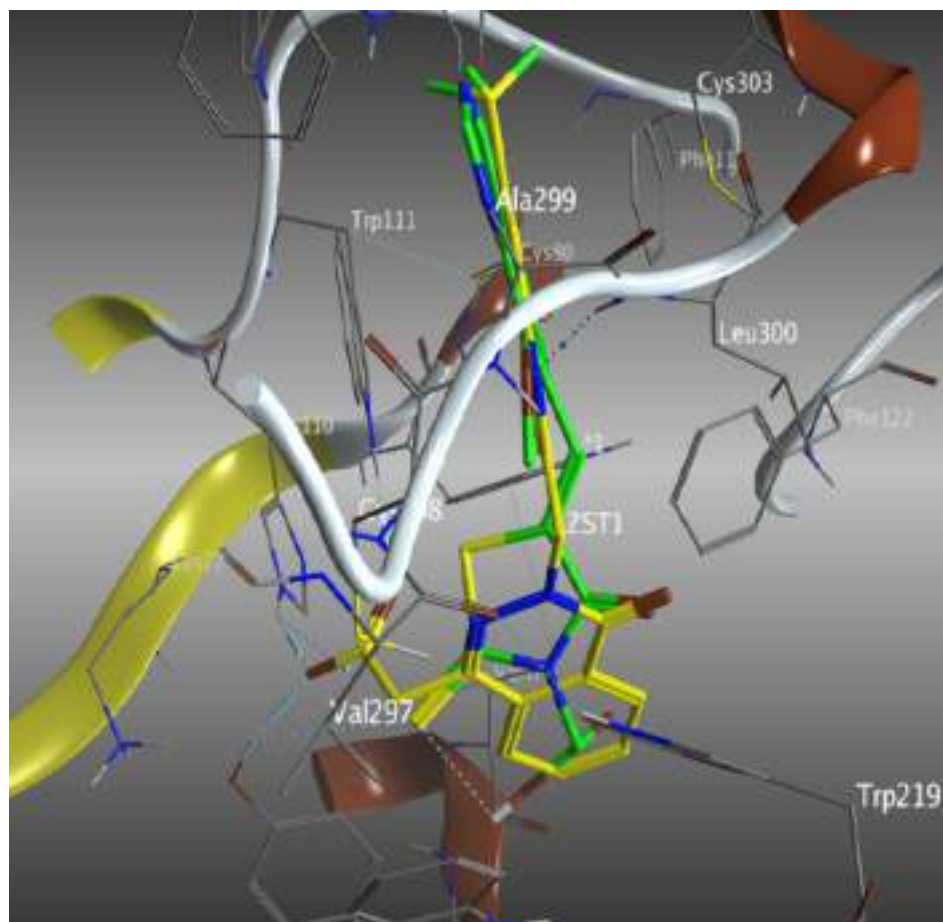


Fig. 17. Depiction of superimposed orientation of scaffold 5 (green) with Pdb-1fzd ligand (yellow) in specificity pocket.

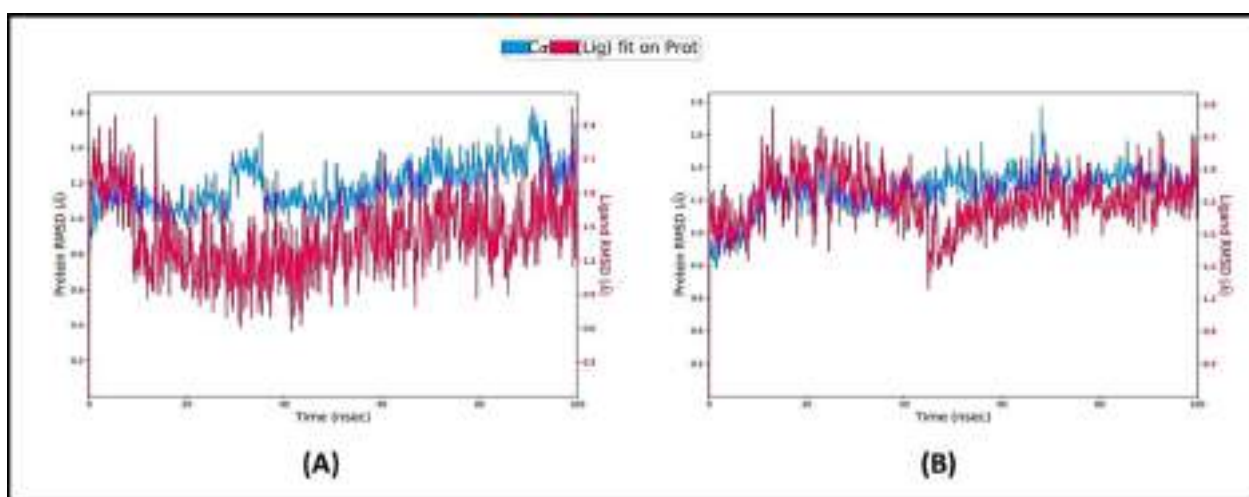


Fig. 18. Presentation of human AR with (A) Molecule 18; (B) scaffold 5, RMSD to measure the average change in displacement of a selection of atoms for a particular frame with respect to a reference frame.

Through π -sulphur interactions, the thioxozolidine ring's sulphur atom forms a 4.90 Å contact with the oxygen atom of tryptophan. As a result, the specificity pocket residue TRP111 (3.57 Å) causes π -sulphure and π - π stacking (3.80 Å) interactions with the thioxozolidine ring's sulphure atom. Likewise another specificity pocket residue PHE122 forms a π -sulphure contact of 4.94 Å with sulphure atom of thioxozolidine ring of Molecule 18. TRP79

residues (5.53 Å) also create π -sulphure contacts with the sulphure atom of the thioxozolidine ring through π - π T shaped hydrophobic interactions with the naphthalene ring. Subsequently, specificity pocket residue LEU300 (5.03 Å) form π -alkyl hydrophobic contact with thioxozolidine ring, CYS303 anchored π -alkyl hydrophobic contact with thioxozolidine ring while VAL47 intercalate with naphthalene ring through π -alkyl hydrophobic interac-

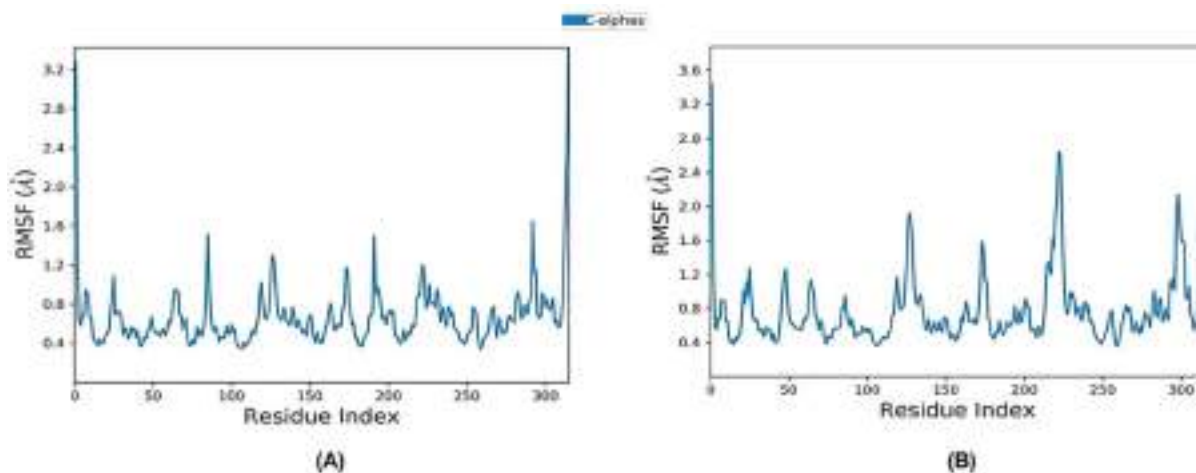


Fig. 19. Presentation of human AR – (A) Molecule 18; (B) scaffold 5 RMSF for characterizing local changes along the protein chain.

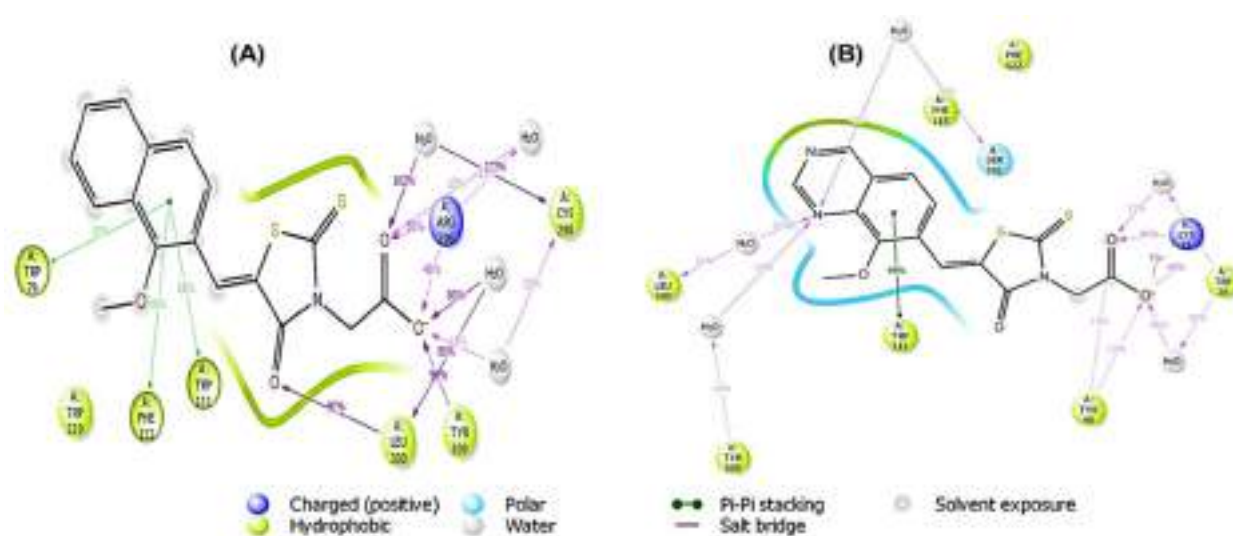


Fig. 20. 2D interaction plots showing ligand interactions of (A) Molecule 18; (B) Scaffold 5 with the binding cavity residues of Aldose Reductase (AR).

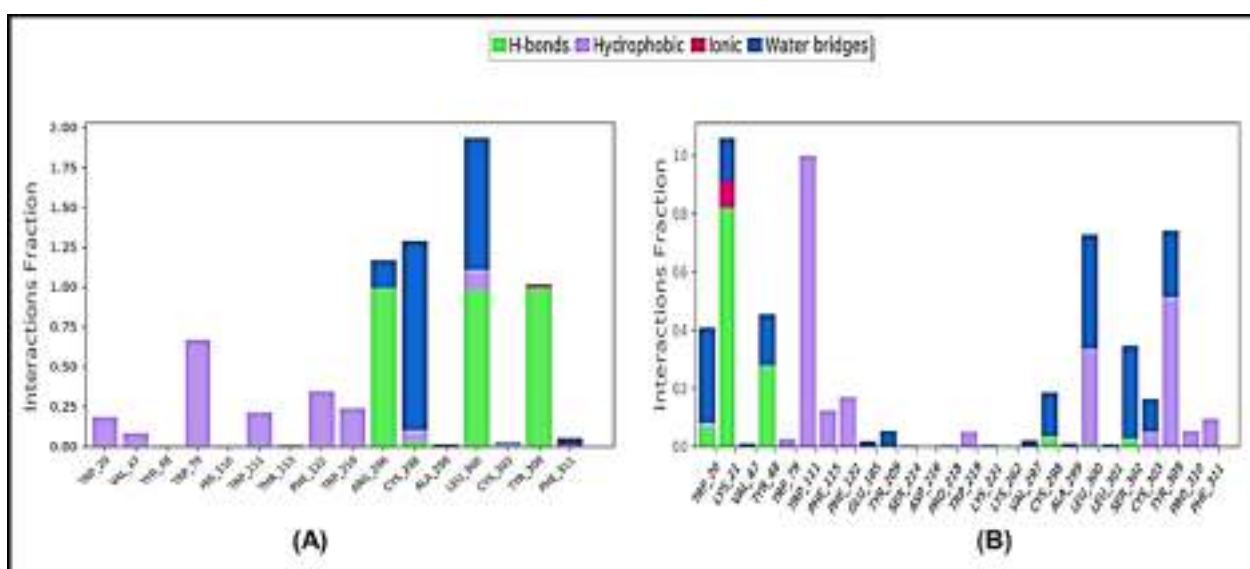


Fig. 21. (A) Molecule 18 contact histogram (H-bonds, Hydrophobic, Ionic, Water bridges) of the ligand, molecule-18 bound with protein recorded in a 100 ns simulation interval; (B) Scaffold 5 contact histogram (H-bonds, Hydrophobic, Ionic, Water bridges) of the ligand, molecule-5 bound with protein recorded in a 100 ns simulation interval.

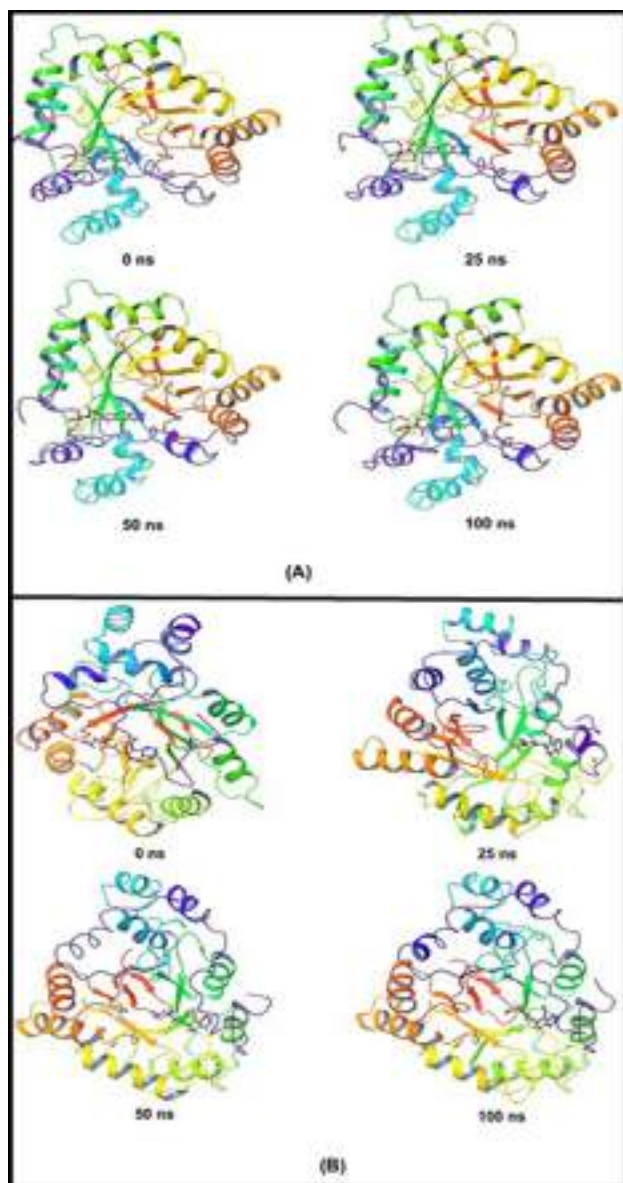


Fig. 22. Stepwise trajectory analysis for every 20 ns displaying the protein, (A) Scaffold 5 and molecule 18 and (B) Scaffold 5 and molecule 18; these show the conformation during 100 ns of simulation scale.

tions. Docking analysis revealed that Molecule 18 bind with specificity pocket residues Trp111, Ala299, Leu300 and Phe122 through conventional hydrogen bond, carbon hydrogen bond, π -sulphur interaction, π - π stacked hydrophobic and π -alkyl hydrophobic interactions. In Molecule 18, thioxozolidine ring, naphthalene ring, terminal carboxyl substituent, methoxy substituent on naphthalene ring plays crucial role for enhancing binding affinity against the human aldose reductase receptor.

4.11. Molecular docking analysis of scaffold 5.

As previously stated, the human AR receptor is divided into two sub-pockets, one containing residues that may be involved in catalysis (Tyr48, Lys77, and His110) that are sideways with the nicotinamide moiety of the cofactor, and the other containing Trp111, Ala299, Leu300, and Phe122 residues.

With a docking score of -8.08 kcal/mol, scaffold 5 achieved the same conformation in the specificity pocket as Pdb-1fzd ligand (see Fig. 16) and was anchored with specificity pocket residues Trp111,

Ala299, and Leu300 via conventional hydrogen bonds, carbon hydrogen bonds, pi-sulfur contact, pi-donor hydrogen bond, π - π stacked hydrophobic contact, and π - π T-shaped contact, pi-alkyl contact and alkyl contacts.(see Fig. 17) According to the results of the molecular docking analysis, scaffold 5 has an excellent binding characteristic and affinity for the AR receptor. It also aligns into the specificity pocket in the same way the co-crystallized pdb ligand does.

The allminus SASA descriptor further emphasizes the need of a solvent-accessible surface. The presence of abundant SASA in the ligand molecule means that the ligand is exposed to the vanderwaals surface of the receptor molecule, increasing hydrophobic interactions. The network of solvent accessible surfaces of the ligand molecules around them remains intact as a result of hydrophobic groups tending to join together. This means that the results of QSAR and docking are complementary and consistent.

Furthermore, the Don_ringC_6Ac descriptor represents a blend of electronic and lipophilic features in the ligand molecule that contribute to hydrogen bonding and hydrophobic interactions. As a result, it is reasonable to conclude that the results of molecular docking and QSAR are complementary and in good agreement. As a result, the descriptor H_ringN_2B emphasizes the role of reduced steric bulk in enhancing receptor surface binding affinity. When we look at the docking score for ligand molecule 18 and scaffold 5, we can see that this observation is correct. As a result, the docking results match the QSAR analysis perfectly.

4.12. Molecular dynamics (MD) simulation study

MD simulation studies were performed in order to define the structural stability and conformational analysis of the screened nitrogen heterocycles with target protein is human aldose reductase (Pdb: 1fzd) which is a 36 kDa sized enzyme displays marked flexibility and malleability with respect to its active site. The Molecule 18 and scaffold 5 anchored with human aldose reductase and docking results were analyzed for 100 ns in MD simulation and displayed in Fig. 18. The conformation of Molecule 18 bound to the human aldose reductase displayed stable and converged conformation at the end of 100 ns (Fig. 18.A). Earlier till 50 ns a little RMSD fluctuations observed but later the system was fully converged with 1 Å deviation (Fig. 18 A, blue). On the other hand, ligand RMSD exhibited quite stable conformation from the beginning to end of the 100 ns simulation suggesting good fit and stable accommodation at the binding site of human aldose reductase protein (Fig. 18 A, red). The scaffold 5 bound to aldose reductase displayed stable conformation with less fluctuations in RMSD of both protein and ligand as displayed in Fig. 18.

The RMSF plots of Molecule 18 and scaffold 5 bound aldose reductase were displayed in Fig. 19. The C- α backbone displayed very less fluctuations in the respective amino acids positions with an average of 0.4 Å (Fig. 19(A)). Whereas, scaffold 5 bound aldose reductase displayed significant fluctuations 1.2–2.0 Å between 110 and 120 residues and 2.8 Å between 210 and 220 residues (Fig. 19 (B)).

The interaction plots (Fig. 20 (A) and (B)) displayed the interaction of binding site residues of aldose reductase with the ligands Molecule 18 and scaffold 5. After 100 ns of simulation Molecule 18 formed conventional H-bonds with the Arg296, Leu300 and Tyr309 (Fig. 20 (A)) to entail into a stable complex. On the other hand, scaffold 5 binds by conventional H-bonds with Lys21 and pi-pi interaction with Trp111 (Fig. 20 (B)).

Throughout the simulation, protein interactions with the ligand can be observed. As seen in the graph above, these interactions can be classified and summarized by type. Hydrogen Bonds, Hydrophobic, Ionic, and Water Bridges are the four forms of protein–ligand interactions (or ‘contacts’). Each interaction type has a number of

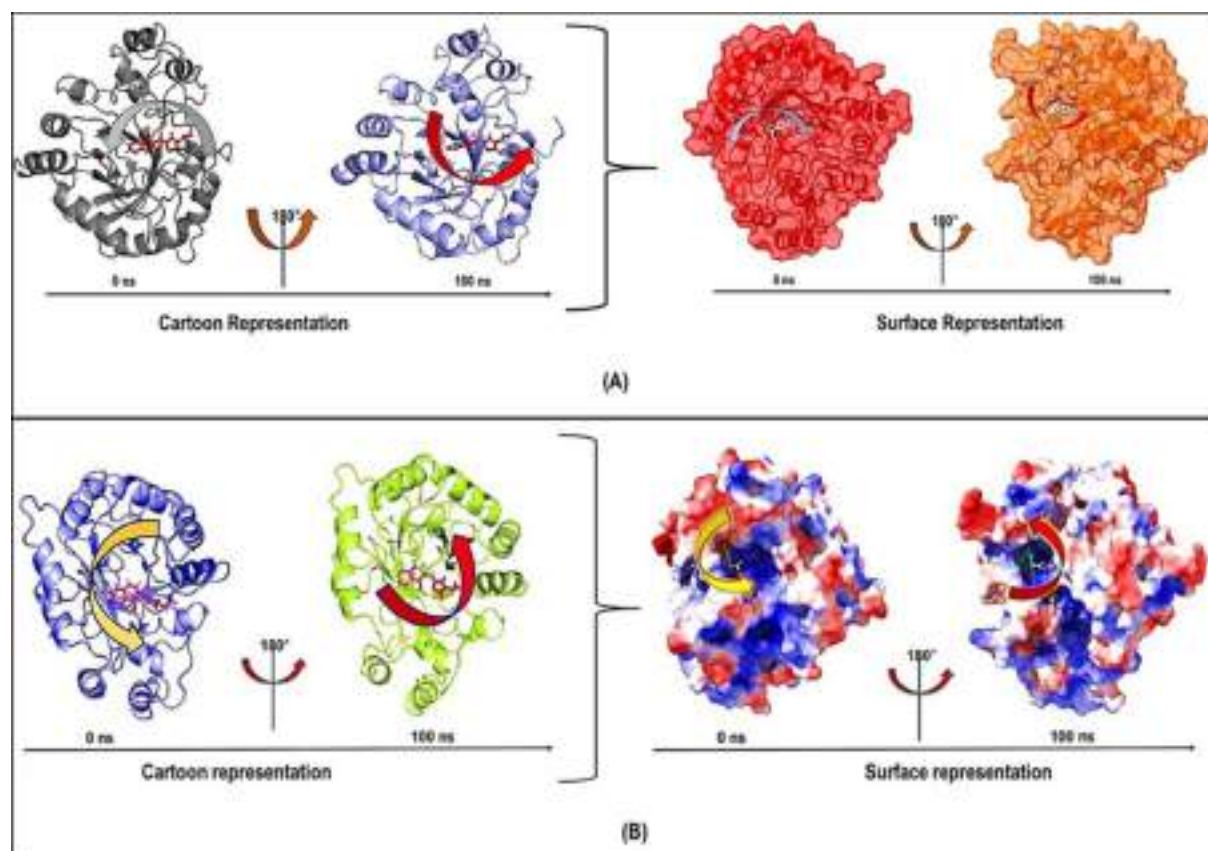


Fig. 23. MMGBSA trajectory (0 ns, before simulation and 100 ns, after simulation) exhibited conformational changes of molecule 18 and scaffold 5 upon binding with the protein, molecule 18 (A) and scaffold 5 (B). The arrows indicate the overall positional variation (movement and pose) of ligands at the binding site cavity.

Table 4

Binding energy calculation of molecule 18 and Scaffold 5 and non-bonded interaction energies from MMGBSA trajectories.

Energies (kcal/mol)	Molecule 18	Scaffold 5
ΔG_{bind}	-61.839 ± 5.673	-55.590 ± 3.705
$\Delta G_{\text{bindLipo}}$	-23.517 ± 1.693	-20.995 ± 1.509
$\Delta G_{\text{bindvdW}}$	-49.459 ± 1.857	-43.737 ± 3.695
$\Delta G_{\text{bindCoulomb}}$	17.834 ± 7.867	9.519 ± 3.212
$\Delta G_{\text{bindHbond}}$	-1.829 ± 0.723	-1.128 ± 0.388
$\Delta G_{\text{bindSolvGB}}$	16.579 ± 8.341	12.866 ± 10.172
$\Delta G_{\text{bindCovalent}}$	1.362 ± 0.728	1.253 ± 0.801

subtypes that can be examined using Maestro's 'Simulation Interactions Diagram' panel, shown in Fig. 21. (A) and (B)). The stacked bar charts are standardized over the course of the trajectory. Some protein residues may make several interactions of the same subtype with the ligand, values above 1.0 are feasible. As seen in Fig. 21; the majority of the significant ligand–protein interactions discovered by MD are hydrogen bonds and hydrophobic interactions.

The stepwise trajectory analysis of every 25 ns of simulation of Scaffold 1 bound to molecule 18 and Scaffold 6 bound to molecule 5 displayed the positional alteration with reference to 0 ns structure [Fig. 22]. It has been observed that the ligand, molecule 18 with Scaffold 5 in Fig. 22 (A) and molecule 18 with Scaffold 5 in Fig. 22 (B) have possessed a structural angular movement at the end frame to achieve its conformational stability and convergence.

4.13. Molecular mechanics generalized born and surface area (MMGBSA) calculations

To assess the binding energy of ligands to protein molecules, the MMGBSA technique is commonly employed. The binding free

energy of each molecule 18 complex and Scaffold 5, as well as the impact of other non-bonded interactions energies, were estimated. With Scaffold 1, the ligand molecule 18 has a binding energy of -61.836 kcal/mol, whereas with Scaffold 6, the ligand molecule 18 has a binding energy of -57.590 kcal/mol. Non-bonded interactions like $G_{\text{bindCoulomb}}$, $G_{\text{bindCovalent}}$, $G_{\text{bindHbond}}$, G_{bindLipo} , $G_{\text{bindSolvGB}}$, and G_{bindvdW} govern G_{bind} . Across all types of interactions, the G_{bindvdW} , G_{bindLipo} , and $G_{\text{bindCoulomb}}$ energies contributed the most to the average binding energy. On the other side, the $G_{\text{bindSolvGB}}$ and $G_{\text{bindCovalent}}$ energies contributed the least to the final average binding energies. Furthermore, the $G_{\text{bindHbond}}$ interaction values of molecule 18 and Scaffold 5 complexes demonstrated stable hydrogen bonds with amino acid residues. In all of the compounds, $G_{\text{bindSolvGB}}$ and $G_{\text{bindCovalent}}$ exhibited unfavorable energy contributions, and so opposed binding. Fig. 23 (A) reveals that between pre-simulation (0 ns) and post-simulation (100 ns), molecule 18 in the binding pocket of Scaffold 1 has undergone a large angular change in the pose (curved to straight) and Fig. 23 (B) reveals that between pre-simulation (0 ns) and post-simulation (100 ns), molecule 5 in the binding pocket of Scaffold 6 has undergone a large angular change in the pose (curved to straight). These conformational changes lead to better binding pocket acquisition and interaction with residues, which leads to enhanced stability and binding energy (shown in Table 4).

Thus, MM-GBSA calculations resulted, from MD simulation trajectories well justified with the binding energy obtained from docking results moreover, the last frame (100 ns) of MMGBSA displayed the positional change of the molecule 18 and scaffold 5 as compared to 0 ns trajectory signify the better binding pose for best fitting in the binding cavity of the protein (see Fig. 23).

4.14. ADMET study of molecule 18 and scaffold 5

ADMET study of molecule 18 and scaffold 5 showed interesting molecular as well as pharmacokinetic properties. Total polar surface area of a molecule should be within the range of 20–130 Å. The molecule 18 TPSA was calculated to be 124.23 and for scaffold 5 TPSA is quite high 150 Å. The next descriptor (LogS, ESol model) was calculated to be water solubility whose range should be < 6. The molecule 18 showed high solubility having –4.5 whereas, scaffold 5 displayed –3.49. Therefore, it can be suggested that both molecule 18 and scaffold 5 have good solubility in water. GI absorption of molecule is found to be significant whereas, scaffold 5 has less GI absorption capacity. For drug likeliness property, molecule 18 displayed significant likeness since none of the properties violate Lipinski, Ghose, Veber, Muegge and Egan filters. Whereas, scaffold 5, except Lipinski and Ghose all other filters are violated. Therefore, suggestively molecule 18 could be a proper drug where as scaffold has drug like properties. Molecule 18 has good capacity to travel through BBB whereas, the possibility for scaffold 5 to pass through BBB is less. Moreover, both molecule 18 and scaffold 5 displayed to have good synthesizing accessibility score 3.37 and 3.2, respectively, which signifies both the lead molecules can be synthesized in the laboratory. As per the recommendation which should be < 6, both the molecule 18 and scaffold 5 are synthesizable. The overall comparative analysis of ADMET parameters suggested that molecule 18 has better chance to become drug molecule instead scaffold 5.

5. Conclusions

In the present investigation, we used QSAR analysis, QSAR-based virtual screening, Molecular docking, and Molecular dynamic simulation to find a new AR receptor inhibitor that is a strong antidiabetic agent. To identify hidden structural features responsible for AR inhibition, a five parametric GA MLR-based full set and divided set QSAR model was developed. Pharmacophoric features such as solvent accessible surface area, partially charged ring carbon atoms within 6 Å of donor, combination of ring nitrogen and hydrogen separated by 2 bonds, and others emerged as projecting features that govern AR inhibition based on the developed QSAR model analysis. We conducted virtual screening using QSAR yielded a novel hit molecule (scaffold 5) with a pIC_{50} of 8.05 M ($IC_{50} = 8.91$ nM). Furthermore, molecular docking analysis of molecule 18 and scaffold 5 demonstrates that both molecules in the specificity binding pocket of AR adopted the same conformation as the pdb ligand. This study offered light on the pharmacophores involved in the binding interactions that inhibit AR. It gives the impression that, molecule 18 and scaffold 5 bind with the specificity pocket residue A: LEU 300 (1.90 Å, ALA299 (2.74 Å), ALA299 (2.62 Å), PRO310 (2.74 Å), TRP 20 (5.21 Å). The QSAR and molecular docking results are consensus and complementary, and the identified pharmacophoric properties must be retained in the development of new and potential AR inhibitors as antidiabetic agents in the future. Finally, the high docking score alongside increase in IC_{50} value (8.91 nM) for scaffold 5 by 0.21 nm shows that it has a higher affinity for the specificity binding pocket of AR receptors, and this study could lead to the development of novel AR inhibitors as new antidiabetic agents. Subsequently, The formation of the significant number of hydrogen bonds in MD simulation corroborated the findings with molecular docking also suggested for a stable complex formation during the MD simulation over 100 ns time scale. MMGBSA is a powerful tool in determining the binding energy of the ligand with its respective protein targets. MMGBSA studies accurately predicted the total binding energy of the Wedelosin at the binding cavity of ALK and BTK and exhibited

a very low binding energy suggesting the capacity of the Scaffold 5 to conform into a stable complex. The binding energies in MMGBSA trajectory supported by van der Waals energy, Lipophilic energy, Coulombic energies and similarly reported elsewhere. The ADMET study revealed that scaffold 5 behave as drug like candidate. The findings of this study may be relevant in the development of novel therapeutic targets for AR as an antidiabetic agent in the future.

Funding

The authors acknowledge the Deanship of Scientific Research at Imam Mohammad Ibn Saud Islamic University, Riyadh, KSA, for its support of this research through the Research group number 21–09–77. Authors are thankful to Dr. Paola Gramatica for providing free copy of QSARINS software.

CRediT authorship contribution statement

Ravindra L. Bakal: Conceptualization. **Rahul D. Jawarkar:** Conceptualization, Writing – original draft, Formal analysis. **J.V. Manwar:** Writing – original draft, Data curation. **M.S. Jaiswal:** Writing – original draft, Data curation. **Arabinda Ghosh:** . **Ajaykumar Gandhi:** . **Magdi E.A. Zaki:** Methodology, Writing – original draft. **Sami Al-Hussain:** . **Abdul Samad:** Writing – review & editing, Validation. **V.H. Masand:** Methodology, Writing – review & editing. **Nobendu Mukerjee:** . **Syed Nasir Abbas Bukhari:** . **Praveenkumar Sharma:** Writing – review & editing. **Israa Lewaa:** .

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors acknowledge the Deanship of Scientific Research, Imam Mohammad Ibn Saud Islamic University, Saudi Arabia, Research Group no. RG-21–09–76. Authors are thankful to Dr. Paola Gramatica, for providing the free copy of QSARINS 2.2.4 Software. Authors are grateful to the R.C. Patel Institute of Pharmacy, Shirpur for providing computational facilities during entire course of research work. The Author, Rahul D. Jawarkar is thankful to The President, Shri Yogendraji Gode, for providing necessary facilities during Research work.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.jsps.2022.04.003>.

References



- Bowers, K. J., D. E. Chow, H. Xu, et al., 2006. Scalable Algorithms for Molecular Dynamics Simulations on Commodity Clusters. ACM/IEEE SC 2006 Conference (SC'06): 43–43.
- Cherkasov, A., Muratov, E.N., Fourches, D., Varnek, A., Baskin, I.I., Cronin, M., Dearden, J., Gramatica, P., Martin, Y.C., Todeschini, R., Consonni, V., Kuz'min, V. E., Cramer, R., Benigni, R., Yang, C., Rathman, J., Terloth, L., Gasteiger, J., Richard, A., Tropsha, A., 2014. QSAR Modeling: Where Have You Been? Where Are You Going To? J. Med. Chem. 57 (12), 4977–5010. <https://doi.org/10.1021/jm4004285>.
- Consonni, V., R. Todeschini, D. Ballabio, et al., 2019. On the Misleading Use of QF32 for QSAR Model Comparison. Molecular Informatics. 38, [Doi: 10.1002/minf.201800029](https://doi.org/10.1002/minf.201800029).

- Dearden, J.C., Cronin, M.T.D., Kaiser, K.L.E., 2009. How not to develop a quantitative structure–activity or structure–property relationship (QSAR/QSPR). *SAR QSAR Environ. Res.* 20 (3–4), 241–266. <https://doi.org/10.1080/10629360902949567>.
- Foppiano, M., Lombardo, G., 1997. Worldwide pharmacovigilance systems and tolrestat withdrawal. *The Lancet.* 349 (9049), 399–400.
- Fujita, T., Winkler, D.A., 2016. Understanding the Roles of the “Two QSARs”. *J. Chem. Inf. Model.* 56 (2), 269–274. <https://doi.org/10.1021/acs.jcim.5b00229>.
- Gramatica, P., 2013. On the Development and Validation of QSAR Models. *Comput. Toxicol.*, 499–526.
- Gramatica, P., 2014. External Evaluation of QSAR Models, in Addition to Cross-Validation: Verification of Predictive Capability on Totally New Chemicals. *Mol. Inf.* 33 (4), 311–314. <https://doi.org/10.1002/minf.201400030>.
- Gramatica, P., 2020. Principles of QSAR Modeling. *Int. J. Quant. Struct. Property Relationships.* 5 (3), 61–97. <https://doi.org/10.4018/IJQSPR.20200701.0a1>.
- Gramatica, P., Cassani, S., Chirico, N., 2014. QSARINS-chem: Insubria datasets and new QSAR/QSPR models for environmental pollutants in QSARINS. *J. Comput. Chem.* 35 (13), 1036–1044. <https://doi.org/10.1002/jcc.23576>.
- Gramatica, P., Chirico, N., Papa, E., Cassani, S., Kovarich, S., 2013. QSARINS: A new software for the development, analysis, and validation of QSAR MLR models. *J. Comput. Chem.* 34 (24), 2121–2132. <https://doi.org/10.1002/jcc.23361>.
- Harit, T., Bellaouchi, R., Asehraou, A., Rahal, M., Bouabdallah, I., Malek, F., 2017. Synthesis, characterization, antimicrobial activity and theoretical studies of new thiophene-based tripodal ligands. *J. Mol. Struct.* 1133, 74–79. <https://doi.org/10.1016/j.molstruc.2016.11.051>.
- Jawarkar, R.D., Bakal, R.L., Zaki, M.E.A., Al-Hussain, S., Ghosh, A., Gandhi, A., Mukerjee, N., Samad, A., Masand, V.H., Lewaa, I., 2021. QSAR Based Virtual screening derived Identification of a Novel Hit as a SARS CoV-229E 3CLpro Inhibitor: GA-MLR QSAR modeling supported by Molecular Docking, Molecular Dynamics Simulation and MMGBSA calculation Approaches. *Arabian J. Chem.* 15 (1), 103499. <https://doi.org/10.1016/j.arabjc.2021.103499>.
- Jez, J.M., Flynn, T.G., Penning, T.M., 1997. A new nomenclature for the aldo-keto reductase superfamily. *Biochem. Pharmacol.* 54 (6), 639–647.
- Jorgensen, W.L., Maxwell, D.S., Tirado-Rives, J., 1996. Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids. *J. Am. Chem. Soc.* 118 (45), 11225–11236. <https://doi.org/10.1021/ja9621760>.
- Krans, H.M.J., 1993. Recent Clinical Experience With Aldose Reductase Inhibitors. *Diabet. Med.* 10, 44S–48S. <https://doi.org/10.1111/j.1464-5491.1993.tb00198.x>.
- Kucerova-Chlupacova, M., Halakova, D., Majekova, M., Tremel, J., Stefek, M., Soltesova Prnova, M., 2020. (4-Oxo-2-thioxothiazolidin-3-yl)acetic acids as potent and selective aldose reductase inhibitors. *Chem. Biol. Interact.* 332, 109286. <https://doi.org/10.1016/j.cbi.2020.109286>.
- Martyna, G.J., Klein, M.L., Tuckerman, M., 1992. Nosé-Hoover chains: The canonical ensemble via continuous dynamics. *J. Chem. Phys.* 97 (4), 2635–2643. <https://doi.org/10.1063/1.463940>.
- Martyna, G.J., Tobias, D.J., Klein, M.L., 1994. Constant pressure molecular dynamics algorithms. *J. Chem. Phys.* 101 (5), 4177–4189. <https://doi.org/10.1063/1.467468>.
- Masand, V.H., Mahajan, D.T., Maldhure, A.K., Rastija, V., 2016. Quantitative structure–activity relationships (QSARs) and pharmacophore modeling for human African trypanosomiasis (HAT) activity of pyridyl benzamides and 3-(oxazolol[4,5-b]pyridin-2-yl)anilides. *Med. Chem. Res.* 25 (10), 2324–2334. <https://doi.org/10.1007/s00044-016-1664-1>.
- Moon, H.-I., Jung, J.-C., Lee, J., 2006. Aldose reductase inhibitory effect by tectorigenin derivatives from *Viola hondoensis*. *Bioorg. Med. Chem.* 14 (22), 7592–7594. <https://doi.org/10.1016/j.bmc.2006.07.002>.
- O’Boyle, N.M., Banck, M., James, C.A., Morley, C., Vandermeersch, T., Hutchison, G.R., 2011. Open Babel: An open chemical toolbox. *J. Cheminf.* 3 (1). <https://doi.org/10.1186/1758-2946-3-33>.
- Pavan, M., Todeschini, R., 2009. Multicriteria Decision-Making Methods. *Comprehensive Chemometrics*, 591–629.
- Pourbasheer, E., Shokouhi Tabar, S., Masand, V.H., Aalizadeh, R., Ganjali, M.R., 2015. 3D-QSAR and docking studies on adenosine A2A receptor antagonists by the CoMFA method. *SAR QSAR Environ. Res.* 26 (6), 461–477.
- Shivakumar, D., Williams, J., Wu, Y., Damm, W., Shelley, J., Sherman, W., 2010. Prediction of Absolute Solvation Free Energies using Molecular Dynamics Free Energy Perturbation and the OPLS Force Field. *J. Chem. Theory Comput.* 6 (5), 1509–1519. <https://doi.org/10.1021/ct900587b>.
- Steele, J.W., Faulds, D., Goa, K.L., 1993. Epalrestat. *Drugs Aging* 3 (6), 532–555. <https://doi.org/10.2165/00002512-199303060-00007>.
- Toukhami, A.Y., Board, J.A., 1996. Ewald summation techniques in perspective: a survey. *Comput. Phys. Commun.* 95 (2–3), 73–92. [https://doi.org/10.1016/0010-4655\(96\)00016-1](https://doi.org/10.1016/0010-4655(96)00016-1).
- Tsai, S.C., Burmakis, T.G., 2016. Aldose Reductase Inhibitors: An Update. *Ann. Pharmacother.* 27 (6), 751–754. <https://doi.org/10.1177/106002809302700616>.
- van Gerven, J.M.A., Tjon-A-Tsien, A.M.L., 1995. The Efficacy of Aldose Reductase Inhibitors in the Management of Diabetic Complications. *Drugs Aging* 6 (1), 9–28. <https://doi.org/10.2165/00002512-199506010-00002>.
- Bikadi, Z., Hazai, E., 2009. Application of the PM6 semi-empirical method to modeling proteins enhances docking accuracy of AutoDock. *J. Cheminf.* 1 (1). <https://doi.org/10.1186/1758-2946-1-15>.
- Gaudreault, F., Morency, L.-P., Najmanovich, R.J., 2015. NRGsuite: a PyMOL plugin to perform docking simulations in real time using FlexAID. *Bioinformatics.* <https://doi.org/10.1093/bioinformatics/btv458>.
- Gaulton, A., Hersey, A., Nowotka, M., Bento, A.P., Chambers, J., Mendez, D., Mutowo, P., Atkinson, F., Bellis, L.J., Cibrián-Uhalte, E., Davies, M., Dedman, N., Karlsson, A., Magariños, M.P., Overington, J.P., Papadatos, G., Smit, I., Leach, A.R., 2017. The ChEMBL database in 2017. *Nucleic Acids Res.* 45 (D1), D945–D954. <https://doi.org/10.1093/nar/gkw1074>.
- Ivanciuc, O., 1996. HyperChem Release 4.5 for Windows. *J. Chem. Inf. Comput. Sci.* 36 (3), 612–614. <https://doi.org/10.1021/ci950190a>.
- Mukerjee, N., Das, A., Maitra, S., Ghosh, A., Khan, P., Alexiou, A., Dey, A., Baishya, D., Ahmad, F., Sachdeva, P., Al-Muhanna, M.K., Kumar, V., 2022. Dynamics of natural product Lupenone as a potential fusion inhibitor against the spike complex of novel Semliki Forest Virus. *PLoS ONE* 17 (2), e0263853. <https://doi.org/10.1371/journal.pone.0263853>.
- Masand, V.H., Rastija, V., 2017. PyDescriptor : A new PyMOL plugin for calculating thousands of easily understandable molecular descriptors. *Chemometr. Intell. Lab. Syst.* 169, 12–18. <https://doi.org/10.1016/j.chemolab.2017.08.003>.



Article

Mechanistic Analysis of Chemically Diverse Bromodomain-4 Inhibitors Using Balanced QSAR Analysis and Supported by X-ray Resolved Crystal Structures

Magdi E. A. Zaki^{1,*}, Sami A. Al-Hussain¹, Aamal A. Al-Mutairi¹, Vijay H. Masand^{2,*}, Abdul Samad³ 
and Rahul D. Jawarkar⁴ 

- ¹ Department of Chemistry, Faculty of Science, Imam Mohammad Ibn Saud Islamic University, Riyadh 13318, Saudi Arabia; sahussain@imamu.edu.sa (S.A.A.-H.); aamutairi@imamu.edu.sa (A.A.A.-M.)
² Department of Chemistry, Vidya Bharati Mahavidyalaya, Amravati 444602, India
³ Department of Pharmaceutical Chemistry, Faculty of Pharmacy, Tishk International University, Erbil 44001, Iraq; abdul.samad@tiu.edu.iq
⁴ Department of Medicinal Chemistry, Dr. Rajendra Gode Institute of Pharmacy, University-Mardi Road, Amravati 444901, India; rahuljawarkar@gmail.com
* Correspondence: mezaki@imamu.edu.sa (M.E.A.Z.); vijaymasand@gmail.com (V.H.M.)

Abstract: Bromodomain-4 (BRD-4) is a key enzyme in post-translational modifications, transcriptional activation, and many other cellular processes. Its inhibitors find their therapeutic usage in cancer, acute heart failure, and inflammation to name a few. In the present study, a dataset of 980 molecules with a significant diversity of structural scaffolds and composition was selected to develop a balanced QSAR model possessing high predictive capability and mechanistic interpretation. The model was built as per the OECD (Organisation for Economic Co-operation and Development) guidelines and fulfills the endorsed threshold values for different validation parameters ($R^2_{tr} = 0.76$, $Q^2_{LMO} = 0.76$, and $R^2_{ex} = 0.76$). The present QSAR analysis identified that anti-BRD-4 activity is associated with structural characters such as the presence of saturated carbocyclic rings, the occurrence of carbon atoms near the center of mass of a molecule, and a specific combination of planer or aromatic nitrogen with ring carbon, donor, and acceptor atoms. The outcomes of the present analysis are also supported by X-ray-resolved crystal structures of compounds with BRD-4. Thus, the QSAR model effectively captured salient as well as unreported hidden pharmacophoric features. Therefore, the present study successfully identified valuable novel pharmacophoric features, which could be beneficial for the future optimization of lead/hit compounds for anti-BRD-4 activity.

Keywords: QSAR; BRD-4; pharmacophoric features; X-ray



Citation: Zaki, M.E.A.; Al-Hussain, S.A.; Al-Mutairi, A.A.; Masand, V.H.; Samad, A.; Jawarkar, R.D. Mechanistic Analysis of Chemically Diverse Bromodomain-4 Inhibitors Using Balanced QSAR Analysis and Supported by X-ray Resolved Crystal Structures. *Pharmaceuticals* **2022**, *15*, 745. <https://doi.org/10.3390/ph15060745>

Academic Editors: Jun Qi and Adam David Durbin

Received: 16 May 2022

Accepted: 7 June 2022

Published: 14 June 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Cancer and heart failure are major causes of mortality [1], health complications, and social and economic problems for millions of people around the globe. Researchers have identified different chemotherapeutic methods to minimise heart failure as well as the onset, growth, and survival of cancer cells [1]. However, different serious health issues initiated or echoed by different anti-cancer and cardiac drugs are of great concerns. Therefore, the quest for a harmless and effective anti-cancer and cardiac drug is an important goal for the research and development laboratories of pharmaceutical companies and academic institutions. For this, researchers generally prefer to inhibit any irregularity occurring during a vital cellular process. A good number of recent studies have confirmed that reversible lysine acetylation (RAL) is a dynamic process responsible for protein post-translational modifications, transcriptional activation, and other cellular processes [2–11]. Therefore, any anomaly with RAL could lead to the initiation of malignancy or its survival [7,12]. RAL is regulated by three types of epigenetic regulatory proteins [12,13]: (1) histone acetyltransferases (HATs) acetylate lysine, (2) Histone deacetylases (HDACs), and (3) bromodomain

(BRD) family of proteins. HATs are responsible for acetylation of lysine residues on histone tails and thereby behave as “writers”, whereas the reverse is true for HDACs and sirtuins, which work as “erasers” that are accountable for the elimination of the acetyl group from acetylated lysine (KAc) [14]. The Bromodomain and Extra-terminal (BET) family selectively recognises and links with acetylated lysine residues in histones H3 and H4 [15–17]; thus, they function as “readers”. The BET proteins, viz., BRD2, BRD3, and BRD-4, and bromodomain testis-specific protein (BRDT) are widely recognised as druggable target proteins for regulating cellular epigenetics [15]. Therefore, intruding interactions between BET proteins and acetylated lysine have attracted many researchers to develop better therapeutics for various human diseases including cancer, acute heart failure, and inflammation [2–11].

BRD-4, also called mitotic chromosomal-associated protein (MCAP), Fshrg4, or Hunk1, is ubiquitously expressed and plays a crucial role in a number of DNA-centered processes [15]. It is generally localised in the nucleus and regulates transcription by RNA polymerase II through a positive transcriptional elongation factor complex [15]. Structurally, it comprises two highly conserved N-terminal bromodomains (BD1 and BD2), an ET domain, and a C-terminal domain (CTD) [7]. Furthermore, BRD-4 contains a set of four helices: αZ , αA , αB , and αC . αZ and αA helices are connected through the ZA loop, whereas the BC loop connects the αB and αC helices [11,18]. Together, the four helices and the two loops create an active acetyl-lysine binding pocket (see Figure 1) [11,18]. The active site also consists of a hydrophobic WPF shelf (Trp81, Pro82, and Phe83), ZA loop, Tyr97, Asn140, and Met149 [11,18]. The majority of BRD-4 inhibitors compete with histone H4 to imitate the interactions with Tyr97 and Asn140 [3]. The WPF shelf is believed to play an important role in deciding the selectivity for BET bromodomains [6].

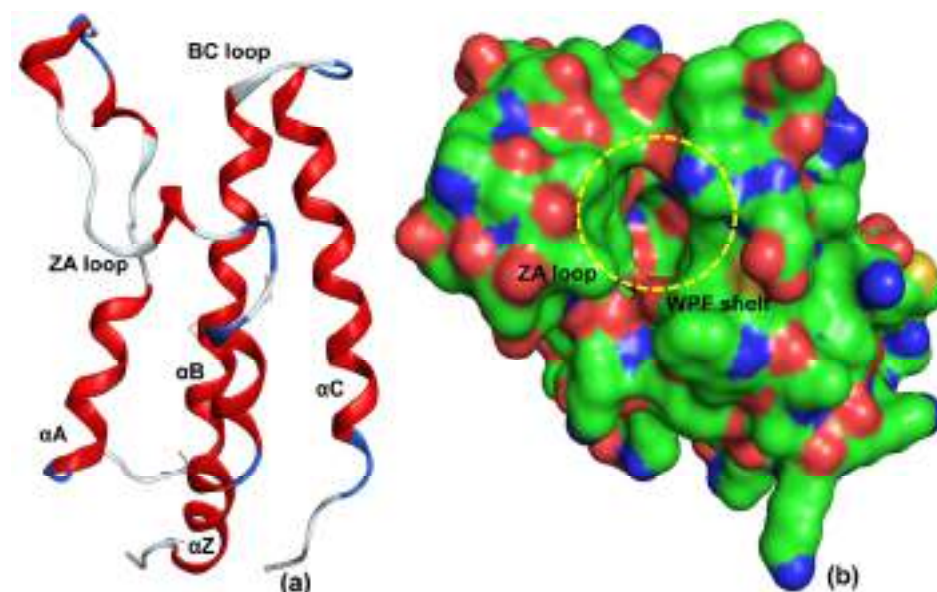


Figure 1. X-ray resolved structure of BRD-4 using pdb 5UVT (a) without molecular surface and (b) with molecular surface (green: carbon; red: oxygen; blue: nitrogen).

Recent studies indicate that the inhibition of BRD-4 is a good strategy, and a good number of BRD-4 inhibitors are in clinical or pre-clinical trials (see Figure 2) [2,19,20].

However, the quest for a safer and effective BRD-4 inhibitor with an optimum ADMET (Absorption, Distribution, Metabolism, Excretion, and Toxicity) profile with a retention of potency is still in progress. For this, it is essential to know the prominent and concealed pharmacophoric features associated with BRD-4 inhibitors. To achieve this goal, a good number of researchers have reported SAR (Structure Activity Relationships) and QSAR (Quantitative SAR) analyses of BRD-4 inhibitors. Tahir et al. [21] developed a CoMSIA (3D-QSAR) model with an R^2_{tr} (coefficient of determination) = 0.982 and R^2_{cv} (or Q^2_{loo}) (cross-validated coefficient of determination for leave-one-out) = 0.500 for a dataset of

60 quinolinone and quinazolinone derivatives as BRD-4 inhibitors. Tong et al. [22] reported four 3D-QSAR models possessing $R^2_{tr} = 0.912$ to 0.963 and $R^2_{cv} = 0.574$ to 0.759 for the BRD-4 inhibitory activity of 4,5-dihydro-[1,2,4]triazolo[4,3-f]pteridine derivatives. Obadawo and co-workers [23] performed QSAR modelling ($R^2_{tr} = 0.93$ and $R^2_{cv} = 0.70$) for 40 different substituted 4-Phenylisoquinolinones as potent BET bromodomain (BRD-4-BD1) inhibitors. Speck-Planche and Scotti [6] performed multi-target QSAR for bromodomain inhibitors using linear discriminant analysis and artificial neural networks. Their binary classification (active/inactive), which is based on a fragment-based topological approach, and analysis led to the identification of a good number of pharmacophoric features. However, the fragment-based topological approach involved the use of SMILES of molecules and thereby lacks the inclusion of 3D information. Thus, even though these studies are successful in identifying easily visible pharmacophoric features, they are based on small datasets with limited variations in structures, binary classification, lack thorough validation and general applicability, and provide partial mechanistic interpretations.

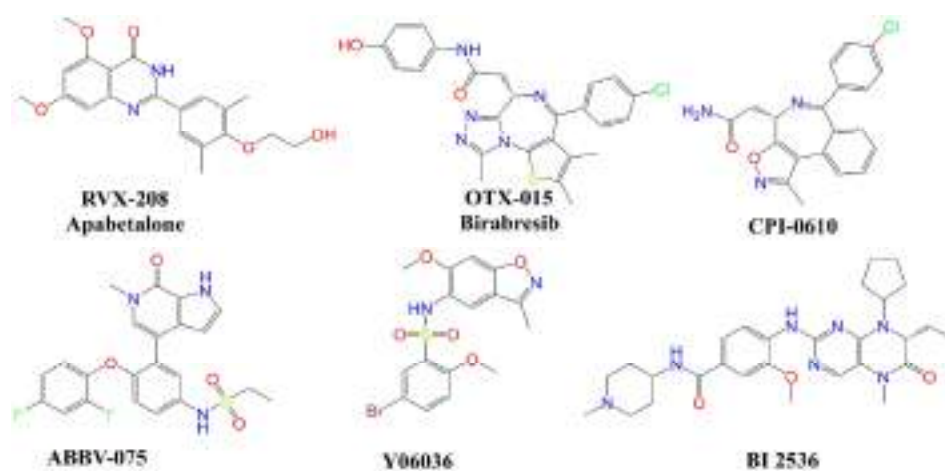


Figure 2. Chemical structures of selected BRD-4 inhibitors.

A literature survey reveals that BRD-4 inhibitors possess structural isomerism (positional, chain, etc.), variations in central scaffolds, and their chemical space is very broad [6,7]; therefore, many concealed or hidden correlations of pharmacophoric features cannot be identified by visual inspection [24]. In such a situation, there is a need to accomplish thorough QSAR analysis using a larger dataset of BRD-4 inhibitors. In the present work, we have performed QSAR analysis of 980 structurally diverse BRD-4 inhibitors. The developed QSAR model possesses a balance of excellent predictive ability with in-depth mechanistic interpretations, which are reinforced by reported X-ray-resolved structures of BRD-4 inhibitors with the target enzyme.

2. Results

The present QSAR analysis is based on a dataset covering a broad chemical space and data range owing to the inclusion of structurally diverse compounds with experimentally measured IC_{50} in the range of 1 nM to 15 μ M. Consequently, this helped us in developing an appropriately validated genetic algorithm multi-linear regression (GA-MLR) model for gathering or extending exhaustive information about the pharmacophoric traits that govern the desired bio-activity (Descriptive QSAR) and also possessing acceptable external predictive capability (Predictive QSAR) [25–27]. The seven variable-based GA-MLR QSAR model (see model-A), along with selected internal and external validation parameters (see Supplementary Material for additional parameters), is as follows.

Model-A: $pIC_{50} = 4.27 (\pm 0.156) + 0.093 (\pm 0.017) * fsp3CringC2B + 0.108 (\pm 0.014) * com_C_4A + 0.391 (\pm 0.066) * Saturated_Carbo_Rings + 0.428 (\pm 0.036) * fsulfonSaroC8B + 0.625 (\pm 0.059) * flipoacc3B + 0.921 (\pm 0.067) * fsp3OaroN6B - 0.367 (\pm 0.063) * fpIaNN4B$.

Validation of Model-A: Method of splitting = Random, No. of descriptors = 7, $N_{\text{training}} = 785$, $N_{\text{test}} = 195$, $R^2_{\text{tr}} = 0.762$, $R^2_{\text{adj.}} = 0.760$, $\text{RMSE}_{\text{tr}} = 0.389$, $\text{MAE}_{\text{tr}} = 0.326$, $\text{CCC}_{\text{tr}} = 0.865$, $s = 0.391$, $F = 355.446$, $R^2_{\text{cv}} (Q^2_{\text{loo}}) = 0.757$, $\text{RMSE}_{\text{cv}} = 0.393$, $\text{MAE}_{\text{cv}} = 0.329$, $\text{CCC}_{\text{cv}} = 0.862$, $Q^2_{\text{LMO}} = 0.756$, $R^2_{\text{Yscr}} = 0.009$, $Q^2_{\text{Yscr}} = -0.012$, $\text{RMSE}_{\text{ex}} = 0.392$, $\text{MAE}_{\text{ex}} = 0.323$, $R^2_{\text{ex}} = 0.762$, $Q^2\text{-F}^1 = 0.762$, $Q^2\text{-F}^2 = 0.760$, $Q^2\text{-F}^3 = 0.758$, $\text{CCC}_{\text{ex}} = 0.860$.

A multitude of statistical validation parameters and analysis of associated graphs has been recommended by different researchers to confirm the statistical robustness and external prediction ability of a QSAR model [28–39]. The same approach has been followed in the present work. A high value of R^2_{tr} , $R^2_{\text{adj.}}$, $R^2_{\text{cv}} (Q^2_{\text{loo}})$, R^2_{ex} , $Q^2\text{-F}^n$, CCC_{ex} , etc., and a small value of LOF (lack-of-fit), RMSE_{tr} , MAE_{tr} , R^2_{Yscr} (R^2 for Y-scrambling), etc., along with different graphs (Figure 3a–d) related to model-A support the external predictive ability, statistical robustness, and point out the lack of chancy correlation for model-A [28–38]. Moreover, the Williams plot [40–44] point out that the majority of molecules (929 molecules) are within the applicability domain; thus, the model is statistically acceptable (see Figure 3b). The outliers with high leverage have been labeled in Figure 3b. Therefore, it fulfills all the Organisation for Economic Co-operation and Development (OECD) endorsed guidelines for generating a thriving QSAR model.

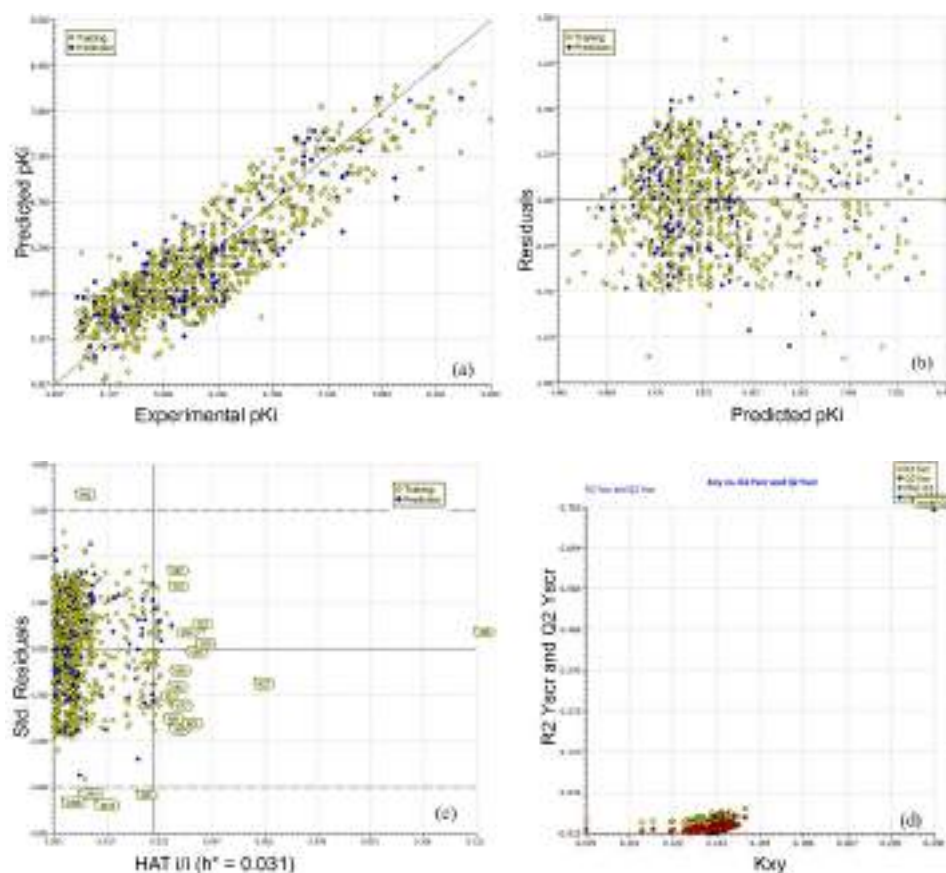


Figure 3. Different graphs associated with the model: (a) experimental vs. predicted pIC_{50} (the solid line represents the regression line); (b) experimental vs. residuals values; (c) Williams plot for applicability domain (the vertical solid line represents $h^* = 0.031$ and horizontal dashed lines represent the upper and lower boundaries for standard residuals); (d) Y-randomization.

The descriptions of seven molecular descriptors constituting model-A have been tabulated in Table 1.

Table 1. Details of molecular descriptors present in model-A.

Variable	Description	Software Used for Calculation
fsp3CringC2B	Frequency of occurrence of ring carbon atoms exactly at 2 bonds from sp ³ -hybridised carbon atoms	PyDescriptor [45]
com_C_4A	Total number of carbon atoms within 4 Å from centre of mass (com) of molecule	PyDescriptor
Saturated_Carbo_Rings	Total number of saturated rings containing carbon atoms only	DataWarrior [46]
fsulfonSaroC8B	Frequency of occurrence of aromatic carbon atoms exactly at 8 bonds from sulphur atoms of Sulfone group	PyDescriptor
fsp3OaroN6B	Frequency of occurrence of aromatic nitrogen atoms exactly at 6 bonds from sp ³ -hybridised oxygen atoms	PyDescriptor
flipoacc3B	Frequency of occurrence of H-bond acceptor atoms exactly at 3 bonds from lipophilic atoms	PyDescriptor
fplaNN4B	Frequency of occurrence of nitrogen atoms exactly at 4 bonds from planer nitrogen atoms	PyDescriptor

Interestingly, five molecular descriptors, viz., com_C_4A, fsp3CringC2B, flipoacc3B, fsulfonSaroC8B, and Saturated_Carbo_Rings, comprise the presence of different types of carbon atoms, which indicates the importance of carbon atoms in deciding BRD-4 inhibitory activity. The same is true for nitrogen, which is a part of three molecular descriptors, viz., flipoacc3B, fplaNN4B, and fsp3OaroN6B. Since, in general, the presence of carbon increases lipophilicity whereas nitrogen is attributed to significantly influence the pharmacological and hydrophilic profile, therefore, a balance of an appropriate number of carbons for lipophilicity and nitrogen is necessary to obtain adequate BRD-4 inhibitory activity. Of the seven descriptors in model-A, six have positive coefficients and only one has a negative coefficient. The effects of descriptors and their role in deciding the BRD-4 inhibitory profile have been discussed in more detail with relevant examples in the Discussion section.

3. Discussion

Mechanistic Interpretation of QSAR Model

An appropriately validated relationship between prominent structural features or molecular descriptors of the molecules with the bioactivity enlarges knowledge about mechanism of action of molecules, reasons for their specificity, and pharmacophoric atoms/groups accountable for the desired bioactivity [20,26,39]. In the present analysis, although we have equated the IC₅₀ values of different molecules in a relationship with a specific molecular descriptor (or feature), a synergistic or reverse effect of other molecular descriptors or unknown factors having a superseding influence in deciding the overall IC₅₀ value of a molecule cannot be ignored. That is, a single molecular descriptor or feature neither decides nor completely explains the experimental IC₅₀ value for such a large and structurally diverse set of molecules. In other words, the effective use of a validated QSAR model depends on the synergetic consideration of constituent molecular descriptors. The newly developed QSAR model-A comprises seven descriptors.

The molecular descriptor fsp3CringC2B represents the frequency of occurrence of ring carbon atoms exactly at two bonds from sp³-hybridised carbon atoms. If the same ring carbon atom was also present at less than two bonds from any other sp³-hybridised carbon atoms, then it was excluded while calculating fsp3CringC2B. Its positive coefficient in model-A and also a correlation of 0.30 with pIC₅₀ indicate that increasing such a combination of ring and sp³-hybridised carbon atoms could lead to better inhibitory activities for BRD-4. For example, a comparison of molecule 736 with 737 indicates the significant influence of ring carbon atoms (shown using green dots in Figure 4a,b) at exactly two bonds from sp³-hybridised carbon atoms. This is further supported by their reported

X-ray resolved structures with BRD-4. Molecule 736 (pdb: 5z1s [47]) has an additional water-mediated interaction with receptors with a distance of 3.37 Å (see Figure 4c) due to the -OCH₃ group present in the benzoxazinone ring. The same -OCH₃ group is responsible for increasing the value of fsp3CringC2B for 736, but it is absent in 737 (pdb: 5z1r [47]). The difference in IC₅₀ for the following pairs of molecules further support the influence of fsp3CringC2B on the activity profile: 255 with 499, 725 with 716, 231 with 240, and 89 with 105, to mention a few.

From this discussion, it appears that ring carbon atoms alone are important. However, replacing fsp3CringC2B by number of ring carbon atoms as a descriptor in model-A reduced the statistical performance from $R^2_{tr} = 0.76$ to 0.73. In addition, the number of ring carbon atoms has a correlation of 0.27 with pIC₅₀. Therefore, fsp3CringC2B is a better descriptor than the number of ring carbon atoms.

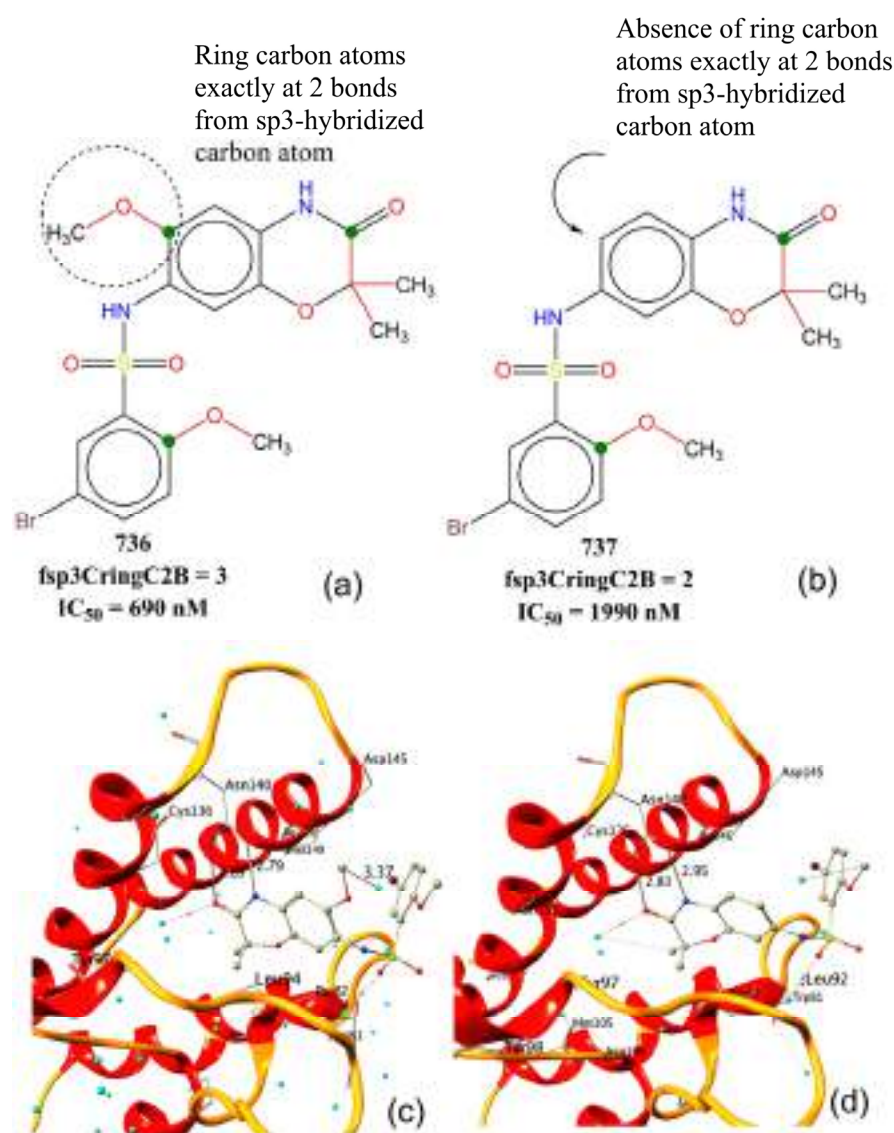


Figure 4. Comparison of BRD-4 inhibitory activity of molecule 736 with 737 with respect to fsp3CringC2B: (a,b) 2D representation of 736 and 737; (c,d) X-ray-resolved structures of 737 (pdb: 5z1r) and 736 (pdb: 5z1s). Cyan coloured spheres represent water molecules, and dashed lines signify interactions and distances in Å.

com_C_4A, which stands for total number of carbon atoms within 4 Å from centre of mass (com) of molecule, has a positive coefficient in model-A. Therefore, increasing the value of com_C_4A leads to better inhibitory activities. This observation is supported

by the fact that it has a correlation of 0.404 with pIC_{50} , and molecules with $\text{IC}_{50} < 10$ nM (29 molecules) possess a high value of com_C_4A . In addition, a simple comparison of the following pairs of the molecules strengthens this observation: 620 with 621, 720 with 710, 724 with 717, 526 with 518, and 691 with 692, and 595 with 596. At first glance, it looks as if com_C_4A is pointing out the importance of the number of carbon atoms. However, nC (number of carbon atoms) has a correlation of 0.29 with pIC_{50} and substituting com_C_4A by nC led to a decrease in statistical performance of model-A from $R^2_{\text{tr}} = 0.76$ to 0.69. Therefore, com_C_4A is a better choice as a variable for model-A.

As the presence of carbon is generally associated with the increased lipophilicity of a molecule, therefore, com_C_4A indicates that the lipophilic part must be concentrated near the com of the molecule for better activities. This in turn provides a crucial hint about the active site of BRD-4. It appears that a significant portion of the active site of BRD-4 is reasonably lipophilic in nature. This is supported by the fact that the active site of BRD-4 consists of a hydrophobic WPF shelf (see Figure 1) [11,18]. Thus, the findings of the present QSAR analysis are supported by the reported X-ray-resolved structure of BRD-4 enzyme. In addition, the pharmacophore model, depicted in Figure 5, generated using most active molecule 297, again points out the presence of a lipophilic region near the com of the molecule.

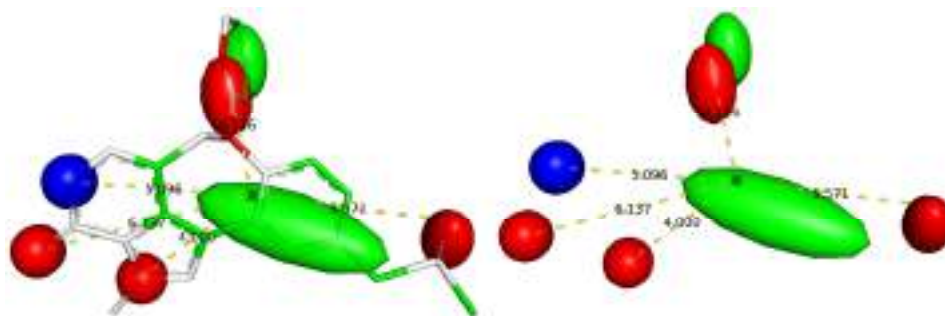


Figure 5. Pharmacophore model using most active molecule 207 (red: H-bond acceptor; blue: H-bond donor; green: hydrophobic region). Distances are shown using yellow dashed lines and figures indicate the distances in Å for different regions from center of mass. The black sphere represents the position of the center of mass of molecule.

Another descriptor that also point outs the importance of lipophilicity of a molecule is flipoacc3B , which represents the frequency of occurrence of H-bond acceptor atoms exactly at three bonds from lipophilic atoms. However, an acceptor atom was excluded while calculating flipoacc3B if it is also present within two or less bonds from same or any other lipophilic atom. Evidently, the lipophilic part of a molecule close to H-bond acceptor moiety (O or N atoms) plays a crucial role in deciding the inhibitory effect for BRD-4. This is once again visible in the pharmacophore model depicted in Figure 6. A simple comparison of the following pairs of molecules supports this observation: 812 with 823, 7 with 8, and 4 with 10.

An easily interpretable and influential molecular descriptor is $\text{Saturated_Carbo_Rings}$, which corresponds to total number of saturated carbocyclic rings. It has positive coefficient in model-A; therefore, increasing such rings is beneficial. A comparison of IC_{50} for 411 with 384 (see Figure 6), 137 with 127, 73 with 67, 131 with 124, 60 with 61, 570 with 573, and 572, 230 with 247 is in favour of this observation.

The importance of $\text{Saturated_Carbo_Rings}$ in model-A indicates that the lipophilicity and flexibility of a molecule are the actual factors governing an activity profile. It is noteworthy that clogP , which represents molecular lipophilicity, has a correlation of 0.193 with pIC_{50} , whereas $\text{Saturated_Carbo_Rings}$ has 0.240. Thus, $\text{Saturated_Carbo_Rings}$ is a better choice, as it pinpoints the specific feature or part of the molecule (saturated carbocyclic rings), which is correlated with the activity due to its lipophilic nature, whereas clogP is a molecular property. A plausible reason could be the crucial role played by hydrophobic zones such as saturated carbocyclic rings at the periphery or outer part of

molecule as a proxy for BRD4 selectivity through their interaction with the WPF shelf [6]. Therefore, saturated carbocyclic rings should be retained in future optimizations for better activity profiles. Thus, the present work is successful in identifying the significance of saturated carbocyclic rings as a novel unreported pharmacophore feature associated with BRD-4 inhibitory activity.

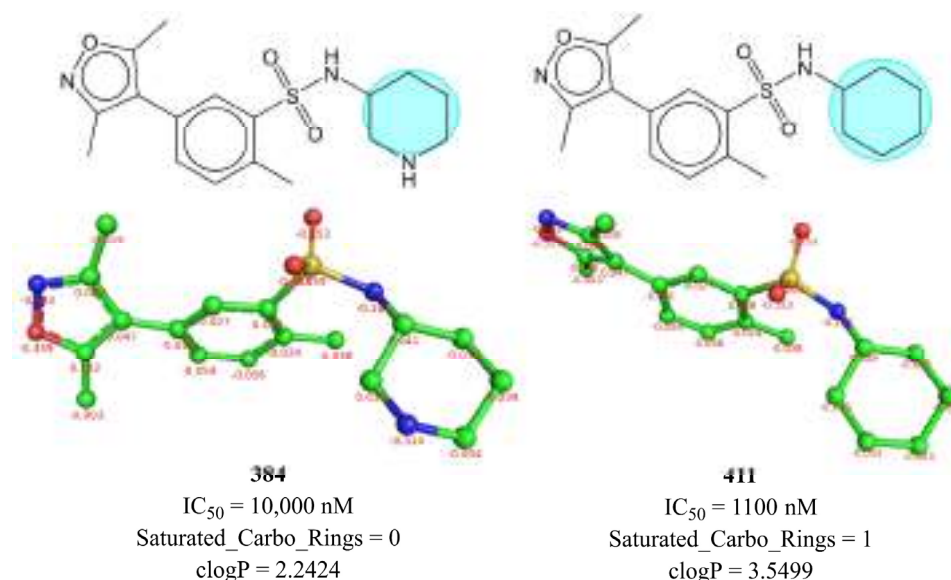


Figure 6. Depiction of influence of Saturated_Carbo_Rings on IC_{50} using molecule 384 and 411 as representative examples (2D and 3D representations with partial charges).

The molecular descriptor *fsulfonSaroC8B* (frequency of occurrence of aromatic carbon atoms exactly at eight bonds from sulphur atoms of Sulfone ($-SO_2-$) group) has a positive coefficient in model-A. Consequently, increasing the value of *fsulfonSaroC8B*, favours binding with BRD-4. It is to be noted that if the same aromatic carbon atom is also present at ≤ 7 bonds from a sulphur atom of same or different Sulfone group through any path, then it was excluded while calculating *fsulfonSaroC8B*. Obviously, this descriptor signifies the importance of the Sulfone group (a highly polar group) and its correlation with aromatic rings (a lipophilic moiety) in deciding the binding with BRD-4. This is clearly reflected in the difference in the activity of the following pairs of molecules: 715 with 718, 723 with 724, 714 with 715, 707 with 716, 936 with 941, and 942 with 943. Recent studies indicate that Sulfone moiety is present in a cleft near the ZA-loop and establishes an H-bond with the $-CONH-$ (amide) of backbone [19].

fsp3OaroN6B stands for the frequency of occurrence of aromatic nitrogen atoms exactly at six bonds from sp^3 -hybridised oxygen atom. If the same aromatic nitrogen atom is also present at ≤ 5 bonds from the same or any other sp^3 -hybridised oxygen atom through any path, then it was excluded while calculating *fsp3OaroN6B*. For example, 79 with 609, 81 with 620, and 614 with 615, to mention a few. In our previous study [20], we identified a similar descriptor *notringO_acc_6B* (total number of all non-ring Oxygen atoms present within a distance of six bonds from H-bond acceptor atoms) as an important pharmacophoric feature that governs the binding affinity (K_i) of a molecule for BRD-4. Thus, a consensus between the previous and the present study indicates that a molecule must have an H-bond acceptor (preferably aromatic nitrogen) at a distance of six bonds from a sp^3 -hybridised oxygen atom (non-ring oxygen favoured). This observation is supported by the difference in the activity of molecule S1, S2, and S3 [48] (see Figure 7). To add further, the sp^3 -hybridised oxygen atom is present as a linker between two aromatic rings or as an $-OR$ (alkoxy) group in a good number of molecules [20].

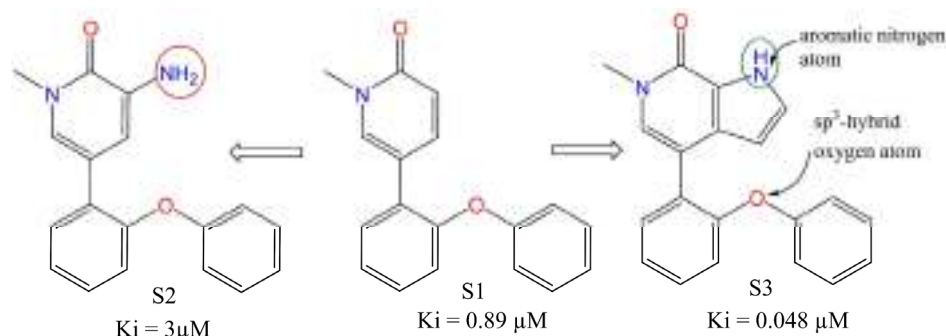


Figure 7. Effect of fsp3OaroN6B on BRD-4 inhibitory activity.

The only descriptor with a negative coefficient in model-A is fplaNN4B, which corresponds to frequency of occurrence of nitrogen atoms exactly at 4 bonds from planer nitrogen atoms. If the same nitrogen atom is also present at ≤ 3 bonds from the same or any other planer nitrogen atom through any path, then it was excluded while calculating fplaNN4B. The following pairs of molecules have a significant difference in their activities, which could be attributed to the presence of fplaNN4B as a structural feature: 945 with 954, 936 with 944, 411 with 385, 170 with 171, 722 with 714, 563 with 577, 737 with 734, 239 with 240. Therefore, such a combination of nitrogen atoms should be avoided to have better inhibition of BRD-4.

4. Materials and Methods

The present work follows the standard procedure recommended by OECD and different researchers to perform QSAR analysis [49–51]. All the software were used with default settings to obtain a QSAR model possessing a balance of predictive ability and mechanistic interpretation; however, some settings were changed, which have been reported at appropriate places. The different steps are as follows.

Step-1: Collection of data and curation: The present work commenced with the collection of a large dataset of 2026 experimentally tested BRD-4 inhibitors from a free and publicly available database BindingDB (<https://www.bindingdb.org/bind/index.jsp>, accessed on 16 March 2022). A QSAR analysis is significantly influenced by the quality of data, its composition, and its appropriate curation before further processing [50,52–54]. Therefore, in the next step, data curation was performed [55], which involved the removal of duplicate entries, organometallic compounds, salts, molecules with ambiguous IC₅₀ values, etc. This reduced the dataset to 980 molecules only. The reduced dataset still consists of molecules with experimental IC₅₀ (nM) in the range 1 nM to 15 μM and the presence of diverse scaffolds such as heterocyclic rings, positional isomers, stereoisomers, etc., enhancing the chemical space and consequently widening the applicability of the newly developed model. The SMILES (Simplified Molecular Input Line Entry System) notations, including experimental IC₅₀ and pIC₅₀ ($= -\log_{10} \text{IC}_{50}$) of all the molecules used in the present work, are available in Supplementary Materials. For the sake of convenience, representative examples have been presented in Figure 8 to depict the structural diversity of the current dataset.

In Table 2, five most and least active molecules have been included as examples only along with their SMILES notation: IC₅₀ (nM) and pIC₅₀ (M).

Step-2: In the next step, SMILES notations were used to develop the optimised 3D structures (semi-empirical PM3 method) of the molecules, accomplished using OpenBabel 2.4 [56] and MOPAC 2012 (openmopac.net) using default settings.

Step-3: A QSAR model achieves a balance of mechanistic interpretation and predictive ability if a good number of diverse molecular descriptors are calculated, followed by adequate pruning to reduce the chances of overfitting from noisy redundant descriptors [57]. The next step involved the calculation of myriad number of 1D- to 3D-molecular descriptors for all molecules. For this, *PyDescriptor* [45] and *DataWarrior* [46] were used, which

generated more than 40,000 molecular descriptors for a single molecule. Obviously, the descriptor pool contained a good number of redundant molecular descriptors; therefore, highly correlated ($|R| > 0.95$) and nearly constant ($>98\%$) variables were removed using QSARINS 2.2.4 [58]. This considerably decreased the size of set of molecular descriptor pool from 30,000 to 4326, which still contained a variety of molecular descriptors.

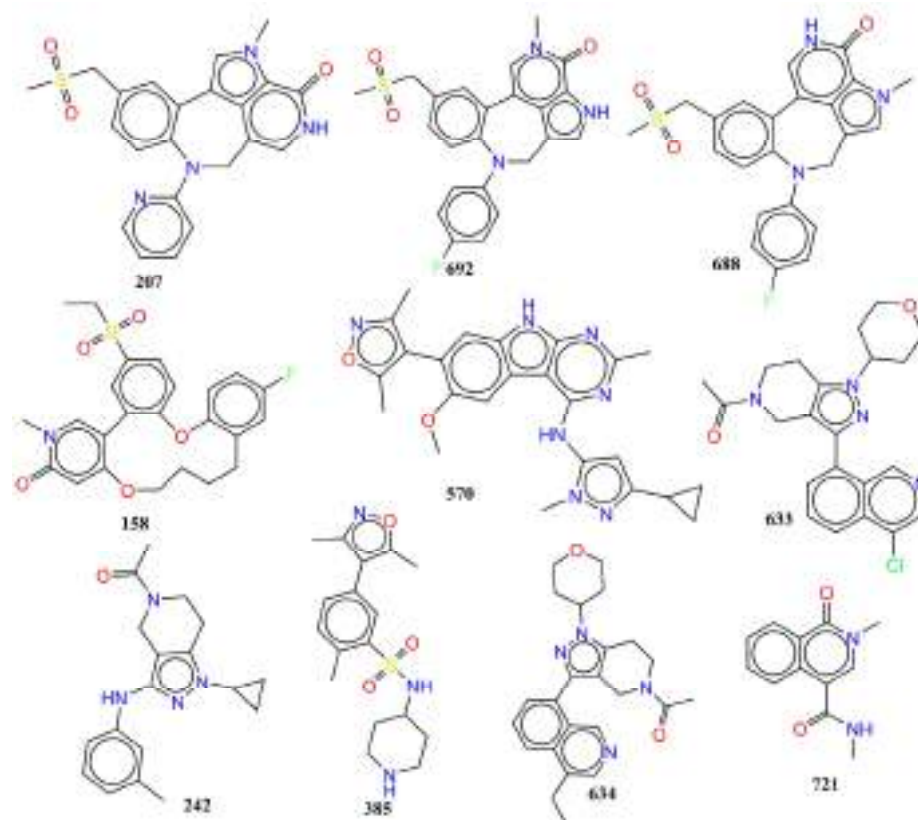


Figure 8. Representative examples to indicate the structural scaffold diversity in the present dataset.

Table 2. SMILES notation and IC_{50} (nM) and pIC_{50} (M) of five most and least active molecules of the selected data set.

SN	Ligand SMILES	IC_{50} (nM)	pIC_{50} (M)
207	<chem>Cn1cc2-c3cc(CS(C)(=O)=O)ccc3N(Cc3c[nH]c(=O)c1c23)c1cccn1</chem>	1	9
692	<chem>Cn1cc2-c3cc(CS(C)(=O)=O)ccc3N(Cc3c[nH]c(=O)c1c23)c1cc(F)cc1</chem>	1.5	8.824
158	<chem>CCS(=O)(=O)c1ccc2Oc3ccc(F)cc3CCCCOc3cc(=O)n(C)cc3-c2c1</chem>	2	8.699
570	<chem>COc1cc2c(cc1-c1c(C)noc1C)[nH]c1nc(C)nc(Nc3cc(nn3C)C3CC3)c21</chem>	2	8.699
688	<chem>Cn1cc2CN(c3ccc(F)cc3)c3ccc(CS(C)(=O)=O)cc3-c3c[nH]c(=O)c1c23</chem>	2.5	8.602
633	<chem>CC(=O)N1CCc2c(C1)c(nn2C1CCOCC1)-c1cccc2c(Cl)cncc12</chem>	14,000	4.854
242	<chem>CC(=O)N1CCc2c(C1)c(Nc1cccc(C)c1)nn2C1CC1</chem>	15,000	4.824
385	<chem>Cc1noc(C)c1-c1ccc(C)c(c1)S(=O)(=O)NC1CCNCC1</chem>	15,000	4.824
634	<chem>CCc1cncc2c(cccc12)-c1nn(C2CCOCC2)c2CCN(Cc12)C(C)=O</chem>	15,000	4.824
721	<chem>CNC(=O)c1cn(C)c(=O)c2cccc12</chem>	15,000	4.824

4.1. Splitting the Data Set into Training and External Sets and Subjective Feature Selection (SFS)

For developing a useful QSAR model and its appropriate validation, it is essential to divide the dataset into training and external (also called as prediction or test set) sets [50,52–54,59]. Consequently, in the present work, the dataset was randomly divided into training (80% = 785 molecules) and external (20% = 195 molecules) sets to minimise any bias. The only purpose of the training set was to choose suitable number of variables

(molecular descriptors), whereas the external set was employed only for validation purpose, i.e., external validation of the model (Predictive QSAR). Genetic Algorithm (GA) and multi-linear regression (MLR) available in QSARINS 2.2.4 were used for model building. For this, Q^2_{LOO} was used as a fitness function, and the number of generations was set to 10,000. A decisive step in QSAR modelling is to select the optimum number of molecular descriptors for model building to avoid over-fitting and to obtain acceptable interpretability. Consequently, the heuristic search involved building multiple models from univariate to multivariate with the successive addition of molecular descriptors until there was an increase in the value of Q^2_{LOO} , which is called the breaking point [39,60]. A 2D graph between the number of molecular descriptors involved in the models and Q^2_{LOO} values has been depicted in Figure 9. The number of variables matching with the breaking point was considered optimal for model building as there was no improvement in the statistical performance of model upon the inclusion of additional molecular descriptors. The analysis led to the matching of breaking points with seven variables. Therefore, QSAR models with more than seven descriptors were excluded.

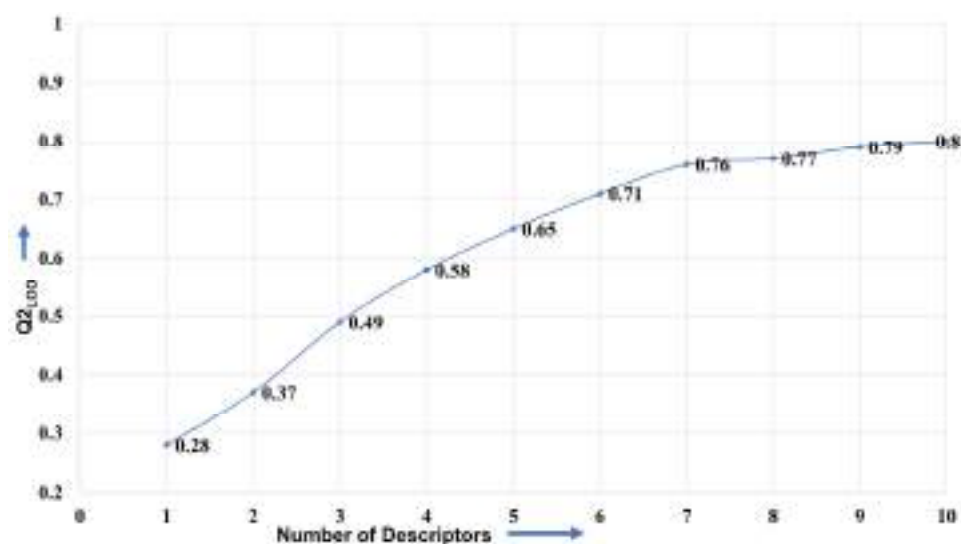


Figure 9. Graph between number of descriptors against leave-one-out coefficient of determination Q^2_{LOO} to identify the optimum number of descriptors.

4.2. Building Regression Model and Its Validation

Appropriate validations involving cross/inter validation, external validation, Y-randomization analysis, and applicability domain (Williams plot) are necessary to estimate the reliability and general applicability of a QSAR model [25,31,33,50,61]. A properly validated QSAR model finds its usage for QSAR-based virtual screening, lead/hit optimization, decision making, etc. The following validation parameters and their recommended threshold values are usually used to assess a model [39,60]: R^2_{tr} (coefficient of determination) ≥ 0.6 ; Q^2_{100} (cross-validated coefficient of determination for leave-one-out) ≥ 0.5 ; Q^2_{LMO} (cross-validated coefficient of determination for leave-many-out) ≥ 0.6 , $R^2 > Q^2$; R^2_{ex} (external coefficient of determination) ≥ 0.6 ; CCC (Concordance Correlation Coefficient) ≥ 0.80 ; $Q^2-F^n \geq 0.60$; high values of external validation parameters R^2_{ex} , Q^2_{F1} , Q^2_{F2} , and Q^2_{F3} , with low values of R^2_{Yscr} (coefficient of determination for Y-randomization); RMSE (Root mean square error); MAE (Mean absolute error); $RMSE_{tr} < RMSE_{cv}$. The formulae for calculating these statistical parameters are available in Supplementary Materials. In the present analysis, a Williams plot was used to assess the applicability domain of the newly developed QSAR model.

4.3. Pharmacophore Model

For pharmacophore modelling, the 3D-optimised structure of the most active molecule, 207, was selected. The model was generated using LIQUID [62,63], a free and easy to use PyMOL plugin, using default settings, except that the contour region was set to 3 for H-bond donor/acceptor and hydrophobic regions.

4.4. Other Experimental Details

The reported X-ray resolved structures (pdb 5z1r and 5z1s) were downloaded from Protein Data Bank (www.rcsb.org accessed on 13 April 2022). PyMOL version 2.4 has been used for the depiction of molecular interactions between the compounds and the protein.

5. Conclusions

In the present study, a seven-descriptor-based and rigorously validated GA-MLR QSAR model with $R^2_{tr} = 0.79$, $Q^2_{LMO} = 0.79$, and $R^2_{ex} = 0.78$ was derived to identify the significant pharmacophoric features that influence BRD-4 inhibitory activity. As mentioned earlier, it is essential to perceive salient and visually unrecognizable pharmacophoric features linked with BRD-4 inhibitory activity for different chemical scaffolds. The analysis indicates that the presence of ring carbon and nitrogen atoms, occurrence of carbon atoms near the center of mass of a molecule, specific combination of planer nitrogen with ring carbon, donor and acceptor atoms, etc., are prominent features to be retained in future optimizations. On the other hand, a combination of nitrogen atoms with planer nitrogen atoms exactly at four bonds should be avoided for better BRD-4 inhibitory activity. The reported crystal structures of BRD-4 inhibitors strengthen these observations. The present study efficaciously captured and reported novel pharmacophoric features and has a good balance of predictive ability and mechanistic interpretations.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/ph15060745/s1>.

Author Contributions: Conceptualization, V.H.M., M.E.A.Z. and A.S.; formal analysis and data curation: V.H.M., R.D.J. and S.A.A.-H.; writing, M.E.A.Z., V.H.M. and A.A.A.-M.; revisions, M.E.A.Z., V.H.M. and A.S.; editing and proofreading, V.H.M., M.E.A.Z., R.D.J. and A.A.A.-M.; funding and resources, S.A.A.-H. and M.E.A.Z. All authors have read and agreed to the published version of the manuscript.

Funding: The authors acknowledge the Deanship of Scientific Research at Imam Mohammad Ibn Saud Islamic University, Riyadh, Saudi Arabia, for its support of this research through research group number RG-21-09-77.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data is contained within the article and Supplementary Material.

Acknowledgments: V. H. Masand is thankful to Paola Gramatica (Italy) and her team for providing the free copy of QSARINS 2.2.4.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

SMILES	Simplified molecular-input line-entry system
GA	Genetic algorithm
MLR	Multiple linear regression
QSAR	Quantitative structure–activity relationship
WHO	World Health Organization
OLS	Ordinary least square
QSARINS	QSAR Insubria
OECD	Organisation for Economic Co-operation and Development

References


1. Boer, R.A.; Meijers, W.C.; Meer, P.; Veldhuisen, D.J. Cancer and heart disease: Associations and relations. *Eur. J. Heart Fail.* **2019**, *21*, 1515–1525. [[CrossRef](#)] [[PubMed](#)]
2. Fidanze, S.D.; Liu, D.; Mantei, R.A.; Hasvold, L.A.; Pratt, J.K.; Sheppard, G.S.; Wang, L.; Holms, J.H.; Dai, Y.; Aguirre, A.; et al. Discovery and optimization of novel constrained pyrrolopyridone BET family inhibitors. *Bioorg. Med. Chem. Lett.* **2018**, *28*, 1804–1810. [[CrossRef](#)] [[PubMed](#)]
3. Guest, E.E.; Pickett, S.D.; Hirst, J.D. Structural variation of protein–ligand complexes of the first bromodomain of BRD4. *Org. Biomol. Chem.* **2021**, *19*, 5632–5641. [[CrossRef](#)] [[PubMed](#)]
4. Liu, Z.; Chen, H.; Wang, P.; Li, Y.; Wold, E.A.; Leonard, P.G.; Joseph, S.; Brasier, A.R.; Tian, B.; Zhou, J. Discovery of Orally Bioavailable Chromone Derivatives as Potent and Selective BRD4 Inhibitors: Scaffold Hopping, Optimization, and Pharmacological Evaluation. *J. Med. Chem.* **2020**, *63*, 5242–5256. [[CrossRef](#)] [[PubMed](#)]
5. Alqahtani, A.; Choucair, K.; Ashraf, M.; Hammouda, D.M.; Alloghbi, A.; Khan, T.; Senzer, N.; Nemunaitis, J. Bromodomain and extra-terminal motif inhibitors: A review of preclinical and clinical advances in cancer therapy. *Future Sci. OA* **2019**, *5*, FSO372. [[CrossRef](#)] [[PubMed](#)]
6. Speck-Planche, A.; Scotti, M.T. BET bromodomain inhibitors: Fragment-based in silico design using multi-target QSAR models. *Mol. Divers* **2018**, *23*, 555–572. [[CrossRef](#)] [[PubMed](#)]
7. Duan, Y.; Guan, Y.; Qin, W.; Zhai, X.; Yu, B.; Liu, H. Targeting Brd4 for cancer therapy: Inhibitors and degraders. *MedChemComm* **2018**, *9*, 1779–1802. [[CrossRef](#)]
8. Zhao, Y.; Bai, L.; Liu, L.; McEachern, D.; Stuckey, J.A.; Meagher, J.L.; Yang, C.-Y.; Ran, X.; Zhou, B.; Hu, Y.; et al. Structure-Based Discovery of 4-(6-Methoxy-2-methyl-4-(quinolin-4-yl)-9H-pyrimido[4,5-b]indol-7-yl)-3,5-dimethylisoxazole (CD161) as a Potent and Orally Bioavailable BET Bromodomain Inhibitor. *J. Med. Chem.* **2017**, *60*, 3887–3901. [[CrossRef](#)]
9. Xing, J.; Lu, W.; Liu, R.; Wang, Y.; Xie, Y.; Zhang, H.; Shi, Z.; Jiang, H.; Liu, Y.-C.; Chen, K.; et al. Machine-Learning-Assisted Approach for Discovering Novel Inhibitors Targeting Bromodomain-Containing Protein 4. *J. Chem. Inf. Model.* **2017**, *57*, 1677–1690. [[CrossRef](#)]
10. Kuang, M.; Zhou, J.; Wang, L.; Liu, Z.; Guo, J.; Wu, R. Binding Kinetics versus Affinities in BRD4 Inhibition. *J. Chem. Inf. Model.* **2015**, *55*, 1926–1935. [[CrossRef](#)]
11. Ember, S.W.J.; Zhu, J.-Y.; Olesen, S.H.; Martin, M.P.; Becker, A.; Berndt, N.; Georg, G.I.; Schönbrunn, E. Acetyl-lysine Binding Site of Bromodomain-Containing Protein 4 (BRD4) Interacts with Diverse Kinase Inhibitors. *ACS Chem. Biol.* **2014**, *9*, 1160–1171. [[CrossRef](#)] [[PubMed](#)]
12. Shorstova, T.; Foulkes, W.D.; Witcher, M. Achieving clinical success with BET inhibitors as anti-cancer agents. *Br. J. Cancer* **2021**, *124*, 1478–1490. [[CrossRef](#)] [[PubMed](#)]
13. Wang, L.; Pratt, J.K.; Soltwedel, T.; Sheppard, G.S.; Fidanze, S.D.; Liu, D.; Hasvold, L.A.; Mantei, R.A.; Holms, J.H.; McClellan, W.J.; et al. Fragment-Based, Structure-Enabled Discovery of Novel Pyridones and Pyridone Macrocycles as Potent Bromodomain and Extra-Terminal Domain (BET) Family Bromodomain Inhibitors. *J. Med. Chem.* **2017**, *60*, 3828–3850. [[CrossRef](#)] [[PubMed](#)]
14. Filippakopoulos, P.; Qi, J.; Picaud, S.; Shen, Y.; Smith, W.B.; Fedorov, O.; Morse, E.M.; Keates, T.; Hickman, T.T.; Felletar, I.; et al. Selective inhibition of BET bromodomains. *Nature* **2010**, *468*, 1067–1073. [[CrossRef](#)] [[PubMed](#)]
15. Liu, Z.; Wang, P.; Chen, H.; Wold, E.A.; Tian, B.; Brasier, A.R.; Zhou, J. Drug Discovery Targeting Bromodomain-Containing Protein 4. *J. Med. Chem.* **2017**, *60*, 4533–4558. [[CrossRef](#)]
16. Filippakopoulos, P.; Picaud, S.; Mangos, M.; Keates, T.; Lambert, J.-P.; Barseyte-Lovejoy, D.; Felletar, I.; Volkmer, R.; Müller, S.; Pawson, T.; et al. Histone Recognition and Large-Scale Structural Analysis of the Human Bromodomain Family. *Cell* **2012**, *149*, 214–231. [[CrossRef](#)]
17. Donati, B.; Lorenzini, E.; Ciarrocchi, A. BRD4 and Cancer: Going beyond transcriptional regulation. *Mol. Cancer* **2018**, *17*, 164. [[CrossRef](#)]
18. Zaware, N.; Zhou, M.-M. Bromodomain biology and drug discovery. *Nat. Struct. Mol. Biol.* **2019**, *26*, 870–879. [[CrossRef](#)]
19. Sheppard, G.S.; Wang, L.; Fidanze, S.D.; Hasvold, L.A.; Liu, D.; Pratt, J.K.; Park, C.H.; Longenecker, K.; Qiu, W.; Torrent, M.; et al. Discovery of N-Ethyl-4-[2-(4-fluoro-2,6-dimethyl-phenoxy)-5-(1-hydroxy-1-methyl-ethyl)phenyl]-6-methyl-7-oxo-1H-pyrrolo[2,3-c]pyridine-2-carboxamide (ABBV-744), a BET Bromodomain Inhibitor with Selectivity for the Second Bromodomain. *J. Med. Chem.* **2020**, *63*, 5585–5623. [[CrossRef](#)]
20. Masand, V.H.; Patil, M.K.; El-Sayed, N.N.E.; Zaki, M.E.A.; Almarhoon, Z.; Al-Hussain, S.A. Balanced QSAR analysis to identify the structural requirements of ABBV-075 (Mivebresib) analogues as bromodomain and extraterminal domain (BET) family bromodomain inhibitor. *J. Mol. Struct.* **2021**, *1229*, 129597. [[CrossRef](#)]
21. Tahir, A.; Alharthy, R.D.; Naseem, S.; Mahmood, N.; Ahmed, M.; Shahzad, K.; Akhtar, M.N.; Hameed, A.; Sadiq, I.; Nawaz, H.; et al. Investigations of Structural Requirements for BRD4 Inhibitors through Ligand- and Structure-Based 3D QSAR Approaches. *Molecules* **2018**, *23*, 1527. [[CrossRef](#)] [[PubMed](#)]
22. Tong, J.-B.; Luo, D.; Feng, Y.; Bian, S.; Zhang, X.; Wang, T.-H. Structural modification of 4, 5-dihydro-[1, 2, 4] triazolo [4, 3-f] pteridine derivatives as BRD4 inhibitors using 2D/3D-QSAR and molecular docking analysis. *Mol. Divers.* **2021**, *25*, 1855–1872. [[CrossRef](#)] [[PubMed](#)]

23. Babatunde Samuel Obadawo, O.E.O. Mayowa Monday Anifowose, Kehinde Henry Fagbohunge, Justinah Solayide Amoko, QSAR modeling of novel substituted 4-Phenylisoquinolinones as potent BET bromodomain (BRD4-BD1) inhibitors. *Biomed. Lett.* **2019**, *5*, 69–78.
24. Zaki, M.E.A.; Al-Hussain, S.A.; Masand, V.H.; Akasapu, S.; Lewaa, I. QSAR and Pharmacophore Modeling of Nitrogen Heterocycles as Potent Human N-Myristoyltransferase (Hs-NMT) Inhibitors. *Molecules* **2021**, *26*, 1834. [[CrossRef](#)]
25. Gramatica, P. Principles of QSAR Modeling. *Int. J. Quant. Struct. Prop. Relatsh.* **2020**, *5*, 61–97. [[CrossRef](#)]
26. Polishchuk, P. Interpretation of Quantitative Structure–Activity Relationship Models: Past, Present, and Future. *J. Chem. Inf. Model.* **2017**, *57*, 2618–2639. [[CrossRef](#)]
27. Fujita, T.; Winkler, D.A. Understanding the Roles of the “Two QSARs”. *J. Chem. Inf. Model.* **2016**, *56*, 269–274. [[CrossRef](#)]
28. Krstajic, D.; Buturovic, L.J.; Leahy, D.E.; Thomas, S. Cross-validation pitfalls when selecting and assessing regression and classification models. *J. Cheminform.* **2014**, *6*, 10. [[CrossRef](#)]
29. Gramatica, P. External Evaluation of QSAR Models, in Addition to Cross-Validation Verification of Predictive Capability on Totally New Chemicals. *Mol. Inf.* **2014**, *33*, 311–314. [[CrossRef](#)]
30. Gütlein, M.; Helma, C.; Karwath, A.; Kramer, S. A Large-Scale Empirical Evaluation of Cross-Validation and External Test Set Validation in (Q)SAR. *Mol. Inf.* **2013**, *32*, 516–528. [[CrossRef](#)]
31. Gramatica, P. On the development and validation of QSAR models. *Methods Mol. Biol.* **2013**, *930*, 499–526. [[CrossRef](#)] [[PubMed](#)]
32. Chirico, N.; Gramatica, P. Real external predictivity of QSAR models. Part 2. New intercomparable thresholds for different validation criteria and the need for scatter plot inspection. *J. Chem. Inf. Model.* **2012**, *52*, 2044–2058. [[CrossRef](#)] [[PubMed](#)]
33. Chirico, N.; Gramatica, P. Real external predictivity of QSAR models: How to evaluate it? Comparison of different validation criteria and proposal of using the concordance correlation coefficient. *J. Chem. Inf. Model.* **2011**, *51*, 2320–2335. [[CrossRef](#)] [[PubMed](#)]
34. Consonni, V.; Ballabio, D.; Todeschini, R. Comments on the definition of the Q₂ parameter for QSAR validation. *J. Chem. Inf. Model.* **2009**, *49*, 1669–1678. [[CrossRef](#)]
35. Rao, R.B.; Fung, G.; Rosales, R. *On the Dangers of Cross-Validation. An Experimental Evaluation*; SIAM: Philadelphia, PA, USA, 2008; pp. 588–596. [[CrossRef](#)]
36. Gramatica, P.; Giani, E.; Papa, E. Statistical external validation and consensus modeling: A QSPR case study for Koc prediction. *J. Mol. Graph. Model.* **2007**, *25*, 755–766. [[CrossRef](#)]
37. Hawkins, D.M.; Basak, S.C.; Mills, D. Assessing model fit by cross-validation. *J. Chem. Inf. Comput. Sci.* **2003**, *43*, 579–586. [[CrossRef](#)]
38. Masand, V.H.; Mahajan, D.T.; Nazeruddin, G.M.; Hadda, T.B.; Rastija, V.; Alfeefy, A.M. Effect of information leakage and method of splitting (rational and random) on external predictive ability and behavior of different statistical parameters of QSAR model. *Med. Chem. Res.* **2015**, *24*, 1241–1264. [[CrossRef](#)]
39. Zaki, M.E.A.; Al-Hussain, S.A.; Bukhari, S.N.A.; Masand, V.H.; Rathore, M.M.; Thakur, S.D.; Patil, V.M. Exploring the Prominent and Concealed Inhibitory Features for Cytoplasmic Isoforms of Hsp90 Using QSAR Analysis. *Pharmaceuticals* **2022**, *15*, 303. [[CrossRef](#)]
40. Kar, S.; Roy, K.; Leszczynski, J. Applicability Domain: A Step Toward Confident Predictions and Decidability for QSAR Modeling. In *Computational Toxicology*; Humana Press: New York, NY, USA, 2018; pp. 141–169.
41. Gramatica, P.; Kovarich, S.; Roy, P.P. Reply to the comment of S. Rayne on “QSAR model reproducibility and applicability: A case study of rate constants of hydroxyl radical reaction models applied to polybrominated diphenyl ethers and (benzo-)triazoles”. *J. Comput. Chem.* **2013**, *34*, 1796. [[CrossRef](#)]
42. Roy, P.P.; Kovarich, S.; Gramatica, P. QSAR model reproducibility and applicability: A case study of rate constants of hydroxyl radical reaction models applied to polybrominated diphenyl ethers and (benzo-)triazoles. *J. Comput. Chem.* **2011**, *32*, 2386–2396. [[CrossRef](#)]
43. Gadaleta, D.; Mangiatordi, G.F.; Catto, M.; Carotti, A.; Nicolotti, O. Applicability Domain for QSAR Models. *Int. J. Quant. Struct. Prop. Relatsh.* **2016**, *1*, 45–63. [[CrossRef](#)]
44. Tropsha, A.; Golbraikh, A. Predictive QSAR modeling workflow, model applicability domains, and virtual screening. *Curr. Pharm. Des.* **2007**, *13*, 3494–3504. [[CrossRef](#)] [[PubMed](#)]
45. Masand, V.H.; Rastija, V. PyDescriptor: A new PyMOL plugin for calculating thousands of easily understandable molecular descriptors. *Chemom. Intell. Lab. Syst.* **2017**, *169*, 12–18. [[CrossRef](#)]
46. Sander, T.; Freyss, J.; von Korff, M.; Rufener, C. DataWarrior: An Open-Source Program For Chemistry Aware Data Visualization And Analysis. *J. Chem. Inf. Model.* **2015**, *55*, 460–473. [[CrossRef](#)] [[PubMed](#)]
47. Xiang, Q.; Zhang, Y.; Li, J.; Xue, X.; Wang, C.; Song, M.; Zhang, C.; Wang, R.; Li, C.; Wu, C.; et al. Y08060: A Selective BET Inhibitor for Treatment of Prostate Cancer. *ACS Med. Chem. Lett.* **2018**, *9*, 262–267. [[CrossRef](#)]
48. McDaniel, K.F.; Wang, L.; Soltwedel, T.; Fidanze, S.D.; Hasvold, L.A.; Liu, D.; Mantei, R.A.; Pratt, J.K.; Sheppard, G.S.; Bui, M.H.; et al. Discovery of N-(4-(2,4-Difluorophenoxy)-3-(6-methyl-7-oxo-6,7-dihydro-1H-pyrrolo[2,3-c]pyridin-4-yl)phenyl)ethanesulfonamide (ABBV-075/Mivebresib), a Potent and Orally Available Bromodomain and Extraterminal Domain (BET) Family Bromodomain Inhibitor. *J. Med. Chem.* **2017**, *60*, 8369–8384. [[CrossRef](#)]
49. Dearden, J.C.; Cronin, M.T.; Kaiser, K.L. How not to develop a quantitative structure-activity or structure-property relationship (QSAR/QSPR). *SAR QSAR Environ. Res.* **2009**, *20*, 241–266. [[CrossRef](#)]

50. Cherkasov, A.; Muratov, E.N.; Fourches, D.; Varnek, A.; Baskin, I.I.; Cronin, M.; Dearden, J.; Gramatica, P.; Martin, Y.C.; Todeschini, R.; et al. QSAR modeling: Where have you been? Where are you going to? *J. Med. Chem.* **2014**, *57*, 4977–5010. [[CrossRef](#)]
51. Huang, J.; Fan, X. Why QSAR fails: An empirical evaluation using conventional computational approach. *Mol. Pharm.* **2011**, *8*, 600–608. [[CrossRef](#)]
52. Muratov, E.N.; Bajorath, J.; Sheridan, R.P.; Tetko, I.V.; Filimonov, D.; Poroikov, V.; Oprea, T.I.; Baskin, I.I.; Varnek, A.; Roitberg, A.; et al. QSAR without borders. *Chem. Soc. Rev.* **2020**, *49*, 3525–3564. [[CrossRef](#)]
53. Golbraikh, A.; Muratov, E.; Fourches, D.; Tropsha, A. Data set modelability by QSAR. *J. Chem. Inf. Model.* **2014**, *54*, 1–4. [[CrossRef](#)] [[PubMed](#)]
54. Martin, T.M.; Harten, P.; Young, D.M.; Muratov, E.N.; Golbraikh, A.; Zhu, H.; Tropsha, A. Does rational selection of training and test sets improve the outcome of QSAR modeling? *J. Chem. Inf. Model.* **2012**, *52*, 2570–2578. [[CrossRef](#)] [[PubMed](#)]
55. Fourches, D.; Muratov, E.; Tropsha, A. Trust, but verify: On the importance of chemical structure curation in cheminformatics and QSAR modeling research. *J. Chem. Inf. Model.* **2010**, *50*, 1189–1204. [[CrossRef](#)] [[PubMed](#)]
56. O’Boyle, N.M.; Banck, M.; James, C.A.; Morley, C.; Vandermeersch, T.; Hutchison, G.R. Open Babel: An open chemical toolbox. *J. Cheminform.* **2011**, *3*, 33. [[CrossRef](#)]
57. Tetko, I.V.; Sushko, I.; Pandey, A.K.; Zhu, H.; Tropsha, A.; Papa, E.; Oberg, T.; Todeschini, R.; Fourches, D.; Varnek, A. Critical assessment of QSAR models of environmental toxicity against *Tetrahymena pyriformis*: Focusing on applicability domain and overfitting by variable selection. *J. Chem. Inf. Model.* **2008**, *48*, 1733–1746. [[CrossRef](#)]
58. Gramatica, P.; Chirico, N.; Papa, E.; Cassani, S.; Kovarich, S. QSARINS: A new software for the development, analysis, and validation of QSAR MLR models. *J. Comput. Chem.* **2013**, *34*, 2121–2132. [[CrossRef](#)]
59. Tropsha, A.; Gramatica, P.; Gombar, V.K. The Importance of Being Earnest Validation is the Absolute Essential for Successful Application and Interpretation of QSPR Models. *QSAR Comb. Sci.* **2003**, *22*, 69–77. [[CrossRef](#)]
60. Zaki, M.E.A.; Al-Hussain, S.A.; Masand, V.H.; Sabnani, M.K.; Samad, A. Mechanistic and Predictive QSAR Analysis of Diverse Molecules to Capture Salient and Hidden Pharmacophores for Anti-Thrombotic Activity. *Int. J. Mol. Sci.* **2021**, *22*, 8352. [[CrossRef](#)]
61. Gramatica, P. Principles of QSAR models validation internal and external. *QSAR Comb. Sci.* **2007**, *26*, 694–701. [[CrossRef](#)]
62. Yuan, S.; Chan, H.C.S.; Hu, Z. Using PyMOL as a platform for computational drug design. *WIREs Comput. Mol. Sci.* **2017**, *7*, e1298. [[CrossRef](#)]
63. Tanrikulu, Y.; Nietert, M.; Scheffer, U.; Proschak, E.; Grabowski, K.; Schneider, P.; Weidlich, M.; Karas, M.; Gobel, M.; Schneider, G. Scaffold hopping by “fuzzy” pharmacophores and its application to RNA targets. *Chembiochem* **2007**, *8*, 1932–1936. [[CrossRef](#)] [[PubMed](#)]

Article

Perceiving the Concealed and Unreported Pharmacophoric Features of the 5-Hydroxytryptamine Receptor Using Balanced QSAR Analysis

Syed Nasir Abbas Bukhari ^{1,*}, Mervat Abdelaziz Elsherif ², Kashaf Junaid ³, Hasan Ejaz ³, Pravej Alam ⁴, Abdul Samad ⁵, Rahul D. Jawarkar ⁶ and Vijay H. Masand ^{7,*}

- ¹ Department of Pharmaceutical Chemistry, College of Pharmacy, Jouf University, Sakaka 72388, Saudi Arabia
² Chemistry Department, College of Science, Jouf University, Sakaka 72388, Saudi Arabia; maelsherif@ju.edu.sa
³ Department of Clinical Laboratory Sciences, College of Applied Medical Sciences, Jouf University, Sakaka 72388, Saudi Arabia; kjunaid@ju.edu.sa (K.J.); hetariq@ju.edu.sa (H.E.)
⁴ Department of Biology, College of Science and Humanities, Prince Sattam Bin Abdulaziz University, Al-Kharj 11942, Saudi Arabia; alamprez@gmail.com
⁵ Department of Pharmaceutical Chemistry, Faculty of Pharmacy, Tishk International University, Erbil 44001, Iraq; abdul.samad@tiu.edu.iq
⁶ Department of Medicinal Chemistry, Dr. Rajendra Gode Institute of Pharmacy, University-Mardi Road, Amravati 444603, Maharashtra, India; rahuljawarkar@gmail.com
⁷ Department of Chemistry, Vidya Bharati Mahavidyalaya, Amravati 444602, Maharashtra, India
* Correspondence: sbukhari@ju.edu.sa (S.N.A.B.); vijaymasand@gmail.com (V.H.M.)



Citation: Bukhari, S.N.A.; Elsherif, M.A.; Junaid, K.; Ejaz, H.; Alam, P.; Samad, A.; Jawarkar, R.D.; Masand, V.H. Perceiving the Concealed and Unreported Pharmacophoric Features of the 5-Hydroxytryptamine Receptor Using Balanced QSAR Analysis. *Pharmaceuticals* **2022**, *15*, 834. <https://doi.org/10.3390/ph15070834>

Academic Editors: Dongsheng Cao and Daniela Catarzi

Received: 23 May 2022

Accepted: 25 June 2022

Published: 5 July 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Abstract: The 5-hydroxytryptamine receptor 6 (5-HT₆) has gained attention as a target for developing therapeutics for Alzheimer's disease, schizophrenia, cognitive dysfunctions, anxiety, and depression, to list a few. In the present analysis, a larger and diverse dataset of 1278 molecules covering a broad chemical and activity space was used to identify visual and concealed structural features associated with binding affinity for 5-HT₆. For this, quantitative structure–activity relationships (QSAR) and molecular docking analyses were executed. This led to the development of a statistically robust QSAR model with a balance of excellent predictivity ($R^2_{tr} = 0.78$, $R^2_{ex} = 0.77$), the identification of unreported aspects of known features, and also novel mechanistic interpretations. Molecular docking and QSAR provided similar as well as complementary results. The present analysis indicates that the partial charges on ring carbons present within four bonds from a sulfur atom, the occurrence of sp³-hybridized carbon atoms bonded with donor atoms, and a conditional occurrence of lipophilic atoms/groups from nitrogen atoms, which are prominent but unreported pharmacophores that should be considered while optimizing a molecule for 5-HT₆. Thus, the present analysis led to identification of some novel unreported structural features that govern the binding affinity of a molecule. The results could be beneficial in optimizing the molecules for 5-HT₆.

Keywords: 5-hydroxytryptamine receptor 6; neurodegeneration; QSAR; molecular docking; pharmacophoric features

1. Introduction

Millions of people are suffering from central nervous system (CNS)-related diseases, such as Alzheimer's disease, schizophrenia, cognitive dysfunctions, anxiety, depression, etc. [1–4], and are facing diverse health and social challenges. Though therapeutic agents are available for the treatment of these diseases, they lack the ability to reduce the continuous loss of cognitive function, as most of them target only the acetylcholine deficit [2,3,5–9]. Therefore, there is a need to develop a therapeutic agent with a different mechanism of action and bio-activity profile to support the existing drug space. Recently, G-protein-coupled receptors (GPCRs), viz., the serotonergic system of serotonin (5-hydroxytryptamine) receptors, have received greater attention due to their vital role in signal transduction pathways

and numerous neurological functions [2,4,7–10]. To add further, 5-hydroxytryptamine receptors (5-HT_{1–7}) play vital roles in many cognitive dysfunctions such as memory loss, reduced learning ability, etc. [4,5,8,9]. Among them, the 5-HT₆ receptor has emerged as a promising target due to its crucial role in the onset of Alzheimer's disease, cognitive processes, mood control, depression, and anxiety [4,5,8,9], to name a few. It is believed that blocking the 5-HT₆ receptor significantly improves learning and memory processes [4,5,8,9]. In addition, it has also emerged as a molecular target for the treatment of obesity and the related metabolic syndrome [11]. An additional advantage associated with 5-HT₆ is its restricted and exclusive occurrence within the CNS, which implies that compounds acting through this receptor could have minimal peripheral side effects [2,3,6,10,12]. Even though several molecules have been identified as promising ligands with high binding affinities for 5-HT₆ (see Figure 1), none of them has cleared the clinical stages or been approved as a drug [1,7–10,12].

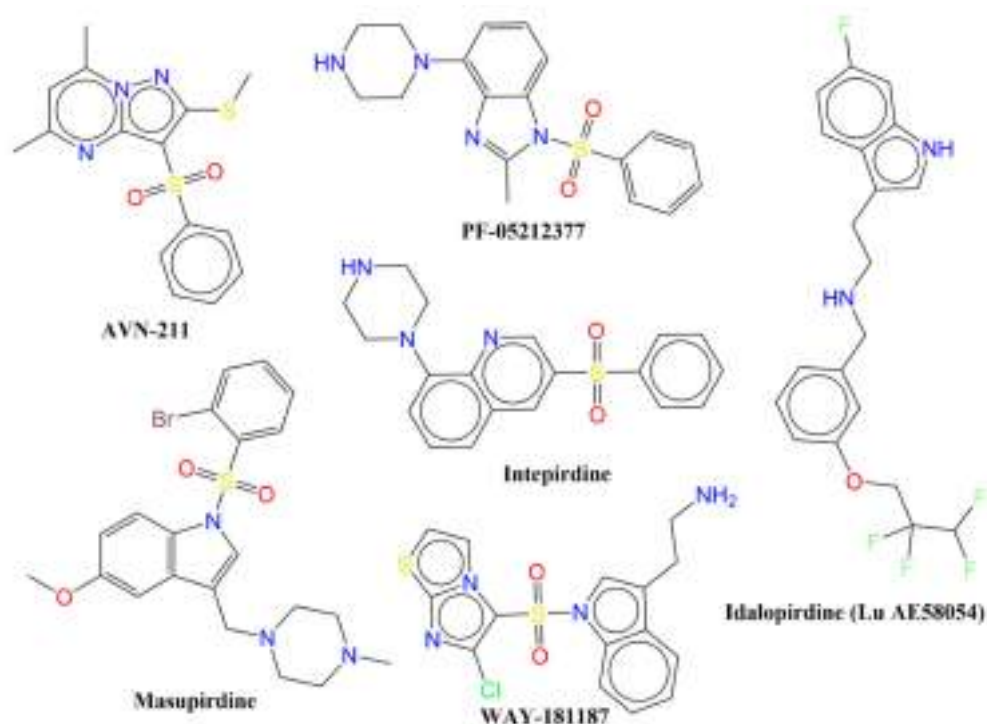


Figure 1. Chemical structures of representative examples of clinically tested 5-HT₆ ligands.

Therefore, there is a need to develop a novel therapeutic agent with a better ADMET (absorption, distribution, metabolism, excretion, and toxicity) profile and the retention of a high binding affinity for 5-HT₆. For this, it is essential to know the prominent and concealed pharmacophoric features associated with binding affinities for 5-HT₆ using a broad, structurally diverse dataset with adequate variations in activity space. To achieve these targets, computer-aided drug designing (CADD) is a contemporary and feasible solution due to its low cost and time efficacy [13,14]. Ligand-based drug designing (LBDD) is a thriving and widely accepted branch of CADD with a higher success rate of identification of key pharmacophoric features. It is the method of choice if the 3D structure of a target protein is not available [14]. Under the terrain of LBDD, the quantitative structure–activity relationship (QSAR) is an effective and multidisciplinary approach that is useful to identify salient and hidden pharmacophoric features [15–18].

Though different types of ligands encompassing diverse scaffolds are known for 5-HT₆, the lack of an X-ray-resolved 3D structure restricts the use of structure-based drug designing (SBDD) approaches. Consequently, many researchers have performed QSAR analysis for 5-HT₆ using different types of scaffolds. Doddareddy and co-workers [1] accomplished a 3D QSAR analysis using a small dataset of 33 N1-arylsulfonylindole

compounds as 5-HT6 antagonists. The analysis reinforced the pharmacophoric features reported by López-Rodríguez et al. [4,19]. Later in 2011, Hao et al. [3] studied relatively a larger dataset of 223 ligands reported for 5-HT6 for QSAR, homology-based molecular docking, and molecular dynamics simulations for 5 ns. Their analysis pointed out that the interaction with the residue Asp106 is important. A 2D and 3D QSAR analysis of arylsulfonamide-derived 5-HT6 receptor antagonists [12] highlighted the importance of the pharmacophore model reported by López-Rodríguez et al. [4,19]. Though these studies contributed to the identification of some important structural features, the developed QSAR models were either based on smaller datasets or specific scaffolds only, thereby lacking general applicability. In addition, poor external predictive ability restricts their usage for a lead optimization pipeline. A QSAR analysis based on a large and diverse set of molecules, with a balance of predictive ability (predictive QSAR) and mechanistic interpretation (mechanistic QSAR) provides an in-depth understanding of the correlation between the structural features and the desired bio-activity [15,20,21]. Therefore, QSAR analysis has been executed in the present work using a large and diverse dataset that covers a broad chemical and activity space to find out the structural features of high importance for 5-HT6 ligands. Molecular docking has been carried out to support and complement the QSAR analysis. Further, an in-depth analysis of a larger dataset comprising diverse scaffolds led to the identification of reported as well as unreported pharmacophoric features, which could be useful in the optimization of molecules during a drug discovery pipeline.

2. Results

In the present work, we have identified reported as well as unreported structural features of 5-HT6 ligands. As stated in the introduction section, the emphasis was on building a genetic algorithm–multilinear regression (GA-MLR) model with a balance of predictive ability and mechanistic interpretations. The newly built six-parametric model is as follows:

$$\text{Model-1: } pK_i = 6.754 (\pm 0.091) - 0.109 (\pm 0.014) * \text{com_Hhyd_3A} - 0.700 (\pm 0.043) * \text{ringC_S_4Bc} - 0.604 (\pm 0.104) * \text{flipo\&S_ringN3B} - 0.528 (\pm 0.075) * \text{KRFPc620} - 0.339 (\pm 0.053) * \text{sp3N_sp2O_8B} + 0.545 (\pm 0.067) * \text{fsp3Cdon1B}$$

The details of the molecular descriptors present in Model-1 have been tabulated in Table 1 and are discussed in detail in the discussion section. Table 2 contains selected validation parameters associated with Model-1.

Table 1. Some details of constituent molecular descriptors present in Model-1.

Molecular Descriptor	Description	Software Used for Calculation	Correlation with pKi
com_Hhyd_3A	Total number of hydrogen atoms with partial charge in the range of ± 0.2 within 3 Å from center of mass of molecule	PyDescriptor	−0.625
ringC_S_4Bc	Sum of partial charges on ring carbon atoms present within four bonds from sulfur atom	PyDescriptor	−0.696
flipo&S_ringN3B	Frequency of occurrence of ring nitrogen atom present exactly at three bonds from lipophilic atom	PyDescriptor	−0.248
sp3N_sp2O_8B	Total number of sp ³ -hybridized nitrogen atoms present within eight bonds from sp ² -hybridized oxygen atoms	PyDescriptor	−0.133
KRFPc620	Nitrogen attached to three CH ₃ CH ₂ - groups	PaDEL	−0.444
fsp3Cdon1B	Frequency of occurrence of H-bond donor atom bonded with sp ³ -hybridized carbon atom	PyDescriptor	0.026

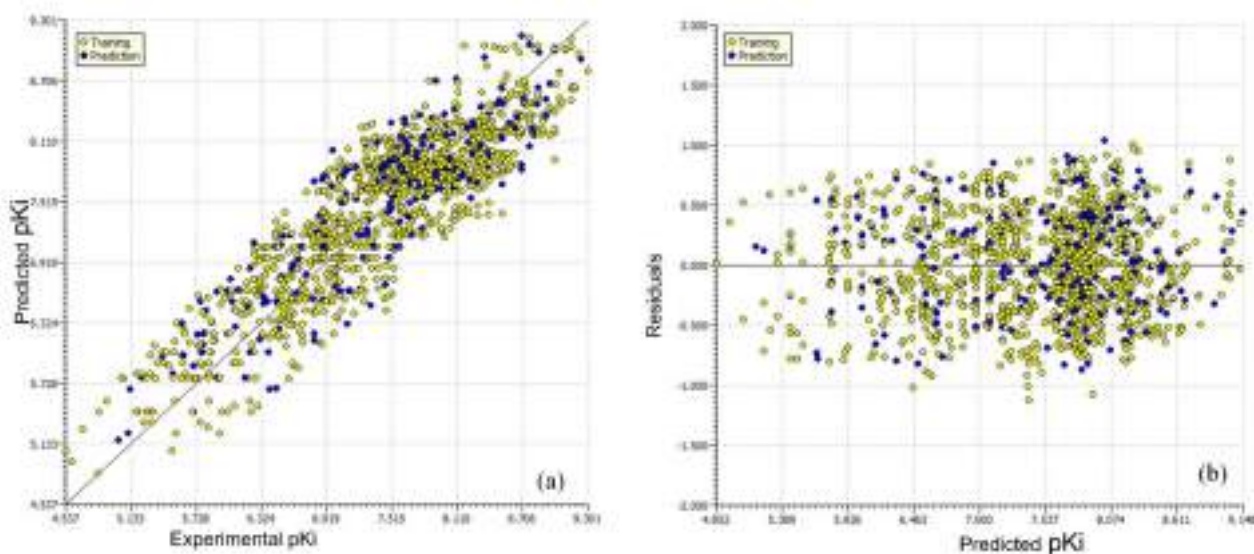
Table 2. Selected statistical validation parameters for Model-1.

Parameter	Value	Parameter	Value
R^2_{tr}	0.783	Q^2_{LMO}	0.779
$R^2_{adj.}$	0.781	R^2_{Yscr}	0.006
$RMSE_{tr}$	0.419	$RMSE_{ex}$	0.425
MAE_{tr}	0.350	MAE_{ex}	0.357
CCC_{tr}	0.878	R^2_{ex}	0.772
R^2_{cv} (Q^2_{loo})	0.780	Q^2-F^1	0.768
$RMSE_{cv}$	0.422	Q^2-F^2	0.768
MAE_{cv}	0.352	Q^2-F^3	0.777
CCC_{cv}	0.876	CCC_{ex}	0.871

Note: tr—Training, cv—Cross-validation, ex—External.

It is evident from above statistical parameters that the model is statistically predictive, with high values of different parameters such as R^2_{tr} (coefficient of determination), $R^2_{adj.}$ (adjusted coefficient of determination), and R^2_{cv} (Q^2_{LOO}) (cross-validated coefficient of determination for leave-one-out), R^2_{ex} (external coefficient of determination), Q^2-F^n , and CCC_{ex} (concordance correlation coefficient for external set), etc., and low values of LOF (lack-of-fit), RMSE (root-mean-square error), MAE (mean absolute error), R^2_{Yscr} (R^2 for Y-scrambling), etc. Thus, the model possesses high external predictive ability, is free from chance correlations, and satisfies the recommended threshold values for various validation parameters [16,22–29]. All validation parameters associated with Model-1 and their formulae are available in the Supplementary Materials.

Figure 2 contains the different graphs associated with Model-1, viz., the experimental vs. predicted pKi (Figure 2a), experimental vs. residuals (Figure 2b), and the Y-randomization plot (Figure 2d). We used a Williams plot to judge the applicability domain of the model (see Figure 2c). Thus, it satisfies all the OECD-endorsed guidelines and criteria for creating a useful QSAR model.

**Figure 2.** Cont.

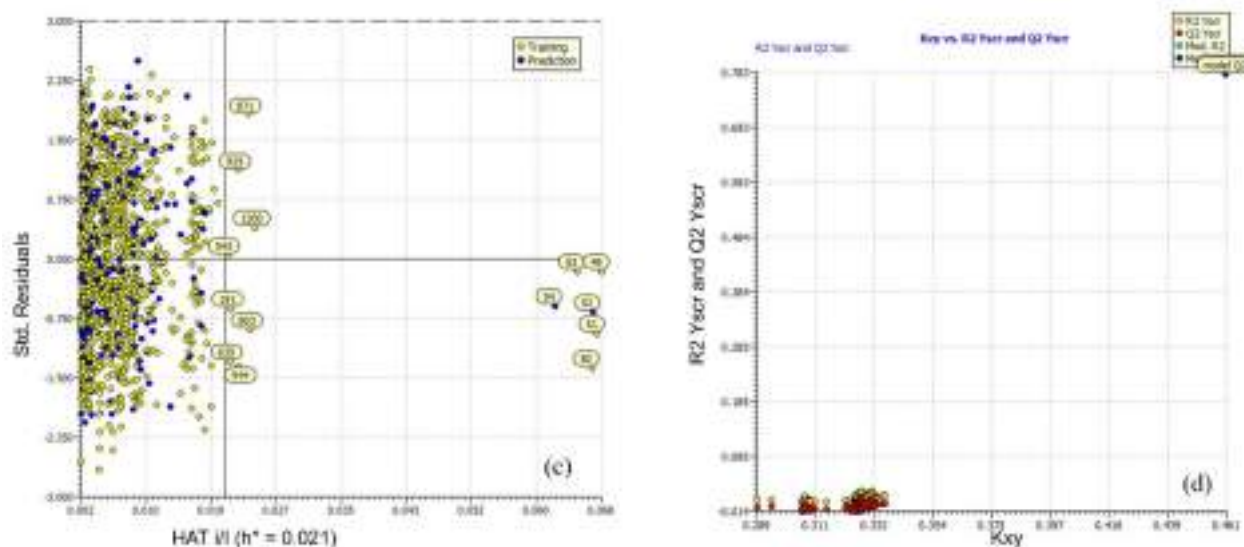


Figure 2. Different graphs associated with the Model-1. (a) Experimental vs. predicted pKi (the solid line represents the regression line); (b) experimental vs. residuals; (c) Williams plot for applicability domain (the vertical solid line represents $h^* = 0.021$ and the horizontal dashed lines represent the upper and lower boundaries of the applicability domain); (d) Y-randomization plot.

The docking scores for all the molecules are incorporated in the Supplementary Materials (see Excel file ‘SupplementaryMaterial-Final.xlsx’, available in the Supplementary Materials).

3. Discussion

The interpretation of a QSAR model using molecular descriptors, also known as mechanistic interpretation, is a crucial aspect to highlight and link various structural features with significant influence in deciding the bio-activity of molecules [15,21]. It is also one of the requirements suggested by the OECD for developing a thriving QSAR model. In the present work, we used a simple approach in which the pKi values of different molecules were compared using a specific molecular descriptor. Nonetheless, we make clear that the final experimental pKi value of a molecule cannot be governed by just a single structural feature (molecular descriptor). To add further, the extending or reverse effect of other molecular descriptors and unknown factors play crucial roles in deciding the pKi value of a molecule. Therefore, a concomitant consideration of all molecular descriptors and their associated structural features is a better strategy for the effective use of a QSAR model. Model-1 encompasses six molecular descriptors.

The molecular descriptor `com_Hhyd_3A` represents the total number of hydrogen atoms with a partial charge in the range of ± 0.200 and also present within 3 Å from the center of mass (com) of the molecule. As the partial charge must be within the range ± 0.200 , this descriptor expresses the role played by non-polar hydrogens [30] when present within 3 Å from the center of mass of the molecule. The descriptor `com_Hhyd_3A` has a negative coefficient in Model-1, which indicates that the lower the value, the better the binding affinity. This could be achieved by replacing non-polar hydrogen atoms with suitable atoms/groups. This indirectly points out that the presence of polar hydrogens, in turn, polar groups nearer to center of mass of a molecule, are beneficial for escalating the pKi value. In addition, hydrogen is smaller than other elements, and replacing it with any other element will increase the steric bulk. Therefore, bulkiness near the center of mass of the molecule is highly favorable for increasing the pKi value. To add further, `com_Hhyd_3A` depends on the location of the center of mass, which changes with the positions of groups/atoms (positional isomers). Therefore, the value of `com_Hhyd_3A` varies for positional isomers, for example, molecules number 331 and 332. Hence, the descriptor effortlessly captures the importance of positional isomerism in deciding the

pKi value. Thus, the descriptor *com_Hhyd_3A* and its negative correlation (correlation coefficient $R = -0.63$) with pKi highlighted the crucial role played by the presence of polar groups and steric bulkiness near the center of mass of a molecule as well as the positional isomerism. This observation was further confirmed by comparing following pairs of molecules: 542 ($K_i = 25520$ nM, *com_Hhyd_3A* = 8) with 547 ($K_i = 2506$ nM, *com_Hhyd_3A* = 6) (depicted in Figure 3), 543 ($K_i = 14650$ nM, *com_Hhyd_3A* = 9) with 545 ($K_i = 8611$ nM, *com_Hhyd_3A* = 6), 1011 ($K_i = 91$ nM, *com_Hhyd_3A* = 6) with 1082 ($K_i = 419$ nM, *com_Hhyd_3A* = 7), 935 ($K_i = 2843$ nM, *com_Hhyd_3A* = 8) vs. 937 ($K_i = 2005$ nM, *com_Hhyd_3A* = 6), and 929 ($K_i = 2427$ nM, *com_Hhyd_3A* = 7) vs. 938 ($K_i = 1540$ nM, *com_Hhyd_3A* = 5), to list a few. Thus, for the first time, *com* has been used as a useful approach to explain the differences in the binding affinities of ligands for 5-HT6. In addition, this novel approach provides a novel justification for the differences in the activity of positional isomers.

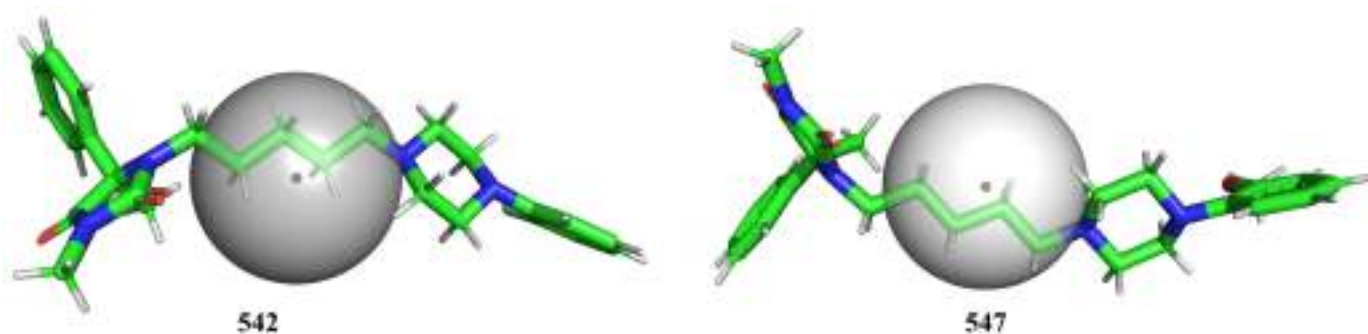


Figure 3. Depiction of *com_Hhyd_3A* using molecules 542 and 547 as representative examples (radius of gray sphere is 3 Å).

ringC_S_4Bc, which signifies the sum of partial charges on ring carbon atoms present within four bonds from a sulfur atom, has a negative coefficient in Model-1. In addition, it has a correlation coefficient of -0.696 with pKi. Therefore, decreasing the value of *ringC_S_4Bc* could lead to a higher pKi value, i.e., a better binding affinity. The descriptor highlights the importance of ring carbon and sulfur atoms; therefore, it looks as if merely ring carbon or sulfur atoms are enough to control the binding affinity. Replacing it with *nringC* (the total number of ring carbon atoms) or *nS* (the total number of sulfur atoms) reduced the statistical performance of Model-1 ($R^2 = 0.74$ and 0.73). Therefore, a combined presence of ring carbons and sulfur atoms within four bonds from each other has more influence. The descriptor is shown using molecules 335 and 339 as representative examples in Figure 4. A comparison of the following pairs of molecules confirmed the observation: 335 with 339, 952 with 953, 954 with 955, 248 with 249, 762 with 763, 221 with 260, 175 with 650, 1089 with 1101, 989 with 1030, etc.

A literature survey revealed that a combination of a H-bond acceptor, such as a sulfonamide or sulfone moiety, along with fused rings such as naphthalene, benzothiophene, or indole, is highly suitable for augmenting the pKi values of 5-HT6 ligands [6] due to their hydrophobic character [5]. Interestingly, the sulfone or sulfonamide groups are directly attached to fused rings in many compounds of the present dataset. In addition, in the present work, we found that the partial charges on the ring carbons of fused rings also influenced the binding affinity of a molecule for 5-HT6. Therefore, the hydrophobic character of fused rings and the partial charges on a ring carbon are equally important. This observation is further supported by the molecular docking analysis. The molecular docking analysis revealed that molecules number 335 and 339 have different poses and types of interactions. Molecule 335 has an established H-bond with Asp87 (distance 2.86 Å) and a pi-pi interaction with Phe-222 (distance 5.17 Å), whereas molecule 339 has only a H-bond formation with Asn225 (distance 2.70 Å) and lacks any pi-pi interactions. Molecule 335 was able to establish a pi-pi interaction due to the presence of a sulfone group, which caused substantial changes in the partial charges of ring carbon atoms and the C-S-C bond

angle (99.8° in 335 and 105.6° in 339) and also changed the solvent-accessible surface areas (2312.8 \AA^2 for 335 and 2098.4 \AA^2 for 339). The docking poses for molecules number 335 and 339 are shown in Figure 4c. The yellow dotted line represents the prominent interaction with distances in Angstrom units. It is clear from the present analysis that not only lipophilic factors but electronic factors associated with ring carbon atoms vicinal to sulfur, in turn aromatic/aliphatic rings, are an important parameter in deciding the inhibitory activity for 5-HT₆. Thus, the present work is successful in identifying novel unreported crucial aspects of a previously known pharmacophoric feature required for better activity for 5-HT₆.

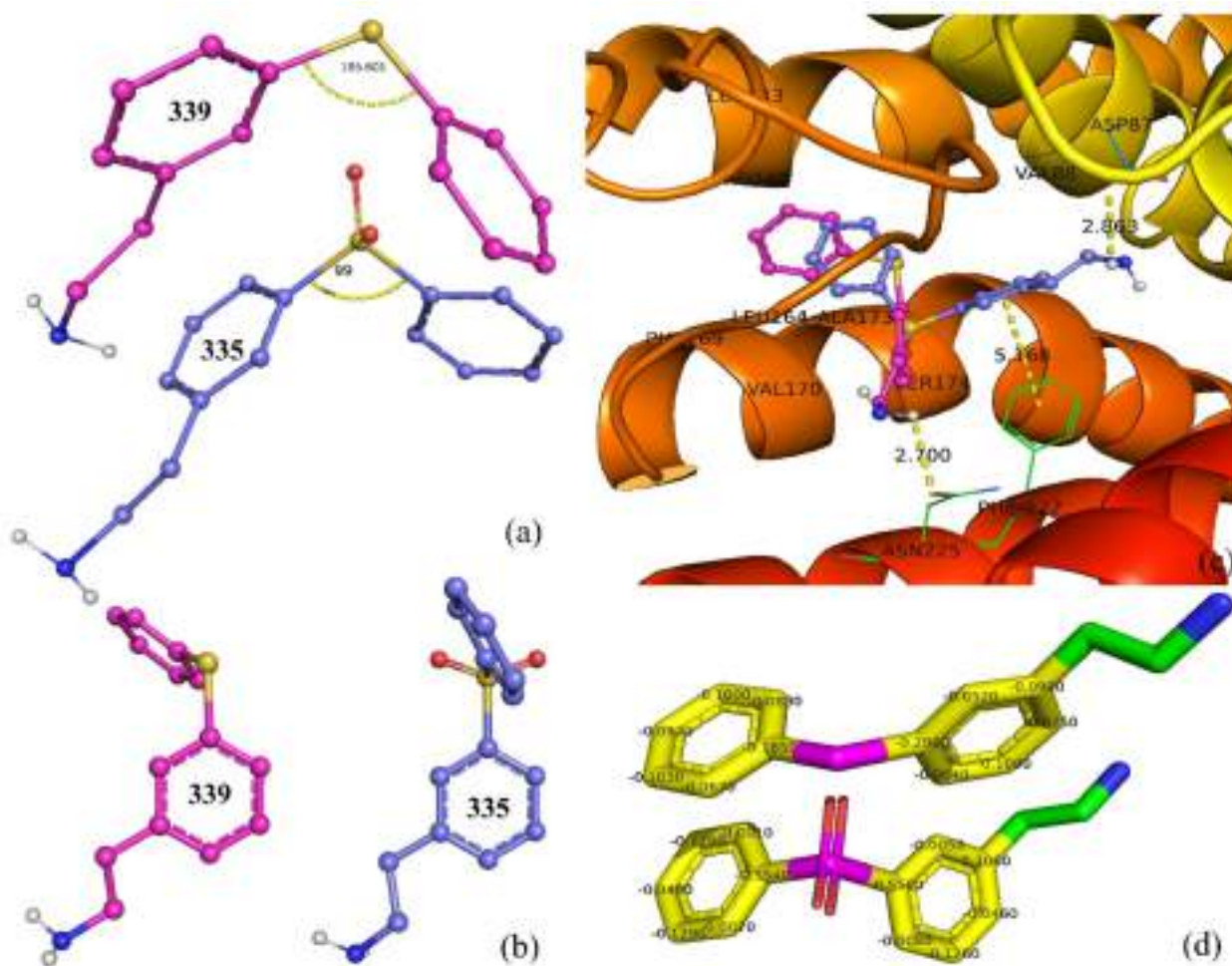


Figure 4. Molecules 335 and 339 as representative examples to depict the effect of descriptor ringC_S_4Bc. (a) Comparison of docking pose of 335 (pink) and 339 (blue), (b) C-S-C bond angle shown using yellow dotted lines, (c) docking pose for 335 (pink) and 339 (blue), (d) carbon atoms (yellow) within four bonds of sulfur (magenta) with respective partial charges.

The molecular descriptor $\text{flipo}\&\text{S_ringN3B}$ represents the frequency of occurrence of a ring nitrogen atom exactly three bonds from a lipophilic atom or sulfur atom. If the same ring nitrogen atom is also present at two or less bonds from any other lipophilic atom or sulfur atom, then it is neglected while calculating $\text{flipo}\&\text{S_ringN3B}$. It has a negative coefficient in Model-1, which suggests that lowering the value of $\text{flipo}\&\text{S_ringN3B}$ could lead to a better binding profile. This observation is justified by the fact that there are 383 molecules with $K_i \leq 10 \text{ nM}$, but only one of them has $\text{flipo}\&\text{S_ringN3B} = 2$ (molecule number 50). Only two molecules (molecules number 569 and 978) possess $\text{flipo}\&\text{S_ringN3B} = 1$, and the remaining 380 molecules have $\text{flipo}\&\text{S_ringN3B} = 0$. A comparison of molecules 870 with 871, 896 with 903, and 892 with 897 and 898 further strengthens the observation. For example, consider molecules 986 and 1024. In the case of

molecule 986, the presence of an additional $-CH_3$ group on the imidazole ring increased the number of lipophilic atoms within three bonds of the ring nitrogen as well as the lipophilic surface area (178.3 \AA^2), thereby having $\text{flipo}\&\text{S_ringN3B} = 0$, while molecule 1024 lacks such a substituent and consequently has $\text{flipo}\&\text{S_ringN3B} = 1$ and a lower lipophilic surface area of 173.5 \AA^2 (See Figure 5).

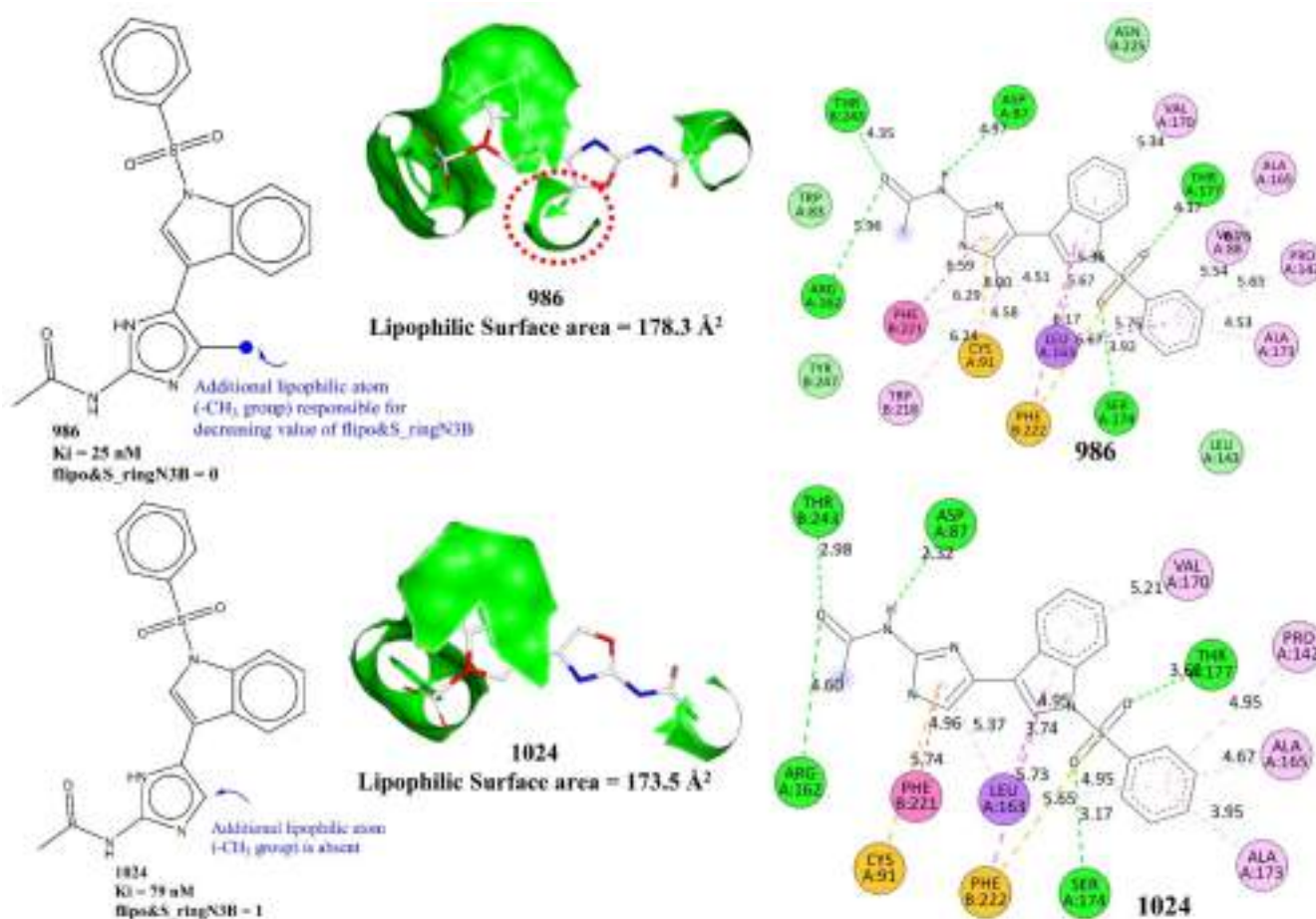


Figure 5. Representation of $\text{flipo}\&\text{S_ringN3B}$ using molecules 986 and 1024 as examples. The green contour in the middle figures depicts the lipophilic surface area.

Likewise, in the case of highly active molecules, a lower value of $\text{flipo}\&\text{S_ringN3B}$ is possible due to a lower number of ring nitrogen (for example molecule 681) or due to the presence of lipophilic atoms within three bonds of a ring nitrogen. Thus, this descriptor signifies that the presence of lipophilic atoms near ring nitrogen atoms decreases the value of $\text{flipo}\&\text{S_ringN3B}$, which in turn increases the pK_i value. In other words, it points out an unreported structural feature of conditional importance of lipophilic atoms as well as ring nitrogen atoms in deciding the binding affinity profile. A plausible reason could be attributed to the fact that increasing the lipophilic environment around a ring nitrogen balances the polarity induced by the highly electronegative nature of nitrogen, which ultimately leads to improved brain penetration [4] due to the better lipophilic characters of the molecule. The molecular docking poses for molecules number 986 and 1024 (see Figure 5) indicate that the additional $-CH_3$ group is responsible for pi-alkyl hydrophobic interactions with the active site residues of 5-HT₆ for molecule number 986. In the case of molecule number 986, the additional $-CH_3$ group interacts with Phe221, Cys91, and Trp218. The interaction with Trp218 is absent in case of molecule number 1024.

The molecular descriptor fsp3Cdon1B , which signifies the frequency of occurrence of a H-bond donor atom directly bonded with an sp^3 -hybridized carbon atom, has a positive coefficient in Model-1. Therefore, increasing such a combination of carbon and H-bond

donor atoms could lead to a higher pKi value. An inspection of molecules with $K_i \leq 10$ nM (383 molecules) reveals that only 16 molecules lack such a combination of donor and carbon, 286 molecules have one such combination, and 82 molecules possess $\text{fsp3Cdon1B} = 2$. A comparison of molecules number 250 ($K_i = 71.1$ nM, $\text{fsp3Cdon1B} = 0$) and 251 ($K_i = 7.2$ nM, $\text{fsp3Cdon1B} = 1$) further supports the positive effect of the presence of fsp3Cdon1B on the pKi value (see Figure 6). The molecular docking poses for molecules number 250 and 251 support this observation (see Figure 6b–d). Both molecules interact with similar residues such as Leu163 (H-bond). However, a prominent difference in the docking poses of molecules number 251 and 250 is the H-bond formation with a distance of 2.52 \AA by the -NH of the piperidine ring with Asp87, which is absent in the case of 250. This is due to the fact that the pyran ring of molecule 250 is oriented away from Asp87, whereas the piperidine ring of 251 is close enough to establish a H-bond with Asp87. Thus, it highlights the importance of the presence of a donor atom in the 5-HT₆ ligand. The -NH of the piperidine ring is responsible for $\text{fsp3Cdon1B} = 1$ for molecule number 250. Thus, the molecular docking and QSAR provided similar and complementary results.

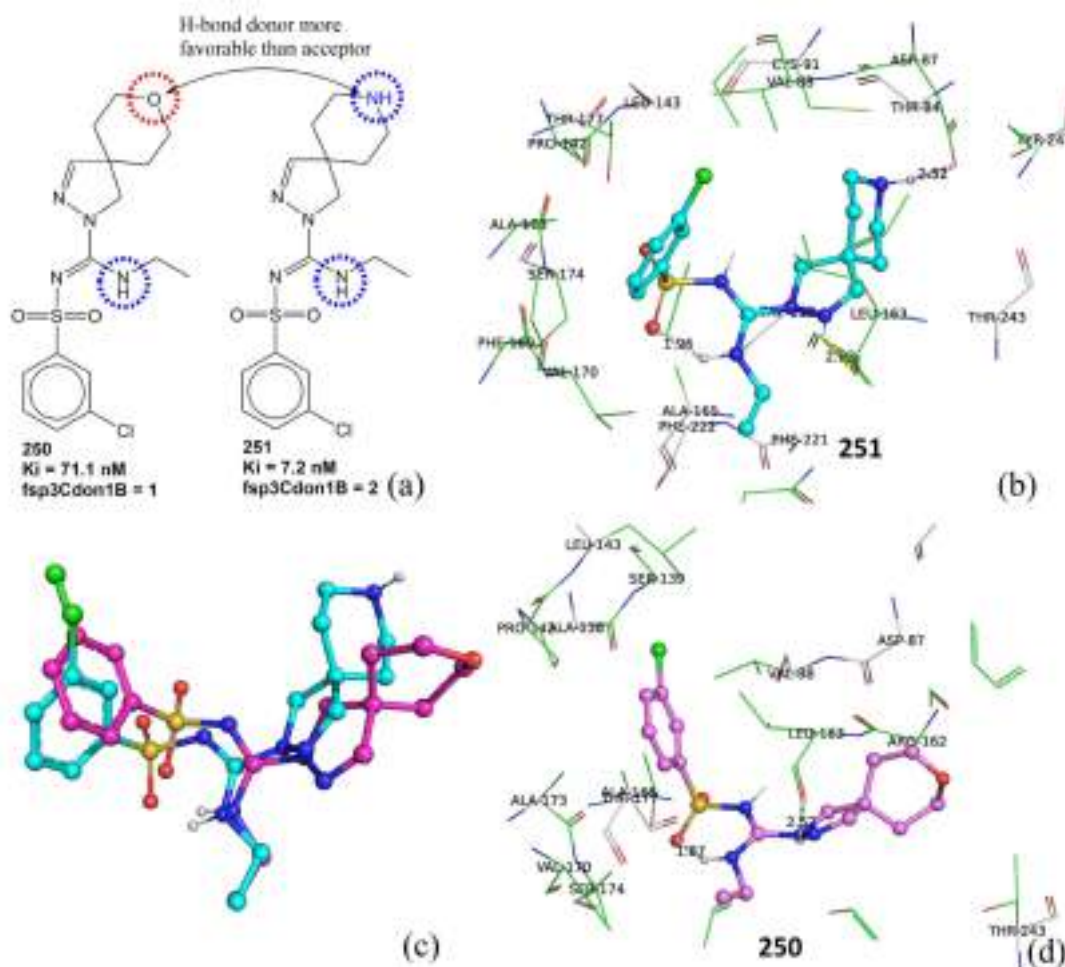


Figure 6. (a) Representation of fsp3Cdon1B using molecules number 250 and 251 as examples, (b) docking pose for molecule number 251 (dotted line represents H-bond formation), (c) overlap of docking pose of 250 (cyan) and 251 (pink), (d) docking pose for 250.

Another molecular descriptor which highlights the importance of nitrogen atoms and their local environment is KRFPC620. The molecular descriptor KRFPC620 considers the count of a fragment that encompasses a planer nitrogen atom attached to three $\text{CH}_3\text{-CH}_2\text{-}$ groups. For better clarification, it is depicted in Figure 7 using representative examples only. The negative coefficient for the descriptor indicates that the pKi value increases with

a decrease in the value of KRFP620. A comparative analysis of the following pairs of molecules supports this trend: 472 with 776, 748 with 749, 757 with 761, 1074 with 1096, 1068 with 1075, 1100 with 1116, 1094 with 1058, etc. To add further, of the 383 molecules with $K_i \leq 10$ nM, 37 molecules have KRFP620 = 1, whereas 346 molecules lack the presence of KRFP620. Therefore, such a combination of nitrogen with $\text{CH}_3\text{-CH}_2\text{-}$ must be avoided for a better pKi value.

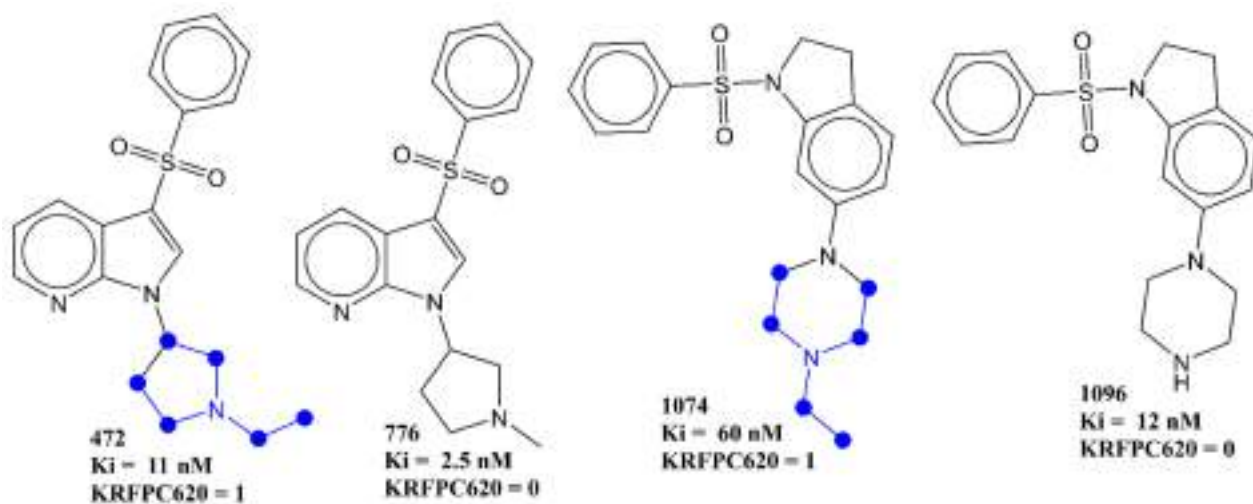


Figure 7. Pictorial representation of KRFP660 (blue colored) using representative examples.

Many researchers have highlighted the importance of the presence of a positive ionizable atom or group such as the nitrogen in piperazine or in a (dimethylamino)ethyl fragment for a better binding affinity for 5-HT₆ [4]. The present work also successfully highlighted the negative impact of a third ethyl group on ionizable nitrogen on binding affinity. Therefore, in agreement with previous studies [4,12], a methyl substituted piperazine, non-substituted piperazine, or a (dimethylamino)ethyl fragment is a better choice to act as a positive ionizable moiety.

Another descriptor whose value must be lowered due to its negative coefficient in Model-1 to have a better pKi value is sp³N_sp²O_8B. It counts the total number of sp³-hybridized nitrogen atoms present within eight bonds of sp²-hybridized oxygen atoms. Of the 383 molecules with $K_i \leq 10$ nM, 181 molecules lack such a combination of nitrogen and oxygen atoms, whereas the remaining 201 molecules possess only one such combination. In addition, the following pairs of molecules, on comparison, support the observation: 624 with 627 (see Figure 8), 557 with 556 and 555, 271 with 274, 116 with 120, and 117 with 123, to mention a few.

To summarize the prominent structural features and the pharmacophoric model, we used the two most active molecules, 681 and 271, as representative examples (see Figure 9). The pharmacophoric pattern consists of two hydrophobic regions (green contour), two H-bond acceptor regions (red contour), and one H-bond acceptor region (blue contour), as depicted in Figure 9a using molecule 681. Thus, the pharmacophoric pattern agrees with previous studies [4,12]. Figure 9b depicts the molecular descriptor ringC_S_4Bc in molecule 271. Figure 9c represents the descriptors fliporingN3B and fsp3Cdon1B. The three hydrophobic -CH_3 groups shown by green dots are responsible for balancing the hydrophilic effect of ring nitrogen atoms, thus enhancing the CNS penetration ability of a molecule. Moreover, the -CH_3 groups labelled as A and B enhance the value of fsp3Cdon1B, as they are directly attached to donor atoms. Thus, the present work not only captured the pharmacophoric features reported by López-Rodríguez et al. [4,12] but also successfully extended it.

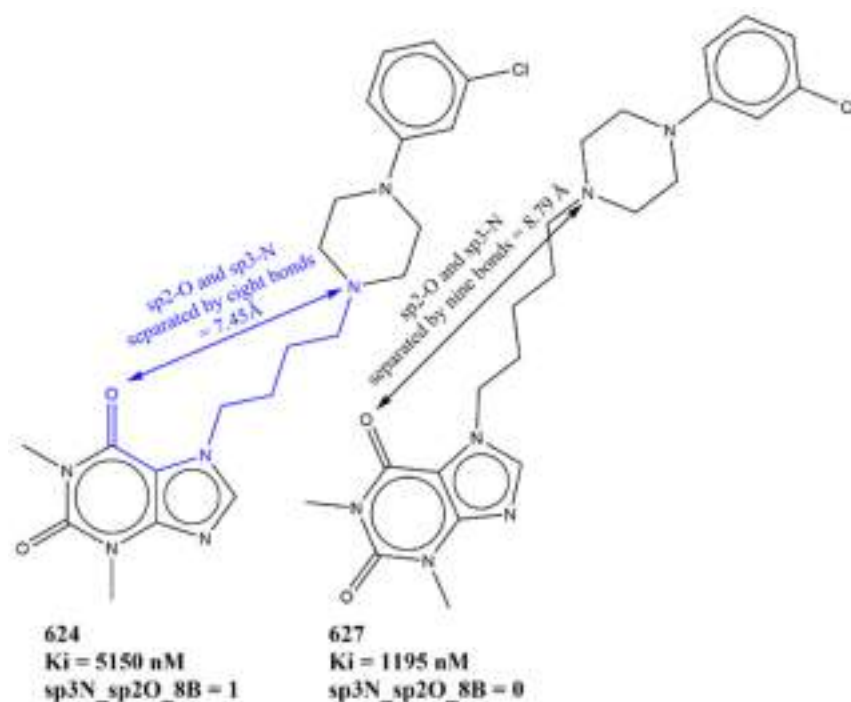


Figure 8. Illustrative examples for representation of sp3N_sp2O_8B.

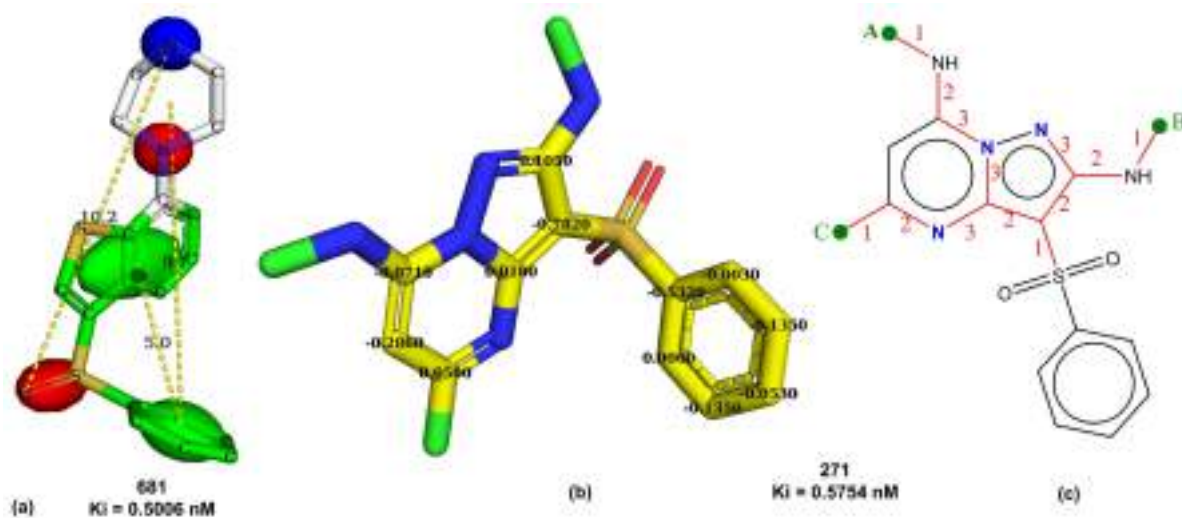


Figure 9. Depiction of (a) pharmacophoric feature using molecule 681 (green: hydrophobic, red: H-bond acceptor, blue: H-bond donor regions, dotted lines show distances in Angstrom units), (b) yellow atoms represent ringC_S_4Bc in molecule 271, (c) fliporingN3B represented by red-colored bonds with numbering using molecule 271.

The molecular docking score spans from -5.7 to -12.2 for molecules in the current dataset (see Supplementary Materials). For the sake of convenience, twenty molecules with the highest and lowest molecular docking scores are tabulated in Table 3. Surprisingly, not only the molecular docking score has a weak correlation of 0.10 with pKi but also many molecules with lower binding affinity values have better docking scores (see Tables 3 and 4). This could be attributed to the large size of the active site of 5-HT₆ (see Figure 10), which allows the adoption of different conformations for molecules. Moreover, recent studies point out that current docking software such as AutoDock, Dock, etc., and respective algorithms for docking scores are inclined toward the flexibility of ligands, which in turn, is associated with the loss of ligand conformational entropy on binding [31]. The molecules

with lower binding affinity for 5-HT₆ have a high degree of flexibility. Thus, all these factors together resulted in an artificially more favorable binding score for the more flexible decoys than for actives.

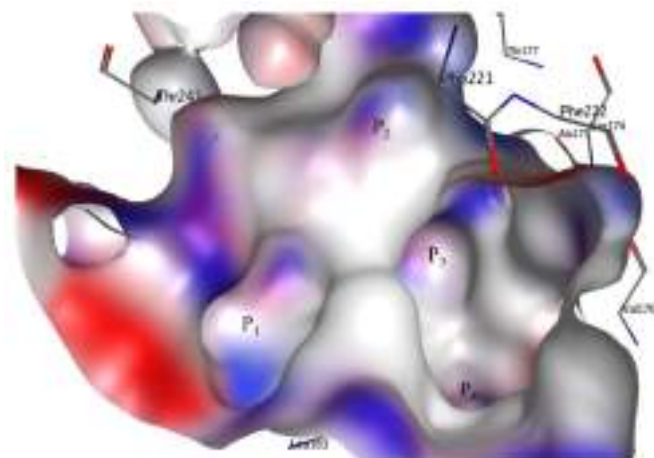


Figure 10. Binding site of 5-HT₆ used in the present work.

Table 3. The molecular docking score and K_i (nM) for ten molecules with the highest and lowest molecular docking scores.

Molecule Number	SMILES	K _i (nM)	Affinity-Docking Score (Kcal/mol)
741	<chem>CC(=O)Nc(cc1)ccc1CCNICC2Cc(c23)ccc4c3ccn4S(=O)(=O)c5ccccc5</chem>	79.43	−12.2
134	<chem>C1NCCCC1C(=O)Nc(ccc2)c(c23)[nH]nc3S(=O)(=O)c4cccc(c45)cccc5</chem>	9.8	−12
489	<chem>c1cccc(c12)[nH]cc2C(C3=O)CC(=O)N3CCCN(CC4)CCC4c5c[nH]c(c56)ccc(c6)OC</chem>	264	−11.9
490	<chem>c1cccc(c12)[nH]cc2C(C3=O)CC(=O)N3CCCN(CC4)CCC4c5c[nH]c(c56)ccc(F)c6</chem>	1146	−11.8
1093	<chem>CC(=O)Nc(n1)[nH]c(c1C)-c2cn(c(c23)cccc3)S(=O)(=O)c4cccc(c45)cccc5</chem>	13	−11.7
668	<chem>Cl(C)CCC1c2c[nH]c(c23)ccc(c3)NS(=O)(=O)c(ccc4)c(c45)nccc5</chem>	21.2	−11.7
133	<chem>C1CCCN1CCC(=O)Nc(ccc2)c(c23)[nH]nc3S(=O)(=O)c4cccc(c45)cccc5</chem>	24	−11.7
381	<chem>FC(F)(F)c1cc(ccc1)S(=O)(=O)n(c(c2c34)CCC(C2)N)c3ccc(c4)OC</chem>	39.1	−11.7
805	<chem>c1cccc(c12)ccc(c2)S(=O)(=O)NCCN(CC3)CC=C3c4c[nH]c(c45)ccc(F)c5</chem>	67	−11.7
628	<chem>c1cccc(c12)CN([C@@H](C2)C(=O)N)C(=O)CCCCN3CCN(CC3)c(cccc4)c4-c5ccccc5</chem>	594	−11.7
1086	<chem>CCn1cncc1-c2c[nH]c(c23)ccc(Br)c3</chem>	1349	−7.3
1016	<chem>CCn1cncc1-c2c[nH]c(c23)cccc3</chem>	3020	−7.3
1203	<chem>N1CCC[C@@H]1Cc2c[nH]c(c23)cccc3</chem>	60	−7.2
1202	<chem>c1cccc(c12)n(cc2)C[C@H]3CCCN3C</chem>	550	−7.2
339	<chem>NCCc1cc(ccc1)Sc2ccccc2</chem>	115	−7.1
93	<chem>CCN(CC)CCc1c[nH]c(c12)cccc2</chem>	575	−6.4
738	<chem>NCCc1c[nH]c(c12)ccc(c2)O</chem>	42.333	−6.2
534	<chem>CC(N)Cc1c[nH]c(c12)cccc2</chem>	910.505	−6.2
444	<chem>NCCc1c[nH]c(c12)ccnc2</chem>	64	−5.7
445	<chem>CN(C)CCc1c[nH]c(c12)ccnc2</chem>	100	−5.7

Table 4. SMILES (simplified molecular input line entry system) notation, Ki (nM), pKi (M), and molecular docking score for the five most and five least active molecules of the selected dataset.

Molecule Number	SMILES (Simplified Molecular Input Line Entry System) Notation	Ki(nM)	pKi(M)	Docking Score (Kcal/mol)
681	<chem>C1CNCCN1c(ccc2)c(c23)sc3S(=O)(=O)c4cccc4</chem>	0.5006	9.301	−9
271	<chem>Cc(n1)cc(NC)n(c12)nc(NC)c2S(=O)(=O)c3cccc3</chem>	0.5754	9.24	−8.3
142	<chem>C1CNCCC1Nc(c2)ccc(c23)[nH]nc3S(=O)(=O)c(c4)ccc(c45)cccc5</chem>	0.6	9.222	−10.8
279	<chem>Cc(n1)IN)c(C)n(c12)nc(NC)c2S(=O)(=O)c3cccc3</chem>	0.6457	9.19	−8.4
428	<chem>CN(C1)CIn2)c1c(C)n(c23)nc(NC)c3S(=O)(=O)c4cccc4</chem>	0.66	9.18	−9.1
622	<chem>O=I(C)c(=O)n(C)c(c12)ncn2CCCCN3CCN(CC3)c4cccc4</chem>	14210	4.847	−9.6
543	<chem>c1cccc(OC)c1N(CC2)CCICCCCI(=O)N(C)C(=O)C3(C)c4cccc4</chem>	14650	4.834	−10.1
267	<chem>c1cccc1N(CICCN2CC(O)CN3C(=O)N(C)C(=O)C3(c4cccc4)c5cccc5</chem>	20410	4.69	−9.7
542	<chem>c1cccc(F)c1CC2)I2CCCCCN3C(=O)N(C)C(=O)C3(C)c4cccc4</chem>	25520	4.593	−10.1
266	<chem>c1cI(OC)c1N(CC2)CCN2CCCN3C(=O)N(C)C(=O)C3(c4cccc4)c5cccc5</chem>	29070	4.537	−10.8

4. Materials and Methods

In the present study, we have followed the OECD (Organization for Economic Co-operation and Development) guidelines and standard protocol, which has been endorsed by different researchers, for an effective QSAR analysis [15,17,18,21,29,32–35]. The different steps for developing a model involved a careful selection of a dataset and data curation, followed by 3D structure generation for all molecules, molecular descriptor calculations and their pruning, model development and its thorough validation (internal and external), and mechanistic interpretation [20,36,37]. These steps were executed sequentially to avoid errors and human bias and to ensure the appropriate validation of the model.

4.1. Selection of Dataset

The size, composition, and structural diversity of the dataset are important characteristics that decide the success and utility of QSAR and molecular docking analysis throughout the pipeline of drug discovery [15–17,21,29,32,34,38–41]. Therefore, a large dataset of 3398 reported ligands for 5-HT₆ was downloaded from BindingDB (<https://www.bindingdb.org/bind/index.jsp>, accessed on 14 January 2022). Then, duplicates, salts, metal derivatives, rule-of-five violators, and molecules with undefined Ki values were removed as a part of data curation [42], which reduced the size of the dataset to only 1278 molecules. The reduced dataset was still diverse enough, with the presence of positional and chain isomers, different heterocyclic and aromatic scaffolds, stereoisomers, etc. The experimental Ki spanned five orders of magnitude (0.5006 nM to 29.07 μM). For a better QSAR analysis, the experimental Ki values were converted to pKi (pKi = −lgKi). To understand the structural variation possessed by the dataset, some highly active and least active molecules are depicted in Figure 11, and their other details are presented in Table 4.

4.2. Calculation of Molecular Descriptors and Objective Feature Selection (OFS)

The next step involved the conversion of all SMILES notations to respective 3D-optimized structures using the appropriate method. For this, OpenBabel 3.1 [43] was used to convert SMILES to SDF (structure data file). After that, MOPAC2016 [44] was used to convert SDF to MOL2 using PM3, a semi-empirical method based on the same formalism and equations as the Austin Model-1 (AM1) method, for structure optimization and partial charge assignment. Then, molecular descriptor calculations were accomplished using PyDescriptor [45] and PaDEL [46], which together provided more than 42,000 molecular descriptors for each molecule. Though the myriad numbers of molecular descriptors increase the possibility of achieving a successful QSAR analysis, they also increase the risk of chance correlations or overfitting from noisy redundant descriptors. Therefore,

OFS was performed using QSARINS 2.2.4 [47], which eliminated near constant (for 90% molecules) and highly intercorrelated ($|R| > 0.90$) molecular descriptors. After OFS, the reduced set of molecular descriptors encompassed only 4186 descriptors, which still covered a broad descriptor space due to the presence of 1D to 3D descriptors, fingerprint, and charged-based descriptors as well as atom-pair descriptors. The remaining descriptors were easily interpretable in terms of structural features, thereby increasing the possibility of mechanistic interpretation of the model.

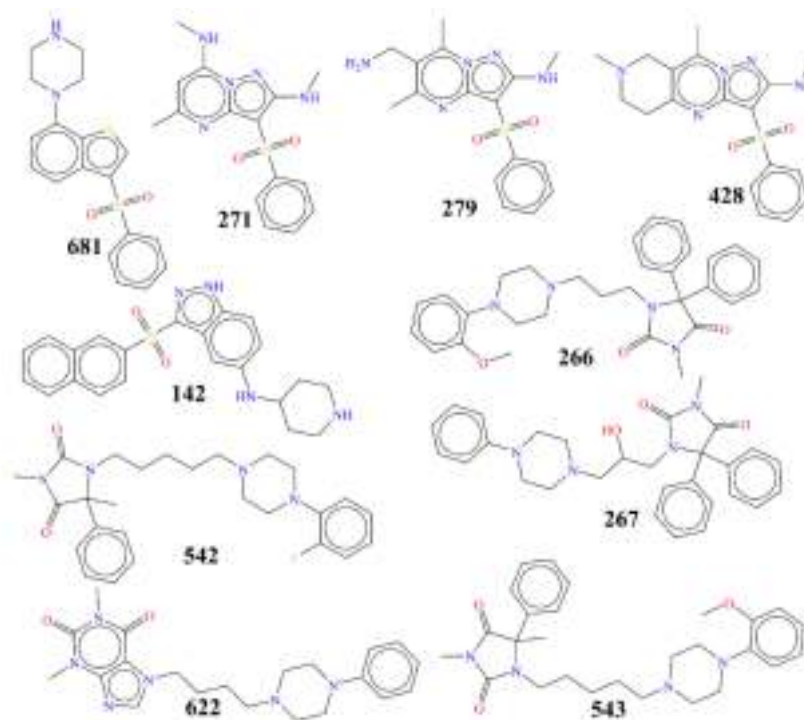


Figure 11. Representative examples from the selected dataset (five most active and five least active molecules) of 5-HT₆ ligands.

4.3. Splitting the Dataset and Subjective Feature Selection (SFS)

SFS is a crucial step during model building to select a suitable number and set of molecular descriptors using a suitable feature selection algorithm such as stepwise regression, genetic algorithm, etc. For the appropriate training and validation of the model, before actual model building, the dataset was randomly divided into a training set (80%, i.e., 1024 molecules) and a prediction set (20%, i.e., 254 molecules). The random division of a dataset into an 80:20 proportion was performed to avoid any bias and minimize information leakage [32,48], to confirm the external predictive ability of the model, and to achieve a better composition of the training and prediction sets [29,34]. The training set was used only for the selection of the appropriate number and set of molecular descriptors. The prediction set, also termed the external validation set, was used only for the validation of the newly developed model.

It is essential to include the optimum number of molecular descriptors in the model to avoid over- and underfitting. Consequently, a simple graphical method was used to identify the optimum number of descriptors for a model. Usually, the successive addition of a variable (molecular descriptor) in a multilinear regression (MLR) model augments the value of Q^2_{LOO} until the final elevation point is achieved [25,29]. After that, there is little or poor augmentation to the value of Q^2_{LOO} . Hence, the number of molecular descriptors matching the elevation point were considered optimum for model building [36]. This is depicted in Figure 12 as a graph. From Figure 12, the final elevation point is matched to six molecular descriptors. Therefore, the heuristic search was confined to including only six molecular descriptors in the QSAR model using a multi-regression analysis. The set

of molecular descriptors were selected using a genetic algorithm as a feature selection algorithm and using Q^2_{LOO} as the fitness parameter along with the genetic algorithm–multilinear regression (GA-MLR) method using QSARINS-2.2.4 [47].

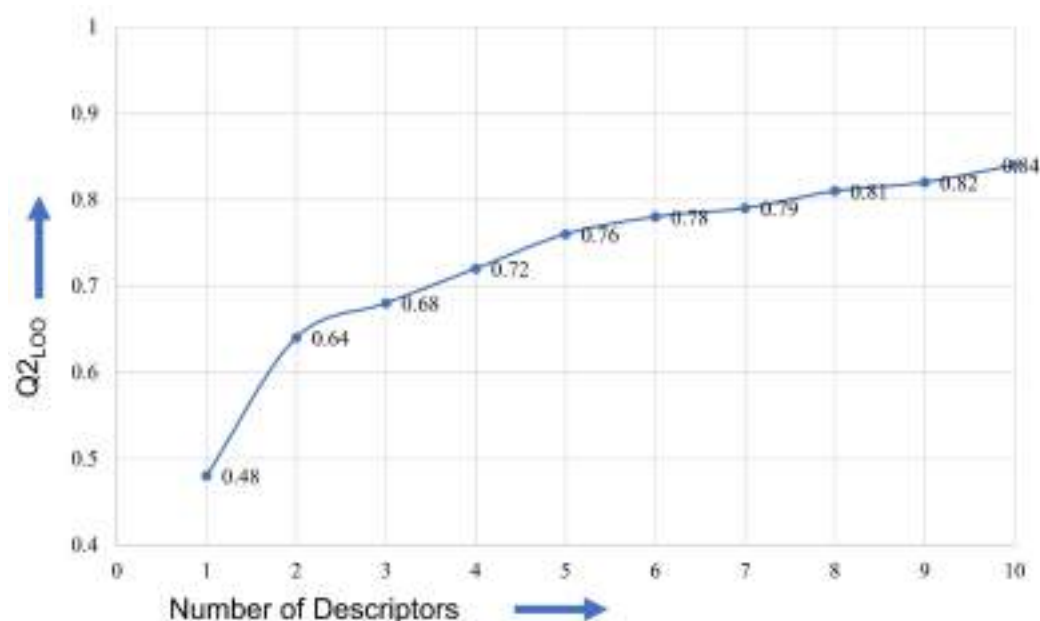


Figure 12. Plot of number of descriptors against leave-one-out coefficient of determination, Q^2_{LOO} , to identify the optimum number of descriptors.

4.4. Building Regression Model and Its Validation

The search for six molecular descriptors for the development model ultimately provided different combinations of different molecular descriptors. However, only one combination of molecular descriptors was selected due to the statistical performance and the satisfaction of the following stringent parameters and criteria, which have been suggested by other researchers [16,22–29]:

$R^2_{tr} \geq 0.6$; $Q^2_{LOO} \geq 0.5$; Q^2_{LMO} (cross-validated coefficient of determination for leave-many-out) ≥ 0.6 ; $R^2_{tr} > Q^2_{LOO}$; $R^2_{ex} \geq 0.6$, CCC (concordance correlation coefficient) ≥ 0.80 ; $Q^2_{F^n} \geq 0.60$; high values of the external validation parameters R^2_{ex} , Q^2_{F1} , Q^2_{F2} , and Q^2_{F3} ; and low values of R^2_{Yscr} (coefficient of determination for Y-randomization), RMSE (root-mean-square error), and MAE (mean absolute error), with $RMSE_{tr} < RMSE_{cv}$.

The formulae and other details of these statistical parameters are available in the Supplementary Materials. For further validation, it was essential to determine the applicability domain of the QSAR model [49,50]. For this, we used a Williams plot (standardized residuals vs. hat values) to evaluate the applicability domain of the QSAR model [47].

4.5. Molecular Docking

Since the X-ray-resolved structure of 5-HT6 is not available, an appropriately validated and reported homology model built using a β_2 receptor template (PDB ID: 4LDE) by Vanda et al. [8] was used for molecular docking [6,7] in the present work. 4LDE (selected template) and 5-HT6 (modelled receptor) possess equivalent positions for the most conserved amino acid in each helix and motifs characteristic for class A GPCRs [6–8]. The sequences of 5-HT6 and 4LDE were retrieved from the UniProtKB/Swiss-Prot database [8]. The protein was prepared using Chimera (www.cgl.ucsf.edu/chimera/, accessed on 14 January 2022) with default settings. Then, for the execution of molecular docking, AutoDock Vina [51], available in DockingApp [52], which allows users to select flexible residues of the input receptor, was used with flexible docking for virtual screening with default settings. The binding site of 5-HT6, which consists of four major grooves, P1–P4, used in the present work, is depicted in Figure 10. The binding site was used for automated

molecular docking for all the molecules. Further details of molecular docking are available in the Supplementary Materials.

5. Conclusions

The present work has resulted in the development of a statistically predictive GA-MLR model with a balance of excellent predictive ability ($R^2_{tr} = 0.78$, $R^2_{ex} = 0.77$) and novel mechanistic interpretations. The present QSAR and molecular docking analysis has successfully identified the importance of steric bulkiness, polar groups, and substitution with respect to the center of mass. It also highlighted the role played by partial charges on ring C atoms present near sulfur atoms. Further, Model-1 successfully identified that a conditional occurrence of lipophilic atoms/groups with respect to nitrogen atoms is better for having a higher pKi value. The crucial role played by the interrelationships between specific types of atoms, such as sp³-hybridized carbon/nitrogen atoms, ring nitrogen atoms, etc., further escalated the expediency of Model-1. The model possesses high external predictive ability, which is evident from its statistical performance. In conclusion, the present work not only effectively identified reported and novel unreported pharmacophoric features associated with 5-HT₆ but also offers a highly predictive QSAR model. It could be beneficial throughout the drug discovery pipeline to optimize the lead compounds to develop a better therapeutic for 5-HT₆.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/ph15070834/s1>: Additional statistical parameters for model-1, Additional details of experimental procedure for molecular docking, Statistical symbols with names and explanations, Molecular descriptors for all molecules used in model-1, their correlation matrix and molecular docking score.

Author Contributions: Conceptualization, V.H.M., S.N.A.B. and M.A.E.; formal analysis and data curation, V.H.M., K.J. and P.A.; writing, V.H.M., H.E., S.N.A.B., R.D.J. and A.S.; editing and proofreading, V.H.M., S.N.A.B., H.E. and M.A.E. All authors have read and agreed to the published version of the manuscript.

Funding: The authors' work was supported through grant number "375213500" from the Deputyship for Research and Innovation, Ministry of Education, in Saudi Arabia.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data is contained within the article and Supplementary Materials.

Acknowledgments: The authors extend their appreciation to the Deputyship for Research and Innovation, Ministry of Education, in Saudi Arabia, and the central laboratory at Jouf University for supporting this study. V.H.M. is thankful to Paola Gramatica and her team for providing a free copy of QSARINS 2.2.4.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

SMILES	Simplified molecular-input line-entry system
GA	Genetic algorithm
MLR	Multiple linear regression
QSAR	Quantitative structure–activity relationship
WHO	World Health Organization
ADMET	Absorption, distribution, metabolism, excretion, and toxicity
OLS	Ordinary least square
QSARINS	QSAR Insubria
OECD	Organisation for Economic Co-operation and Development

References

1. Doddareddy, M.R.; Cho, Y.S.; Koh, H.Y.; Pae, A.N. CoMFA and CoMSIA 3D QSAR analysis on N1-arylsulfonylindole compounds as 5-HT6 antagonists. *Bioorganic Med. Chem.* **2004**, *12*, 3977–3985. [[CrossRef](#)] [[PubMed](#)]
2. Codony, X.; Burgueño, J.; Ramírez, M.J.; Vela, J.M. 5-HT6 Receptor Signal Transduction. In *Pharmacology of 5-HT6 Receptors—Part 1*; Academic Press: Cambridge, MA, USA, 2010; pp. 89–110.
3. Hao, M.; Li, Y.; Li, H.; Zhang, S. Investigation of the Structure Requirement for 5-HT6 Binding Affinity of Arylsulfonyl Derivatives: A Computational Study. *Int. J. Mol. Sci.* **2011**, *12*, 5011–5030. [[CrossRef](#)] [[PubMed](#)]
4. Benhamú, B.; Martín-Fontecha, M.; Vázquez-Villa, H.; Pardo, L.; López-Rodríguez, M.L. Serotonin 5-HT6 Receptor Antagonists for the Treatment of Cognitive Deficiency in Alzheimer's Disease. *J. Med. Chem.* **2014**, *57*, 7160–7181. [[CrossRef](#)]
5. Karila, D.; Freret, T.; Bouet, V.; Boulouard, M.; Dallemagne, P.; Rochais, C. Therapeutic Potential of 5-HT6 Receptor Agonists. *J. Med. Chem.* **2015**, *58*, 7901–7912. [[CrossRef](#)] [[PubMed](#)]
6. Łażewska, D.; Kurczab, R.; Wićcek, M.; Kamińska, K.; Satała, G.; Jastrzębska-Więsek, M.; Partyka, A.; Bojarski, A.J.; Wesolowska, A.; Kieć-Kononowicz, K.; et al. The computer-aided discovery of novel family of the 5-HT 6 serotonin receptor ligands among derivatives of 4-benzyl-1,3,5-triazine. *Eur. J. Med. Chem.* **2017**, *135*, 117–124. [[CrossRef](#)] [[PubMed](#)]
7. Kucwaj-Brysz, K.; Baltrukevich, H.; Czarnota, K.; Handzlik, J. Chemical update on the potential for serotonin 5-HT6 and 5-HT7 receptor agents in the treatment of Alzheimer's disease. *Bioorganic Med. Chem. Lett.* **2021**, *49*, 128275. [[CrossRef](#)] [[PubMed](#)]
8. Vanda, D.; Canale, V.; Chaumont-Dubel, S.; Kurczab, R.; Satała, G.; Koczurkiewicz-Adamczyk, P.; Krawczyk, M.; Pietruś, W.; Blicharz, K.; Pękała, E.; et al. Imidazopyridine-Based 5-HT6 Receptor Neutral Antagonists: Impact of N1-Benzyl and N1-Phenylsulfonyl Fragments on Different Receptor Conformational States. *J. Med. Chem.* **2021**, *64*, 1180–1196. [[CrossRef](#)] [[PubMed](#)]
9. Zajdel, P.; Grychowska, K.; Mogilski, S.; Kurczab, R.; Satała, G.; Bugno, R.; Kos, T.; Gołębiowska, J.; Malikowska-Racia, N.; Nikiforuk, A.; et al. Structure-Based Design and Optimization of FPPQ, a Dual-Acting 5-HT3 and 5-HT6 Receptor Antagonist with Antipsychotic and Pro-cognitive Properties. *J. Med. Chem.* **2021**, *64*, 13279–13298. [[CrossRef](#)]
10. Staroń, J.; Kurczab, R.; Warszycki, D.; Satała, G.; Krawczyk, M.; Bugno, R.; Lenda, T.; Popik, P.; Hogendorf, A.S.; Hogendorf, A.; et al. Virtual screening-driven discovery of dual 5-HT6/5-HT2A receptor ligands with pro-cognitive properties. *Eur. J. Med. Chem.* **2020**, *185*, 111857. [[CrossRef](#)]
11. Heal, D.J.; Smith, S.L.; Fisas, A.; Codony, X.; Buschmann, H. Selective 5-HT6 receptor ligands: Progress in the development of a novel pharmacological approach to the treatment of obesity and related metabolic disorders. *Pharmacol. Ther.* **2008**, *117*, 207–231. [[CrossRef](#)]
12. Da Silva, A.P.; de Angelo, R.M.; de Paula, H.; Honório, K.M.; da Silva, A.B.F. Drug design of new 5-HT6 antagonists: A QSAR study of arylsulfonamide derivatives. *Struct. Chem.* **2020**, *31*, 1585–1597. [[CrossRef](#)]
13. Imam, S.S.; Gilani, S.J. Computer Aided Drug Design: A Novel Loom to Drug Discovery. *Org. Med. Chem.* **2017**, *1*, 113–118. [[CrossRef](#)]
14. Baig, M.H.; Ahmad, K.; Roy, S.; Ashraf, J.M.; Adil, M.; Siddiqui, M.H.; Khan, S.; Kamal, M.A.; Provaznik, I.; Choi, I. Computer Aided Drug Design: Success and Limitations. *Curr. Pharm. Des.* **2016**, *22*, 572–581. [[CrossRef](#)]
15. Gramatica, P. Principles of QSAR Modeling. *Int. J. Quant. Struct. Prop. Relatsh.* **2020**, *5*, 61–97. [[CrossRef](#)]
16. Gramatica, P. On the development and validation of QSAR models. *Methods Mol. Biol.* **2013**, *930*, 499–526. [[CrossRef](#)]
17. Gramatica, P.; Cassani, S.; Roy, P.P.; Kovarich, S.; Yap, C.W.; Papa, E. QSAR Modeling is not Push a Button and Find a Correlation: A Case Study of Toxicity of (Benzo-)triazoles on Algae. *Mol. Inform.* **2012**, *31*, 817–835. [[CrossRef](#)]
18. Huang, J.; Fan, X. Why QSAR fails: An empirical evaluation using conventional computational approach. *Mol. Pharm.* **2011**, *8*, 600–608. [[CrossRef](#)]
19. López-Rodríguez, M.L.; Benhamú, B.; de la Fuente, T.; Sanz, A.; Pardo, L.; Campillo, M. A Three-Dimensional Pharmacophore Model for 5-Hydroxytryptamine6 (5-HT6) Receptor Antagonists. *J. Med. Chem.* **2005**, *48*, 4216–4219. [[CrossRef](#)]
20. Masand, V.H.; Patil, M.K.; El-Sayed, N.N.E.; Zaki, M.E.A.; Almarhoon, Z.; Al-Hussain, S.A. Balanced QSAR analysis to identify the structural requirements of ABBV-075 (Mivebresib) analogues as bromodomain and extraterminal domain (BET) family bromodomain inhibitor. *J. Mol. Struct.* **2021**, *1229*, 129597. [[CrossRef](#)]
21. Fujita, T.; Winkler, D.A. Understanding the Roles of the “Two QSARs”. *J. Chem. Inf. Model.* **2016**, *56*, 269–274. [[CrossRef](#)]
22. Krstajic, D.; Buturovic, L.J.; Leahy, D.E.; Thomas, S. Cross-validation pitfalls when selecting and assessing regression and classification models. *J. Cheminform.* **2014**, *6*, 10. [[CrossRef](#)]
23. Chirico, N.; Gramatica, P. Real external predictivity of QSAR models. Part 2. New intercomparable thresholds for different validation criteria and the need for scatter plot inspection. *J. Chem. Inf. Model.* **2012**, *52*, 2044–2058. [[CrossRef](#)]
24. Chirico, N.; Gramatica, P. Real external predictivity of QSAR models: How to evaluate it? Comparison of different validation criteria and proposal of using the concordance correlation coefficient. *J. Chem. Inf. Model.* **2011**, *51*, 2320–2335. [[CrossRef](#)]
25. Consonni, V.; Ballabio, D.; Todeschini, R. Comments on the definition of the Q2 parameter for QSAR validation. *J. Chem. Inf. Model.* **2009**, *49*, 1669–1678. [[CrossRef](#)]
26. Rao, R.B.; Fung, G.; Rosales, R. On the Dangers of Cross-Validation. An Experimental Evaluation. In Proceedings of the 2008 SIAM International Conference on Data Mining (SDM), Atlanta, GA, USA, 24–26 April 2008; pp. 588–596; ISBN 978-0-89871-654-2. [[CrossRef](#)]
27. Gramatica, P.; Giani, E.; Papa, E. Statistical external validation and consensus modeling: A QSPR case study for Koc prediction. *J. Mol. Graph. Model.* **2007**, *25*, 755–766. [[CrossRef](#)]

28. Tropsha, A.; Gramatica, P.; Gombar, V.K. The Importance of Being Earnest Validation is the Absolute Essential for Successful Application and Interpretation of QSPR Models. *QSAR Comb. Sci.* **2003**, *22*, 69–77. [[CrossRef](#)]
29. Cherkasov, A.; Muratov, E.N.; Fourches, D.; Varnek, A.; Baskin, I.I.; Cronin, M.; Dearden, J.; Gramatica, P.; Martin, Y.C.; Todeschini, R.; et al. QSAR modeling: Where have you been? Where are you going to? *J. Med. Chem.* **2014**, *57*, 4977–5010. [[CrossRef](#)]
30. Todeschini, R.; Consonni, V. *Molecular Descriptors for Chemoinformatics*; Wiley-VCH: Weinheim, Germany, 2009; Volumes I and II.
31. Huang, S.-Y. Comprehensive assessment of flexible-ligand docking algorithms: Current effectiveness and challenges. *Brief. Bioinform.* **2018**, *19*, 982–994. [[CrossRef](#)]
32. Masand, V.H.; Mahajan, D.T.; Nazeruddin, G.M.; Hadda, T.B.; Rastija, V.; Alfeefy, A.M. Effect of information leakage and method of splitting (rational and random) on external predictive ability and behavior of different statistical parameters of QSAR model. *Med. Chem. Res.* **2015**, *24*, 1241–1264. [[CrossRef](#)]
33. Polishchuk, P. Interpretation of Quantitative Structure–Activity Relationship Models: Past, Present, and Future. *J. Chem. Inf. Model.* **2017**, *57*, 2618–2639. [[CrossRef](#)]
34. Golbraikh, A.; Muratov, E.; Fourches, D.; Tropsha, A. Data set modelability by QSAR. *J. Chem. Inf. Model.* **2014**, *54*, 1–4. [[CrossRef](#)] [[PubMed](#)]
35. Li, J.; Gramatica, P. The importance of molecular structures, endpoints' values, and predictivity parameters in QSAR research: QSAR analysis of a series of estrogen receptor binders. *Mol. Divers* **2010**, *14*, 687–696. [[CrossRef](#)] [[PubMed](#)]
36. Zaki, M.E.A.; Al-Hussain, S.A.; Masand, V.H.; Akasapu, S.; Lewaa, I. QSAR and Pharmacophore Modeling of Nitrogen Heterocycles as Potent Human N-Myristoyltransferase (Hs-NMT) Inhibitors. *Molecules* **2021**, *26*, 1834. [[CrossRef](#)] [[PubMed](#)]
37. Zaki, M.E.A.; Al-Hussain, S.A.; Masand, V.H.; Akasapu, S.; Bajaj, S.O.; El-Sayed, N.N.E.; Ghosh, A.; Lewaa, I. Identification of Anti-SARS-CoV-2 Compounds from Food Using QSAR-Based Virtual Screening, Molecular Docking, and Molecular Dynamics Simulation Analysis. *Pharmaceuticals* **2021**, *14*, 357. [[CrossRef](#)]
38. Masand, V.H.; Mahajan, D.T.; Hadda, T.B.; Jawarkar, R.D.; Alafeefy, A.M.; Rastija, V.; Ali, M.A. Does tautomerism influence the outcome of QSAR modeling? *Med. Chem. Res.* **2014**, *23*, 1742–1757. [[CrossRef](#)]
39. Masand, V.H.; Mahajan, D.T.; Gramatica, P.; Barlow, J. Tautomerism and multiple modelling enhance the efficacy of QSAR: Antimalarial activity of phosphoramidate and phosphorothioamidate analogues of amiprofos methyl. *Med. Chem. Res.* **2014**, *23*, 4825–4835. [[CrossRef](#)]
40. Gramatica, P. External Evaluation of QSAR Models, in Addition to Cross-Validation Verification of Predictive Capability on Totally New Chemicals. *Mol. Inform.* **2014**, *33*, 311–314. [[CrossRef](#)]
41. Tropsha, A. Recent trends in statistical QSAR modeling of environmental chemical toxicity. In *Molecular, Clinical and Environmental Toxicology*; Springer: Berlin/Heidelberg, Germany, 2012; Volume 101, pp. 381–411. [[CrossRef](#)]
42. Fourches, D.; Muratov, E.; Tropsha, A. Trust, but verify: On the importance of chemical structure curation in cheminformatics and QSAR modeling research. *J. Chem. Inf. Model.* **2010**, *50*, 1189–1204. [[CrossRef](#)]
43. O'Boyle, N.M.; Banck, M.; James, C.A.; Morley, C.; Vandermeersch, T.; Hutchison, G.R. Open Babel: An open chemical toolbox. *J. Cheminform.* **2011**, *3*, 33. [[CrossRef](#)]
44. Stewart, J.J.P. MOPAC: A semiempirical molecular orbital program. *J. Comput. Aided Mol. Des.* **1990**, *4*, 1–103. [[CrossRef](#)]
45. Masand, V.H.; Rastija, V. PyDescriptor: A new PyMOL plugin for calculating thousands of easily understandable molecular descriptors. *Chemom. Intell. Lab. Syst.* **2017**, *169*, 12–18. [[CrossRef](#)]
46. Yap, C.W. PaDEL-descriptor: An open source software to calculate molecular descriptors and fingerprints. *J. Comput. Chem.* **2011**, *32*, 1466–1474. [[CrossRef](#)]
47. Gramatica, P.; Chirico, N.; Papa, E.; Cassani, S.; Kovarich, S. QSARINS: A new software for the development, analysis, and validation of QSAR MLR models. *J. Comput. Chem.* **2013**, *34*, 2121–2132. [[CrossRef](#)]
48. Martin, T.M.; Harten, P.; Young, D.M.; Muratov, E.N.; Golbraikh, A.; Zhu, H.; Tropsha, A. Does rational selection of training and test sets improve the outcome of QSAR modeling? *J. Chem. Inf. Model.* **2012**, *52*, 2570–2578. [[CrossRef](#)]
49. Kar, S.; Roy, K.; Leszczynski, J. Applicability Domain: A Step Toward Confident Predictions and Decidability for QSAR Modeling. In *Computational Toxicology*; Humana Press: New York, NY, USA, 2018; pp. 141–169.
50. Netzeva, T.I.; Worth, A.; Aldenberg, T.; Benigni, R.; Cronin, M.T.; Gramatica, P.; Jaworska, J.S.; Kahn, S.; Klopman, G.; Marchant, C.A.; et al. Current status of methods for defining the applicability domain of (quantitative) structure-activity relationships. The report and recommendations of ECVAM Workshop 52. *Altern. Lab. Anim. ATLA* **2005**, *33*, 155–173. [[CrossRef](#)]
51. Trott, O.; Olson, A.J. AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.* **2009**, *31*, 455–461. [[CrossRef](#)]
52. Di Muzio, E.; Toti, D.; Polticelli, F. DockingApp: A user friendly interface for facilitated docking simulations with AutoDock Vina. *J. Comput. Aided Mol. Des.* **2017**, *31*, 213–218. [[CrossRef](#)]



Article

Synthesis and Evaluation of Some New 4*H*-Pyran Derivatives as Antioxidant, Antibacterial and Anti-HCT-116 Cells of CRC, with Molecular Docking, Antiproliferative, Apoptotic and ADME Investigations

Nahed N. E. El-Sayed ^{1,*}, Magdi E. A. Zaki ^{2,*}, Sami A. Al-Hussain ², Abir Ben Bacha ^{3,4}, Malika Berredjem ⁵, Vijay H. Masand ⁶, Zainab M. Almarhoon ⁷ and Hanaa S. Omar ^{8,9}

- ¹ National Organization for Drug Control and Research, Egyptian Drug Authority (EDA), 51 Wezaret El-Zeraa St., Giza 35521, Egypt
 - ² Department of Chemistry, Faculty of Sciences, Imam Mohammad Ibn Saud Islamic University, Riyadh 13318, Saudi Arabia; sahusain@imamu.edu.sa
 - ³ Biochemistry Department, College of Sciences, King Saud University, Riyadh 11495, Saudi Arabia; aalghanouchi@ksu.edu.sa
 - ⁴ Laboratory of Plant Biotechnology Applied to Crop Improvement, Faculty of Science of Sfax, University of Sfax, Sfax 3038, Tunisia
 - ⁵ Laboratory of Applied Organic Chemistry LCOA, Synthesis of Biomolecules and Molecular Modeling Group, Badji-Mokhtar-Annaba University, Annaba 23000, Algeria; mberredjem@yahoo.fr
 - ⁶ Department of Chemistry, Vidya Bharati College, Amravati 444602, Maharashtra, India; vijaymasand@gmail.com
 - ⁷ Department of Chemistry, College of Science, King Saud University, Riyadh 11451, Saudi Arabia; zalmarhoon@ksu.edu.sa
 - ⁸ Department of Genetics, Faculty of Agriculture, Cairo University, Giza 12613, Egypt; hanaa.omar@agr.cu.edu.eg
 - ⁹ GMO Laboratory, Faculty of Agriculture, Cairo University, Research Park, Giza 12613, Egypt
- * Correspondence: nnelsayed@gmail.com or nahed.elsayed@edaegypt.gov.eg (N.N.E.E.-S.); mezaki@imamu.edu.sa (M.E.A.Z.)



Citation: El-Sayed, N.N.E.; Zaki, M.E.A.; Al-Hussain, S.A.; Ben Bacha, A.; Berredjem, M.; Masand, V.H.; Almarhoon, Z.M.; Omar, H.S. Synthesis and Evaluation of Some New 4*H*-Pyran Derivatives as Antioxidant, Antibacterial and Anti-HCT-116 Cells of CRC, with Molecular Docking, Antiproliferative, Apoptotic and ADME Investigations. *Pharmaceuticals* **2022**, *15*, 891. <https://doi.org/10.3390/ph15070891>

Academic Editor: Luis M. T. Frija

Received: 20 May 2022

Accepted: 2 July 2022

Published: 19 July 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Abstract: Colorectal cancer oncogenesis is linked to dysbiosis, oxidative stress and overexpression of CDK2. The 4*H*-pyran scaffold is considered an antitumoral, antibacterial and antioxidant lead as well as a CDK2 inhibitor. Herein, certain 4*H*-pyran derivatives were evaluated as antibacterial, antioxidant and cytotoxic agents against HCT-116 cells. Derivatives **4g** and **4j** inhibited all the tested Gram-positive isolates, except for *B. cereus* (ATCC 14579), with lower IC₅₀ values (μM) than ampicillin. In addition, **4g** and **4j** demonstrated the strongest DPPH scavenging and reducing potencies, with **4j** being more efficient than BHT. In cell viability assays, **4d** and **4k** suppressed the proliferation of HCT-116 cells, with the lowest IC₅₀ values being 75.1 and 85.88 μM, respectively. The results of molecular docking simulations of **4d** and **4k**, inhibitory kinase assays against CDK2, along with determination of CDK2 protein concentration and the expression level of CDK2 gene in the lysates of HCT-116 treated cells, suggested that these analogues blocked the proliferation of HCT-116 cells by inhibiting kinase activity and downregulating expression levels of CDK2 protein and gene. Moreover, **4d** and **4k** were found to induce apoptosis in HCT-116 cells via activation of the caspase-3 gene. Lastly, compounds **4g**, **4j**, **4d** and **4k** were predicted to comply with Lipinski's rule of five, and they are expected to possess excellent physicochemical and pharmacokinetic properties suitable for in vivo bioavailability, as predicted by the SwissADME web tool.

Keywords: colorectal cancer; dysbiosis; oxidative stress; 4*H*-pyran; CDK2; caspases; ADME

1. Introduction

Colorectal cancer (CRC) is a huge international health burden. It is currently ranked as the third most common cancer and the fourth most common cause of cancer-related death

globally [1]. According to estimates from GLOBOCAN, in 2020 there were 1.9 million new CRC incident cases and 0.9 million deaths worldwide. The global projections for 2040 of new CRC cases are estimated to reach approximately 3.2 million [2].

Colorectal tumorigenesis is a highly complex process (Figure 1). Although little is known about the exact causes of sporadic cancer (with no family history of genetic predisposition), it is initiated by a number of carcinogenic events, which lead to the accumulation of genetic mutations in oncogenes and tumor suppressor genes, in addition to epigenetic modifications. These events drive the transformation of normal cells into uncontrolled adenomas and eventually to malignant carcinomas [3]. A good understanding of the factors promoting or accompanying these carcinogenic events would enable better policies for prevention and targeted therapy. Oxidative stress and dysbiosis of colon microflora constitute risk factors that drive these carcinogenic events, while deregulation of cell-cycle related proteins is a key feature of all cancer cells [4].

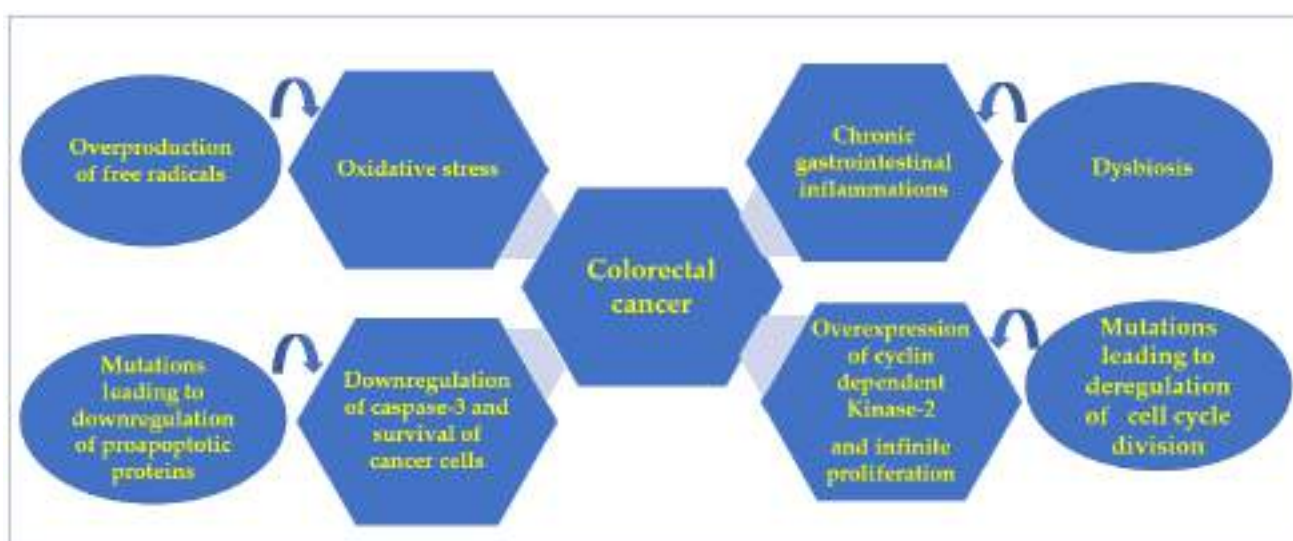


Figure 1. Background of this work.

Oxidative stress is caused by an overwhelming production of reactive oxygen species (ROS) accompanied by the downregulation of antioxidant enzymes within the cell. These reactive molecular species may be generated from endogenous sources (intrinsic), including mitochondria, inflammatory cells and several enzymatic cellular complexes. In addition, they can originate from external sources (extrinsic), such as toxins, certain drugs, radiation and tobacco smoking. It is now well established that uncontrolled increased levels of reactive oxygen species contribute to the etiology of CRC via different mechanisms. One mechanism involves lipid peroxidation via attacking polyunsaturated fatty acids, which lead to the production of malonaldehyde, hexanal, acrolein and 4-hydroxy-2-nonenal, all of which are known mutagenic products capable of inducing persistent destabilization of chromosomes [5]. A second mechanism involves DAN modification, which results in mutations particularly in the p53 gene and mitochondrial DNA [5]. A third mechanism comprises induction of chronic inflammatory bowel diseases (IBDs), including, Crohn's disease (CD) and ulcerative colitis (UC) [6]. Moreover, deregulation of ROS is accompanied by the generation of an inflammatory environment, which suppresses apoptosis and activates proliferation and angiogenesis, thus eventually leading to the initiation of neoplastic transformation [7]. In addition, several studies have also proved the involvement of ROS in the migration and invasion of cancerous cells [8]. Furthermore, ROS play a prominent role in chemotherapy resistance toward drugs, such as 5-fluorouracil, vinblastine, doxorubicin and tamoxifen [8]. Consequently, trapping free radicals with antioxidants seems to be a possible approach to prevent inflammation and cancer. Additionally, antioxidants are reported to exhibit anticancer activities through activating immune response, inhibiting angiogenesis, downregulating oncogenes and stimulating tumor suppressor genes [9].

Experimental evidence has established that dysbiosis of certain intestinal bacterial strains can promote chronic gastrointestinal inflammations, which eventually will lead to the development and spread of CRC [10,11]. The Gram-negative strains, *Escherichia coli* [12] and *Klebsiella pneumonia* belong to this class of virulent pathogens [13]. Additionally, certain studies have hypothesized that *Pseudomonas* bacteria are directly pro-oncogenic and promoters for CRC [14].

Additionally, animal model studies have shown that *Enterococcus faecalis* caused colon inflammation after infection, which promoted CRC occurrence [15]. In addition to this, some case studies indicated that the mean copy number of *Enterococcus faecalis* in people with colorectal cancer was significantly higher than in those with polyps and healthy people, which implies the ability of this strain to induce colorectal carcinogenesis [16].

Despite the recent advances in novel chemotherapy, such as immune checkpoint inhibitors, limitations due to chemotherapeutic resistance remain. Many preclinical and clinical studies have shown that microbiota modulate cancer treatment responses via influencing pharmacokinetic properties, such as metabolism, enzymatic degradation and pharmacodynamics (i.e., immunomodulation), in cancer therapy. Therefore, some new insights into cancer treatment are proposed based on manipulating the gut microbiota to augment cancer therapeutic responses [17].

Cancer patients who are undergoing anticancer treatment have severely weakened immune systems; thus, they are more vulnerable to bacterial infections and their life-threatening complications. A previous retrospective surveillance study revealed that the most common associated bacterial pathogens were *Klebsiella pneumonia* and *Enterococcus faecium* (bloodstream), *Pseudomonas aeruginosa* (respiratory tract), *Escherichia coli* (in urinary tract), *Staphylococcus aureus* and *S. epidermidis* (skin and soft tissues) [18]. It is worth noting that some of these microbes are rather difficult to treat, as they are considered to be drug-resistant isolates [19]. Furthermore, several reports have associated antibiotic resistance with oxidative stress [20]. Therefore, molecules exerting antioxidant [21] and/or antibacterial [22] effects would be considered as attractive candidates that could provide protection against the development of CRC and its complications.

Up to 85% of CRC cases exhibit chromosomal instability (CIN), featuring changes in chromosome numbers and structure. CIN affects the expression of tumor-associated genes and/or genes regulating cell-cycle division [23]. Decades of research in diverse fields, including cell biology, yeast genetics, biochemistry and genetic engineering, have identified the regulators of the cell division cycle [24]. Normally, the transitions of the cell cycle of the mammalian cell are strictly mediated by many cyclins and their effectors, cyclin dependent kinases (CDKs), which are a group of conserved serine/threonine protein kinases, each of which regulates a specific stage of the cell cycle [25]. In human cells, 20 CDKs and 29 cyclins have been identified. The CDK family is comprised of CDK1–7 and CDK14–18, which mediate cell-cycle transitions and cell division, whereas CDK8–13 and CDK19–20 regulate gene transcription [26]. Thus, any deregulation in the function or change in the level of these enzymes can result in the induction of tumorigenesis, including CRC [27]. In fact, previous reports have confirmed that the expression of CDK2 is dramatically increased in various colon cancer cell lines [28], including HCT-116, HCT-15 [29], LoVo and DLD-1 [30], compared with “normal” human colon epithelial cells (HCECs). Thus, targeting CDK2 has been recognized as a potential therapeutic approach for colon cancer treatment [31]. In addition, CDK2 inhibitors are likewise expected to have effectiveness in combination with other drugs or where synthetic lethality can be identified [32].

Taken together, the discovery, design and development of novel effective molecules for neutralizing free radicals and eradicating virulent bacterial strains and inhibiting CKD2 are to be considered as preventive and therapeutic approaches against CRC.

Consequently, in a quest to design new chemical compounds to meet all these health needs, the pyran scaffold can be considered a potential synthetic target. The 4*H*-Pyran motif is embedded in bioactive 4*H*-chromene derivatives which display antitumoral and antibacterial effects, such as compounds **Ia,b** [33], and antioxidant activities, such as

derivatives **IIa,b** [34]. Moreover, compelling evidence has indicated that CDK2 is a valid anticancer target of pyran-containing compounds **III–VI**, as shown in Figure 2 [35–37], due to its pro-tumorigenic role via interference with cell division cycle.

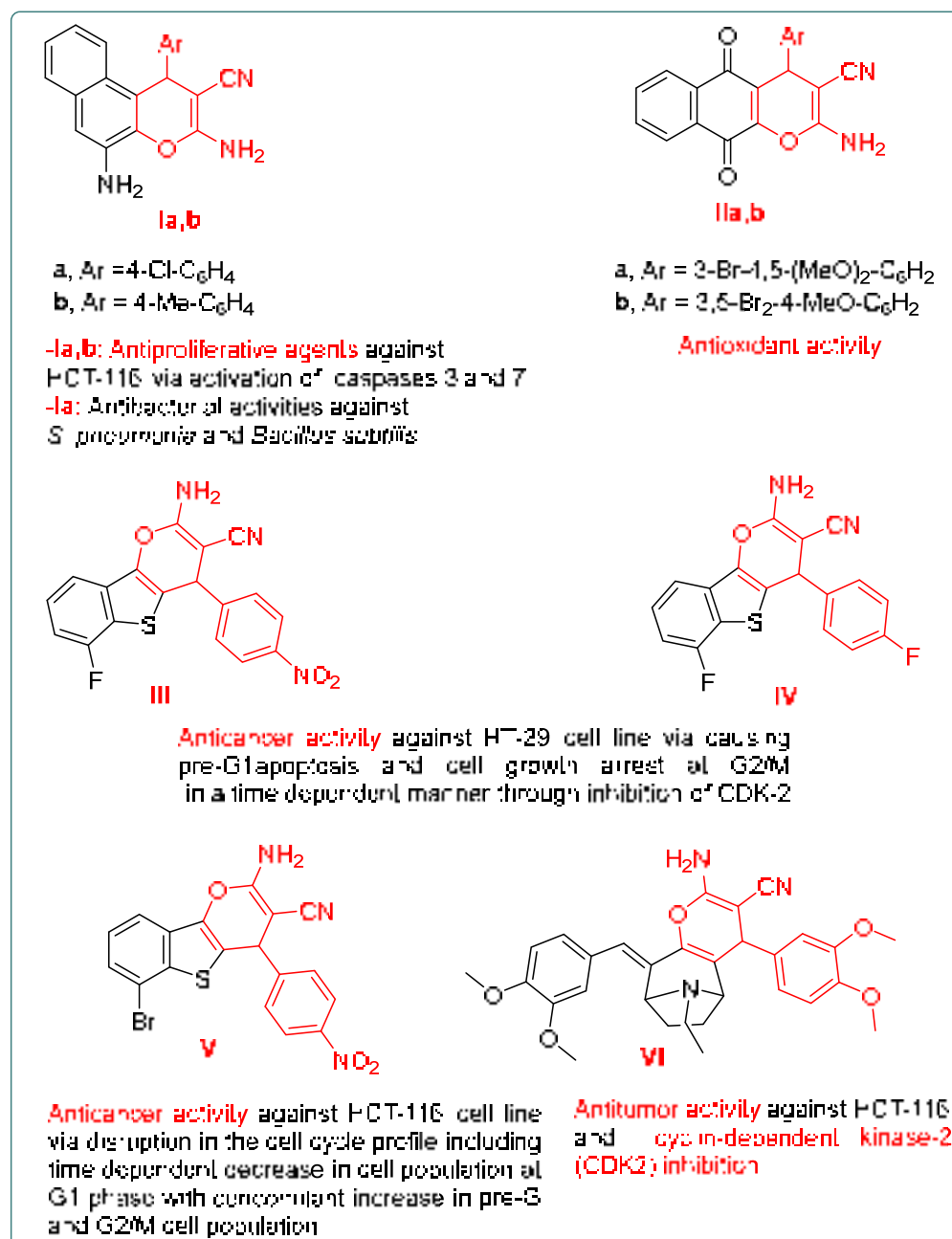


Figure 2. Examples of pyran-containing compounds with antibacterial [33], antioxidant [34] and anticancer activities via inhibition of CDK2 [35–37].

Synthetically, 4*H*-pyrans and pyrano[2,3-*c*]pyrazoles derivatives are easily accessible, using a variety of catalysts [38,39] at ambient temperature, or under microwave or ultrasound irradiations [40–42]. These compounds are proven to display antibacterial [43–45], antitumoral [46–48] and antioxidant activities [49].

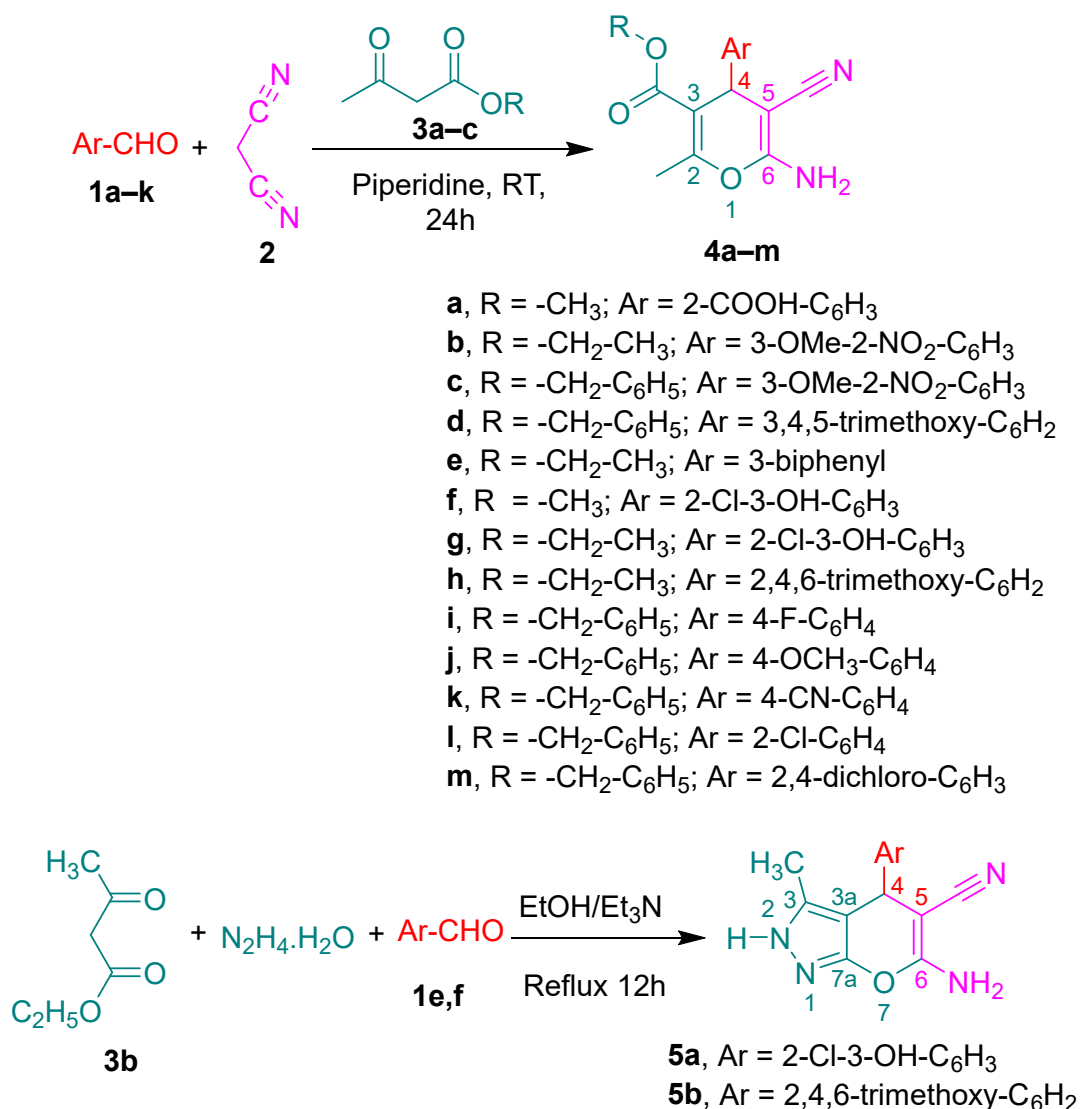
Based on the above-stated facts and in continuation of our research program of anticancer drug discovery, we designed and synthesized some new 4*H*-pyran and pyrano[2,3-*c*]pyrazole derivatives for in vitro biological screening as antioxidant, antibacterial and antiproliferative agents against HCT-116 cells of human colorectal cancer [50,51]. The underlying mechanism of the antiproliferative actions of the promising cytotoxic 4*H*-

pyran derivatives was studied using molecular docking simulations in the ATP binding pocket of CDK2 and through the performance of a CDK2 inhibitory assay. Moreover, the concentration of CDK2 protein and the expression level of CDK2 gene in the lysates of HCT-116 cells treated with the cytotoxic candidates were measured and compared with positive and negative control experiments. Furthermore, quantitative real time PCR experiments for the caspase-3 gene were carried out to investigate the apoptotic activity of these candidates. Lastly, the expected physiochemical and pharmacokinetic properties of the active candidates were predicted using a Lipinski's rule of five filter and BOILED-EGG and bioavailability radar charts.

2. Results and Discussion

2.1. Chemistry

Initially, the 4*H*-pyran templates **4a–m** were prepared using one-pot, tandem cascade Knoevenagel condensation/Michael addition and cyclization reactions of diverse aryl aldehydes **1a–k**, malononitrile **2** and various β -ketoesters **3a–c** in ethanolic piperidine solutions at room temperature. Similarly, the synthesis of pyrano[2,3-*c*]pyrazole derivatives **5a,b** was efficiently achieved via one-pot, four-component reactions of ethyl acetoacetate **3b**, hydrazine hydrate, aromatic aldehydes **1e,f** and malononitrile **2** under thermal heating and in the presence of ethanol as a solvent and triethyl amine as a catalyst (Scheme 1).



Scheme 1. Synthesis of 4*H*-pyrans and 4*H*-pyrano[2,3-*c*]pyrazole derivatives.

The structures of all synthesized compounds were confirmed by their spectroscopic data, in addition to the results of elemental analyses. Thus, the IR spectra of derivatives **4a–m** showed the characteristic stretching absorption bands in the frequency range of ν_{\max} (KBr)/ cm^{-1} 3478~3416 (asymmetric vibration— NH_2) and 3349~3197 (symmetric vibration— NH_2), 2203~2188 ($\text{C}\equiv\text{N}$) and 1724~1689 ($\text{C}=\text{O}$); whereas the IR spectra of **5a,b** indicated the presence of the stretching absorption bands at ν_{\max} (KBr)/ cm^{-1} 3377 and 3309 (asymmetric vibration— NH_2), 3381 and 3357 (symmetric vibration— NH_2), 3159 and 3183 (NH) and 2170 and 2183 ($\text{C}\equiv\text{N}$).

The detailed ^1H NMR and ^{13}C NMR analytical data for derivatives **4a–m** and **5a,b** are presented in the experimental section. In addition, the mass spectrum (EI) of each compound displayed the anticipated molecular ion peak. The spectra of some representative examples are provided in the Supplementary Materials.

Finally, the results of the elemental analyses for C, H and N of all the synthesized compounds were found to be within the permissible limits.

2.2. Biological Screening

2.2.1. Antioxidant Assays

Antioxidants are substances capable of free radical scavenging, quenching singlet oxygen, inactivating peroxides and other reactive oxygen species (ROS), chelating metal ions, quenching secondary oxidation products and inhibiting pro-oxidative enzymes. Therefore, the antioxidant potential of a compound can be evaluated using a variety of chemical assays with different mechanisms [52]. Since the synthesized compounds are electron-rich, with substituents capable of electron donation and H-bond donation as well, in this study, two radical scavenging methods were employed, including the 1,1 diphenyl-2-picrylhydrazyl (DPPH) assay and the ferrous-reducing antioxidant capacity (FRAC) assay.

Initially, the DPPH radical scavenging potency assay was performed at various concentrations of the studied compounds ranging from 0.03 to 1 mg/mL. The IC_{50} values for DPPH were calculated from the plot of the scavenging effect against the compound concentration.

The obtained results (Table 1) showed that the compounds **4g**, **4j**, **4l**, **4m** and **4d** exhibited the highest scavenging activity at 1 mg/mL, with scavenging potencies of 90.50, 88.00, 70.20, 69.00 and 65.25 %, respectively, as compared to 95.30 % for BHT. At the same concentration, the remaining compounds demonstrated poor to moderate DPPH radical scavenging activities ranging from 23.25 to 57.10%. Determination of the half-maximal inhibitory concentrations indicated that compounds **4g** and **4j** displayed the most promising free radical scavenging activities, with the lowest IC_{50} values of 0.329 and 0.1941 mM, respectively, as compared to BHT, which exhibited 0.245 mM.

Secondly, FRAC was determined over the concentration range of 0.0–1 mg/mL via monitoring Fe^{3+} – Fe^{2+} transformation in the presence of the studied compounds and the reference antioxidant BHT to estimate the values of efficient concentration (EC_{50}), expressed in mM. As depicted in Table 1, the pyrans **4g**, **4j** and BHT exhibited comparable tendencies with respect to their reducing power, with EC_{50} values of 0.072, 0.074 and 0.089 mM, respectively, whereas the rest of the tested compounds were less effective, with EC_{50} values in the range of 0.149–1.562 mM.

Table 1. Antioxidant activities of the studied 4*H*-pyran and 4*H*-pyrano[2,3-*c*]pyrazole derivatives as determined by calculation of IC₅₀ (mM) and EC₅₀ (mM) values. Experiments were performed in triplicate and are reported as the means ± standard deviations.

Comp. No.	IC ₅₀ (mM) for DPPH Scavenging Potency	DPPH Radical Scavenging Efficiency (%) at 1 mg/mL	Reducing Power EC ₅₀ (mM)
4a	2.291 ± 0.134	24.15 ± 1.62	0.827 ± 0.045
4b	2.560 ± 0.117	23.25 ± 1.06	1.197 ± 0.079
4c	2.243 ± 0.050	57.10 ± 1.56	1.234 ± 0.067
4d	0.504 ± 0.032	65.25 ± 1.77	0.149 ± 0.016
4e	1.651 ± 0.097	39.25 ± 1.06	0.305 ± 0.039
4f	3.336 ± 0.131	43.75 ± 1.77	0.398 ± 0.045
4g	0.329 ± 0.042	90.50 ± 2.12	0.072 ± 0.004
4h	1.309 ± 0.075	33.50 ± 2.12	0.221 ± 0.011
4i	1.784 ± 0.115	26.05 ± 1.06	0.274 ± 0.039
4j	0.194 ± 0.011	88.00 ± 1.41	0.074 ± 0.004
4k	1.050 ± 0.075	40.55 ± 2.19	1.562 ± 0.102
4l	0.919 ± 0.026	70.20 ± 1.70	0.801 ± 0.019
4m	0.421 ± 0.017	69.0 ± 2.83	0.506 ± 0.034
5a	0.826 ± 0.092	45.50 ± 2.12	0.363 ± 0.047
5b	1.781 ± 0.082	37.50 ± 2.12	0.467 ± 0.041
BHT	0.245 ± 0.027	95.30 ± 0.42	0.089 ± 0.003

2.2.2. Antibacterial Assays

The results for the in vitro antibacterial properties of the synthesized compounds against five Gram-positive and four Gram-negative human-pathogenic strains, which were assessed by both the agar diffusion method (based on inhibition zone (IZ)) and the broth dilution bioassay (for the determination of the half-maximal inhibitory concentration (IC₅₀) values), using ampicillin as the reference antibiotic drug, are shown in Tables 2 and 3.

Some of the tested compounds were found to be more effective against Gram-positive bacteria than Gram-negative isolates by producing larger IZs and lowering IC₅₀ against certain strains. Reduced sensitivity against Gram-negative pathogens can be attributed to the restricted permeability of the chemical inhibitors across their cell walls due to the existence of an outer cell envelope, which acts as a selective barrier [53].

The results against Gram-positive strains identified compound 4g as the most active agent against *B. subtilis* (ATCC 6633), with a lowered IC₅₀ value (25.69 μM) and a bigger inhibition zone of 28.00 mm, compared to the results obtained for ampicillin: IC₅₀ = 37.20 μM, IZ = 25.50 mm. Despite pyrans 4f, 4i, 4j, 4l and 4m displaying lower IC₅₀ values (ranging from 36.79 to 23.64 μM), their IZs were bigger, in the range of 21.50 (4j) to 9.50 mm (4i).

Considering *S. aureus* (ATCC 25923) and *S. epidermidis* (ATCC 14990), bigger IZs (mm) were displayed by 4g (27.50, 29.75) and 4j (25.25, 26.65) as compared to the results of 21.50, 22.50 obtained for ampicillin, respectively. In addition, these derivatives exhibited lowered IC₅₀ (μM) values (4g: 27.78, 30.32) and (4j: 33.34, 33.34) as compared to 38.64 and 50.09 produced by ampicillin against these isolates, respectively.

Table 2. Antibacterial evaluation of the studied compounds against some Gram-positive bacterial strains by measurement of the diameters of the inhibition zones (IZs, mm) and determination of the half-maximal inhibitory concentrations (IC₅₀s, μM). Ampicillin was used as the reference antibiotic. Results were obtained from three independent experiments.

Comp. No.	<i>B. cereus</i> (ATCC 14579)		<i>B. subtilis</i> (ATCC 6633)		<i>E. faecalis</i> (ATCC 29122)		<i>S. aureus</i> (ATCC 25923)		<i>S. epidermidis</i> (ATCC 14990)	
	IZ (mm)	IC ₅₀ (μM)	IZ (mm)	IC ₅₀ (μM)	IZ (mm)	IC ₅₀ (μM)	IZ (mm)	IC ₅₀ (μM)	IZ (mm)	IC ₅₀ (μM)
4a	7.50 ± 0.71	71.75 ± 2.04	5.20 ± 0.28	64.43 ± 3.37	9.00 ± 1.41	49.64 ± 1.81	8.00 ± 1.41	41.20 ± 2.48	9.50 ± 0.71	42.32 ± 1.33
4b	16.50 ± 0.71	43.14 ± 1.98	19.50 ± 0.71	49.12 ± 1.39	15.50 ± 0.71	56.77 ± 1.59	20.50 ± 0.71	50.23 ± 1.39	23.50 ± 0.71	32.42 ± 1.39
4c	14.00 ± 1.41	44.50 ± 0.83	11.00 ± 1.41	46.27 ± 1.68	10.00 ± 1.41	41.05 ± 2.18	15.50 ± 0.71	37.01 ± 2.02	16.65 ± 0.50	30.37 ± 1.00
4d	9.50 ± 0.71	39.52 ± 0.80	12.00 ± 1.41	40.32 ± 1.95	14.50 ± 0.71	29.10 ± 2.27	9.00 ± 1.41	24.40 ± 1.15	8.50 ± 0.71	26.69 ± 1.15
4e	23.00 ± 1.41	33.57 ± 1.58	20.50 ± 0.71	37.73 ± 1.58	15.50 ± 0.71	37.32 ± 2.16	16.50 ± 2.10	34.82 ± 2.16	20.50 ± 2.12	30.80 ± 1.58
4f	22.50 ± 0.71	42.40 ± 2.65	19.50 ± 0.71	36.79 ± 0.87	19.00 ± 1.41	34.61 ± 0.59	16.50 ± 0.71	38.35 ± 1.31	14.50 ± 0.71	33.05 ± 2.65
4g	29.00 ± 1.41	29.42 ± 1.49	28.00 ± 1.41	25.69 ± 1.70	24.50 ± 0.71	31.82 ± 1.49	27.50 ± 0.71	27.78 ± 1.25	29.75 ± 0.35	30.32 ± 1.49
4h	4.50 ± 0.71	46.48 ± 1.52	6.00 ± 1.41	50.08 ± 2.83	8.00 ± 1.41	45.94 ± 0.75	17.50 ± 0.71	32.72 ± 0.93	16.50 ± 0.71	35.79 ± 1.52
4i	14.50 ± 0.71	38.01 ± 1.37	9.50 ± 0.71	31.56 ± 1.95	6.50 ± 0.71	37.46 ± 1.37	9.50 ± 2.12	40.21 ± 1.37	11.50 ± 0.71	45.97 ± 0.96
4j	24.50 ± 0.71	27.63 ± 1.51	21.50 ± 0.71	23.64 ± 0.37	23.20 ± 0.57	28.16 ± 1.51	25.25 ± 1.06	33.34 ± 1.70	26.65 ± 0.92	33.34 ± 1.70
4k	7.00 ± 1.41	48.06 ± 1.35	4.50 ± 0.71	42.14 ± 2.48	5.50 ± 0.71	45.37 ± 0.57	5.75 ± 0.35	43.89 ± 0.57	4.50 ± 0.71	50.22 ± 2.48
4l	10.00 ± 1.41	33.22 ± 2.42	11.50 ± 0.71	32.56 ± 1.50	6.00 ± 1.41	37.81 ± 1.50	9.00 ± 1.41	31.12 ± 0.55	10.50 ± 0.71	33.35 ± 1.10
4m	18.50 ± 0.71	35.76 ± 1.20	21.50 ± 0.71	32.75 ± 2.05	20.50 ± 0.71	35.28 ± 2.22	17.50 ± 0.71	33.23 ± 1.01	18.50 ± 0.71	35.04 ± 1.88
5a	12.50 ± 0.71	62.43 ± 0.46	12.50 ± 0.71	62.43 ± 0.46	10.50 ± 0.71	51.86 ± 1.39	14.00 ± 1.40	43.93 ± 1.39	13.50 ± 0.71	54.84 ± 1.88
5b	8.00 ± 1.41	42.65 ± 1.66	8.50 ± 0.71	41.77 ± 1.23	9.75 ± 1.06	45.13 ± 1.87	5.50 ± 0.71	39.14 ± 1.66	6.60 ± 0.57	50.05 ± 0.61
Ampicillin	23.50 ± 0.71	25.76 ± 0.57	25.50 ± 0.71	37.20 ± 1.43	24.50 ± 0.71	36.49 ± 3.03	21.50 ± 0.71	38.64 ± 2.00	22.50 ± 0.71	50.09 ± 3.43

Table 3. Antibacterial evaluation for the studied compounds against some Gram-negative bacterial strains by measurement of the diameters of the inhibition zones (IZs, mm) and determination of the half-maximal inhibitory concentrations (IC₅₀s, µM). Ampicillin was used as the reference antibiotic. Results were obtained from three independent experiments.

Comp. No.	<i>E. coli</i> (ATCC 25966)		<i>K. pneumonia</i> (ATCC 700603)		<i>P. aeruginosa</i> (ATCC 27853)		<i>S. enteric</i> (ATCC 43972)	
	IZ (mm)	IC ₅₀ (µM)	IZ (mm)	IC ₅₀ (µM)	IZ (mm)	IC ₅₀ (µM)	IZ (mm)	IC ₅₀ (µM)
4a	4.5 ± 0.71	138.41 ± 6.68	2.5 ± 0.7	131.09 ± 3.50	5.5 ± 0.7	37.6 ± 0.5	4.5 ± 0.7	119.63 ± 1.91
4b	2.5 ± 0.71	100.19 ± 3.90	1.5 ± 0.7	99.07 ± 2.50	2.5 ± 0.7	84.05 ± 5.01	1.5 ± 0.7	118.14 ± 3.34
4c	2.5 ± 0.71	93.74 ± 4.98	3.5 ± 0.7	112.25 ± 1.42	7.5 ± 0.7	85.90 ± 2.61	8.5 ± 0.7	100.26 ± 2.61
4d	3.5 ± 0.71	88.67 ± 2.52	4.5 ± 0.7	72.63 ± 0.69	6.5 ± 0.7	67.36 ± 1.37	7.0 ± 0.0	65.77 ± 2.29
4e	2.7 ± 0.35	103.22 ± 3.05	2.5 ± 0.7	90.73 ± 2.78	4.5 ± 0.7	87.96 ± 2.78	3.5 ± 0.7	96.14 ± 2.50
4f	7.5 ± 0.71	95.72 ± 5.61	4.5 ± 0.7	82.00 ± 3.43	6.0 ± 1.4	104.76 ± 2.81	6.5 ± 0.7	94.16 ± 3.43
4g	11.0 ± 1.41	61.84 ± 5.38	11.5 ± 0.7	61.84 ± 2.99	13.5 ± 0.7	54.07 ± 4.48	11.5 ± 0.7	53.02 ± 3.29
4h	1.5 ± 0.71	114.05 ± 0.80	3.5 ± 0.7	118.33 ± 2.67	4.5 ± 0.7	89.48 ± 2.94	5.5 ± 0.7	111.65 ± 2.94
4i	5.5 ± 0.71	82.57 ± 1.65	6.5 ± 0.7	79.86 ± 3.02	8.5 ± 0.7	62.02 ± 2.20	8.0 ± 1.4	91.52 ± 2.74
4j	10.5 ± 0.71	57.12 ± 0.80	12.0 ± 1.4	48.35 ± 3.99	14.5 ± 0.7	49.68 ± 4.78	14.5 ± 0.7	51.81 ± 2.92
4k	4.5 ± 0.71	80.24 ± 4.31	2.5 ± 0.7	77.82 ± 2.96	3.5 ± 0.7	73.24 ± 3.50	2.5 ± 0.7	75.53 ± 2.42
4l	3.5 ± 0.7	90.07 ± 2.63	2.5 ± 0.7	77.46 ± 7.08	4.5 ± 0.7	72.74 ± 2.63	4.0 ± 0.0	77.60 ± 1.58
4m	5.5 ± 0.71	69.35 ± 2.89	6.5 ± 0.7	66.70 ± 2.41	9.5 ± 0.7	60.92 ± 2.17	8.5 ± 0.7	61.77 ± 2.17
5a	2.8 ± 0.21	109.67 ± 5.95	6.0 ± 1.4	96.46 ± 4.96	5.5 ± 0.7	85.89 ± 3.30	4.5 ± 0.7	104.72 ± 1.32
5b	3.5 ± 0.71	116.84 ± 4.09	4.5 ± 0.7	109.54 ± 1.75	6.5 ± 0.7	92.30 ± 1.46	7.5 ± 0.7	106.03 ± 2.92
Ampicillin	23.5 ± 0.71	57.96 ± 4.87	21.0 ± 1.4	32.91 ± 2.29	21.5 ± 0.7	41.50 ± 1.72	19.5 ± 0.7	14.88 ± 0.34

Moreover, derivative **4b** selectively exerted a larger IZ (23.50 mm) and a lower IC₅₀ value (32.42 µM) against *S. epidermidis*, whereas compounds **4a**, **4c**, **4d**, **4e**, **4f**, **4l**, **4m** and **5b** inhibited it, with lowered IC₅₀ values (µM) spanning from 50.05 (**5b**) to 26.69 (**4d**), though they displayed smaller inhibition zones than the reference, ranging from 6.60 (**5b**) to 20.50 (**4e**). Therefore, *S. epidermidis* can be considered the most sensible strain to the synthesized compounds. Similarly, compounds **4c**, **4d**, **4e**, **4f**, **4h**, **4l** and **4m** suppressed the growth of *S. aureus*, with lowered IC₅₀ values spanning from 24.40 (**4d**) to 38.35 (**4f**) and smaller IZs ranging from 9.00 (**4d** and **4l**) to 17.50 (**4m** and **4h**).

With regard to *E. faecalis* (ATCC 29122), pyran **4g** and the standard drug produced equal inhibition zones of 24.50 mm; however, **4g** was capable of exhibiting a lowered IC₅₀ value of 31.82 µM in comparison with that of 36.49 µM for ampicillin.

Contrarily, pyrans **4d**, **4f**, **4j** and **4m** showed lower inhibition zones, ranging from 14.50 (**4d**) to 23.20 mm (**4j**), and lower IC₅₀ values, ranging from 28.16 (**4j**) to 35.28 µM (**4m**), than ampicillin.

Lastly, although the same compounds **4g** and **4j** exerted the biggest inhibition zones of 29.00 and 24.50 mm relative to that of 23.50 mm exerted by ampicillin against *B. cereus*, they exhibited higher IC₅₀ values (29.42 and 27.63 µM, respectively) as compared to ampicillin (25.76 µM).

Collectively, **4g** showed promising antibacterial efficiency (lower IC₅₀ values and bigger IZs) against all of the tested isolates as compared to the reference antibiotic, except for *B. cereus*, which was the microbe least sensitive to the synthesized compounds. Similarly, **4j** was very effective against *S. aureus* (ATCC 25923) and *S. epidermidis* (ATCC 14990).

Contrarily, the anti-Gram-negative evaluation revealed that none of the studied compounds was capable of inhibiting the microbes with bigger IZs or lower IC₅₀ values, except for **4a**, which inhibited *P. aeruginosa* (ATCC 27853), with an IC₅₀ of 37.6 µM, which is lower than those of ampicillin (41.50 µM) and **4j**, which suppressed *E. coli* (ATCC 25966) by a comparable concentration of 57.12 µM to that of ampicillin (57.96 µM).

Furthermore, the lowest IC₅₀ values against *K. pneumonia* (ATCC 700603) and *S. enteric* (ATCC 43972) were displayed by the derivatives **4g** (61.84 and 54.07 µM) and **4j** (48.35 and 49.68 µM) as compared to 32.91 and 41.50 µM, respectively, for ampicillin.

In summary, although none of the tested compounds showed significant activities against Gram-negative microbes, the observed anti-Gram-positive potential of compounds **4g** and **4j** is of great significance. This is due to the fact that Gram-positive strains are reported to be common causes of bloodstream, skin, soft-tissue and intra-abdominal in-

fections [54], which result in severe manifestations, such as sepsis [55] and/or endocarditis [56], especially in immunosuppressed cancer patients. Furthermore, recent reports have indicated that some of these bacterial isolates showed multidrug-resistant profiles [57], particularly *E. faecalis* and *S. aureus*, which amplify their serious complications and increase health care expenditures.

The observed broad-spectrum anti-Gram-positive profiles of **4g** and **4j** as compared to the other analogues can be attributed to their improved antioxidant capabilities. Indeed, the positive correlation between these two activities is well-documented. Catechin, which is a well-known potent free radical scavenger, has been reported to display anti-Gram positive and anti-Gram negative activities as well. Several mechanisms have been reported to account for the toxicity of catechin towards bacteria. Oxidative damage through membrane permeabilization was found to be the inhibitory mechanism of catechin against *B. subtilis* [58]. In addition, some experimental data showed a significant decrease in superoxide dismutase (SOD) and catalase (CAT) activity after treatment of the Gram-positive *S. aureus* strain with catechin at its minimum inhibitory concentration (MIC). These inhibitions could be attributed to the capability of catechin as an antioxidant agent to chelate the metal cofactors (Zn, Fe and Mn) at SOD catalytic sites [59]. Inhibiting superoxide dismutase would increase the bacterial sensitivity to the reactive oxygen species, resulting in inhibition of bacterial growth [60]. Other alternative proposals for the bactericidal actions of catechin include its ability to generate hydrogen peroxide, which causes cell membrane damage [61]. Hydrogen peroxide generation by antioxidants was reportedly attributable to antioxidant autoxidation in culture media [62]. In the view of the aforementioned facts, it can be suggested that the antibacterial effects of the antioxidant candidates **4g** and **4j** might be attributed to one or more of the indicated mechanisms, which could be investigated in future.

2.2.3. Cell Viability Assays

The synthesized compounds were also examined at four concentrations (10, 25, 50 and 100 µg/mL) for their cytotoxic effects on human colorectal cancer HCT-116 cells after their incubation together for 24 h. The results, which were expressed in terms of percent of viable cells at each concentration as plotted in Figure 3, revealed that analogs **4g** and **4j** were not able to inhibit the growth of the cancerous cells, even at the highest used concentration (100 µg/mL). Conversely, they displayed the highest detectable percentages of viable cells: 86.0 and 94.0% at 100 µg/mL, respectively. Meanwhile, eleven compounds, namely, **4k**, **4b**, **4d**, **6b**, **4l**, **4m**, **4c**, **4e**, **5a**, **4h** and **4i**, exhibited potent cytotoxic activity at a concentration of 100 µg/mL by reducing cells viability to the extent of 11.0–32.5. The remaining compounds in this series, namely, **4a** and **4f**, displayed the least cytotoxic effects with percent cell viability measures of 46.0 and 55.5, respectively.

Based on these results, the IC₅₀ values for the promising cytotoxic candidates were determined, as shown in Table 4.

The calculated IC₅₀ values (µM) ranged from 75.10 (**4d**) to 332.59. The smallest median concentrations were exhibited by compounds **4d** (75.10), **4k** (85.88), **4m** (89.33), **4l** (93.35) and **5b** (97.91).

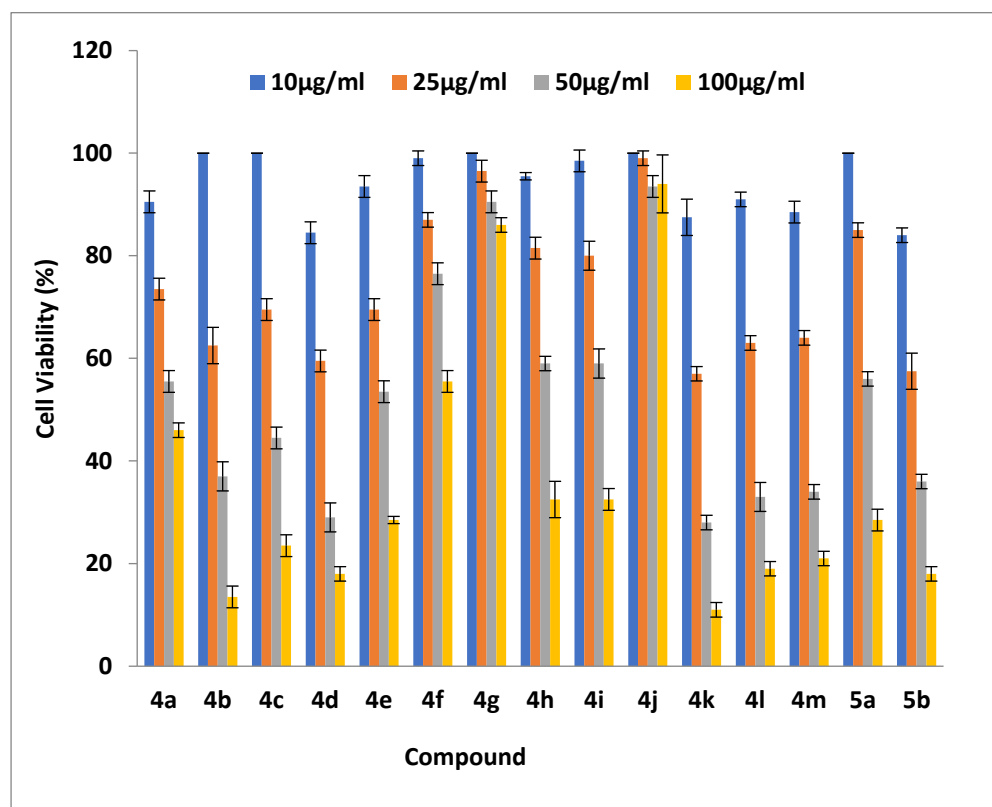


Figure 3. Cytotoxic potency of the studied compounds on HCT-116 cells expressed as percentages of viable cells at 10, 25, 50 and 100 µg/mL concentrations. Results were obtained from three independent experiments.

Table 4. IC₅₀ values of cytotoxicity for the studied 4*H*-pyran and pyrano[2,3-*c*]pyrazole derivatives on HCT-116 cells expressed in µg/mL and µM. Experiments were performed in triplicate and are reported as the means ± standard deviations.

Comp. No.	Mean IC ₅₀ (µg/mL)	Mean IC ₅₀ (µM)
4a	51.99 ± 4.54	165.43 ± 14.44
4b	36.51 ± 3.18	101.60 ± 8.83
4c	43.01 ± 2.03	102.07 ± 4.82
4d	32.78 ± 2.22	75.10 ± 5.09
4e	49.71 ± 3.53	137.92 ± 9.78
4f	106.67 ± 4.29	332.59 ± 13.38
4g	ND ¹	ND ¹
4h	69.21 ± 4.29	184.89 ± 11.46
4i	65.95 ± 4.53	180.99 ± 12.42
4j	ND ¹	ND ¹
4k	31.90 ± 2.83	85.88 ± 7.61
4l	35.55 ± 1.77	93.35 ± 4.63
4m	37.10 ± 2.10	89.33 ± 5.05
5a	60.00 ± 3.28	189.21 ± 10.84
5b	33.52 ± 2.42	97.91 ± 7.06

¹ ND: Not determined.

2.2.4. Investigation of the Underlying Mechanism of Action for the Antiproliferative Candidates

2.2.4.1. Molecular Docking Simulations against CDK2 as a Potential Molecular Target

The molecular docking approach has become a powerful tool for identifying compounds showing potential anticancer activity against a specific target and with a selective

inhibition mechanism [63]. Cancer is increasingly viewed as a cell-cycle disease. Cyclin-dependent kinases (CDKs) are a crucial protein family involved in cell proliferation through regulating the progression of the cell cycle and transcription. Animal model studies and molecular analyses of human tumors have indicated that many of these regulators are altered in cancer, particularly, CDK4, CDK6 and CDK2 and their substrates, which control progression through the G1/S phases of cell-cycle division. A literature survey confirmed that CDK2 is a valid target for 4*H*-pyran derivatives (Figure 2) with anticancer activities [33–37]. Therefore, to investigate whether or not the cellular mechanism by which the most active cytotoxic derivatives **4d** and **4k** suppressed the proliferation of HCT-116 cells is related to the inhibition of CDK2, the induced-fit docking approach in the ATP active site of the enzyme using the Glide program of Schrodinger-Maestro 11.2 was employed prior to the in vitro inhibitory kinase assays. The ATP-binding pocket of CDK2 is considered to be the site most commonly targeted by the kinase inhibitors. They compete with ATP to bind at the kinase site by forming hydrogen bonds with backbone amino acid residues and by establishing hydrophobic interactions with side chains of surrounding residues [64], leading to the inhibition of kinase phosphorylation and suppression of CDK2 hyperactivation, thus, holding back infinite cell proliferation [65].

In this study, the docking procedures were initially validated using redocking approach. Thus, the CDK2-inhibitor 4-[3-hydroxyanilino]-6,7-dimethoxyquinazoline (co-crystallized ligand, **DTQ**) was removed, after which it was redocked to reproduce the reported complex (PDB code: 1DI8). The significant residues in the active site of the protein are Lys33, Glu51 and Asp145. Figure 4 shows the superimposition of the predicted and the co-crystal structure of the reference inhibitor in the ATP active site of the protein. Moreover, the value of the root-mean-square deviation (RMSD) between the redocked conformation of **DTQ** and that observed in the X-ray crystallographic complex was found to be 0.34 Å, which is less than the cutoff value (2 Å) for the correct docking procedures. These observations confirmed that the redocked ligand was closely bound to the true conformation, indicating the reliability of the docking protocols.

Afterwards, a comparative in silico molecular modeling study using the most active cytotoxic derivatives **4d** and **4k**, along with the least active analogue, **4f**, and the well-known CDK2 inhibitor **BMS-265246**, which was used in the in vitro inhibitory kinase assays, as well as the co-crystallized ligand (**DTQ**), was performed. Upon completion of each docking calculation, a maximum of 100 poses per ligand were generated and the final best-docked conformations were ranked using an XP Glide score function and Glide energy.

According to the data presented in Table 5, the cocrystal ligand **DTQ** and **BMS-265246** showed better scoring functions than the synthesized compounds. Both **4d** and **4k** showed high affinity to the target protein, with **4d** demonstrating comparable binding strength (Glide energy = −48.700 Kcal/mol) to the co-crystal ligand (Glide energy = −49.122 Kcal/mol) and **BMS-265246** (Glide energy = −47.340 Kcal/mol), whereas the inactive analog **4f** exhibited the lowest affinity to the enzyme with the highest Glide energy value (−30.726 Kcal/mol).

Lastly, the lowest-energy docked complexes of the co-crystallized ligand **DTQ**, the reference inhibitor **BMS-265246**, **4d**, **4k** and **4f** with the protein were used to identify the crucial interactions for inhibition of CDK2. The ligand–protein interactions (hydrogen bonds, as well as hydrophobic interactions) are shown in Figures 5–14.

Table 5. Binding free energy of **4d**, **4k**, **4f**, **BMS-265246** and **DTQ** at the active site of CDK2.

Ligands	Glide Score XP	Glide Energy Kcal/mol
4k	−6.383	−43.311
4d	−6.224	−48.700
4f	−4.160	−30.726
Co-crystallized ligand (DTQ¹)	−9.649	−49.122
BMS-265246	−7.187	−47.340

¹ **DTQ**: 4-[3-hydroxyanilino]-6,7-dimethoxyquinazoline.

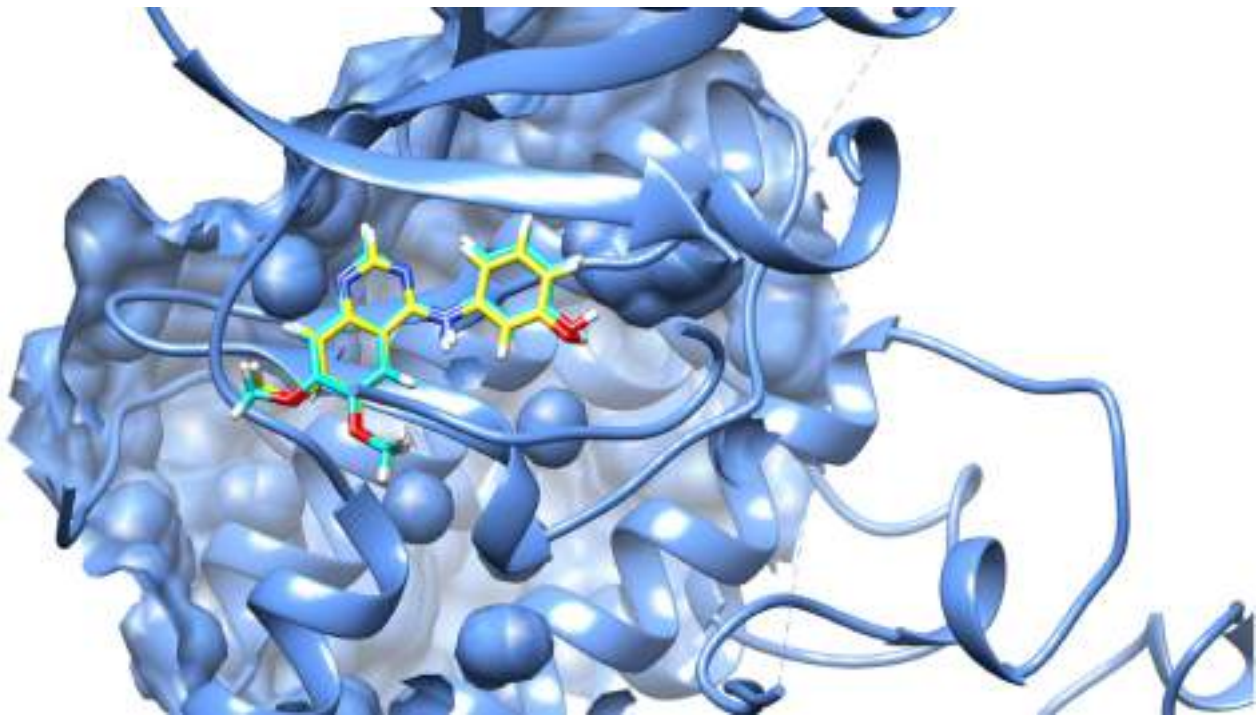


Figure 4. Redocked (yellow) and co-crystallized (baby blue) ligand (DTQ) in the ATP binding pocket of CDK2 after self-docking calculations.

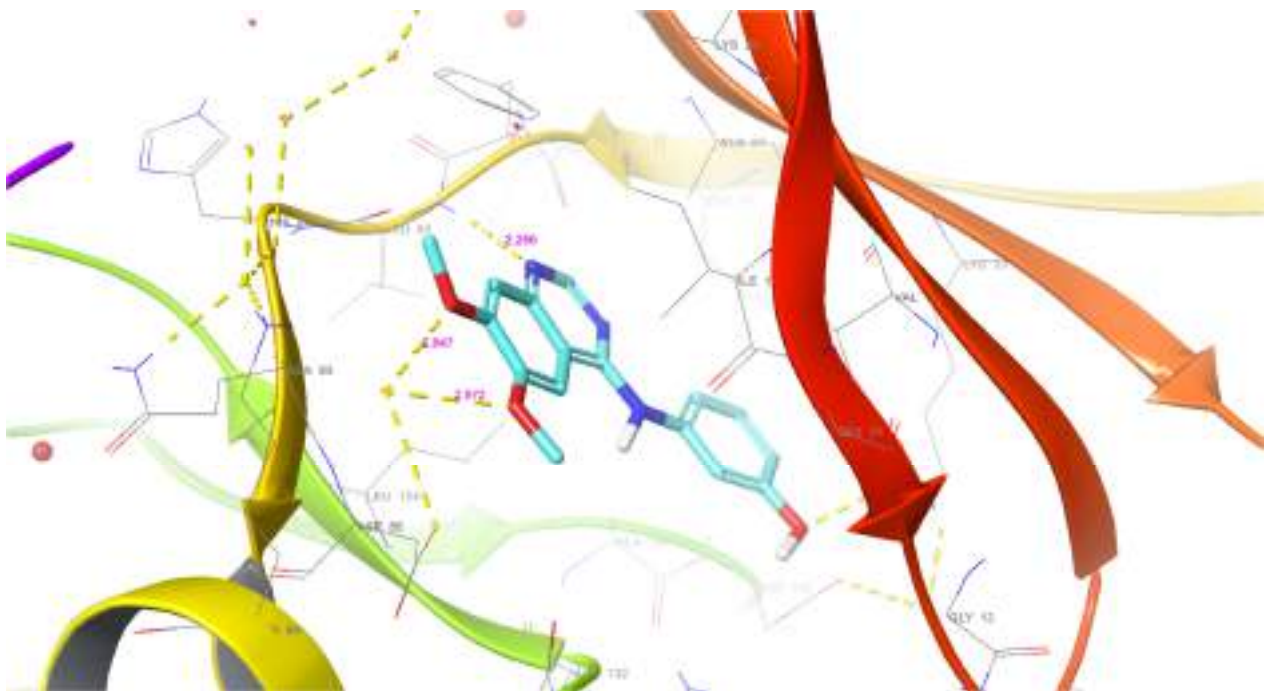


Figure 5. Three-dimensional binding interactions of DTQ within the ATP binding pocket of CDK2. Hydrogen bonds (yellow dotted lines), hydrogen atoms (white), nitrogen atoms (blue), and oxygen atoms (red).

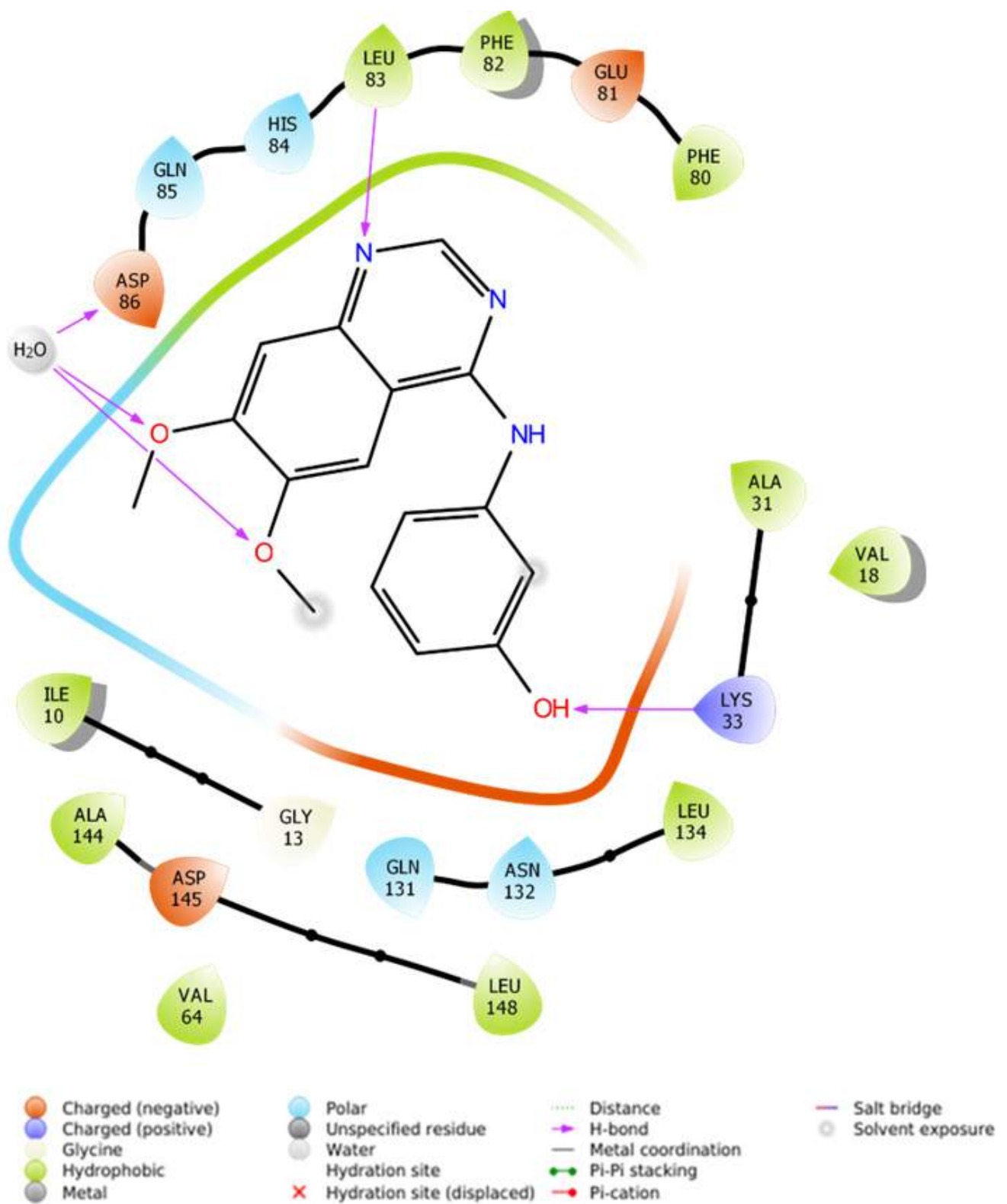


Figure 6. Two-dimensional binding interactions of DTQ within the ATP binding pocket of CDK2.

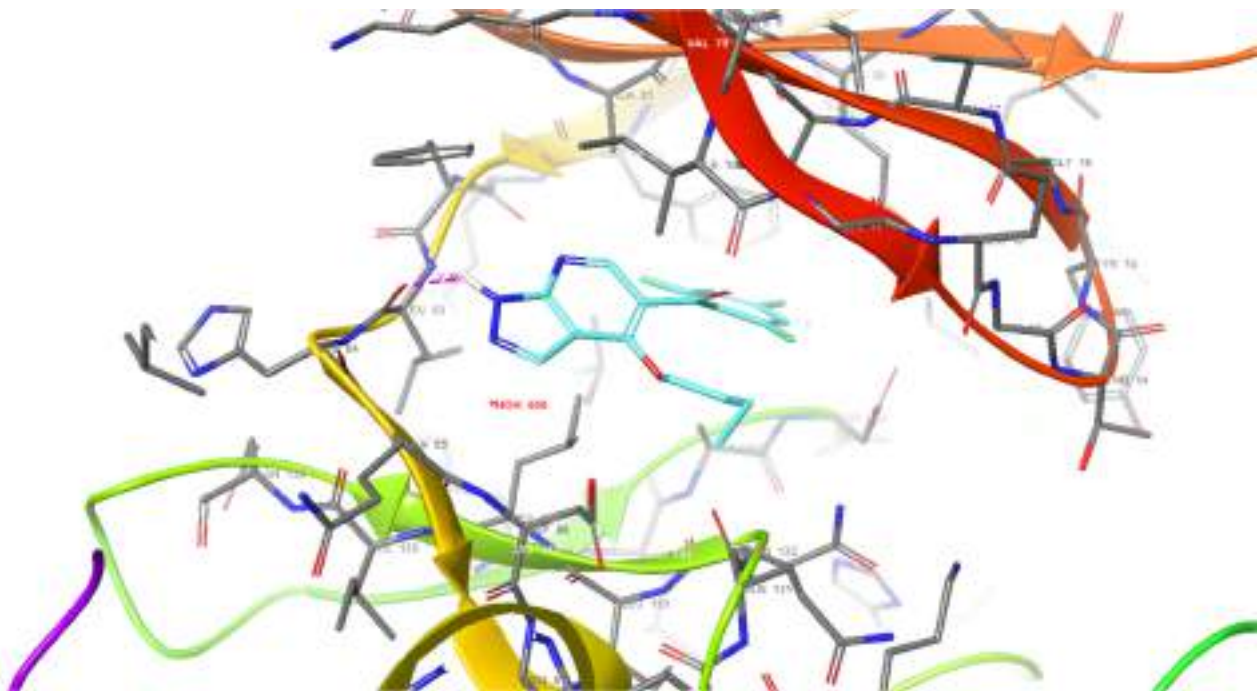


Figure 7. Three-dimensional binding interactions of **BMS-265246** within the ATP binding pocket of CDK2. Hydrogen bond (purple dotted line), hydrogen atoms (white), nitrogen atoms (blue), fluorine atoms (light green), and oxygen atoms (red).

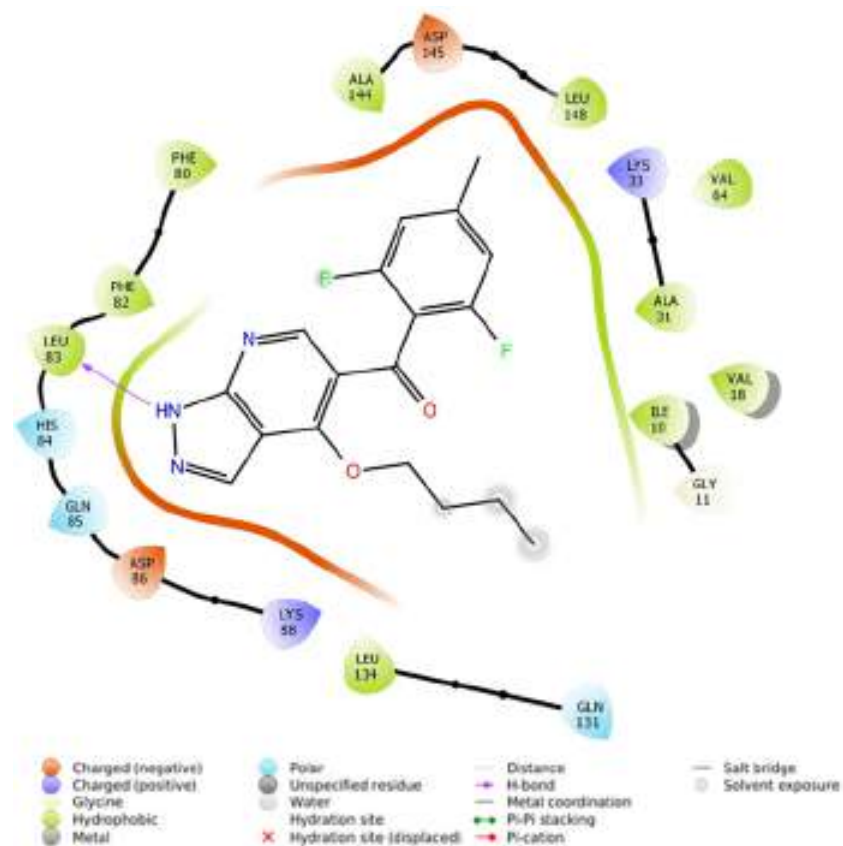


Figure 8. Two-dimensional interactions of **BMS-265246** within the ATP binding pocket of CDK2 kinase.

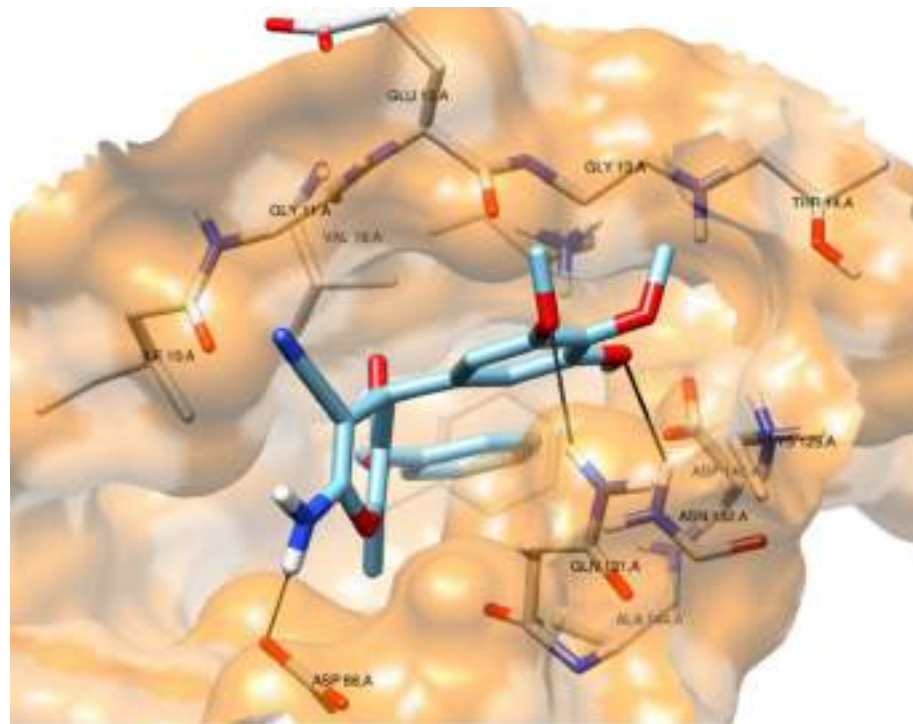


Figure 9. Three-dimensional model of binding interactions of compound **4d** after docking calculations in the ATP binding pocket of CDK2. Hydrogen bond (black lines), hydrogen atoms (white), nitrogen atoms (blue), and oxygen atoms (red).

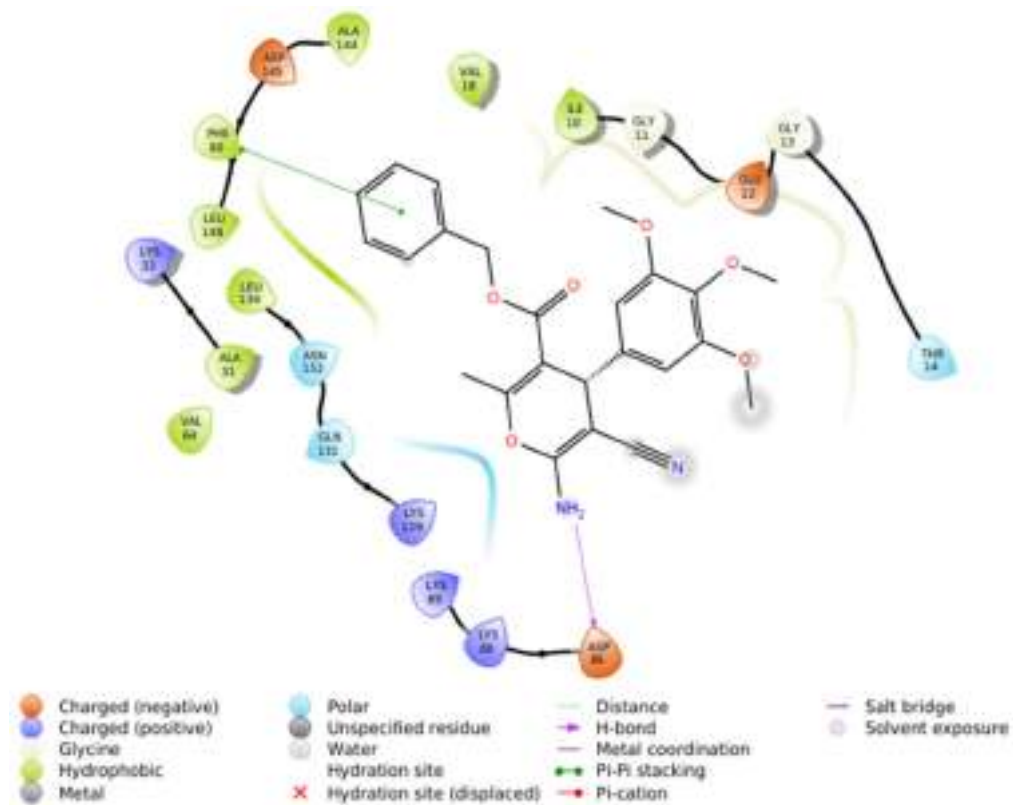


Figure 10. Two-dimensional model of binding interactions of compound **4d** after docking calculations in the ATP binding pocket of CDK2.

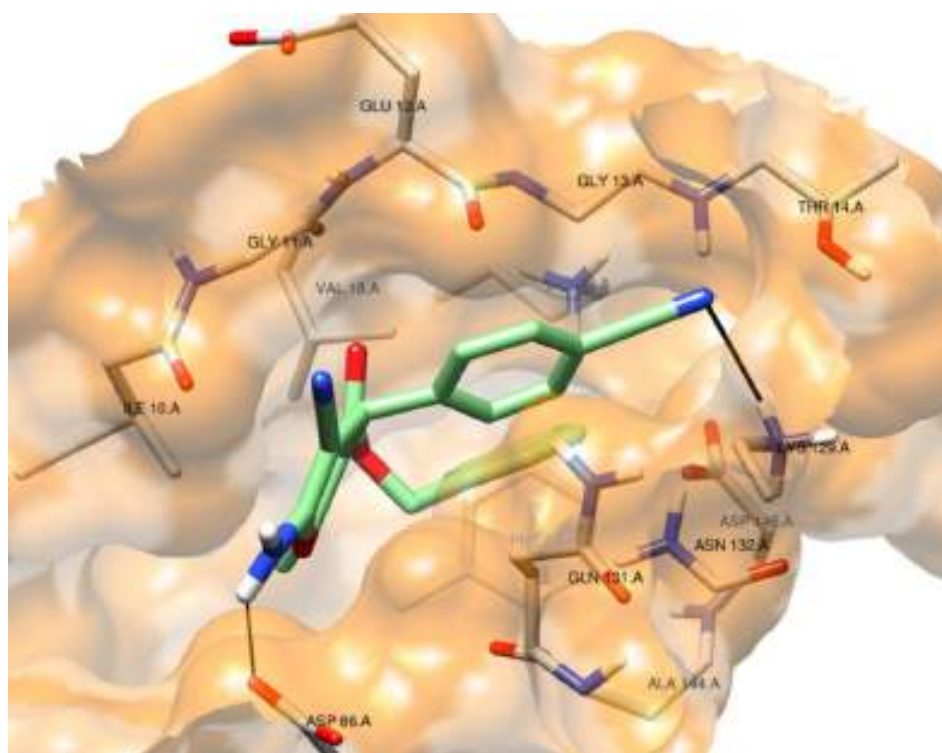


Figure 11. Three-dimensional model of binding interactions of compound **4K** after docking calculations in the ATP binding pocket of CDK2. Hydrogen bond (black lines), hydrogen atoms (white), nitrogen atoms (blue), and oxygen atoms (red).

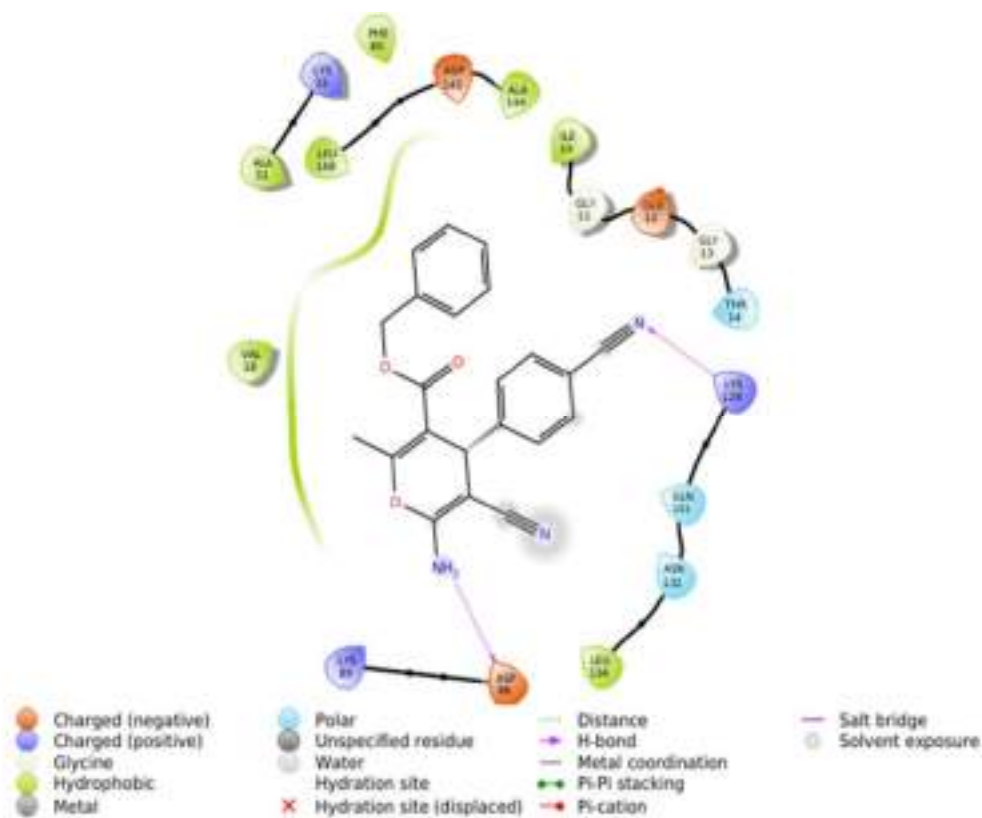


Figure 12. Two-dimensional model of binding interactions of compound **4K** after docking calculations in the ATP binding pocket of CDK2.

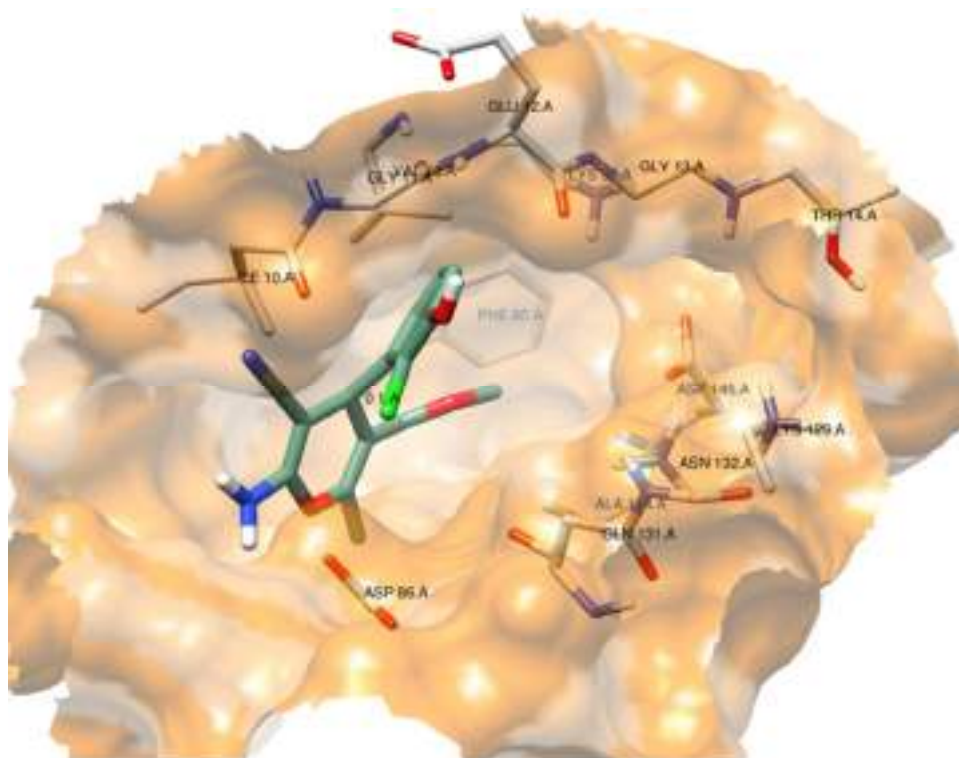


Figure 13. Three-dimensional model of binding interactions of compound **4f** after docking calculations in the ATP binding pocket of CDK2. Hydrogen atoms (white), nitrogen atoms (blue), chlorine atom (green), and oxygen atoms (red).

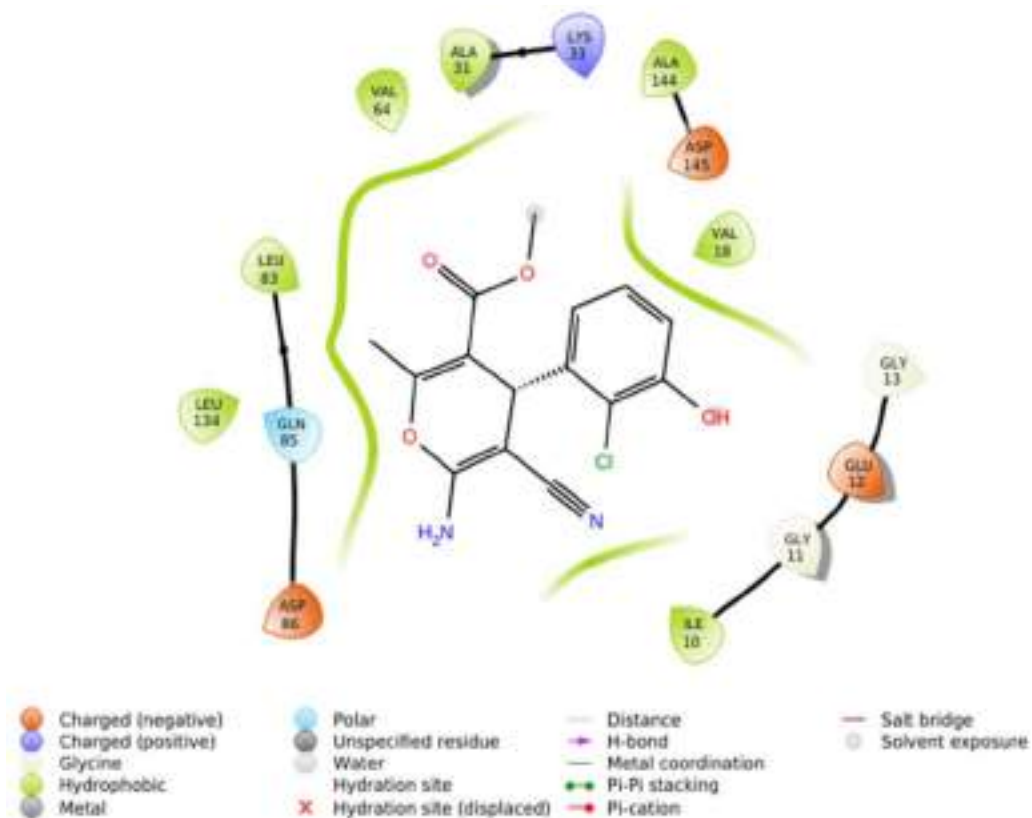


Figure 14. Two-dimensional model of binding interactions of compound **4f** after docking calculations in the ATP binding pocket of CDK2.

Docking pose for DTQ:

Through examination of molecular docking, the pose of **DTQ** is shown to be very interesting, as it involves an active site water molecule as a bridge for H-bond establishment [66]. In addition to this, the phenolic OH and N-1 of the pyrimidine ring of **DTQ** are responsible for H-bond formation with Lys33 and Leu83, respectively (Figures 5 and 6).

It is noteworthy that the H-bonding interactions with Lys33 and Leu83 on the protein backbone are important for potent inhibitory activity [67].

Docking pose for BMS-265246:

The docking pose for **BMS-265246** indicates that it coordinated different residues of the enzyme due to hydrophobic and polar interactions. The Leu83 residue and the NH of the pyrazole ring of **BMS-265246** are responsible for a strong H-bond formation with a distance of 2.15 Å. The N-2 atom of the same pyrazole ring interacted with Asp86 due to mild polar interactions. The 2,6-difluoro-4-methyl phenyl moiety established hydrophobic interactions with the nearby residues, viz. Ala144, Asp145 and Leu148 (Figures 7 and 8). From these docking results, the obtained pose of **BMS-265246** established similar important molecular interactions, in particular, H-bonding with Leu83, as documented before [68].

Docking pose for 4d:

This pyran derivative formed three hydrogen bonds—with amino acid residue ASP86 through its amino group as a donor, with amino acid GLN131 through the 3-methoxyl group as an acceptor and the last with ASN132 as an acceptor through the 5-methoxyl group. These binding interactions confirmed the importance of the methoxy groups at the *m*-positions (3,5-positions) as well as the amino group for further optimization. Additionally, this ligand showed lipophilic interactions with ILE10, VAL18, ILE10, PHE80 and LEU148. Moreover, the phenyl ring was involved in pi–pi stacking interactions with PHE80 (Figures 9 and 10)

Docking pose for 4k:

Likewise, the analysis of the binding interactions between CDK2 and compound **4k** (Figures 11 and 12) revealed the formation of two hydrogen bonds between the amino acid residues ASP86 and LYS129 in the active pocket and amino (as a donor) and nitrile (as an acceptor) substituents of **4k**. Additionally, the phenyl ring of the benzyl moiety showed lipophilic interactions with ILE10, VAL18, ILE10, PHE80, LEU148, LEU134 and VAL64.

Docking pose for 4f:

Contrarily, the inactive derivative **4f** did not form any hydrogen bonds inside the active site of CDK2. However, it showed lipophilic interactions with ILE10, VAL18, LEU83 and Val64, as shown in Figures 13 and 14.

Overall, it can be concluded that there is a positive correlation between the docking simulations and the results of the cell viability assays, which support our claim that CDK2 is a potential target which could be responsible for the observed antiproliferative activities of compounds **4d** and **4k**.

Moreover, the docking analyses of **4d**, **4k** and **4f** successfully provided information on possible structural modifications to improve kinase inhibitory efficiency. These analyses highlighted that an inhibitor for CDK2 requires a bulkier ring to be attached to the main scaffold (the pyran ring in the present case) through a flexible chain of two to three atoms. Furthermore, it should be present at the periphery of the molecule to penetrate well inside the ATP binding site to establish lipophilic and mild polar interactions [69]. In addition, docking analyses indicated that H-bond donors or acceptors on the central scaffold could substantially enhance binding with the receptor.

2.2.4.2. In Vitro CDK2 Inhibitory Assays

To further substantiate the promising results of the docking analyses of pyrans **4d** and **4k** in the ATP active site of CDK2, their in vitro potencies to inhibit the enzyme activity were evaluated over a concentration range of 0.01–10 µM. The results of kinase assays (Figure 15) revealed that both pyrans strongly affected CDK2 activity in a dose-dependent manner. Thus, compound **4d**, at 0.01, 0.1, 1, or 10 µM, decreased CDK2 activity by 26.9, 41.5,

72.5 and 89%, respectively, while compound **4k** inhibited kinase efficiency by 21.9, 42.4, 66.3 and 83.6%, respectively, at the same concentrations. On the other hand, the positive control **BMS-265246** was found to be more effective and reduced CDK2 activity by 92.3% at the highest concentration used (10 μ M).

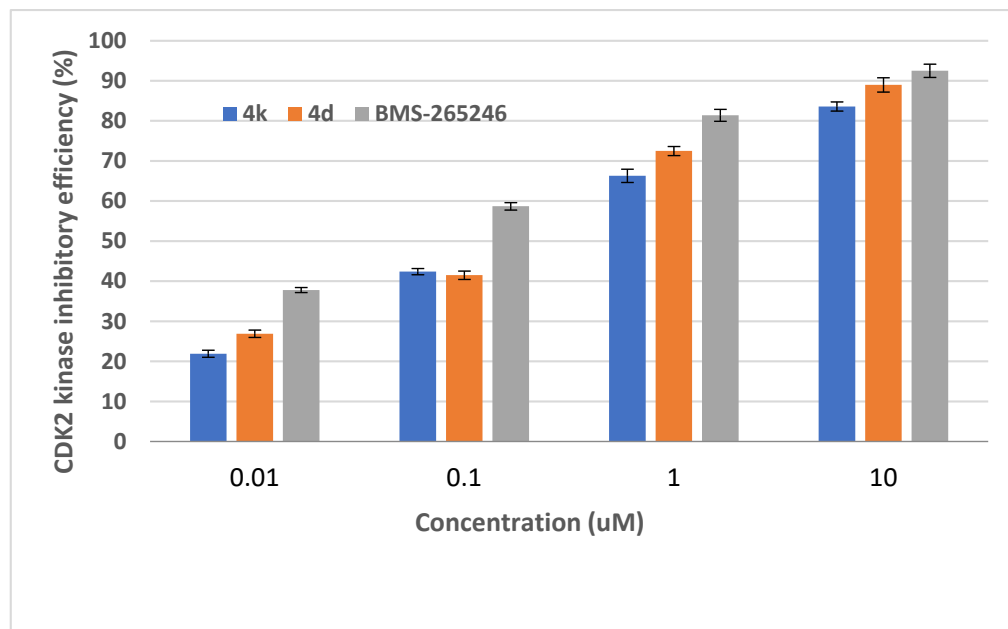


Figure 15. In vitro evaluation of the CDK2 inhibitory efficiency of pyrans **4d** and **4K** as compared to the reference inhibitor **BMS-265246** over a concentration range of 0.01–10 μ M. Results were obtained from three independent experiments.

Moreover, the IC_{50} values for **4k**, **4d** and **BMS-265246** were deduced (Figure 16) and they were found to be 0.214, 0.143 and 0.036 μ M, respectively.

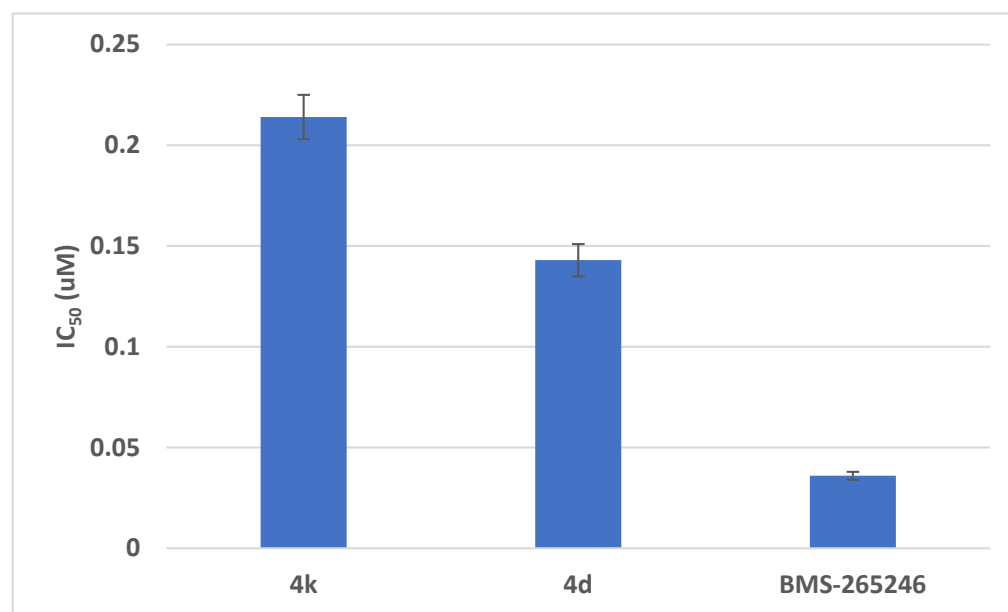


Figure 16. IC_{50} values of **4d**, **4K** and **BMS-265246** against CDK2.

Thus, our results for the kinase inhibitory assays are in agreement with the previously reported studies, which indicated that tropane–pyran hybrid structures could be considered as promising core scaffolds for developing new anticancer agents acting as CDK2 inhibitors [37].

It is worth noting that inhibition of CDK2 activity is considered to be a good approach for preventing chemotherapy-induced alopecia (CIA) by arresting the cell cycle without sensitization of the epithelium [70].

Collectively, the results of the docking simulations, the cell viability inhibition assays in HCT-116 colon cancer cells and the in vitro CDK2 inhibitory assays confirmed the potential of pyrans **4k** and **4d** as antiproliferative agents via the induction of cell-cycle arrest through CDK-2 inhibition.

2.2.4.3. Further Mechanistic Studies via Quantitative Determination of the Concentration of CDK2 Protein and the Expression Profile of the CDK2 Gene in HCT-116 Cells Treated with the Cytotoxic Candidates

During the pathogenesis of cancer in humans, the enzymatic machinery, including human cyclins and their kinase (CDK) partners, which control the decisions to progress from a resting state (G0) to the cell cycle (G0-to-G1 transition) and/or to progress from the G1 (Gap-1) phase (in which the cell prepares for DNA synthesis) to the S phase (DNA synthesis phase), was found to be dysregulated, resulting in unrestrained proliferation. Furthermore, mutations have been observed in genes encoding CDKs; thus, the growth promoter and activator of G1 phase progression (the CDK2 gene) has been reported to be overexpressed in some human cancer subtypes, including colorectal, leukemia, pancreatic and renal carcinomas [24,71]. Moreover, overexpression of CDK2 in primary CRC tumors is linked to lymph node metastasis [72]. Therefore, different strategies for therapeutic intervention have been proposed based on inhibiting CDK activity and silencing the co-expressed genes in colon cancer cells to prevent them from proliferating and to induce their death [73].

Based on these facts, the effects of **4d** and **4k** on the expression levels of the CDK2 protein as well as the CDK2 gene were studied in a two-step protocol.

First, the commercial Human CDK2 ELISA Kit was used to estimate concentrations of the protein in the lysate of HCT-116 cells treated with **4d** and **4k**. The results were compared with the positive (cells treated with the reference CDK inhibitor, **BMS-265246**) and negative control experiments. The data presented in Figure 17 show that the concentration of CDK2 was estimated to be 7.466 ng/mL in the negative control experiment, while it was reduced to 2.229, 3.932 and 4.634 ng/mL in the positive control, **4k**- and **4d**-treated samples, respectively. These results implied that compounds **4d** and **4k** downregulated the expression of the CDK2 protein 1.9- and 1.6-fold, respectively, as compared to the negative control.

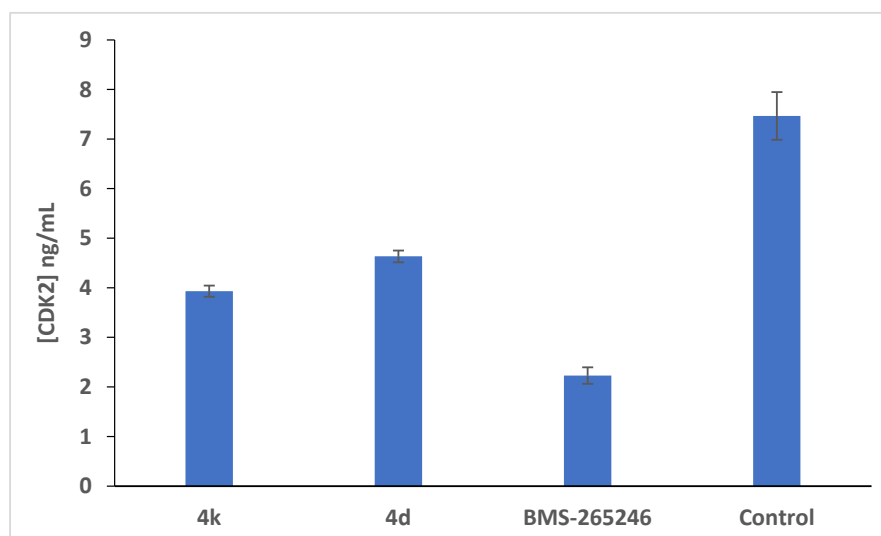


Figure 17. In vitro quantitative determination of the concentrations of CDK2 (ng/mL) in HCT116 cells treated with pyrans **4d** and **4k** compared with the positive control **BMS-265246** and the negative control samples.

Second, the expression profiles of the CDK2 gene in HCT-116 cells treated with compounds **4d** and **4k** were determined. The results were compared with the positive control (cells treated with the well-known cytotoxic drug 5-fluorouracil) and the negative control (untreated cells).

As shown in Figure 18, the CDK2 transcript levels were decreased significantly, 0.29-, 0.45- and 0.50-fold, by 5-FU, **4K** and **4d**, respectively, as compared to the untreated cells after 24 h of the treatment.

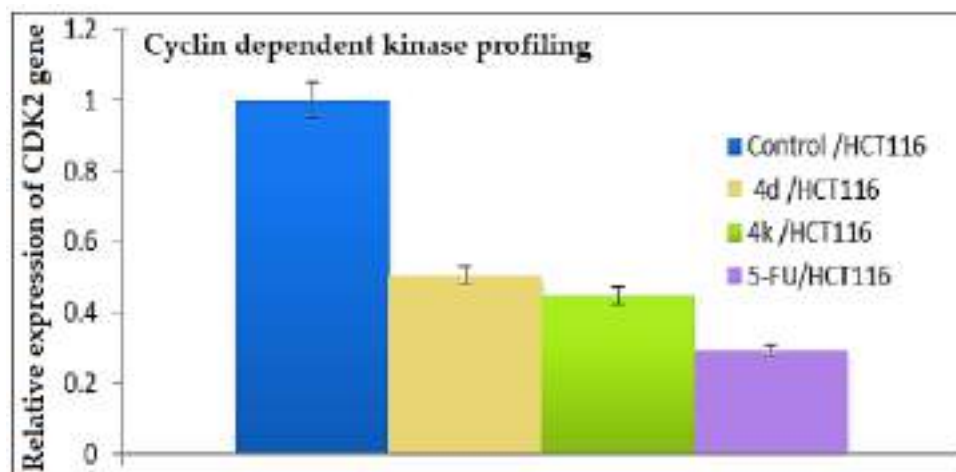


Figure 18. The expression profiles of the CDK2 gene in the lysate of HCT-116 cancer cells treated with compounds **4d** and **4k**.

Accordingly, our results imply that pyrans **4K** and **4d** might cause growth arrest in HCT-116 cancer cells via downregulation of CDK2 gene and protein as well.

2.2.5. Investigation of the Apoptotic Potential via Real-Time PCR Determination of the Expression Profile of the Caspase-3 Gene in HCT-116 Cells Treated with Pyrans **4d** and **4k**

Normally, apoptosis, or programmed cell death, is considered as a safeguard mechanism for managing stress and maintaining tissue homeostasis. It proceeds through a series of well-ordered biochemical cascades of events, which are regulated by a network of proteins. However, due to genetic and epigenetic mutations, deregulation of these proteins occurs, which results in evasion of apoptosis. This oncogenic transformation is a hallmark of human cancer, and it contributes to other tumorigenic events, including uncontrolled growth, accumulation of further genetic mutations, tumor angiogenesis and metastasis and chemotherapeutic resistance [74]. Therefore, reinitiating selective cancer cell death is a fundamental goal in anticancer drug development. Research in cancer biology has provided fundamental insights into the apoptotic pathways and their deregulation in colon cancer. Moreover, it has identified the molecular targets for reinitiating selective tumor cell death, including antiapoptotic Bcl-2 family proteins, apoptosis inhibitors and activators of the TRAIL death receptor signaling pathway, as well as other apoptotic markers, such as caspases. Among the caspases, which play critical roles in apoptosis, are the initiators, including caspases-8 or 9, and the effectors, including caspase-3 and caspase-7. After being activated by the initiators, the executioner caspases cleave numerous vital structural and regulatory proteins, leading to cell death [75]. Previous studies have demonstrated that expression level of caspase-3 is downregulated in cancer; therefore, enhancing its activity with natural and synthetic compounds was suggested as a possible strategy for cancer therapy [76,77]. In fact, pyran-containing, 4*H*-chromene derivatives, such as 3-amino-1-(4-fluorophenyl)-1*H*-benzo[*f*]chromene-2-carbonitrile [78] and 3-amino-1-(4-iodophenyl)-8-methoxy-1*H*-benzo[*f*]chromene-2-carbonitrile [79], were found to be effective against HT-29 cells of CRC and -MDA-MB-231 cells of breast cancer, respectively, via a 3-caspase-dependent apoptosis mechanism; therefore, we anticipated that our closely related pyran analogues **4d** and **4k** might be caspase-3 inducers as well. Thus, the real-time

quantitative PCR analyses of caspase-3-gene levels following treatment with compounds **4d** and **4k** at a concentration of 10 mg/mL for 24 h, were performed to investigate their potential as proapoptotic agents.

As shown in Figure 19, the treated HCT-116 cultures showed significant expression of caspase-3 gene as compared to the positive (5-FU-treated) and the negative (pure HCT-116 cells) control cultures. The transcription levels of the caspase-3 gene in HCT-116 in treated cultures were increased 7.31-, 4.62- and 2.55-fold by 5-FU, **4k** and **4d**, respectively, as compared to the negative control. These results indicated a concentration of 10 mg/mL for increasing the expression of caspase-3 gene and subsequent induction of mitochondrial apoptosis of the HCT-116 cells.

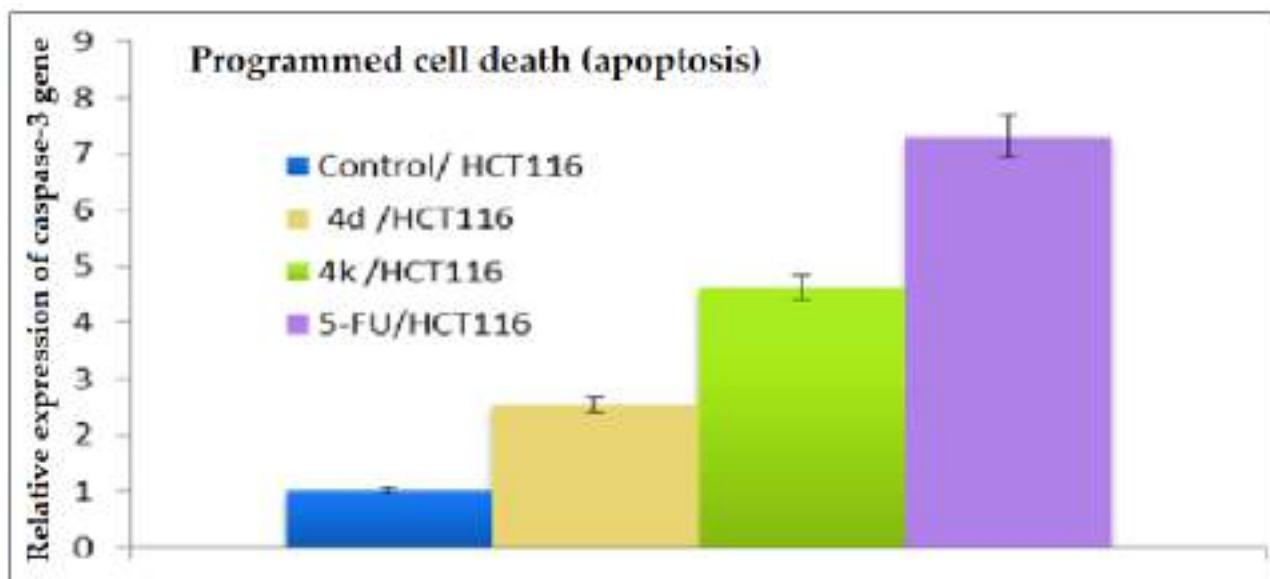


Figure 19. The expression profiles of caspase-3 gene in lysates of HCT-116 cells treated with 5-FU and pyrans **4d** and **4k** as compared to the untreated HCT-116 cells (negative control).

Overall, pyrans **4d** and **4k** can be considered interesting candidates to be subjected to further studies in order to investigate their usefulness in developing new anti-CRC agents capable of inhibiting the proliferation of HCT-116 cells through inducing cell-cycle arrest at the G1/S boundary by targeting CDK2 and enhancing the apoptosis of cancerous cells via activation of the caspase-3 gene.

2.3. Physicochemical and Pharmacokinetic Properties and Lipinski's Rule of Five

Six physicochemical parameters, comprising Lipophilicity (LIPO), Size, Polarity (POLAR), Insolubility (INSOL), Unsaturation (UNSAT) and Flexibility (FLEX) were predicted for the tested compounds using the bioavailability radar chart [80], in which the pink-colored zone (Figure 20) indicates the suitability of physicochemical properties to have good in vivo bioavailability.

The results indicated that, compounds **4g** and **4d** did not violate any parameters in the radar chart, while **4j** and **4k** derivatives violated only the UNSAT parameter, which is acceptable. The calculated log *p* values were intermediate (2.27 for **4g**, 2.99 for **4j**, 2.98 for **4d** and 2.80 for **4k**). Based on these data, it can be predicted that all the tested derivatives would have good bioavailability profiles, especially the molecules **4g** and **4d**.

The SwissADME server provides a BOILED-EGG chart to indicate human intestinal absorption (white part), blood–brain barrier penetration (yellow part) and the probability of the tested compound acting as a substrate (PGP⁺, blue color) or not a substrate (PGP⁻, red color) for permeability glycoprotein (PGP).

The results (summarized in Figure 21) showed that all the tested derivatives can be expected to be of good oral absorption, which is indicated by their presence in the white

area. Moreover, they would not penetrate the blood–brain barrier (BBB, egg yolk area); thus, they would not cause CNS toxicity. Finally, the four compounds are red-colored, which means that they are not anticipated to be substrates for PGP; therefore, they would not suffer from restricted entry to the target cells through efflux pumps [81].

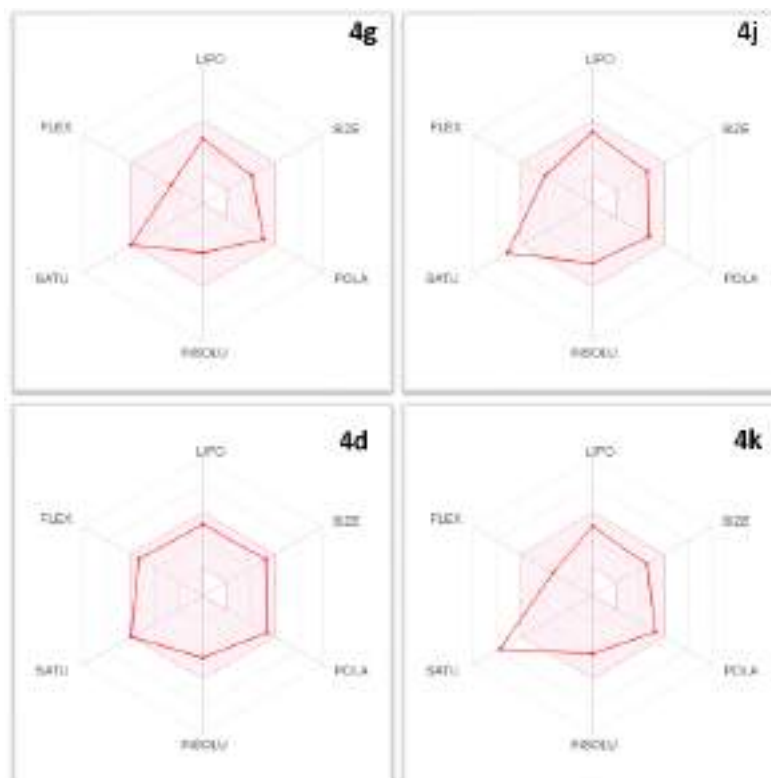


Figure 20. The bioavailability radar plots for the tested compounds (the colored zone is the suitable physicochemical space for oral bioavailability).

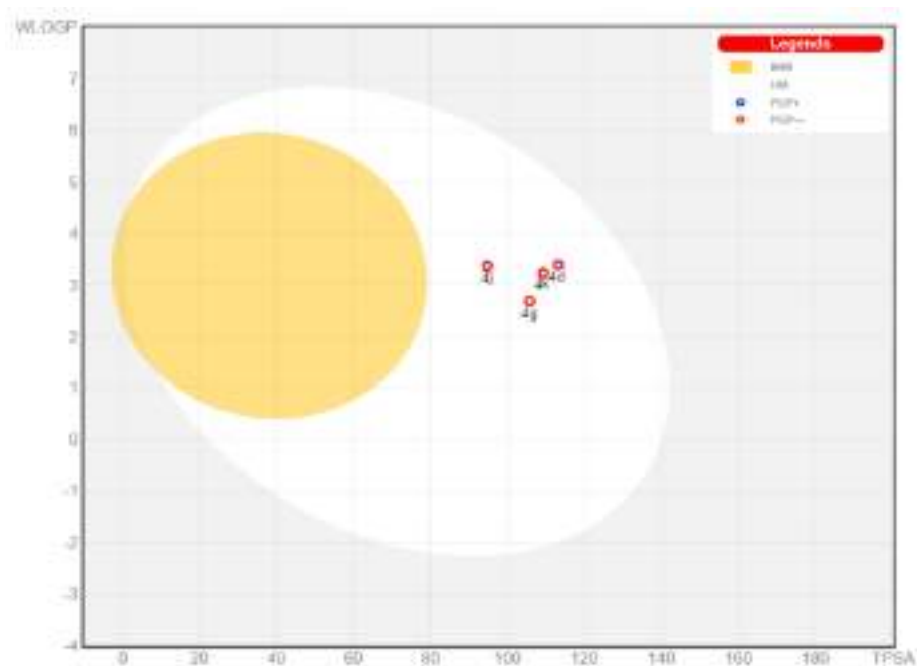


Figure 21. The BOILED-EGG chart for the studied compounds.

Eventually, the drug-likeness and good oral bioavailability characteristics for the studied compounds were predicted using the Lipinski's rule of five filter [82]. As shown in Table 6 no compound did violate any parameter of the rule, which states that no violation or even one violation indicates expected good oral bioavailability.

Table 6. Drug-likeness of the tested compounds.

Comp. No.	Properties	Comment
4g	<ul style="list-style-type: none"> • Log P = 2.27 (<5) • Molecular weight = 334.75 g/mol (<500) • No. of H-bond donor groups (OHs + NHs) = 3 (≤ 5) • No. of H-bond acceptor atoms (Os + Ns) = 6 (≤ 10) 	No violation
4j	<ul style="list-style-type: none"> • Log P = 2.99 (<5) • Molecular weight = 376.41 g/mol (<500) • No. of H-bond donor groups (OHs + NHs) = 2 (≤ 5) • No. of H-bond acceptor atoms (Os + Ns) = 6 (≤ 10) 	No violation
4d	<ul style="list-style-type: none"> • Log P = 2.99 (<5) • Molecular weight = 436.46 g/mol (<500) • No. of H-bond donor groups (OHs + NHs) = 2 (≤ 5) • No. of H-bond acceptor atoms (Os + Ns) = 8 (≤ 10) 	No violation
4k	<ul style="list-style-type: none"> • Log P = 2.80 (<5) • Molecular weight = 371.39 g/mol (<500) • No. of H-bond donor groups (OHs + NHs) = 2 (≤ 5) • No. of H-bond acceptor atoms (Os + Ns) = 6 (≤ 10) 	No violation

Although the in-silico predictions are promising, these results are preliminary and should be confirmed experimentally.

3. Materials and Methods

3.1. Chemistry

3.1.1. General

All reagents were used as supplied commercially. The IR spectra were obtained using potassium bromide (KBr) discs and recorded on a PerkinElmer FTIR spectrophotometer, Spectrum BX 1000 in wave number (cm^{-1}). Nuclear magnetic resonance spectroscopy was performed on a Bruker Avance 500 spectrometer, operating at 500 MHz for ^1H and 125 MHz for ^{13}C at 25 °C (Research Unit, College of Pharmacy, Prince Sattam Bin Abdulaziz University, Al-Kharj, KSA), or on an Eclipse 300 FT NMR spectrometer operating at 300 MHz spectrometer for ^1H and at 75 MHz for ^{13}C at 25 °C (College of Science, King Saud University, Riyadh, KSA). The chemical shifts are expressed in ppm and tetramethyl silane (TMS) was used as internal standard; mass spectra were recorded on a Shimadzu Qp-2010 Plus Mass Spectrometer using Ionization Mode: EI (Micro Analytical Center, Cairo University, Egypt). Percentages of C, H and N in the new compounds were evaluated at the regional center for Mycology and Biotechnology (RCMB), Al-Azhar University.

3.1.2. General Procedures for the Synthesis of 6-Amino-5-cyano-4-(aryl)-2-methyl-4H-pyran-3-carboxylic Acid Esters **4a–m**

To an ethanolic mixture (20 mL) of benzaldehyde derivative **1a–k** (0.004 mol), namely, 2-formyl-benzoic acid, 3-methoxy-2-nitro benzaldehyde, 3,4,5-trimethoxy benzaldehyde, biphenyl-3-carbaldehyde, 2-chloro-3-hydroxy-benzaldehyde, 2,4,6-trimethoxy benzaldehyde, 4-fluorobenzaldehyde, 4-methoxybenzaldehyde, 4-formyl-benzonitrile, 2-chloro-benzaldehyde, or 2,4-dichlorobenzaldehyde, and malononitrile **2** (0.004 mol, 0.264 g), piperidine (0.8 mL) and the appropriate β -ketoester derivative **3a–c** (0.004 mol) namely;

methyl acetoacetate, ethyl acetoacetate or benzyl acetoacetate were added. The resulting reaction mixture in each was stirred at room temperature for 24 h, then it was worked up as described below to obtain the desired product.

6-Amino-4-(2-carboxy-phenyl)-5-cyano-2-methyl-4H-pyran-3-carboxylic acid methyl ester **4a**

Filtration of the obtained reaction suspension afforded a yellow solid, which was washed with H₂O/dil. HCl, air dried and recrystallized from EtOH to give shiny white crystals, yield (63%), m.p. 216–218 °C; ν_{\max} (KBr)/cm⁻¹ 3328 (broad OH and NH₂), 3193 (CH-Ar), 2895 (CH-aliphatic), 2656, 2544, 2196 (CN), 1716 (C=O-ester), 1675 (C=O-carboxylic), 1605 (C=C), 1445 1405, 1377, 1325, 1264, 1226, 1182, 1124, 1068, 1042, 955, 876, 842, 803, 749, 728, 648, 624, 518, 442; δ_{H} (500 MHz; DMSO-*d*₆) 7.73 (1H, dd, *J* = 7.7, 1.1 Hz, CH-Ar), 7.50 (1H, t, *J* = 7.3 Hz, CH-Ar), 7.29–7.25 (2H, m, 2 × CH-Ar), 6.85 (2H, br. s, NH₂), 5.63 (1H, s, C⁴H-pyran), 3.46 (3H, s, O-CH₃), 2.31 (3H, s, CH₃); δ_{C} (125 MHz; DMSO-*d*₆), 168.60 (C=O), 165.83 (C=O), 158.73, 157.04, 146.03, 131.95, 130.48, 129.36, 129.01, 126.39, 119.27, 107.59 (4 × CH-Ar, 2 × C_q-Ar, 3 × C_q-pyran, CN), 56.05 and 51.29 (C⁵-pyran and O-CH₃), 32.72 (C⁴H-pyran), 18.40 (CH₃); MS (EI) *m/z* (%) [M⁺] 314.85 (0.24) for C₁₆H₁₄N₂O₅, 268.90 (39.95), 252.85 (14.51), 236.90 (100.00), 208.90 (11.83), 192.90 (10.28), 133.00 (31.62), 116.05 (21.80), 105.00 (41.68), 89.00 (27.86), 77.05 (77.39), 66.00 (29.21), 57.05 (15.98). Anal. Calcd. for C₁₆H₁₄N₂O₅: C, 61.14; H, 4.49; N, 8.91. Found; C, 61.36; H, 4.70; N, 9.07.

6-Amino-5-cyano-4-(3-methoxy-2-nitro-phenyl)-2-methyl-4H-pyran-3-carboxylic acid ethyl ester **4b**

The excess solvent was evaporated and the resulting yellow solid was washed with water, air dried and recrystallized from EtOH to afford a pure product as a pale yellow solid, yield (85%), m.p. 220–222 °C; ν_{\max} (KBr)/cm⁻¹ 3475 (asymmetric NH₂), 3351 (symmetric NH₂), 3168 and 3098 (CH-Ar), 2989, 2940, 2898 and 2847 (CH-aliphatic), 2192 (CN), 1679 (C=O), 1629 and 1586 (C=C), 1530 (NO₂), 1478, 1449, 1372, 1332, 1277, 1210, 1122, 1056, 1018, 956, 908, 851, 802, 655, 613, 562, 512, 443; δ_{H} (500 MHz; DMSO-*d*₆) 7.51 (1H, t, *J* = 8.2 Hz, CH-Ar), 7.18 (1H, d, *J* = 8.4 Hz, CH-Ar), 7.06 (2H, br. s, NH₂), 6.88 (1H, d, *J* = 7.9 Hz, CH-Ar), 4.35 (1H, s, C⁴H-pyran), 3.96 (2H, q, *J* = 7.1 Hz, O-CH₂-CH₃), 3.87 (3H, s, O-CH₃), 2.32 (3H, s, CH₃), 0.99 (3H, t, *J* = 7.1 Hz, O-CH₂-CH₃); δ_{C} (125 MHz; DMSO-*d*₆), 164.76 (C=O), 158.89, 158.08, 149.97, 139.63, 137.79, 131.78, 120.83, 118.79, 111.67, 105.54 (3 × CH-Ar, 3 × C_q-Ar, 3 × C_q-pyran, CN), 60.21 (O-CH₂), 56.55 and 55.80 (C⁵-pyran and O-CH₃), 34.43 (C⁴H-pyran), 18.22 (2 × CH₃); MS (EI) *m/z* (%) [M⁺] 359.80 (0.40) for C₁₇H₁₇N₃O₆, 341.85 (35.07), 324.85 (5.56), 311.90 (6.19), 296.85 (4.09), 283.85 (8.98), 272.90 (12.99), 297.90 (43.20), 251.90 (11.35), 237.90 (10.52), 226.90 (34.93), 206.90 (11.00), 178.90 (12.93), 160.95 (11.26), 140.05 (6.60), 127.05 (9.20), 115.05 (20.68), 102.00 (23.05), 90.00 (29.84), 76.05 (69.38), 67.00 (100.00), 59.01 (38.59), 51.05 (33.51); Anal. Calcd. for C₁₇H₁₇N₃O₆: C, 56.82; H, 4.77; N, 11.69. Found: C, 57.11; H, 4.83; N, 11.97.

6-Amino-5-cyano-4-(3-methoxy-2-nitro-phenyl)-2-methyl-4H-pyran-3-carboxylic acid benzyl ester **4c**

Addition of ice/H₂O to the reaction mixture and filtration gave the crude product as a yellow precipitate which was air dried and recrystallized from EtOH/petroleum ether to obtain the pure product as a yellow powder, yield (64%), m.p. 190–192 °C; ν_{\max} (KBr)/cm⁻¹ 3434 (asymmetric NH₂), 3321 (symmetric NH₂), 3221; 3190 and 3090, 3031 (CH-Ar), 2942 and 2846 (CH-aliphatic), 2195 (CN), 1716 (C=O), 1673 and 1603 (C=C), 1528 (NO₂), 1479, 1449, 1406, 1370, 1315, 1285, 1225, 1174, 1181, 1133, 1063, 963, 912, 847, 804, 756, 695, 662, 566, 524, 435; δ_{H} (300 MHz; DMSO-*d*₆) 7.55–7.46 (2H, m, 2 × CH-Ar), 7.30–7.17 (3H, m, 3 × CH-Ar), 7.08 (2H, br. s, NH₂), 7.07–7.01 (2H, m, 2 × CH-Ar), 6.87 (1H, d, *J* = 7.5 Hz, CH-Ar), 5.07 and 4.98 (2H, AB q, *J* = 12.6 Hz, O-CH₂-Ph), 4.39 (1H, s, C⁴H-pyran), 3.88 (3H, s, O-CH₃), 2.32 (3H, s, CH₃); δ_{C} (75 MHz; DMSO-*d*₆), 164.11 (C=O), 158.80, 158.06, 150.11, 139.42, 137.64, 135.98, 131.82, 128.22, 127.79, 127.48, 120.89, 118.67, 111.77, 105.49 (8 × CH-Ar, 4 × C_q-Ar, 3 × C_q-pyran, CN), 65.56 (O-CH₂-Ph), 56.64, 55.89 (C⁵-pyran

and O-CH₃), 34.65 (C⁴H-pyran), 18.48 (CH₃); MS (EI) *m/z* (%) [M⁺ + 1] 422.10 (0.25) for C₂₂H₁₉N₃O₆, 404.00 (1.48), 360.10 (1.07), 343.05 (0.02), 309.10 (0.76), 291.10 (1.16), 283.05 (1.27), 268.05 (3.94), 252.05 (1.83), 227.00 (4.78), 201.00 (2.05), 153.05 (0.43), 140.05 (0.78), 107.10 (2.60), 92.05 (8.02), 91.05 (100.00), 77.05 (3.46), 65.00 (12.07), 51.00 (3.09); Anal. Calcd. for C₂₂H₁₉N₃O₆: C, 62.70; H, 4.54; N, 9.97; O, 22.78. Found: C, 62.59; H, 4.68; N, 10.19.

6-Amino-5-cyano-2-methyl-4-(3,4,5-trimethoxy-phenyl)-4H-pyran-3-carboxylic acid benzyl ester **4d**

Filtration of the resulting mixture afforded a yellowish-white powder. Recrystallization from MeOH/petroleum ether afforded shiny white crystals, yield (57%), m.p. 158–160 °C; ν_{\max} (KBr)/cm⁻¹ 3430 (asymmetric NH₂), 3327 (symmetric NH₂), 3181 (CH-Ar), 2940 and 2887 (CH-aliphatic), 2373, 2192 (CN), 1711 (C=O), 1689, 1641 and 1593 (C=C), 1503, 1463, 1427, 1378, 1322, 1252, 1222, 1120, 1059, 995, 888, 856, 802, 759, 719, 693, 662, 578, 515 cm⁻¹; δ_{H} (300 MHz; DMSO-*d*₆) 7.37–7.25 (3H, m, 3 × CH-Ar of benzyl group), 7.18–7.11 (2H, m, 2 × CH-Ar of benzyl group), 6.92 (2H, br. s, NH₂), 6.37 (2H, s, 2 × CH-Ar of trimethoxyphenyl group), 5.10 and 5.02 (2H, ABq, *J* = 12.9 Hz, O-CH₂-Ph), 4.32 (1H, s, C⁴H-pyran), 3.68 (6H, s, 2 × O-CH₃), 3.64 (3H, s, O-CH₃), 2.35 (3H, s, CH₃); δ_{C} (75 MHz; DMSO-*d*₆), 165.45 (C=O), 158.44, 157.36, 152.89, 140.52, 136.47, 135.89, 140.52, 136.47, 135.89, 128.27, 127.88, 127.56, 119.76, 106.73, 104.351 (7 × CH-Ar, 5 × C_q-Ar, 3 × C_q-pyran, CN), 65.77 (O-CH₂-Ph), 59.99, 57.29 and 55.76 (C⁵-pyran and 3 × O-CH₃), 39.30 C⁴H-pyran, 18.30 (CH₃); MS (EI) *m/z* (%) [M⁺ + 1] 437.05 (6.98) for C₂₄H₂₄N₂O₆, [M⁺] 436.05 (24.88), 345.00 (36.17), 327.00 (2.08), 301.05 (3.55), 284.00 (5.49), 269.05 (9.70), 244.00 (3.34), 221.00 (2.36), 168.00 (2.28), 134.05 (1.57), 128.05 (0.90), 109.05 (1.08), 91.05 (100.00), 65.00 (9.29), 51.00 (2.45). Anal. Calcd. for C₂₄H₂₄N₂O₆: C, 66.04; H, 5.54; N, 6.42. Found: C, 66.21; H, 5.59; N, 6.64.

6-Amino-4-biphenyl-3-yl-5-cyano-2-methyl-4H-pyran-3-carboxylic acid ethyl ester **4e**

The resulting mixture was poured onto ice/H₂O and filtered to give a pale-yellow powder, air dried and recrystallized from EtOH/petroleum ether to afford white crystals, yield (73%); m.p. 186–190 °C; ν_{\max} (KBr)/cm⁻¹ 3374 (asymmetric NH₂), 3213 (asymmetric NH₂), 3060 and 3032 (CH-Ar), 2976 and 2922 (CH-aliphatic), 2368, 2340, 2188 (CN), 1708 (C=O), 1676 and 1602 (C=C), 1475, 1419, 1373, 1321, 1262, 1211, 1180, 1132, 1069, 951, 902, 868, 774, 746, 698, 647, 545, 513; δ_{H} (500 MHz; DMSO-*d*₆) 7.62 (2H, d, *J* = 7.5 Hz, 2 × CH-Ar), 7.52 (1H, d, *J* = 7.9 Hz, CH-Ar), 7.48 (2H, t, *J* = 7.55 Hz, 2 × CH-Ar), 7.44–7.41 (2H, m, 2 × CH-Ar), 7.38 (1H, t, *J* = 7.3 Hz, CH-Ar), 7.17 (1H, d, *J* = 7.8 Hz, CH-Ar), 6.98 (2H, br. s, NH₂), 4.43 (1H, s, C⁴H-pyran), 3.98 (2H, q, *J* = 7.1 Hz, O-CH₂-CH₃), 2.35 (3H, s, CH₃), 1.02 (3H, t, *J* = 7.1 Hz, O-CH₂-CH₃); δ_{C} (125 MHz; DMSO-*d*₆), 165.46 (C=O), 158.54, 156.81, 145.60, 140.21, 140.04, 129.19, 128.97, 127.48, 126.58, 126.21, 125.51, 125.28, 119.76, 107.07 (9 × CH-Ar, 3 × C_q-Ar, 3 × C_q-pyran, CN), 60.16 (O-CH₂-Ph), 57.15 (C⁵-pyran), 38.85 (C⁴H-pyran), 18.22 (CH₃), 13.67 (CH₃); MS (EI) *m/z* (%) [M⁺ + 1] 360.85 (9.65) and [M⁺] 359.90 (36.27) for C₂₂H₂₀N₂O₃, 330.85 (26.30), 313.90 (24.30), 293.90 (4.77), 286.90 (15.54), 270.90 (8.75), 228.90 (10.19), 206.90 (100.00), 178.90 (24.65), 160.95 (15.76), 133.05 (6.67), 115.00 (10.63), 105.05 (15.92), 88.00 (16.94), 77.05 (22.94), 67.00 (44.11), 51.00 (14.94); Anal. Calcd. for C₂₂H₂₀N₂O₃: C, 73.32; H, 5.59; N, 7.77. Found: C, 73.59; H, 5.71; N, 7.98.

6-Amino-4-(2-chloro-3-hydroxy-phenyl)-5-cyano-2-methyl-4H-pyran-3-carboxylic acid methyl ester **4f**

The resulting mixture was poured on ice–dil. HCl mixture, then filtered and the obtained solid was air dried to obtain an off-white powder. Recrystallization from EtOH/petroleum ether gave a shiny white powder yield (85%), m.p. 238–240 °C; ν_{\max} (KBr)/cm⁻¹ 3478 (OH), 3319 (NH₂), 3015 (CH-Ar), 2946 and 2843 (CH-aliphatic), 2366, 2200 (CN), 1713 (C=O), 1674, 1634 and 1593 (C=C), 1462, 1403, 1374, 1349, 1261, 1227.7, 1187, 1133, 1074, 976, 853, 801, 759, 726, 667, 547, 500, 443; δ_{H} (500 MHz; DMSO-*d*₆) 10.09 (1H, s, OH), 7.08 (1H, t, *J* = 7.9 Hz, CH-Ar), 6.85 (2H, br. s, NH₂), 6.81 (1H, dd, *J* = 8.1, 1.1 Hz, CH-Ar), 6.60 (1H, dd, *J* = 7.6, 1.1 Hz, CH-Ar), 4.93 (1H, s, C⁴H-pyran), 3.49 (3H, s, O-CH₃), 2.34 (3H, s, CH₃); δ_{C}

(125 MHz; DMSO- d_6), 165.82 (C=O), 158.60, 157.52, 152.94, 143.39, 127.38, 119.63, 119.50, 119.30, 119.23, 114.55, 106.22 (3 × CH-Ar, 3 × C_q-Ar, 3 × C_q-pyran, CN), 56.45 and 51.41 (C⁵-pyran and O-CH₃), 35.28 (C⁴H-pyran), 18.12 (CH₃); MS (EI) m/z (%) [$M^+ + 1$, ³⁷Cl] 322.80 (0.92); [M^+ , ³⁷Cl] 321.80 (4.70); [$M^+ + 1$, ³⁵Cl] 320.80 (3.09) and [M^+ , ³⁵Cl] 319.80 (13.67) for C₁₅H₁₃ClN₂O₄, 306.80 (1.28), 304.80 (3.81), 285.85 (10.38), 284.90 (58.22), 272.85 (1.03), 270.85 (0.36), 252.85 (5.53), 246.85 (1.34), 244.85 (3.08), 224.90 (4.03), 222.90 (1.39), 193.90 (10.96), 192.90 (100.00), 175.90 (9.94), 160.95 (28.16), 127.05 (6.44), 114.00 (10.61), 105.05 (20.05), 99.00 (12.74), 89.00 (23.42), 77.05 (19.61), 67.00 (54.84), 59.05 (37.40), 57.05 (11.01), 51.05 (10.88). Anal. Calcd. for C₁₅H₁₃ClN₂O₄: C, 56.17; H, 4.09; N, 8.73. Found: C, 56.44; H, 4.32; N, 8.99.

6-Amino-4-(2-chloro-3-hydroxy-phenyl)-5-cyano-2-methyl-4H-pyran-3-carboxylic acid ethyl ester **4g**

The resulting reaction mixture was poured onto ice-dil. HCl mixture, then filtered and the obtained solid was air dried to afford an off-white powder. Recrystallization from EtOH/petroleum ether gave a shiny off white powder (65%), m.p. 212–213 °C; ν_{\max} (KBr)/cm⁻¹ 3454 (OH), 3350 (NH₂), 3197 (CH-Ar), 2974 and 2853 (CH-aliphatic), 2642, 2574, 2372, 2203 (CN), 1713 (C=O), 1669 and 1606 (C=C), 1583, 1465, 1408, 1375, 1322, 1293, 1234, 1185, 1134, 1073, 975, 862, 807, 763, 727, 670, 560, 487; δ_H (300 MHz; DMSO- d_6) 10.10 (1H, s, OH), 7.08 (1H, t, $J = 7.1$ Hz, CH-Ar), 6.95 (2H, br. s, NH₂), 6.83 (1H, d, $J = 7.8$ Hz, CH-Ar), 6.60 (1H, d, $J = 7.8$ Hz, CH-Ar), 4.90 (1H, s, C⁴H-pyran), 3.94 (2H, q, $J = 7.2$ Hz, O-CH₂-CH₃), 2.38 (3H, s, CH₃), 0.98 (3H, t, $J = 7.2$ Hz, O-CH₂-CH₃); δ_C (75 MHz; DMSO- d_6), 165.28 (C=O), 158.55, 157.39, 152.98, 143.49, 127.42, 119.48, 119.29, 114.48, 106.37 (3 × CH-Ar, 3 × C_q-Ar, 3 × C_q-pyran, CN), 60.09 (O-CH₂), 56.26 (C_q-pyran), 34.6 (C⁴H-pyran), 18.08 (CH₃), 13.62 (CH₃); MS (EI) m/z (%) [$M^+ + 1$, ³⁷Cl] 337.00 (0.63), [M^+ , ³⁷Cl] 336.00 (3.06), [$M^+ + 1$, ³⁵Cl] 335.00 (2.34) and [M^+ , ³⁵Cl] 334.00 (8.94) for C₁₆H₁₅ClN₂O₄, 307.00 (5.17), 305.00 (14.95), 300.05 (15.37), 299.10 (81.62), 291.00 (92.77), 289.00 (7.87), 271.05 (3.09), 261.00 (9.05), 225.00 (6.39), 208.00 (12.99), 207.00 (100.00), 179.00 (31.94), 161.00 (17.35), 133.05 (6.00), 127.05 (4.38), 115.05 (3.79), 105.05 (5.32), 89.05 (6.04), 77.05 (3.88), 67.00 (15.90), 63.00 (11.37), 57.05 (3.59), 51.00 (4.07); Anal. Calcd. for C₁₆H₁₅ClN₂O₄: C, 57.41; H, 4.52; N, 8.37. Found: C, 57.57; H, 4.41; N, 8.56.

6-Amino-5-cyano-2-methyl-4-(2,4,6-trimethoxy-phenyl)-4H-pyran-3-carboxylic acid ethyl ester **4h**

Filtration of the resulting reaction mixture afforded a yellow solid, which was washed with water and air dried. Recrystallization from EtOH afforded the pure product as a yellow powder, yield (60%), m.p. 196–198 °C; ν_{\max} (KBr)/cm⁻¹ 3465 (asymmetric NH₂), 3330 (symmetric NH₂), 3228 and 3192 (CH-Ar), 2940 and 2839 (CH-aliphatic), 2371, 2198 (CN), 1709 (C=O), 1688 and 1602 (C=C), 1460, 1404, 1330, 1226, 1150, 1119, 1062, 948, 922, 816, 785, 717, 649, 606, 500; δ_H (300 MHz; DMSO- d_6) 6.54 (2H, br. s, NH₂), 6.18 (2H, s, 2 × CH-Ar), 4.88 (1H, s, C⁴H-pyran), 3.90 (2H, q, $J = 6.9$ Hz, O-CH₂-CH₃), 3.75 (3H, s, O-CH₃), 3.71 (6H, s, 2 × O-CH₃), 2.23 (3H, s, CH₃), 1.02 (3H, t, $J = 6.9$ Hz, O-CH₂-CH₃); δ_C (75 MHz; DMSO- d_6) 166.31 (C=O), 159.85, 159.700, 159.13, 156.94, 120.54, 112.39, 105.44, 91.41 (2 × CH-Ar, 4 × C_q-Ar, 3 × C_q-pyran, CN), 59.78, 56.36, 56.12, 55.96, 55.27 (O-CH₂, C⁵-Pyran, 3 × O-CH₃), 27.26 (C⁴H-pyran), 18.08 (CH₃), 13.91 (CH₃); MS (EI) m/z (%) [$M^+ + 1$] 375.00 (5.63) and [M^+] 374.05 (24.69) for C₁₉H₂₂N₂O₆, 345.00 (17.76), 343.00 (29.18), 329.05 (6.42), 313.00 (4.86), 297.00 (100.00), 285.05 (18.39), 269.05 (13.20), 257.05 (21.80), 241.05 (4.48), 227.00 (4.35), 207.00 (11.32), 179.00 (12.32), 161.00 (11.25), 134.05 (10.03), 77.00 (8.03), 66.95 (15.11), 57.00 (5.43), 51.00 (3.59). Anal. Calcd. for C₁₉H₂₂N₂O₆: C, 60.95; H, 5.92; N, 7.48. Found: C, 61.17; H, 6.08; N, 7.71.

6-Amino-5-cyano-4-(4-fluoro-phenyl)-2-methyl-4H-pyran-3-carboxylic acid benzyl ester **4i**

The resulting reaction mixture was poured onto iced water and the precipitated solid was filtered off to obtain a yellow powder, which was washed with plenty of water, air dried, recrystallized from EtOH/ petroleum ether to give a pale yellow powder, yield (80%),

m.p. 124–126 °C; ν_{\max} (KBr)/ cm^{-1} 3458 (asymmetric NH_2), 3324 (symmetric NH_2), 3238 and 3095 (CH-Ar), 2939 (CH-aliphatic), 2370, 2199 (CN), 1721 (C=O), 1680 and 1602 (C=C), 1504, 1450, 1405, 1376, 1321, 1219, 1177, 1119, 1056, 958, 852, 763, 742, 696, 595, 521, 466; δ_{H} (300 MHz; $\text{DMSO-}d_6$) 7.31–7.27 (3H, m, 3 \times CH-Ar), 7.13–7.09 (6H, m, 6 \times CH-Ar), 6.94 (2H, br. s, NH_2), 5.07 and 4.98 (2H, ABq, $J = 16.4$ Hz, O- CH_2 -Ph), 4.34 (1H, s, C^4H -pyran), 2.32 (3H, s, CH_3); δ_{C} (75 MHz; $\text{DMSO-}d_6$), 164.65 (C=O), 157.80, 156.91, 140.17, 135.43, 128.80, 128.72, 127.98, 127.63, 127.42, 115.06, 114.78, 106.18 (9 \times CH-Ar, 3 \times C_q -Ar, 3 \times C_q -pyran, CN), 65.48 (O- CH_2 -Ph), 56.99 (C^5 -pyran), 38.74 (C^4H -pyran), 18.04 (CH_3); MS (EI) m/z (%) [$\text{M}^+ + 1$] 365.05 (0.42) and [M^+] 364.05 (1.41) for $\text{C}_{21}\text{H}_{17}\text{FN}_2\text{O}_3$, 273.05 (44.27), 269.10 (9.28), 251.05 (0.71), 229.00 (3.41), 213.00 (2.77), 200.00 (0.74), 184.00 (1.45), 158.00 (1.24), 145.05 (1.10), 133.05 (1.33), 121.05 (0.83), 91.05 (100.00), 77.00 (2.43), 65.00 (13.54), 51.00 (3.00).

6-Amino-5-cyano-4-(4-methoxy-phenyl)-2-methyl-4H-pyran-3-carboxylic acid benzyl ester **4j**

The resulting reaction mixture was poured onto iced water and the precipitate was filtered off to obtain a yellow solid, which was washed with plenty of water, air dried and recrystallized from EtOH–petroleum ether to give off-white shiny crystals, yield (87%), m.p. 170–172 °C; ν_{\max} (KBr)/ cm^{-1} 3407 (asymmetric NH_2), 3328 (symmetric NH_2), 3263, 3219 and 3009 (CH-Ar), 2958 (CH-aliphatic), 2370, 2189 (CN), 1698 (C=O), 1678, 1646 and 1603 (C=C), 1511, 1417, 1379, 1334, 1261, 1208, 1174, 1113, 1059, 1032, 1004, 948, 852, 813, 771, 692, 601, 568, 527; δ_{H} (300 MHz; $\text{DMSO-}d_6$) 7.31–7.29 (3H, m, 3 \times CH-Ar), 7.12–7.09 (2H, m, 2 \times CH-Ar), 7.00 (2H, dd, $J = 8.4, 2.7$ Hz, 2 \times CH-Ar), 6.94 (2H, br. s, NH_2), 6.84 (2H, dd, $J = 8.4, 2.72$ Hz, 2 \times CH-Ar), 5.05 and 5.3 (2H, ABq, $J = 12.6$ Hz, O- CH_2 -Ph), 4.27 (1H, s, C^4H -pyran), 3.74 (3H, s, O- CH_3), 2.31 (3H, s, CH_3); δ_{C} (75 MHz; $\text{DMSO-}d_6$), 165.39 (C=O), 158.25, 158.14, 156.76, 136.82, 135.79, 128.24, 127.84, 127.63, 119.74, 113.83, 107.18 (9 \times CH-Ar, 3 \times C_q -Ar, 4 \times C_q -pyran, CN), 65.67 (O- CH_2 -Ph), 57.57, 55.04 (C^5 -pyran and O- CH_3), 38.92 (C^4H -pyran), 18.23 (CH_3); MS (EI) m/z (%) [$\text{M}^+ + 1$] 377.05 (0.99) and [M^+] 376.05 (3.61) for $\text{C}_{22}\text{H}_{20}\text{N}_2\text{O}_4$, 310.05 (0.77), 285.05 (51.33), 269.05 (6.84), 241.05 (6.47), 225.00 (4.71), 212.00 (0.68), 184.00 (3.11), 161.00 (4.63), 135.05 (2.31), 108.05 (1.55), 91.05 (100.00), 77.00 (4.75), 65.00 (13.86), 51.00 (3.36).

6-Amino-5-cyano-4-(4-cyano-phenyl)-2-methyl-4H-pyran-3-carboxylic acid benzyl ester **4k**

Evaporation of the excess solvent afforded a sticky, dark-red residue, which was treated with iced water, to resolve it into a dark-red precipitate. Filtration, air drying and recrystallization from EtOH gave shiny golden crystals, yield (58%), m.p. 156–158 °C; ν_{\max} (KBr)/ cm^{-1} 3407 (asymmetric NH_2), 3301 (symmetric NH_2), 3265, 3224, 3196 and 3061 (CH-Ar), 2964 and 2887 (CH-aliphatic), 2373, 2203 (CN), 1720 (C=O), 1685 and 1607 (C=C), 1497, 1449, 1405, 1376, 1319, 1258, 1216, 1182, 1129, 1057, 913, 864, 797, 752, 696, 615, 549, 507, 484, 437; δ_{H} (300 MHz; $\text{DMSO-}d_6$) 7.76 (2H, d, $J = 8.1$ Hz, 2 \times CH-Ar), 7.33–7.27 (5H, m, 5 \times CH-Ar), 7.08–7.02 (4H, m, 2 \times CH-Ar and NH_2), 5.08 and 4.96 (2H, ABq, $J = 12.6$ Hz, O- CH_2 -Ph), 4.44 (1H, s, C^4H -pyran), 2.34 (3H, s, CH_3); δ_{C} (75 MHz; $\text{DMSO-}d_6$), 164.98 (C=O), 158.50, 158.38, 150.35, 135.64, 132.57, 128.23, 127.92, 127.74, 119.31, 118.76, 109.61, 105.66 (9 \times CH-Ar, 3 \times C_q -Ar, 3 \times C_q -pyran, 2 \times CN), 65.79 (O- CH_2 -Ph), 56.33 (C^5 -pyran), 34.85 (C^4H -pyran), 18.43 (CH_3); MS (EI) m/z (%) [M^+] 371.00 (0.94) for $\text{C}_{22}\text{H}_{17}\text{N}_3\text{O}_3$, 280.00 (55.65), 269.05 (16.00), 236.00 (2.30), 178.95 (2.02), 128.05 (1.20), 107.05 (2.71), 91.05 (6.47), 65.00 (12.33), 51.00 (3.64).

6-Amino-4-(2-chloro-phenyl)-5-cyano-2-methyl-4H-pyran-3-carboxylic acid benzyl ester **4l**

Filtration of the resulting reaction suspension afforded a pale-yellow solid that was washed with plenty of water, air dried and recrystallized from MeOH–petroleum ether to give white crystals, yield (70%), m.p. 162–164 °C; ν_{\max} (KBr)/ cm^{-1} 3467 (asymmetric NH_2), 3321 (symmetric NH_2), 3217, 3180, 3058 and 3030 (CH-Ar), 2955 (CH-aliphatic), 2368, 2200 (CN), 1724 (C=O), 1678 and 1598 (C=C), 1467, 1440, 1402, 1378, 1321, 1229, 1177, 1127, 1060, 958, 830, 745, 698, 674, 616, 577, 499, 420; δ_{H} (300 MHz; $\text{DMSO-}d_6$) 7.37–7.14 (7H, m, 7 \times CH-Ar), 7.02–6.90 (4H, m, 2 \times CH-Ar and NH_2), 5.07 and 4.98 (2H, ABq,

$J = 12.6$ Hz, O-CH₂-Ph), 4.91 (1H, s, C⁴H-pyran), 2.36 (3H, s, CH₃); δ_C (75 MHz; DMSO-*d*₆), 165.34 (C=O), 158.59, 158.56, 140.06, 135.77, 132.06, 129.74, 129.49, 128.42, 128.27, 127.82, 127.57, 119.19, 105.76 (9 × CH-Ar, 3 × C_q-Ar, 3 × C_q-pyran, CN), 65.76 (O-CH₂-Ph), 56.19 (C⁵-Pyran), 35.02 (C⁴H-pyran), 18.39 (CH₃); MS (EI) m/z (%) [M^+ , ³⁷Cl] 382.05 (0.64) and [M^+ , ³⁵Cl] 380.05 (1.17) for C₂₁H₁₇ClN₂O₃, 291.00 (8.32), 289.00 (24.11), 271.05 (0.47), 269.10 (12.27), 247.00 (0.67), 245.00 (3.25), 209.00 (1.63), 207.00 (0.46), 181.00 (1.10), 179.00 (0.92), 167.00 (0.95), 165.00 (1.23), 136.05 (0.29), 134.10 (0.90), 91.05 (100.00), 77.05 (3.48), 65.00 (16.45), 51.00 (4.41). Anal. Calcd. for C₂₁H₁₇ClN₂O₃: C, 66.23; H, 4.50; N, 7.36. Found: C, 66.48; H, 4.73; N, 7.57.

6-Amino-5-cyano-4-(2,4-dichloro-phenyl)-2-methyl-4H-pyran-3-carboxylic acid benzyl ester **4m**

Filtration of the reaction suspension and recrystallization of the obtained solid from ethanol gave a white powder, yield (57%), m.p. 146–148 °C; ν_{\max} (KBr)/cm⁻¹ 3466 (asymmetric NH₂), 3322 (symmetric NH₂), 3223 and 3190 (CH-Ar), 2957 (CH-aliphatic), 2371, 2341, 2202 (CN), 1720 (C=O), 1683, 1642 and 1603 (C=C), 1462, 1402, 1375, 1223, 1176, 1124, 1098, 1060, 1005, 956, 844, 745, 698, 647, 584, 517, 455; δ_H (300 MHz; DMSO-*d*₆) 7.45 (1H, d, $J = 2.1$ Hz, CH-Ar), 7.35 (1H, dd, $J = 8.4, 2.1$ Hz, CH-Ar), 7.28–7.18 (4H, m, 4 × CH-Ar), 7.05–6.93 (4H, m, 2 × CH-Ar, NH₂), 5.09 (1H, d, $J = 12.6$ Hz, one of O-CH₂-Ph), 4.92 (1H, d, $J = 12.6$ Hz, one of O-CH₂-Ph), 4.87 (1H, s, C⁴H-pyran), 2.36 (3H, s, CH₃); δ_C (75 MHz; DMSO-*d*₆), 164.50 (C=O), 158.46, 158.09, 140.88, 135.27, 132.54, 131.57, 130.72, 128.37, 127.78, 127.50, 127.37, 118.61, 104.89 (8 × CH-Ar, 4 × C_q-Ar, 3 × C_q-pyran, CN), 65.37 (O-CH₂), 55.30 (C⁵-pyran), 34.64 (C⁴H-pyran), 18.01 (CH₃); MS (EI) m/z (%) [M^+ , ³⁷Cl] 416.00 (0.58), [M^+ , ³⁵Cl] 414.00 (0.79) for C₂₁H₁₆Cl₂N₂O₃, 324.90 (19.35), 322.90 (29.45), 307.90 (0.29), 305.90 (0.37), 280.95 (1.65), 279.95 (0.60), 269.05 (13.68), 270.00 (4.80), 243.00 (1.44), 221.90 (0.76), 207.95 (1.20), 186.95 (0.94), 164.00 (1.00), 134.05 (1.62), 92.05 (8.02), 91.05 (100.00), 79.05 (2.57), 65.00 (14.26), 57.05 (5.88), 51.00 (3.42). Anal. Calcd. for C₂₁H₁₆Cl₂N₂O₃: C, 60.74; H, 3.88; N, 6.75. Found: C, 60.95; H, 4.12; N, 6.93.

3.1.3. General Procedures for Synthesis of 6-Amino-4-(aryl)-3-methyl-2,4-dihydro-pyrano[2,3-*c*]pyrazole-5-carbonitrile **5a,b**

To a mixture of ethyl acetoacetate **3b** (0.003, 0.39 g, 0.38 mL) and hydrazine hydrate (2.1 equiv., 0.0062, 0.31 g, 0.3 mL), the appropriate aromatic aldehyde **1e** or **1f** (0.003 mol) was added, followed by malononitrile **2** (0.003, 0.2 g), EtOH (20 mL) and Et₃N (0.0007 mol, 0.07 g, 0.1 mL). The resulting reaction mixture was heated under reflux for 12 h.

6-Amino-4-(2-chloro-3-hydroxy-phenyl)-3-methyl-2,4-dihydro-pyrano[2,3-*c*]pyrazole-5-carbonitrile **5a**

Evaporation of the excess solvent and recrystallization of the remaining product from EtOH afforded a beige solid, yield (65%), m.p. 240–242 °C; ν_{\max} (KBr)/cm⁻¹ 3377 and 3309 (OH, NH₂ and NH), 3159 (CH-Ar), 2929 (CH-aliphatic), 2698, 2368, 2171 (CN), 1650 and 1597 (C=C), 1515, 1486, 1458, 1412, 1344, 1291, 1257, 1189, 1161, 1102, 1051, 962, 863, 834, 782, 724, 670, 604, 559, 462; δ_H (300 MHz; DMSO-*d*₆) 12.12 (1H, s, NH), 10.18 (1H, br. s, OH), 7.11 (1H, t, $J = 7.8$ Hz, Ar-H), 6.90 (2H, br. s, NH₂), 6.84 (1H, dd, $J = 8.0, 1.1$ Hz, Ar-H), 6.58 (1H, apparent d, $J = 7.5$ Hz, Ar-H), 5.06 (1H, s, C⁴H-pyran), 1.82 (3H, s, CH₃); δ_C (75 MHz; DMSO-*d*₆) 161.19, 154.92, 153.01, 135.28, 127.4, 120.47, 119.26, 114.52, 97.57, (3 × CH-Ar, 3 × C_q-Ar, 3 × C_q-pyran, C_q-pyrazole, CN), 55.99 (C⁵-pyran), 30.67 (C⁴H-pyran), 9.55 (CH₃); MS (EI) m/z (%) [$M^+ + 1$ ³⁷Cl] 304.95 (1.37), [M^+ , ³⁷Cl], 304.00 (7.36), [$M^+ + 1$ ³⁵Cl] 303.00 (5.84) and [M^+ , ³⁵Cl] 302.00 (22.49) for C₁₄H₁₁ClN₄O₂, 278.00 (1.13), 276.00 (3.31), 269.05 (0.21), 267.05 (1.22), 237.00 (1.99), 201.00 (18.55), 176.00 (10.65), 175.00 (100.00), 128.05 (1.25), 115.10 (5.92), 105.05 (1.99), 99.00 (2.82), 89.00 (4.58), 77.00 (3.02), 66.00 (7.52), 63.00 (8.23), 57.05 (3.77), 51.00 (4.44). Anal. Calcd. for C₁₄H₁₁ClN₄O₂: C, 55.55; H, 3.66; N, 18.51. Found: C, 55.79; H, 3.78; N, 18.68.

6-Amino-3-methyl-4-(2,4,6-trimethoxy-phenyl)-2,4-dihydro-pyran[2,3-*c*]pyrazole-5-carbonitrile **5b**

The resulting reaction mixture was cooled down and poured onto ice/water to obtain a pale-yellow precipitate. Recrystallization from EtOH/DMF afforded a yellow/mustard solid (70%), m.p. 182–184 °C; ν_{\max} (KBr)/ cm^{-1} 3183 and 3183 (NH₂ and NH), 3002 (CH-Ar), 2940 and 2841 (CH-aliphatic), 2696, 2372, 2184 (CN), 1655 and 1599 (C=C), 1490, 1462, 1409, 1334, 1281, 1226, 1149, 1118, 1040, 951, 922, 828, 759, 733, 650, 581, 526; δ_{H} (300 MHz; DMSO-*d*₆) 11.75 (1H, s, NH), 6.54 (2H, s, 2 × Ar-H), 6.14 (2H, br. s, NH₂), 5.02 (1H, s, CH-pyran), 3.72 (3H, s, OCH₃), 3.64 (6H, s, 2 × OCH₃), 1.86 (3H, s, CH₃); δ_{C} (75 MHz; DMSO-*d*₆) 161.68, 159.29, 155.09, 133.60, 120.28, 111.57, 97.63, (2 × CH-Ar, 4 × C_q-Ar, 3 × C_q-pyran, C³-pyrazole, CN), 55.80, 55.24, 54.79 (C⁵-pyran and 3 × O-CH₃), 24.3 (C⁴H-pyran), 8.89 (CH₃); MS (EI) *m/z* (%) [M⁺] 302.00 (22.49) for C₁₇H₁₈N₄O₄, 342.05 (15.25), 327.05 (29.93), 311.00 (100.00), 295.00 (5.77), 276.05 (14.53), 245.00 (54.35), 230.00 (8.58), 201.00 (4.44), 175.00 (26.46), 139.10 (8.45), 121.05 (6.39), 98.05 (11.78), 77.05 (10.21), 65.05 (7.86), 57.05 (15.75).

3.2. Biology

3.2.1. Antioxidant Assays

The in vitro antioxidant potential of the synthesized compounds was assessed by two different techniques and compared to butylated hydroxytoluene (BHT), a widely used commercial antioxidant which was taken as the positive control.

DPPH Radical Scavenging Assay

The compounds' capacities to scavenge the stable radical 1,1 diphenyl-2-picrylhydrazyl (DPPH) formed in solution by donation of a hydrogen atom or an electron was investigated according to the method described by Bersuder and coworkers [83]. Due to the fact that the initial blue/purple solution of diphenyl picrylhydrazine changes to yellow in the presence of compounds with a capacity to scavenge DPPH free radicals, this reaction is used as a measure of a compound's ability to scavenge any free radical. Briefly, a 0.5 mL volume of DPPH ethanolic solution was mixed with an equal volume of each sample concentration (0.03 to 1 mg/mL), shaken strongly and incubated at room temperature for 1 h in darkness. The absorbance of the residual DPPH radicals was measured at 519 nm and compared to that of the control (without the tested compound). DPPH radical scavenging was calculated using the following formula: Scavenging effect (%) = $(1 - A_{\text{compound}}/A_{\text{Control}}) \times 100$, where A_{compound} and A_{control} are the absorbances of the tested compound and of the control, respectively. A plot of the scavenging effect (%) versus the sample concentration was also performed to determine the compound concentration providing 50% inhibition (IC₅₀).

Reducing Power Assay

The reducing power of the studied compounds was assessed according to Oyaizu's method [84]. Briefly, different concentrations (ranging from 0 to 1 mg/mL) of each tested compound were first mixed with 1 mL of 0.2 M sodium phosphate buffer (pH 6.6) and 1 mL of 1% potassium ferricyanide (K₃Fe(CN)₆) and incubated at 50 °C for 20 min. After the addition of 1.25 mL of 20% trichloroacetic acid (TCA), the mixture was centrifuged for 10 min at 3000 rpm and the upper layer solution was then mixed with 0.5 mL of 0.1% fresh ferric chloride and an equal volume of deionized water. Finally, the absorption of the resulting mixture was measured at 700 nm using a UV spectrophotometer against distilled water as the blank, while butylated hydroxytoluene (BHT) was used as a positive control. The sample ferric reducing power capability was indicated by increased absorbance.

3.2.2. Antibacterial Activity

The antibacterial activity of all synthesized compounds was evaluated against the Gram-negative strains *E. coli* (ATCC 25966), *K. pneumonia* (ATCC 700603), *P. aeruginosa* (ATCC 27853) and *S. enteric* (ATCC 43972) and against the Gram-positive strains *B. cereus*

(ATCC 14579), *B. subtilis* (ATCC 6633), *E. faecalis* (ATCC 29122), *S. aureus* (ATCC 25923) and *S. epidermidis* (ATCC 14990), using the agar diffusion method through measuring the appearance of the inhibition zone on the surface of the top agar, as reported by Berghe and Vlietinck [85]. Ampicillin (10 µg/well) was used as the positive reference standard. Bacterial viability was also investigated by determining the colony-forming ability (CFU) of bacteria incubated at different time intervals without or with appropriate amounts of the compound, which were mixed with 2×10^7 CFU/mL in sterile BHI and incubated under shaking for 60 min at 37 °C. Samples were serially diluted into sterile BHI, streaked onto media agar plates and incubated for 24 h at 37 °C. The antibacterial potency of the tested compounds was expressed as the residual number of CFUs with reference to the initial inoculums. The results presented as the half-maximal (50 %) inhibitory concentration (IC₅₀) values are the means of three different measurements.

3.2.3. Cell Culture

Cytotoxic potency was examined for a human colon cancer cell line HCT-116 (American Type Culture Collection; Manassas, VA, USA) using various amounts of each compound (10, 25, 50 and 100 µg). Samples were diluted in Dulbecco's Modified Eagle's Medium, consisting of 10% Fetal Bovine Serum, and added to cells grown and cultured for 24 h in a 5% CO₂-humidified incubator at 37 °C. Then, the activity of lactate dehydrogenase released from damaged cells was determined in the collected supernatant aliquots using an ELISA endpoint assay (Benchmark Plus, Bio-Rad, Hercules, CA, USA). As positive and negative controls, respectively, 0.1% Triton X-100 in the assay medium and the assay medium only were used. Cell viability, shown as mean values ± SDs (n = 3), was expressed as a relative percentage of the OD values determined at 600 nm in compound-treated cells and the control.

3.2.4. CDK2 Inhibitory Assay

CDK-2 inhibition activity of the studied compounds **4d** and **4k** was evaluated using a commercial CDK2 ELISA Kit (cat. no.: 79599; BPS Biosciences, San Diego, CA, USA) following the manufacturer's instructions. Briefly, different concentrations of each compound or the positive control **BMS-265246** (0.01, 0.1, 1 or 10 µM) were incubated with 20 µL of diluted CDK2/CyclinA2 enzyme at 30 °C for 45 min. Following the addition of 50 µL Kinase-Glo[®] Reagent to each well, the plate was incubated at room temperature for 15 min. Then, the luminescence signal was measured using the microplate reader. The CDK2 inhibitory activity was expressed as inhibition percentage, which was determined by comparison with a control experiment for comparative purposes. IC₅₀ values were deduced from the curves. All measurements were performed in triplicate.

3.2.5. In Vitro Quantitative Determination of CDK2 Concentration in HCT-116 Cells

The in vitro quantitative measurement of concentration of CDK2 in lysates of HCT-116 cells was carried out using a commercial Human CDK2 ELISA Kit (cat. no.: LS-F22176; LifeSpan Biosciences Inc., Seattle, WA, USA) in the presence or absence of the studied compounds (**4d** and **4k**), following the manufacturer's instructions. Briefly, standards, blanks or samples were first incubated in the corresponding wells for 90 min at 37 °C. Then, biotinylated detection antibody was added to each well and the plate was incubated for 1 h at 37 °C followed by the addition of an Avidin–Horseradish Peroxidase (HRP) conjugate which binds to the biotin. After incubation of the mixture for 30 min at 37 °C, unbound Avidin–HRP conjugate was washed away and a TMB substrate was then added which reacts with the HRP enzyme, resulting in color development. After that, the reaction was stopped using a sulfuric acid stop solution to terminate the color-development reaction and the optical density of each well was measured using a microplate reader at 450 nm. **BMS-265246** was used as the standard drug for inhibition of CDK2. All measurements were performed in triplicate.

3.2.6. Gene Expression Profiles

Design of the Primer

Primers specific for the cyclin-dependent kinase-2 (CDK2) and caspase-3 genes were designed using primer blast (<https://www.ncbi.nlm.nih.gov/tools/primer-blast/>, accessed on 31 October 2021) or primer3 (<https://primer3.ut.ee/>, accessed on 31 October 2021) software. The selected genes covered the two main groups of the cyclin-dependent kinase-2 (CDK2) gene and the caspase-3 gene. Primers were used in RT-qPCR analysis to amplify fragments of 100–200 bp in length (Table 7).

Table 7. Oligonucleotide primer pairs used for quantitative real-time polymerase chain reaction (RT-qPCR) analysis.

Gene Name	Forward Primer	Reverse Primer
Caspase-3	F 5'-GGAAGCGAATCAATGGACTCTGG-3'	R 5'-GCATCGACATCTGTACCAGACC-3'
Housekeeping gene (GAPDH)	F 5'-GCACCGTCAAGGCTGAGAAC-3'	R 5'-ATGGTGGTGAAGACGCCAGT-3'

RNA Isolation and Reverse Transcription

Total RNA was extracted from each cell culture flask using the RNeasy extraction kit (RNeasy micro kit, cat. no. 74004). Up to 1×10^6 cells, depending on the cell line, were disrupted in buffer RLT and homogenized and disrupted. Ethanol was then added to the lysate, creating conditions that promote the selective binding of RNA to the RNeasy membrane. The samples were then applied to the RNeasy Mini spin column. Total RNA binds to the membrane, contaminants were efficiently washed away and high-quality RNA was eluted in RNase-free water. All bind, wash and elution steps were performed by centrifugation in a micro-centrifuge, with DNase I treatment. The amount of extracted RNA was quantified by estimating the absorbance at 260 nm. The purity of the RNA was checked by measuring the ratio of the absorbance at 260 and 280 nm, where a ratio ranging from 1.8 to 2.0 was taken to be pure. The absence of degradation of the RNA was verified by RNA electrophoresis on a 1.5% agarose gel containing ethidium bromide. First-strand cDNA was generated from 1 µg of each flask using the High-Capacity cDNA Archive Kit, Model for One-Step RT-PCR procedures (RT-PCR kit- BioRad-USA, cat. no. 345-0412), according to the manufacturer's protocol.

Quantitative Real-Time PCR (qRT-PCR)

Quantitative real-time PCR was carried out with the Thermal Cycler Rotorgene Real-Time System II (Rotorgene, South Korea) with the SYBR kit. The primer sequences are shown in Table 5. The PCR reaction was carried out in triplicate in 96-well plates. The mixture included 12.5 µL SYBR premix ExTag, 1 µL of 60 ng cDNA as the template, 5 µL of 2 µmol/L primer premix and 6.5 µL of DNase-free nuclease water at a total volume of 25 µL. The thermal profile of the real-time system was one step at 95 °C for 30 s, followed by 30 to 45 cycles at 95 °C for 10 s (denaturation) and at 55 °C for 30 s (annealing and extension), followed by an added dissociation pattern. The actin gene was used as an internal control gene, was abundant and remained constant, and GAPDH was used as an internal standard (housekeeping gene).

Data Analysis

The relative expression levels through the average cycle threshold (CT) were successfully detected. Average CT values were calculated from the triplicate experiment conducted for each gene; the CT value was detected by subtracting the average CT value of genes from the CT value of actin and GAPDH genes. The relative expression levels of the target genes were calculated using the $\Delta\Delta C_t$ method [86] and the reference genes GAPDH for the cancer cell line. Finally, a fold change equation (2^{-1}) was used to estimate relative expression levels, while the standard deviation was calculated from the replicated experimental data.

3.3. In Silico Studies

The cyclin-dependent kinase 2 (CDK2) enzyme was selected as the molecular target to investigate the underlying molecular mechanism of the antiproliferative effects based on similarity to previously reported antitumoral pyran derivatives, which induced cell-cycle arrest via targeting CDK2 enzyme [35–37]. The structure of the target protein was retrieved from the Protein Data Bank (www.rcsb.org accessed on 10 January 2022). Thus, the three-dimensional crystal structure (PDB code: 1DI8) of CDK2 in complex with 4-[3-hydroxy anilino]-6,7-dimethoxyquinazoline (**DTQ**) was used.

The protein was prepared with a protein preparation wizard available in the Schrodinger suite, using the standard protocol recommended by Schrodinger, which involved removal of water/solvent molecules, fixing of non-standard residues and adding hydrogens and partial charges.

The structures of the studied compounds (ligands), including **4d**, **4f**, **4k**, **BMS-265246** and **DTQ**, were sketched using the built-in panel of Maestro 11.

Ligand preparation is a capability of the Schrodinger software suite that combines tools for generating 3D structures and 2D (SDF) representations as well as performing ligand geometry minimization. By employing the Ligprep protocol, all the ligands were organized using OPLS3 with default settings, and the output file was saved in maegz format automatically [87].

The Glide module [88] in extra precision (XP) mode [89] was used for flexible molecular docking of all the molecules inside the active site (ATP binding site of the kinase) [90], using default settings, as recommended in the manual for the Glide module from Schrodinger. The residues Lys33, Glu51 and Asp145, constituting the active site, were set as flexible residues. For validation of the docking protocol, the co-crystallized ligand **DTQ** (4-[3-hydroxyanilino]-6,7-dimethoxyquinazoline) was redocked, with an RMSD of 0.34 Å.

Upon completion of each docking calculation, a maximum of 100 poses per ligand were generated, which were passed through a series of filters, and the final best docked structures were ranked using a Glide score function and Glide energy. The Glide score of the predicted poses, which is the scoring energy for the best pose (lowest energy) of each ligand in the binding site, and the Glide energy were used to quantify the binding strength of the different compounds to the target protein. The protein–ligand complexes were analyzed to examine various types of interactions. For the best-scored ligands, the 2D and 3D plots of molecular ligand–receptor interactions were analyzed for hydrogen bonds and halogen bonds.

ADME Evaluation

SwissADME [91] is a free web tool for evaluating the pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules.

It was accessed on 10 January 2022 to predict the ADME properties of the bioactive synthesized compounds **4g**, **4j**, **4d** and **4k**.

4. Conclusions

Eleven new derivatives of 2-amino-4H-pyran and one 2,4-dihydropyrano[2,3-c]pyrazole were synthesized, fully characterized and examined for antioxidant and antibacterial activities as well as cytotoxicity on a HCT-116 cell line of colorectal cancer. Derivatives **4g** and **4j** showed promising antioxidant potencies as compared to BHT. Moreover, these analogues were more potent than ampicillin, displaying lowered IC₅₀ values against the Gram-positive strains *B. subtilis* (ATCC 6633), *S. aureus* (ATCC 25923) and *S. epidermidis* (ATCC 14990) and *E. faecalis* (ATCC 29122). The cell viability assays showed that compounds **4d** and **4k** exhibited the strongest antiproliferative activities, whereas **4g** and **4j** were found to be inactive. Thus, the latter analogues would be suitable candidates for further toxicity studies to evaluate their potential as safe antioxidant and anti-Gram-positive bacterial agents that could be used as preventive and adjuvant therapeutic agents against CRC. The antiproliferative mechanism of action of **4d** and **4k** was investigated using molecular

docking simulations within the ATP binding site of CDK-2. The docking results revealed that **4d** and **4k** would inhibit CDK2 activity by competing with ATP to bind on the kinase site through establishing hydrogen bonds with backbone amino acid residues and by establishing hydrophobic interactions with side chains of surrounding residues. In addition, the docking results revealed that a more powerful inhibitor for CDK2 requires a bulkier ring to be attached to the pyran ring through a flexible chain of two to three atoms. Moreover, H-bond donors or acceptors on the central scaffold could substantially enhance binding with the receptor. Further mechanistic studies, including kinase inhibitory assays, quantitative measurement of CDK-2 protein and real-time PCR profiling of CDK-2 gene in HCT-116 treated cells, confirmed that the antiproliferative actions of **4d** and **4k** could be attributed to inhibiting the activity and downregulating the expression level of CDK-2 protein and gene as well. Consequently, these derivatives would be useful leads for the generation of new anti-CRC agents that do not cause chemotherapy-induced alopecia and arrest the cell cycle without sensitization of the epithelium. Finally, investigation of the proapoptotic potential of these analogues using real-time PCR profiling of the caspase-3 gene in HCT-116 treated cells indicated that the concentration of 10 mg/mL of **4d** or **4k** is optimal for inducing mitochondrial apoptosis of the HCT-116 cells via upregulating the expression of caspase-3; therefore, further studies on this topic will be pursued in the future. Despite the in-silico predictions of the ADME profiles and drug-likeness properties of the bioactive candidates **4g**, **4j**, **4d** and **4k** using the bioavailability radar plots, the BOILED-EGG chart and Lipinski's rule of five filter provided a first glance at the potential of these derivatives to be orally bioactive, though more in vivo investigations using animal models are needed to confirm the validity of these predictions.

Supplementary Materials: The following supporting information can be downloaded at <https://www.mdpi.com/article/10.3390/ph15070891/s1>: Representative spectra of the synthesized compounds.

Author Contributions: Conceptualization, N.N.E.E.-S., Z.M.A. and M.E.A.Z.; methodology and analysis, N.N.E.E.-S., Z.M.A. (chemistry), A.B.B. (biological screening), M.B., V.H.M. (molecular docking and ADME profiling) and H.S.O. (gene expression); writing—original draft preparation, N.N.E.E.-S., A.B.B. and M.B.; funding acquisition, M.E.A.Z. and S.A.A.-H.; writing—review and editing, N.N.E.E.-S., M.E.A.Z. and S.A.A.-H. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Deanship of Scientific Research at Imam Mohammad bin Saud Islamic University, Riyadh, Saudi Arabia, through research group no.: RG.21-09-76.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data is contained within the article.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Bray, F.; Ferlay, J.; Soerjomataram, I.; Siegel, R.L.; Torre, L.A.; Jemal, A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.* **2018**, *68*, 394–424. [[CrossRef](#)] [[PubMed](#)]
2. Xi, Y.; Xu, P. Global colorectal cancer burden in 2020 and projections to 2040. *Transl. Oncol.* **2021**, *14*, 101174–101181. [[CrossRef](#)]
3. Curraiss, P.; Rosa, I.; Claro, I. Colorectal cancer carcinogenesis: From bench to bedside. *World J. Gastrointest. Oncol.* **2022**, *14*, 654–663. [[CrossRef](#)] [[PubMed](#)]
4. Malumbres, M.; Carnero, A. Cell cycle deregulation: A common motif in cancer. *Prog. Cell Cycle Res.* **2003**, *5*, 5–18.
5. Pizzimenti, S.; Toaldo, C.; Pettazzoni, P.; Dianzani, M.U.; Barrera, G. The “two-faced” effects of reactive oxygen species and the lipid peroxidation product 4-hydroxynonenal in the hallmarks of cancer. *Cancers* **2010**, *2*, 338–363. [[CrossRef](#)] [[PubMed](#)]
6. Tian, T.; Wang, Z.; Zhang, J. Pathomechanisms of oxidative stress in inflammatory bowel disease and potential antioxidant therapies. *Oxidative Med. Cell. Longev.* **2017**, *2017*, 4535194. [[CrossRef](#)] [[PubMed](#)]
7. Carini, F.; Mazzola, M.; Rappa, F.; Jurjus, A.; Geagea, A.G.; Al Kattar, S.; Bou-Assi, T.; Jurjus, R.; Damiani, P.; Leone, A.; et al. Colorectal carcinogenesis: Role of oxidative stress and antioxidants. *Anticancer Res.* **2017**, *37*, 4759–4766. [[CrossRef](#)]

8. Aggarwal, V.; Tuli, H.S.; Varol, A.; Thakral, F.; Yerer, M.B.; Sak, K.; Varol, M.; Jain, A.; Khan, M.; Sethi, G. Role of reactive oxygen species in cancer progression: Molecular mechanisms and recent advancements. *Biomolecules* **2019**, *9*, 735. [[CrossRef](#)]
9. Zahra, K.F.; Lefter, R.; Ali, A.; Abdellah, E.C.; Trus, C.; Ciobica, A.; Timofte, D. The involvement of the oxidative stress status in cancer pathology: A double view on the role of the antioxidants. *Oxidative Med. Cell. Longev.* **2021**, *2021*, 9965916. [[CrossRef](#)]
10. Park, C.H.; Eun, C.S.; Han, D.S. Intestinal microbiota, chronic inflammation, and colorectal cancer. *Intest. Res.* **2018**, *16*, 338–345. [[CrossRef](#)]
11. Zou, S.; Fang, L.; Lee, M.H. Dysbiosis of gut microbiota in promoting the development of colorectal cancer. *Gastroenterol. Rep.* **2018**, *6*, 1–12. [[CrossRef](#)] [[PubMed](#)]
12. Leung, A.; Tsoi, H.; Yu, J. Fusobacterium and escherichia: Models of colorectal cancer driven by microbiota and the utility of microbiota in colorectal cancer screening. *Expert Rev. Gastroenterol. Hepatol.* **2015**, *9*, 651–657. [[CrossRef](#)] [[PubMed](#)]
13. Antonic, V.; Stojadinovic, A.; Kester, K.E.; Weina, P.J.; Brücher, B.L.; Protic, M.; Avital, I.; Izadjoo, M. Significance of infectious agents in colorectal cancer development. *J. Cancer* **2013**, *4*, 227–240. [[CrossRef](#)] [[PubMed](#)]
14. Gao, Z.; Guo, B.; Gao, R.; Zhu, Q.; Qin, H. Microbiota dysbiosis is associated with colorectal cancer. *Front. Microbiol.* **2015**, *6*, 20. [[CrossRef](#)]
15. Dai, Z.; Zhang, J.; Wu, Q.; Chen, J.; Liu, J.; Wang, L.; Chen, C.; Xu, J.; Zhang, H.; Shi, C.; et al. The role of microbiota in the development of colorectal cancer. *Int. J. Cancer* **2019**, *145*, 2032–2041. [[CrossRef](#)]
16. Geravand, M.; Fallah, P.; Yaghoobi, M.H.; Soleimanifar, F.; Farid, M.; Zinatizadeh, N.; Yaslianifard, S. Investigation of enterococcus faecalis population in patients with polyp and colorectal cancer in comparison of healthy individuals. *Arq. Gastroenterol.* **2019**, *56*, 141–145. [[CrossRef](#)]
17. Ting, N.L.N.; Lau, H.C.H.; Yu, J. Cancer pharmacomicrobiomics: Targeting microbiota to optimise cancer therapy outcomes. *Gut* **2022**, *71*, 1412–1425. [[CrossRef](#)]
18. Bhat, S.; Muthunatarajan, S.; Mulki, S.S.; Archana Bhat, K.; Kotian, K.H. Bacterial infection among cancer patients: Analysis of isolates and antibiotic sensitivity pattern. *Int. J. Microbiol.* **2021**, *2021*, 8883700. [[CrossRef](#)]
19. Benítez-Chao, D.F.; León-Buitimea, A.; Lerma-Escalera, J.A.; Morones-Ramírez, J.R. Bacteriocins: An overview of antimicrobial, toxicity, and biosafety assessment by in vivo models. *Front. Microbiol.* **2021**, *12*, 677–695. [[CrossRef](#)] [[PubMed](#)]
20. Martelli, G.; Giacomini, D. Antibacterial and antioxidant activities for natural and synthetic dual-active compounds. *Eur. J. Med. Chem.* **2018**, *158*, 91–105. [[CrossRef](#)]
21. Stone, W.L.; Krishnan, K.; Campbell, S.E.; Palau, V.E. The role of antioxidants and pro-oxidants in colon cancer. *World J. Gastrointest. Oncol.* **2014**, *6*, 55–66. [[CrossRef](#)] [[PubMed](#)]
22. Lamb, R.; Ozsvari, B.; Lisanti, C.L.; Tanowitz, H.B.; Howell, A.; Martinez-Outschoorn, U.E.; Sotgia, F.; Michael, P.; Lisanti, M.P. Antibiotics that target mitochondria effectively eradicate cancer stem cells, across multiple tumor types: Treating cancer like an infectious disease. *Oncotarget* **2015**, *6*, 4569–4584. [[CrossRef](#)]
23. Nguyen, H.T.; Duong, H.Q. The molecular characteristics of colorectal cancer: Implications for diagnosis and therapy. *Oncol. Lett.* **2018**, *16*, 9–18. [[CrossRef](#)] [[PubMed](#)]
24. Gordon, E.M.; Ravicz, J.R.; Liu, S.; Chawla, S.P.; Hall, F.L. Cell cycle checkpoint control: The cyclin G1/Mdm2/p53 axis emerges as a strategic target for broad-spectrum cancer gene therapy—A review of molecular mechanisms for oncologists. *Mol. Clin. Oncol.* **2018**, *9*, 115–134. [[CrossRef](#)] [[PubMed](#)]
25. Mani, S.; Wang, C.; Wu, K.; Francis, R.; Pestell, R. Cyclin-dependent kinase inhibitors: Novel anticancer agents. *Expert Opin. Investig. Drugs* **2000**, *9*, 1849–1870. [[CrossRef](#)]
26. Ding, L.; Cao, J.; Lin, W.; Chen, H.; Xiong, X.; Ao, H.; Yu, M.; Lin, J.; Cui, Q. The roles of cyclin-dependent kinases in cell-cycle progression and therapeutic strategies in human breast cancer. *Int. J. Mol. Sci.* **2020**, *21*, 1960. [[CrossRef](#)]
27. Coxon, C.R.; Anscombe, E.; Harnor, S.J.; Martin, M.P.; Carbain, B.; Golding, B.T.; Hardcastle, I.R.; Harlow, L.K.; Korolchuk, S.; Matheson, C.J.; et al. Cyclin-dependent kinase (CDK) inhibitors: Structure-activity relationships and insights into the CDK-2 selectivity of 6-substituted 2-arylamino-purines. *J. Med. Chem.* **2017**, *60*, 1746–1767. [[CrossRef](#)]
28. Cam, W.R.; Masaki, T.; Shiratori, T.Y.; Kato, N.; Okamoto, M.; Yamaji, Y.; Igarashi, K.; Sano, T.; Omata, M. Activation of cyclin e-dependent kinase activity in colorectal cancer. *Dig. Dis. Sci.* **2001**, *46*, 2187–2198. [[CrossRef](#)]
29. Lim, T.G.; Lee, S.Y.; Huang, Z.; Lim, D.Y.; Chen, H.; Jung, S.K.; Bode, A.M.; Lee, K.W.; Dong, Z. Curcumin suppresses proliferation of colon cancer cells by targeting CDK2. *Cancer Prev. Res.* **2014**, *7*, 466–474. [[CrossRef](#)]
30. Shi, X.N.; Li, H.; Yao, H.; Liu, X.; Li, L.; Leung, K.S.; Kung, H.F.; Lin, M.C. Adapalene inhibits the activity of cyclin-dependent kinase 2 in colorectal carcinoma. *Mol. Med. Rep.* **2015**, *12*, 6501–6508. [[CrossRef](#)]
31. Ikwu, F.A.; Isyaku, Y.; Obadawo, B.S.; Lawal, H.A.; Ajibowu, S.A. In silico design and molecular docking study of CDK2 inhibitors with potent cytotoxic activity against HCT116 colorectal cancer cell line. *J. Genet. Eng. Biotechnol.* **2020**, *18*, 51–63. [[CrossRef](#)] [[PubMed](#)]
32. Li, K.; You, J.; Wu, Q.; Meng, W.; He, Q.; Yang, B.; Zhu, C.; Cao, J. Cyclin-dependent kinases-based synthetic lethality: Evidence, concept, and strategy. *Acta Pharm. Sin. B* **2021**, *11*, 2738–2748. [[CrossRef](#)]
33. Afifi, T.H.; Okasha, R.M.; Ahmed, H.E.; Ilaš, J.; Saleh, T.; Abd-El-Aziz, A.S. Structure-activity relationships and molecular docking studies of chromene and chromene based azo chromophores: A novel series of potent antimicrobial and anticancer agents. *EXCLI J.* **2017**, *16*, 868–902. [[CrossRef](#)] [[PubMed](#)]
34. Al Nasr, I.S.; Jentzsch, J.; Shaikh, A.; Singh Shuveksh, P.; Koko, W.S.; Khan, T.A.; Ahmed, K.; Schobert, R.; Ersfeld, K.; Biersack, B. New pyrano-4H-benzo[g]chromene-5, 10-diones with antiparasitic and antioxidant activities. *Chem. Biodivers.* **2021**, *18*, e2000839–e2000851. [[CrossRef](#)] [[PubMed](#)]

35. Mouineer, A.; Zaher, A.; El-Malah, A.; Sobh, E.A. Design, synthesis, antitumor activity, cell cycle analysis and ELISA assay for cyclin dependant kinase-2 of a new (4-aryl-6-fluoro-4H-benzo [4,5] thieno [3,2-b] pyran) derivatives. *Mediterr. J. Chem.* **2017**, *6*, 165–179. [[CrossRef](#)]
36. Zaher, A.F.; Abuel-Maaty, S.M.; El-Nassan, H.B.; Amer, S.A.; Abdelghany, T.M. Synthesis, antitumor screening and cell cycle analysis of novel benzothieno [3,2-b] pyran derivatives. *J. Enzym. Inhib. Med. Chem.* **2016**, *31*, 145–153. [[CrossRef](#)] [[PubMed](#)]
37. Samir, N.; George, R.F.; Elrazaz, E.Z.; Ayoub, I.M.; Shalaby, E.M.; Plaisier, J.R.; Demitri, N.; Wink, M. Synthesis of some tropane-based compounds targeting colon cancer. *Future Med. Chem.* **2020**, *12*, 2123–2140. [[CrossRef](#)] [[PubMed](#)]
38. Yang, Z.J.; Gong, Q.T.; Wang, Y.; Yu, Y.; Liu, Y.H.; Wang, N.; Yu, X.Q. Biocatalytic tandem multicomponent reactions for one-pot synthesis of 2-Amino-4H-Pyran library and in vitro biological evaluation. *Mol. Catal.* **2020**, *491*, 110983–110992. [[CrossRef](#)]
39. de Souza Siqueira, M.; da Silva-Filho, L.C. NbCl₅-promoted the synthesis of 4H-pyrans through multicomponent reaction. *Tetrahedron Lett.* **2016**, *57*, 5050–5052. [[CrossRef](#)]
40. Rahman, N.; Nongthombam, G.S.; Rani, J.W.; Nongrum, R.; Kharmawlong, G.K.; Nongkhaw, R. An environment-friendly magnetic organo-nanomaterial as a potent catalyst in synthesis of pyranopyrazole derivatives. *Curr. Organocatal.* **2018**, *5*, 150–161. [[CrossRef](#)]
41. Peng, Y.; Song, G. Amino-functionalized ionic liquid as catalytically active solvent for microwave-assisted synthesis of 4H-pyrans. *Catal. Commun.* **2007**, *8*, 111–114. [[CrossRef](#)]
42. Al-Kadasi, M.A.; Osman, H.A.; Nazeruddin, G.M. Silica ammonium acetate as an efficient and recyclable heterogeneous catalyst for synthesis of 4H-pyran derivatives under ultrasound irradiation at ambient conditions. *Am. Chem. Sci. J.* **2014**, *4*, 587–599. [[CrossRef](#)]
43. Shehab, W.S.; Ghoneim, A.A. Synthesis and biological activities of some fused pyran derivatives. *Arab. J. Chem.* **2016**, *9*, S966–S970. [[CrossRef](#)]
44. Shahbazi, S.; Ghasemzadeh, M.A.; Shakib, P.; Zolfaghari, M.R.; Bahmani, M. Synthesis and antimicrobial study of 1, 4-dihydropyrano [2,3-c] pyrazole derivatives in the presence of amino-functionalized silica-coated cobalt oxide nanostructures as catalyst. *Polyhedron* **2019**, *170*, 172–179. [[CrossRef](#)]
45. Kumar, D.; Reddy, V.B.; Sharad, S.; Dube, U.; Kapur, S. A facile one-pot green synthesis and antibacterial activity of 2-amino-4H-pyrans and 2-amino-5-oxo-5,6,7,8-tetrahydro-4H-chromenes. *Eur. J. Med. Chem.* **2009**, *44*, 3805–3809. [[CrossRef](#)]
46. Shukla, P.; Sharma, A.; Fageria, L.; Chowdhury, R. Novel spiro/nonspiro pyranopyrazoles: Eco-friendly synthesis, in-vitro anticancer activity, DNA binding, and in-silico docking studies. *Curr. Bioact. Compd.* **2019**, *15*, 257–267. [[CrossRef](#)]
47. Wang, D.C.; Xie, Y.M.; Fan, C.; Yao, S.; Song, H. Efficient and mild cyclization procedures for the synthesis of novel 2-amino-4H-pyran derivatives with potential antitumor activity. *Chin. Chem. Lett.* **2014**, *25*, 1011–1013. [[CrossRef](#)]
48. Saleh, N.M.; El-Gazzar, M.G.; Aly, H.M.; Othman, R.A. Novel anticancer fused pyrazole derivatives as EGFR and VEGFR-2 dual TK inhibitors. *Front. Chem.* **2020**, *7*, 917–929. [[CrossRef](#)]
49. El-Husseiny, W.M.; El-Sayed, M.A.A.; Abdel-Aziz, N.I.; El-Azab, A.S.; Ahmed, E.R.; Abdel-Aziz, A.A.M. Synthesis, antitumor and antioxidant activities of novel α , β -unsaturated ketones and related heterocyclic analogues: EGFR inhibition and molecular modelling study. *J. Enzyme Inhib. Med. Chem.* **2018**, *33*, 507–518. [[CrossRef](#)]
50. El-Sayed, N.N.; Al-Otaibi, T.M.; Alonazi, M.; Masand, V.H.; Barakat, A.; Almarhoon, Z.M.; Ben Bacha, A. Synthesis and characterization of some new quinoxalin-2(1H) one and 2-methyl-3H-quinazolin-4-one derivatives targeting the onset and progression of CRC with SAR, molecular docking, and ADMET analyses. *Molecules* **2021**, *26*, 3121. [[CrossRef](#)]
51. El-Sayed, N.N.; Almanai, N.M.; Ben Bacha, A.; Al-Obeed, O.; Ahmad, R.; Abdulla, M.; Alafeefy, A.M. Synthesis and evaluation of anticancer, antiphospholipases, antiproteases, and antimetabolic syndrome activities of some 3H-quinazolin-4-one derivatives. *J. Enzyme Inhib. Med. Chem.* **2019**, *34*, 672–683. [[CrossRef](#)] [[PubMed](#)]
52. Shahidi, F.; Zhong, Y. Measurement of antioxidant activity. *J. Funct. Foods* **2015**, *18*, 757–781. [[CrossRef](#)]
53. Champlin, F.R.; Ellison, M.L.; Bullard, J.W.; Conrad, R.S. Effect of outer membrane permeabilisation on intrinsic resistance to low triclosan levels in pseudomonas aeruginosa. *Int. J. Antimicrob. Agents* **2005**, *26*, 159–164. [[CrossRef](#)] [[PubMed](#)]
54. Kulkarni, A.P.; Nagvekar, V.C.; Veeraraghavan, B.; Warriar, A.R.; Ts, D.; Ahdal, J.; Jain, R. Current perspectives on treatment of gram-positive infections in India: What is the way forward? *Interdiscip. Perspect. Infect. Dis.* **2019**, *2019*, 7601847. [[CrossRef](#)] [[PubMed](#)]
55. Arnaout, M.K.; Tamburro, R.F.; Bodner, S.M.; Sandlund, J.T.; Rivera, G.K.; Pui, C.H.; Ribeiro, R.C. Bacillus cereus causing fulminant sepsis and hemolysis in two patients with acute leukemia. *J. Pediatr. Hematol. Oncol.* **1999**, *21*, 431–435. [[CrossRef](#)]
56. Olmos, C.; Vilacosta, I.; López, J.; Sáez, C.; Anguita, M.; García-Granja, P.E.; Sarriá, C.; Silva, J.; Álvarez-Álvarez, B.; Martínez-Monzonis, M.A.; et al. Short-course antibiotic regimen compared to conventional antibiotic treatment for gram-positive cocci infective endocarditis: Randomized clinical trial (SATIE). *BMC Infect. Dis.* **2020**, *20*, 417–424. [[CrossRef](#)]
57. Cascioferro, S.; Parrino, B.; Carbone, D.; Pecoraro, C.; Diana, P. Novel strategies in the war against antibiotic resistance. *Future Med. Chem.* **2021**, *13*, 529–531. [[CrossRef](#)]
58. Fathima, A.; Rao, J.R. Selective toxicity of Catechin—A natural flavonoid towards bacteria. *Appl. Microbiol. Biotechnol.* **2016**, *100*, 6395–6402. [[CrossRef](#)]
59. Damo, S.; Chazin, W.J.; Skaar, E.P.; Kehl-Fie, T.E. Inhibition of bacterial superoxide defense: A new front in the struggle between host and pathogen. *Virulence* **2012**, *3*, 325–328. [[CrossRef](#)]
60. Donadio, G.; Mensitieri, F.; Santoro, V.; Parisi, V.; Bellone, M.L.; De Tommasi, N.; Izzo, V.; Dal Piaz, F. Interactions with microbial proteins driving the antibacterial activity of flavonoids. *Pharmaceuticals* **2021**, *13*, 660. [[CrossRef](#)]

61. Arakawa, H.; Maeda, M.; Okubo, S.; Shimamura, T. Role of hydrogen peroxide in bactericidal action of catechin. *Biol. Pharm. Bull.* **2004**, *27*, 277–281. [[CrossRef](#)] [[PubMed](#)]
62. Grzesik, M.; Bartosz, G.; Stefaniuk, I.; Pichla, M.; Namieśnik, J.; Sadowska-Bartosz, I. Dietary antioxidants as a source of hydrogen peroxide. *Food Chem.* **2019**, *278*, 692–699. [[CrossRef](#)] [[PubMed](#)]
63. Zubair, M.S.; Anam, S.; Khumaidi, A.; Susanto, Y.; Hidayat, M.; Ridhay, A. Molecular docking approach to identify potential anticancer compounds from Begonia (*Begonia* sp.). In Proceedings of the 1st International Conference on Science and Technology 2015 (ICST-2015), Universitas Gadjah Mada, Indonesia, 21 July 2016; Volume 1755, pp. 080005–080007.
64. Chohan, T.A.; Qian, H.; Pan, Y.; Chen, J.Z. Cyclin-dependent kinase-2 as a target for cancer therapy: Progress in the development of CDK2 inhibitors as anti-cancer agents. *Curr. Med. Chem.* **2015**, *22*, 237–263. [[CrossRef](#)] [[PubMed](#)]
65. Li, Y.; Zhang, J.; Gao, W.; Zhang, L.; Pan, Y.; Zhang, S.; Wang, Y. Insights on structural characteristics and ligand binding mechanisms of CDK2. *Int. J. Mol. Sci.* **2015**, *16*, 9314–9340. [[CrossRef](#)]
66. García-Sosa, A.T.; Mancera, R.L. The effect of a tightly bound water molecule on scaffold diversity in the computer-aided de novo ligand design of CDK2 inhibitors. *J. Mol. Model.* **2006**, *12*, 422–431. [[CrossRef](#)]
67. Bramson, H.N.; Corona, J.; Davis, S.T.; Dickerson, S.H.; Edelstein, M.; Frye, S.V.; Gampe, R.T.; Harris, P.A.; Hassell, A.; Holmes, W.D.; et al. Oxindole-based inhibitors of cyclin-dependent kinase 2 (CDK2): Design, synthesis, enzymatic activities, and X-ray crystallographic analysis. *J. Med. Chem.* **2001**, *44*, 4339–4358. [[CrossRef](#)]
68. Misra, R.N.; Xiao, H.Y.; Rawlins, D.B.; Shan, W.; Kellar, K.A.; Mulheron, J.G.; Sack, J.S.; Tokarski, J.S.; Kimball, S.D.; Webster, K.R. 1*H*-pyrazolo [3,4-*b*] pyridine inhibitors of cyclin-dependent kinases: Highly potent 2,6-difluorophenacyl analogues. *Bioorg. Med. Chem. Lett.* **2003**, *13*, 2405–2408. [[CrossRef](#)]
69. Łukasik, P.; Baranowska-Bosiacka, I.; Kulczycka, K.; Gutowska, I. Inhibitors of cyclin-dependent kinases: Types and their mechanism of action. *Int. J. Mol. Sci.* **2021**, *22*, 2806. [[CrossRef](#)]
70. Davis, S.T.; Benson, B.G.; Bramson, H.N.; Chapman, D.E.; Dickerson, S.H.; Dold, K.M.; Eberwein, D.J.; Edelstein, M.; Frye, S.V.; Gampe, R.T.J.; et al. Prevention of chemotherapy-induced alopecia in rats by CDK inhibitors. *Science* **2001**, *291*, 134–137. [[CrossRef](#)]
71. Deshpande, A.; Sicinski, P.; Hinds, P.W. Cyclins and cdks in development and cancer: A perspective. *Oncogene* **2005**, *24*, 2909–2915. [[CrossRef](#)]
72. Li, J.-Q.; Miki, H.; Ohmori, M.; Wu, F.; Funamoto, Y. Expression of cyclin E and cyclin-dependent kinase 2 correlates with metastasis and prognosis in colorectal carcinoma. *Hum. Pathol.* **2001**, *32*, 945–953. [[CrossRef](#)] [[PubMed](#)]
73. Thoma, O.M.; Neurath, M.F.; Waldner, M.J. Cyclin-dependent kinase inhibitors and their therapeutic potential in colorectal cancer treatment. *Front. Pharmacol.* **2021**, *12*, 757120–757129. [[CrossRef](#)] [[PubMed](#)]
74. Zhang, L.; Yu, J. Role of apoptosis in colon cancer biology, therapy, and prevention. *Curr. Colorectal. Cancer Rep.* **2013**, *9*, 331–340. [[CrossRef](#)] [[PubMed](#)]
75. Abdullah, M.; Syam, A.F.; Meilany, S.; Laksono, B.; Prabu, O.G.; Bekti, H.S.; Indrawati, L.; Makmun, D. The value of caspase-3 after the application of annona muricata leaf extract in COLO-205 colorectal cancer cell line. *Gastro. Res. Prac.* **2017**, *2017*, 4357165. [[CrossRef](#)]
76. Hosseini, S.S.; Hajikhani, B.; Faghihloo, E.; Goudarzi, H. Increased expression of caspase genes in colorectal cancer cell line by nisin. *Arch. Clin. Infect. Dis.* **2020**, *15*, e97734. [[CrossRef](#)]
77. Yadav, P.; Yadav, R.; Jain, S.; Vaidya, A. Caspase-3: A primary target for natural and synthetic compounds for cancer therapy. *Chem. Biol. Drug Des.* **2021**, *98*, 144–165. [[CrossRef](#)]
78. Ahagh, M.H.; Dehghan, G.; Mehdipour, M.; Teimuri-Mofrad, R.; Payami, E.; Sheibani, N.; Ghaffari, M.; Asadi, M. Synthesis, characterization, anti-proliferative properties and DNA binding of benzochromene derivatives: Increased Bax/Bcl-2 ratio and caspase-dependent apoptosis in colorectal cancer cell line. *Bioorg. Chem.* **2019**, *93*, 103329–103344. [[CrossRef](#)]
79. Elgaafary, M.; Fouda, A.M.; Mohamed, H.M.; Hamed, A.; El-Mawgoud, H.K.; Jin, L.; Ulrich, J.; Simmet, T.; Syrovets, T.; El-Agrody, A.M. Synthesis of β -enamionitrile-linked 8-methoxy-1*H*-benzo [*f*] chromene moieties and analysis of their antitumor mechanisms. *Front. Chem.* **2021**, *9*, 759148. [[CrossRef](#)]
80. Daina, A.; Michielin, O.; Zoete, V. SwissADME: A free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. *Sci. Rep.* **2017**, *7*, 42717. [[CrossRef](#)]
81. Sharom, F.J. The P-glycoprotein efflux pump: How does it transport drugs? *J. Memb. Bio.* **1997**, *160*, 161–175. [[CrossRef](#)]
82. Giménez, B.G.; Santos, M.S.; Ferrarini, M.; Fernandes, J.P.S. Evaluation of blockbuster drugs under the rule-of-five. *Die Pharm.* **2010**, *65*, 148–152. [[CrossRef](#)]
83. Bersuder, P.; Hole, M.; Smith, G. Antioxidants from a heated histidine-glucose model system. I: Investigation of the antioxidant role of histidine and isolation of antioxidants by high-performance liquid chromatography. *J. Am. Oil Chem. Soc.* **1998**, *75*, 181–187. [[CrossRef](#)]
84. Oyaizu, M. Studies on products of browning reaction antioxidative activities of products of browning reaction prepared from glucosamine. *Jpn. J. Nutr. Diet.* **1986**, *44*, 307–315. [[CrossRef](#)]
85. Vanden Berghe, D.A.; Vlietinck, A.J. Screening methods for antibacterial agents from higher plants. In *Methods in Plant Biochemistry: Assays for Bioactivity*; Dey, P.M., Harborne, J.B., Hostettman, K., Eds.; Academic Press: London, England, 1991; Volume 6, pp. 47–69.
86. Gomez, K.A.; Gomez, A.A. *Statistical Procedures for Agricultural Research*, 2nd ed.; John Wiley & Sons: Hoboken, NJ, USA, 1984; p. 680.
87. Schrödinger, LLC. *Release 2015-2 (2015) LigPrep*; Version 3.4; Schrödinger, LLC: New York, NY, USA, 2015; p. 26400175.

88. Friesner, R.A.; Banks, J.L.; Murphy, R.B.; Halgren, T.A.; Klicic, J.J.; Mainz, D.T.; Repasky, M.P.; Knoll, E.H.; Shelley, M.; Perry, J.K.; et al. Glide: A new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *J. Med. Chem.* **2004**, *47*, 1739–1749. [[CrossRef](#)]
89. Friesner, R.A.; Murphy, R.B.; Repasky, M.P.; Frye, L.L.; Greenwood, J.R.; Halgren, T.A.; Sanschagrin, P.C.; Mainz, D.T. Extra precision glide: Docking and scoring incorporating a model of hydrophobic enclosure for protein—Ligand complexes. *J. Med. Chem.* **2006**, *49*, 6177–6196. [[CrossRef](#)]
90. Sumirtanurdin, R.; Sungkar, S.; Hisprastin, Y.; Sidharta, K.D.; Nurhikmah, D.D. Molecular docking simulation studies of curcumin and its derivatives as cyclin-dependent kinase 2 inhibitors. *Turk. J. Pharm. Sci.* **2020**, *17*, 417–423. [[CrossRef](#)]
91. Available online: <http://www.swissadme.ch/index.php> (accessed on 10 January 2022).

Article

QSAR, Molecular Docking, MD Simulation and MMGBSA Calculations Approaches to Recognize Concealed Pharmacophoric Features Requisite for the Optimization of ALK Tyrosine Kinase Inhibitors as Anticancer Leads

Rahul D. Jawarkar ^{1,*}, Praveen Sharma ¹, Neetesh Jain ¹, Ajaykumar Gandhi ², Nobendu Mukerjee ³, Aamal A. Al-Mutairi ⁴, Magdi E. A. Zaki ^{4,*}, Sami A. Al-Hussain ⁴, Abdul Samad ⁵, Vijay H. Masand ⁶, Arabinda Ghosh ⁷ and Ravindra L. Bakal ⁸

- ¹ Faculty of Pharmacy, Oriental University, Indore 453555, Madhya Pradesh, India; praveen140581@gmail.com (P.S.); drnkjain9781@gmail.com (N.J.)
- ² Department of Chemistry, Government College of Arts and Science, Aurangabad 431004, Maharashtra, India; gascajay18@gmail.com
- ³ Department of Microbiology, Ramakrishna Mission Vivekananda Centenary College, Kolkata 700118, West Bengal, India; nabendu21@rkmvccrahara.org
- ⁴ Department of Chemistry, Faculty of Science, Imam Mohammad Ibn Saud Islamic University, Riyadh 13318, Saudi Arabia; aamutairi@imamu.edu.sa (A.A.A.-M.); sahussain@imamu.edu.sa (S.A.A.-H.)
- ⁵ Department of Pharmaceutical Chemistry, Faculty of Pharmacy, Tishk International University, Erbil 44001, Kurdistan Region, Iraq; abdul.samad@tiu.edu.iq
- ⁶ Department of Chemistry, Vidyabharati Mahavidyalaya, Camp Road, Amravati 444602, Maharashtra, India; vijaymasand@gmail.com
- ⁷ Microbiology Division, Department of Botany, Gauhati University, Guwahati 781014, Assam, India; dra.ghosh@gauhati.ac.in
- ⁸ Department of Medicinal Chemistry, Dr. Rajendra Gode Institute of Pharmacy, University-Mardi Road, Amravati 444603, Maharashtra, India; rlbakal@gmail.com
- * Correspondence: rahuljawarkar@gmail.com (R.D.J.); mezaki@imamu.edu.sa (M.E.A.Z.); Tel.: +91-7385178762 (R.D.J.)



Citation: Jawarkar, R.D.; Sharma, P.; Jain, N.; Gandhi, A.; Mukerjee, N.; Al-Mutairi, A.A.; Zaki, M.E.A.; Al-Hussain, S.A.; Samad, A.; Masand, V.H.; et al. QSAR, Molecular Docking, MD Simulation and MMGBSA Calculations Approaches to Recognize Concealed Pharmacophoric Features Requisite for the Optimization of ALK Tyrosine Kinase Inhibitors as Anticancer Leads. *Molecules* **2022**, *27*, 4951. <https://doi.org/10.3390/molecules27154951>

Academic Editors: Kulandaisamy Arulsamy, Muralikannan Maruthamuthu, Nagendran Tharmalingam and Irini Doytchinova

Received: 30 May 2022

Accepted: 22 July 2022

Published: 3 August 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Abstract: ALK tyrosine kinase ALK TK is an important target in the development of anticancer drugs. In the present work, we have performed a QSAR analysis on a dataset of 224 molecules in order to quickly predict anticancer activity on query compounds. Double cross validation assigns an upward plunge to the genetic algorithm–multi linear regression (GA-MLR) based on robust univariate and multivariate QSAR models with high statistical performance reflected in various parameters like, fitting parameters; $R^2 = 0.69\text{--}0.87$, $F = 403.46\text{--}292.11$, etc., internal validation parameters; $Q^2_{\text{LOO}} = 0.69\text{--}0.86$, $Q^2_{\text{LMO}} = 0.69\text{--}0.86$, $\text{CCC}_{\text{cv}} = 0.82\text{--}0.93$, etc., or external validation parameters $Q^2_{\text{F1}} = 0.64\text{--}0.82$, $Q^2_{\text{F2}} = 0.63\text{--}0.82$, $Q^2_{\text{F3}} = 0.65\text{--}0.81$, $R^2_{\text{ext}} = 0.65\text{--}0.83$ including $\text{RMSE}_{\text{tr}} < \text{RMSE}_{\text{cv}}$. The present QSAR evaluation successfully identified certain distinct structural features responsible for ALK TK inhibitory potency, such as planar Nitrogen within four bonds from the Nitrogen atom, Fluorine atom within five bonds beside the non-ring Oxygen atom, lipophilic atoms within two bonds from the ring Carbon atoms. Molecular docking, MD simulation, and MMGBSA computation results are in consensus with and complementary to the QSAR evaluations. As a result, the current study assists medicinal chemists in prioritizing compounds for experimental detection of anticancer activity, as well as their optimization towards more potent ALK tyrosine kinase inhibitor.

Keywords: ALK tyrosine kinase inhibitors; QSAR; anticancer; molecular docking; MD simulation; MMGBSA

1. Introduction

The cancer kinome is currently acknowledged as a powerful target for the treatment of cancer; it comprises over 500 protein kinases, however only a few of them possess

therapeutic activity. The term ALK was coined from a chromosomal rearrangement inside anaplastic large cell lymphoma (ALCL) that was described as a front companion but discovered in 1994 [1]. Anaplastic lymphoma kinase (ALK) is a kind of oncogenic protein that is often expressed in the brain, small intestine, or testis but not in normal lymphoid cells [2]. The main physiologic aspect inhibited by the anaplastic lymphoma kinase (ALK) gene is brain development, which can keep many cancers altered, including non-small-cell lung cancer (NSCLC) or anaplastic large cell lymphomas (ALCL) [3].

Moreover, ALK gene activation appears to contribute to the initiation of carcinogenesis in a variety of human cancers such as anaplastic large cell lymphoma, lung cancer, inflammatory myofibroblastic tumors, and neuroblastoma, with the end result of fusion with additional oncogenes (NPM, EML4, TIM, etc.) and gene amplification, mutation, and protein overexpression [4]. As a basis for the researchers' involvement with the tyrosine kinase as a specific target in cancer treatment, the ALK fusion protein was created. TK plays a major role in signal transduction and is classified as a protein kinase up to the point where it is separated from the phosphate group by a tyrosine residue [5]. ALK is a recognized molecular target in a variety of ALK mutated malignancies, including non-small cell lung cancer. On the other hand, the rise of drug resistance has almost completely restricted the scientific advantage of targeting ALK with tyrosine kinase inhibitors (TKI) [6]. Furthermore, because of the treatment of ALK rearranged cancer, ALK has been suggested as a therapeutic target protein.

Academic institutions and the pharmaceutical sector are working hard to develop potent ALK inhibitors. Currently, the US Food and Drug Administration (US FDA) has approved crizotinib, Entrectinib, ceritinib, or alectinib for the treatment of patients with metastatic "ALK-positive" NSCLC [7–12]. Small-molecular inhibitors of ALK, such as AP26113 [13] and lorlatinib (PF-06463922) [14], are currently being evaluated in clinical studies (See Figure 1). On the other hand, irreversible drug resistance is rapidly spreading over the world, endangering the efficacy of chemotherapy containing these drugs. Alectinib [2], ceritinib [3], ASP3026 [4], PF-0643922, X-396, AP26113, and TSR-011 are examples of small molecules of ALK inhibitors that have been developed and are now being tested in clinical studies) [15–17].

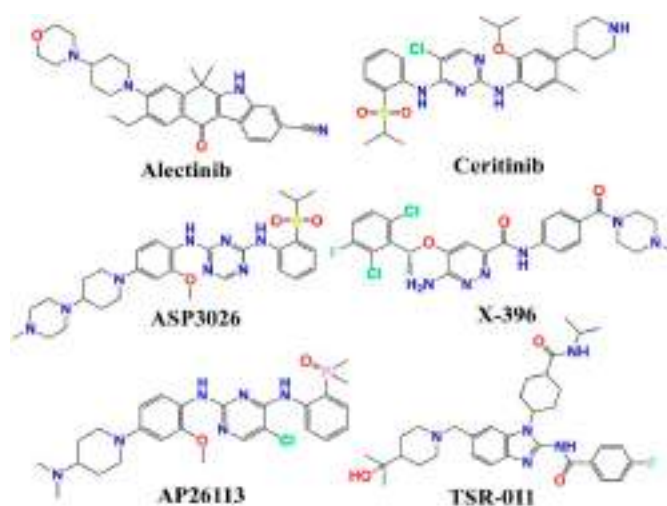


Figure 1. Demonstration of some ALK Tyrosine kinase inhibitors presently under clinical trials.

We conducted a quantitative structure activity relationship (QSAR) investigation on a dataset of 224 compounds, including clinically installed ALK tyrosine kinase inhibitory activity (K_i), in order to evaluate the critical structural and physicochemical requirements for ALK inhibitors as potent anticancer agents.

Following are the most common QSAR modelling steps: (I) selection of a dataset of molecules that cover a wide chemical space along with verified bio-activity expressed in terms of K_i , IC_{50} or EC_{50} ; (II) generation of 3D-structures of the molecules followed

by their optimization using appropriate molecular mechanics; (III) molecular descriptor calculation and data pruning using an acceptable statistical method, if required; (IV) QSAR model development using an appropriate technique that recommends promising molecular descriptor selection; and (V) double cross validation of the developed QSAR models. Statistical QSAR evaluates the bioactivity of a chemical based on its *in vitro* modification and *in vivo* testing in a wet lab. Furthermore, illustrative QSAR legitimately vary along with the statistical parameters between QSAR models that provide deep understanding for the pharmacokinetic then optimization of the lead drug [18–26].

Therefore, in the present study, we have attempted to create a QSAR model by utilizing a dataset of 224 structurally diverse compounds whose ALK tyrosine kinase inhibitory activity was previously determined experimentally (Ki). Additionally, the most active compounds in the current dataset have been examined for molecular docking studies, which have been followed by MD simulation and MMGBSA calculations. The primary goal of the current work is to uncover the array of Pharmacophoric characteristics involved in the binding affinity or stability of the drug-ALK complex. Furthermore, the stability on the drug receptor complex was observed and analyzed using MD simulation or MMGBSA techniques. The QSAR model developed in this work should provide useful information to the synthetic chemists in the discovery and development of leads to more powerful ALK Tyrosine kinase inhibitors.

2. Results

All the statistical parameters associated with fitting, double validation, or Y-scrambling for the generation of de-novo-QSAR models 1.1–1.2, along with the respective threshold values for half of the parameters are displayed in Table 1 (at the bottom of the table).

Table 1. The statistical parameters connected with the fitting, double validation and Y-scrambling for models 1.1 and 1.2.

Statistical Parameters	Model-1.1 (Univariate Dividedset Model)	Model-1.2 (Multivariate DividedSet Model)
Fitting		
R ²	0.699	0.86
R ² _{adj}	0.692	0.86
R ² -R ² _{adj}	0.001	0.003
LOF	0.57	0.25
K _{xx}	0.00	0.29
Delta K	0.83	0.14
RMSE _{tr}	0.74	0.48
MAE _{tr}	0.57	0.38
RSS _{tr}	101.03	42.10
CCC _{tr}	0.81	0.93
S	0.75	0.49
F	403.4	292.1
Internal Validation		
Q ² _{LOO}	0.68	0.86
R ² -Q ² _{LOO}	0.005	0.007
RMSE _{cv}	0.75	0.49
MAE _{cv}	0.58	0.39
PRESS _{cv}	102.9	44.6
CCC _{cv}	0.81	0.92
Q ² _{LMO}	0.68	0.86
R ² _{Yscr}	0.005	0.02
RMSE AV _{Yscr}	1.35	1.3
Q ² _{Yscr}	-0.01	0.03
External Validation		
RMSE _{ext}	0.79	0.57
MAE _{ext}	0.61	0.45
PRESS _{ext}	27.8	14.3
R ² _{ext}	0.65	0.83

Table 1. Cont.

Statistical Parameters	Model-1.1 (Univariate Dividedset Model)	Model-1.2 (Multivariate DividedSet Model)
Q^2_{-F1}	0.64	0.82
Q^2_{-F2}	0.63	0.82
Q^2_{-F3}	0.65	0.81
CCC_{ext}	0.80	0.90
r^2m aver.	0.52	0.75
r^2m delta	0.10	0.15
k'	0.99	0.99
K	0.98	0.99
$Clos'$	0.10	0.04
$Clos$	0.01	0.0

Fitting parameters such as R^2 , R^2_{adj} , and CCC_{tr} , among others, have achieved values well above the mentioned threshold limits, indicating the statistical acceptability of comprehensive QSAR models with a wide variety of chemical descriptors in them. Q^2_{LOO} , Q^2_{LMO} , and so on are internal validation parameters whose values indicate the statistical robustness of QSAR models. High values for external validation parameters R^2_{ext} , Q^2_{-Fn} , and so on indicate external predictability for both models which is reflected in a graph of experimental endpoint versus model predicted endpoint (Figure 2A,C). Williams plots that support applicability domain (AD) on the developed QSAR models are shown in Figure 2. To exclude the possibility of occasional improvement in QSAR models, appropriate threshold values based on sufficient parameters and minimal correlation among molecular descriptors must be maintained (See Supplementary Materials File S5 for the detailed formulas for the calculation of various QSAR model performance parameters). Statistical robustness and high external predictability are strong arguments in favor of these models.

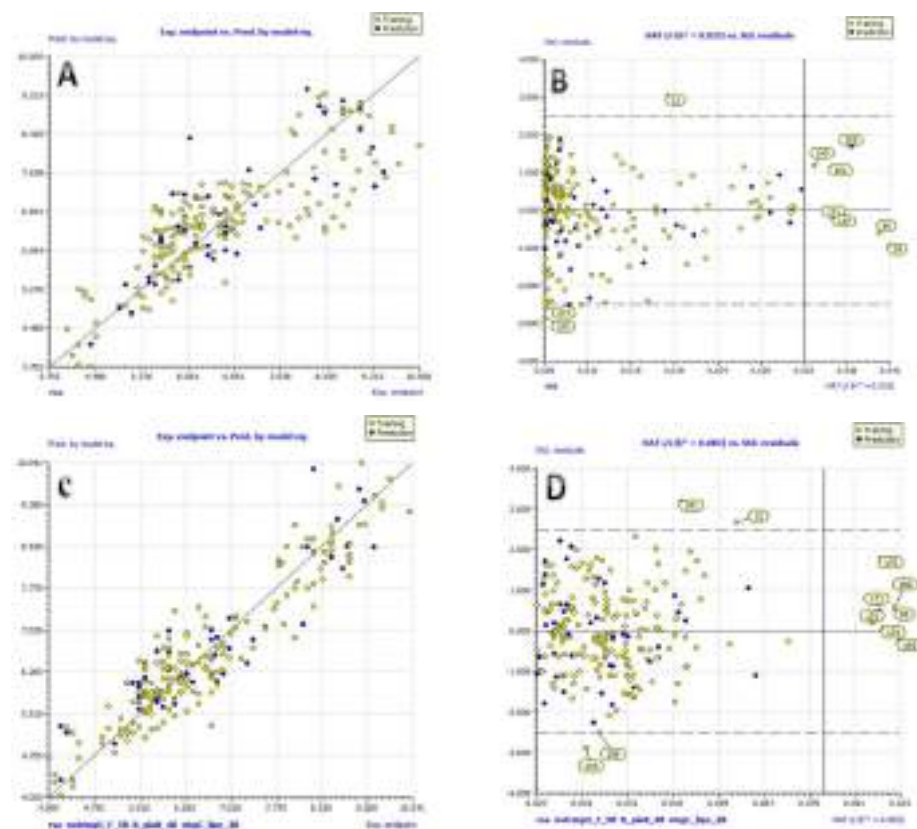


Figure 2. (A) Graph of experimental vs. Predicted pKi values for model 1.1; (B) Williams plot for model 1.1; (C) Graph of experimental vs. Predicted pKi values for model 1.2; (D) Williams plot for model 1.2.

2.1. Model. 1.1 (Univariate Analysis)

The univariate statistical analysis of the divided dataset QSAR model revealed that the dataset molecules have a gradual then spectacular outset ($R^2 = 0.690$) with the descriptor *rsa*, which stands for ratio of surface area (ratio of molecular surface area to the solvent accessible surface area). The developed univariate QSAR model is as follow:

$$pKi = -8.9 (\pm 1.5) + 25.0 (\pm 2.4) \times rsa \quad (1)$$

2.2. Model. 1.2 (Multivariate Analysis)

The another QSAR model with multiple variable is given below;

$$pKi = -6.7 (\pm 1.09) + 21.6 (\pm 2.1) \times rsa + 1.08 (\pm 0.16) \times \text{notringO_F_5B} + 0.06 (\pm 0.03) \times \text{N_plaN_4B} + -0.03 (\pm 0.02) \times \text{ringC_lipo_2B} \quad (2)$$

(Supplementary Materials Tables S3 and S4.)

3. Discussion

A precisely validated correlation between visible features of molecules, as embodied by molecular descriptors and their ALK TK inhibitory potency, amplifies the records about mechanistic features of molecules, as well as specificity and volume (presence and even absence) of various structural characteristics for favorable anticancer activity. In a broad sense, the ALK TK inhibitory efficacy of the compounds in the current dataset is the aggregate of four chemical descriptors that emerged in the developed univariate and multivariate QSAR models. Molecular descriptors may be classified into two groups based on their sign between sophisticated QSAR models.

The molecular descriptors *rsa*, *notringO_F_5B*, and *N_plaN_4B* performed well in the established QSAR models. Amplification of the values of these chemical descriptors can also contribute to an increase in the compound's ALK TK inhibitory efficacy.

The raised QSAR models include the molecular descriptor *ringC_lipo_2B* with a negative coefficient, and by decreasing the value of that molecular descriptor the compound's ALK TK inhibitory efficacy may be increased. The value of these four molecular descriptors is highlighted in the next section by comparing the variation in ALK TK inhibitory potency of the molecules (expressed in terms of *Ki* and *pKi*) with the shift in the values of the molecular descriptors seen in the QSAR models. However, the compound's bioactivity is the result of the combined action of several molecular descriptors that may or may not have been included in QSAR models.

3.1. Mechanistic Interpretation

3.1.1. *rsa*

The present QSAR evaluation performed on a given dataset marked the ratio of molecular surface area (*All_MSA*, molecular surface area) to the solvent accessible surface area (*All_SASA*) encoded by *rsa* (ratio of surface area) as one of the best performing molecular descriptors with a positive relationship with ALK TK inhibitory potency of the molecule. The chemical descriptor *rsa* (ratio of surface area) encrypts information on the molecular surface area (*All_MSA*: molecular surface area of the molecule (all atoms)) to the solvent accessible surface area (*All_SASA*: solvent accessible surface area of the molecule (all atoms)) ratio, and has a positive relationship with the molecule's ALK TK inhibitory potency. A small change in *rsa* results in a big change in the inhibitory activity of ALK TK. Because *rsa* is the ratio of the value of *All_MSA* to the value of *All_SASA*, the large possible value of *All_SASA* to the small value of *All_SASA* will set the *rsa* to the larger value, thus increasing the molecule's ALK tyrosine kinase inhibitory efficacy. This is demonstrated by comparing molecule 178 (*pKi* = 10, *rsa* = 0.68, *All_MSA* = 400.5, *All_SASA* = 587.6) to molecule 181 (*pKi* = 9.252, *rsa* = 0.66, *All_MSA* = 426.6, *All_SASA* = 639.9). The significance of *rsa* may be demonstrated by another pair of molecules 110 (*pKi* = 4.5, *rsa* = 0.528) and 47 (*pKi* = 5.4, *rsa* = 0.575) that also corroborate the observation (see Figure 3a). A triad of

compound 54 (pKi = 5.9, rsa = 0.573), compound 65 (pKi = 6.0, rsa = 0.589), and compound 144 (pKi = 6.9, rsa = 0.610) also highlights the importance of high value of ratio of surface area (see Figure 3b).

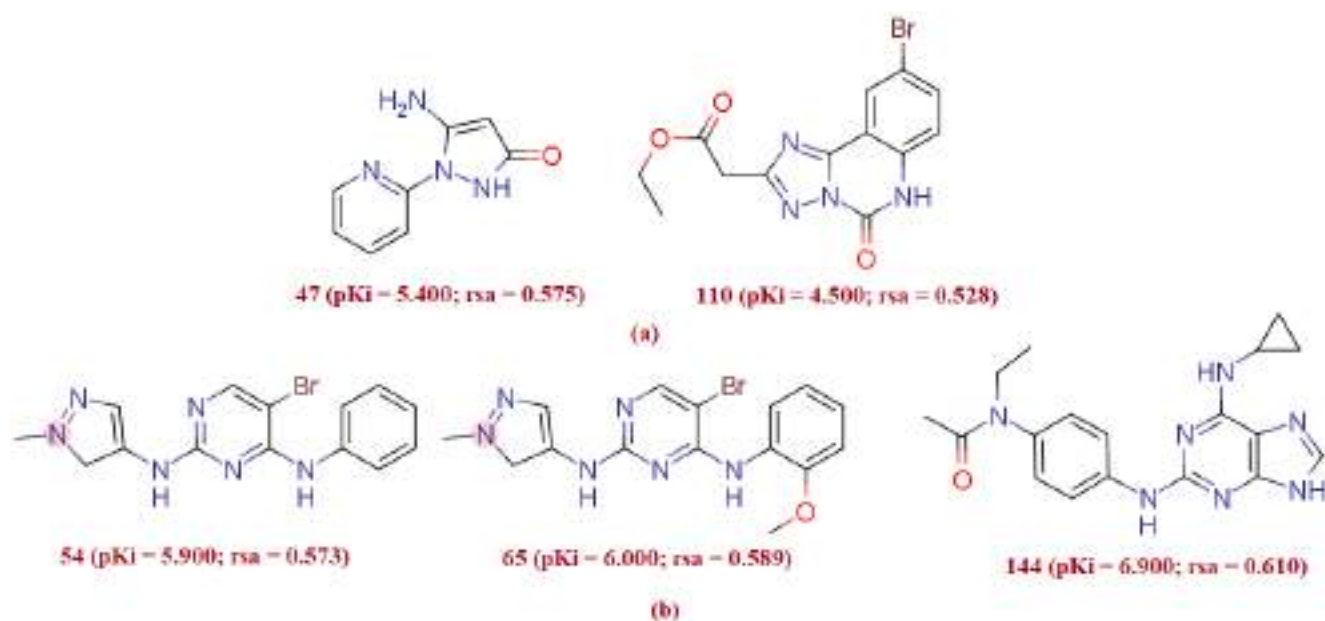


Figure 3. Illustration of the molecular descriptor *rsa* for the molecular pair 47 and 110 (a), and for the molecular pair; 54, 65 and 144 (b) only.

3.1.2. N_{plaN}_4B

The molecular descriptor N_{plaN}_4B represents the number of Nitrogen atoms with four bonds from the planar Nitrogen atom and it has a positive coefficient in the developed QSAR model. A significant number of such booster pairs of Nitrogen and planar Nitrogen may provide a more powerful ALK TK inhibitory activity. This observation is reinforced by comparing the molecule 92 (pKi = 6.8, N_{plaN}_4B = 5) with five booster Nitrogen pairs to the molecule 78 (pKi = 5.9, N_{plaN}_4B = 2) with just two booster Nitrogen pairs.

Moreover, in clinical trial agent AP26113 has such booster pairs, i.e. planar Nitrogen within four bonds from the Nitrogen atoms. The present observation confirmed that the QSAR model has successfully identified similar Pharmacophoric traits which are also present in clinical trial agents AP26113. Therefore, the planar Nitrogen within 4 bonds from the Nitrogen atom is mandatory for enhancing the affinity of ALK tyrosine kinase inhibitors (see Figure 4).



Figure 4. Illustration of Molecular descriptor N_{plaN}_4B in clinical trial molecule AP26113 (IC₅₀ = 0.07 nM).

Additionally, replacement of the molecular descriptor N_{plaN}_4B with *fplaN*_4B (represent the frequency of occurrence of Nitrogen atom exactly at four bonds from the planar Nitrogen atom) ($Q^2_{100} = 0.85$, $R^2 = 0.85$) and *ringN*_plaN_6B (occurrence of planar Nitrogen within six bonds from the ring Nitrogen atom) ($Q^2_{100} = 0.85$, $R^2 = 0.85$) led to the

diminution in the statistical presentation of the model. Thus, it can be concluded that the molecular descriptor N_plaN_4B is the better choice for predicting the ALK TK inhibitory potency. Consequently, the optimal value of distance between planar nitrogen and nitrogen atom is four bonds.

3.1.3. notringO_F_5B

The molecular descriptor $notringO_F_5B$ represents the number of Fluorine atoms within five bonds from the non-ring Oxygen atom. This molecular descriptor has a positive relationship with the ALK TK inhibitory activity of the compound, and therefore augmenting its value could offer a more potent ALK TK inhibitor. The significance of the presence and large value of a pair of Fluorine within five bonds from non-ring Oxygen can be rationalized from the fact that in the present dataset, the relatively least active compounds with $pKi \geq 7.400$ (with very few exceptions) either Fluorine atoms itself absent or such booster pair of Fluorine and non-ring Oxygen is absent, i.e. $notringO_F_5B = 0$. Whereas, in most active compounds with $pKi \geq 9.155$ at least one such Fluorine is five bonds away from the non-ring Oxygen atom ($notringO_F_5B \geq 1$). In addition to this, there are 36 such diverse sets of compounds in the entire dataset which comprises one to two such a pair of oxygen atom and fluorine atom present within five bonds. Moreover, the compounds such as, 174 ($pKi = 9.24$, $notringO_F_5B = 2$), 175 $pKi = 9.20$, $notringO_F_5B = 2$), 165, and 167 were present in prediction set while; the rest of 32 active compounds were exist in training set. Around 16% of the molecule comprises this molecular descriptor. The occurrence of the molecular descriptor $notringO_F_5B$ was not only limited to the series of homologues molecules, but it occurs in the diverse set of molecules like 161 and 167 also. Additional evidence in support is the molecule 161 ($pKi = 9.398$, $notringO_F_5B = 1$) with the molecule 173 ($pKi = 9.420$, $notringO_F_5B = 2$) (see Figure 5).

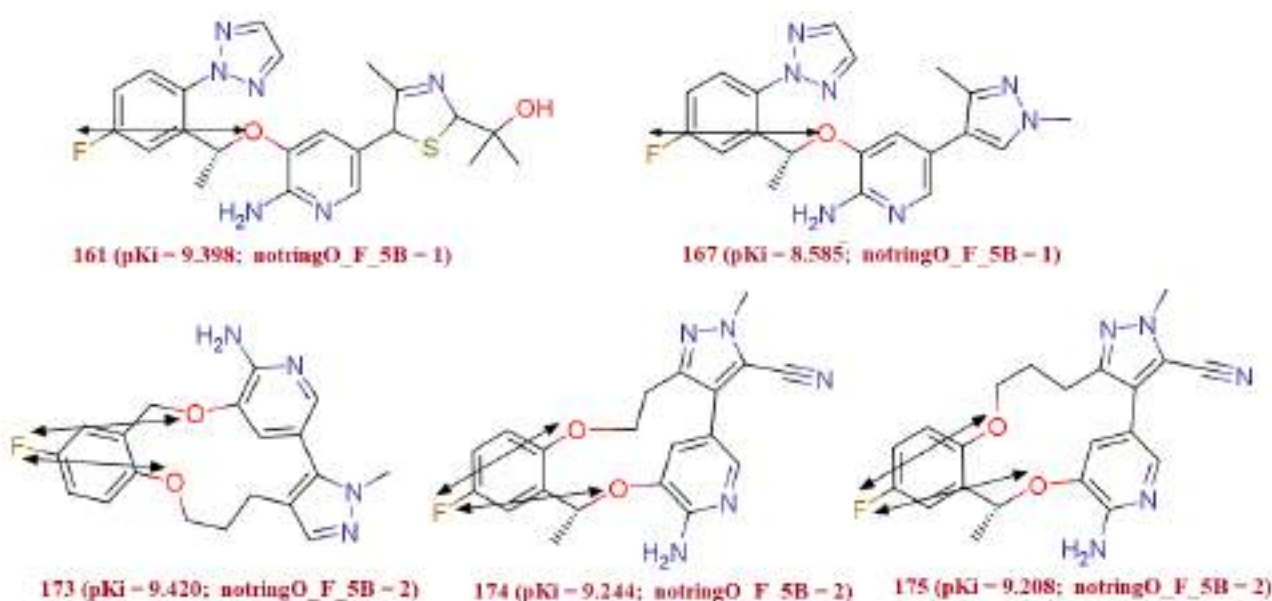


Figure 5. Illustration of molecular descriptor $notringO_F_5B$ for the molecules 161 and 173, 174 and 175, and 167.

From this observation it is revealed that the combination of Fluorine atom with non-ring Oxygen atom is independently important for inhibitory potency of ALK TK; but shifting a fluorine atom with any sulfur atom [i.e., $notringO_S_5B$ ($Q^2_{LOO} = 0.7219$, $R^2 = 0.7344$)] that represent the occurrence of Sulfur atom within five bonds from the non-ring Oxygen atom] or any acceptor atom [i.e., $notringO_Acc_5B$ ($Q^2_{LOO} = 0.7244$, $R^2 = 0.7350$)] that represent the occurrence of acceptor atom within five bonds from the non-ring Oxygen atom] significantly diminishes the statistical presentation of the QSAR model. Therefore, the presence of a fluorine atom has good correlation with the Ki value.

3.1.4. ringC_lipo_2B

The molecular descriptor ringC_lipo_2B encodes information on the occurrence of the ring carbon atoms within two bonds from lipophilic atoms. This observation is supported by comparing the pKi value of the molecule 156 (pKi = 8.55, ringC_lipo_2B = 14) with the molecule 162 (pKi = 8.30, ringC_lipo_2B = 19), for which decrease in the value of the molecular descriptor ringC_lipo_2B for the molecule 162 to 14 resulted into an increase in the pKi value by about 0.25 per unit. The triad of the molecules 180 (pKi = 8.699, ringC_lipo_2B = 16), 179 (pKi = 9.155, ringC_lipo_2B = 14), 178 (pKi = 10, ringC_lipo_2B = 11) also signifies the importance of the molecular with this Pharmacophoric future (see Figure 6a,b). This is obvious as the macrolides and aromatic rings are quite abundant in the present dataset molecules.

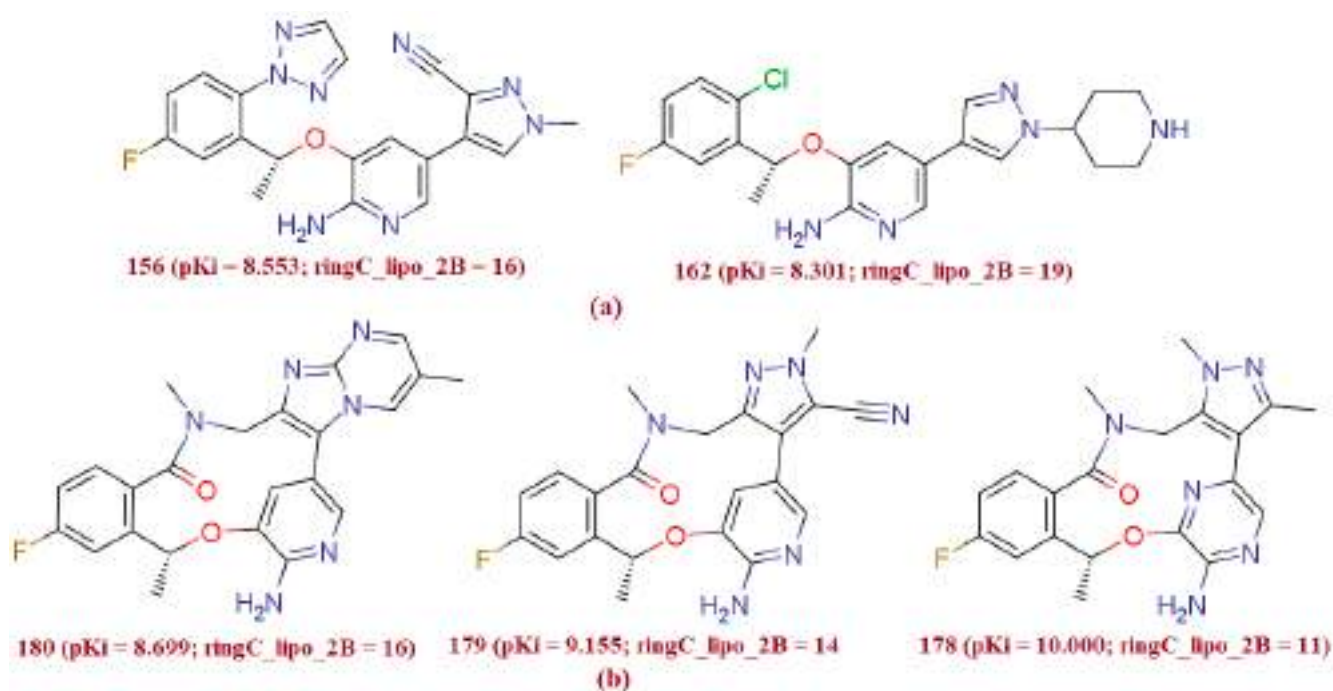


Figure 6. Presentation of the molecular descriptor ringC_lipo_2B for the molecular pair 156 and 162 (a) and for another molecular pair; 180, 179 and 178 (b) only (the ring carbon atoms within 2 bonds from the lipophilic carbon atoms are highlighted by blue ball).

On the other hand, when we have shifted molecular descriptor ringC_lipo_2B with the descriptors ringC_lipo_1B and fringCnotringC1B in which the statistical performance of the QSAR model was meaningfully improved with the molecular descriptor fringCnotringC1B ($Q^2_{\text{LOO}} = 0.87$, $R^2 = 0.87$); while performance slightly goes down with the descriptor ringC_lipo_1B ($Q^2_{\text{LOO}} = 0.84$, $R^2 = 0.85$). Therefore, from the present observation, it is revealed that with an increase in numbers of the non-ring carbon atoms attached directly to the ring carbon atoms, TK inhibitory potency could increase. Based on this observation, the optimal distance between non ring carbon atom/lipophilic atom and ring carbon atom must be one. Moreover, we have highlighted the structure of molecule 156 and 162 to rationalize the impact of the molecular descriptor ringC_lipo_2B. Absence of the triazole ring, methyl group on the pyrazole ring and carbonitrile group significantly affects the TK inhibitory potency, and may be the possible reason for the decline in the potency of the molecule 162.

3.2. Molecular Docking

ALK was discovered to be a new receptor tyrosine kinase (RTK) with an external ligand-binding domain (1030 amino acids), a transmembrane domain (28 amino acids), and an intracellular tyrosine kinase domain based on the amino acid sequences (561 amino

acids) [26,27]. While the human ALK tyrosine kinase domain is very comparable to the insulin receptor, its extracellular domain is unique among the RTK family in that it comprises two MAM domains (meprin, A5 protein, as well as receptor protein tyrosine phosphatase mu), an LDLa domain (low-density lipoprotein receptor class A), and a glycine-rich region [27,28].

ALK's ATP binding site has 27 residues, and to boost selectivity against other kinases, residues that differ from ALK were targeted. The ALK Leu1198 residue is preserved in 26 percent of the kinome and is typically Phe or Tyr in other kinases. By expanding into this pocket and bumping against the bigger Phe and Tyr residues, this smaller Leu residue might potentially provide selectivity against the majority of kinases (60 percent) [14].

The protein data bank provided the ALK tyrosine kinase pdb file (pdb id-5fto, Resolution 1.7 Å). The pdb 5fto was chosen for its X-ray resolution or sequel completeness. The protein 5ft was prepared by UCF Chimera chimera-1.16-win64 software (<https://www.cgl.ucsf.edu/Visitors/index.html>, Oakland, California, accessed on 2 March 2022). During protein preparation, we have retained water molecules. High affinity for a protein target must be attained as it is a crucial component of drug design. Although there are statistical mechanics-based formal mathematical equations that can be used to calculate binding free energies, doing so in practice is quite challenging, especially when the effect is caused by a single water molecule rather than the bulk properties of water and it is impossible to capture solvation effects. It is impossible to avoid these granular effects; a review of PDB structures reveals that each ligand-protein combination contains 4–6 ligand-bound water molecules. Furthermore, water not only stabilizes ligand interactions but plays a biological role in dictating specificity. Therefore, the improved protein with water molecules was appropriate for docking analysis. Before the docking investigation, the natural ligand (Entrectinib) was removed; in the present study, the binding site for native ligand, namely the active site, has been studied. As a result, the compounds were docked between the active site, where the native ligand was originally bound orthosterically along ALK tyrosine kinase, and the docking pose for the most active molecules 172 and 178, and an example is shown here for convenience (see Figure 7a,b). Based on the activity profile, we have carried out molecular docking analysis of the compounds 172 and 178 only. The docking analysis of compound 172 into the ALK tyrosine kinase binding pocket revealed conventional hydrogen bonding, carbon hydrogen bonding, pi-pi stacked, and pi-alkyl hydrophobic interactions (See Figure 7c), with a docking score of -8.009 kcal/mol (RMSD = 0.84 Å) and binding energy of -77.42 kcal/mol (see Table 2). In the binding pocket of the ALK tyrosine kinase, compound 172 adopts the same collapsed conformation as the co-crystallized ligand Entrectinib (See Figure 7d). The hydrogen atom on the N1 nitrogen between the pyrazole ring performs conventional hydrogen bonding with the oxygen atom of the residue GLU1197 forming the hinge region with the interatomic association of 2.87 Å, where the oxygen atom of the stated residue acted as hydrogen bond acceptor, and the hydrogen atom on the N1 nitrogen atom acted as hydrogen bond donor. Alongside, another conventional hydrogen bond was discovered in the hydrogen atom on the N2 pyrazole nitrogen with the residue MET1199 of the hinge region (interatomic distance 2.17 Å), within which pyrazole nitrogen appeared as like a hydrogen bond acceptor along MET1199 residue emerged as a hydrogen bond donor in the current composite 172-ALK tyrosine kinase complex. Table 2 displays the full docking results for the composite 172. Furthermore, the gatekeeper residue LEU1198 was attached with N2 nitrogen over the pyrazole ring through carbon hydrogen bonding along an interatomic distance of 2.82 Å. It was reported that the LEU1198 decide the selectivity of the ligand against variety of the kinases. The present observation in the docking analysis supports this fact.

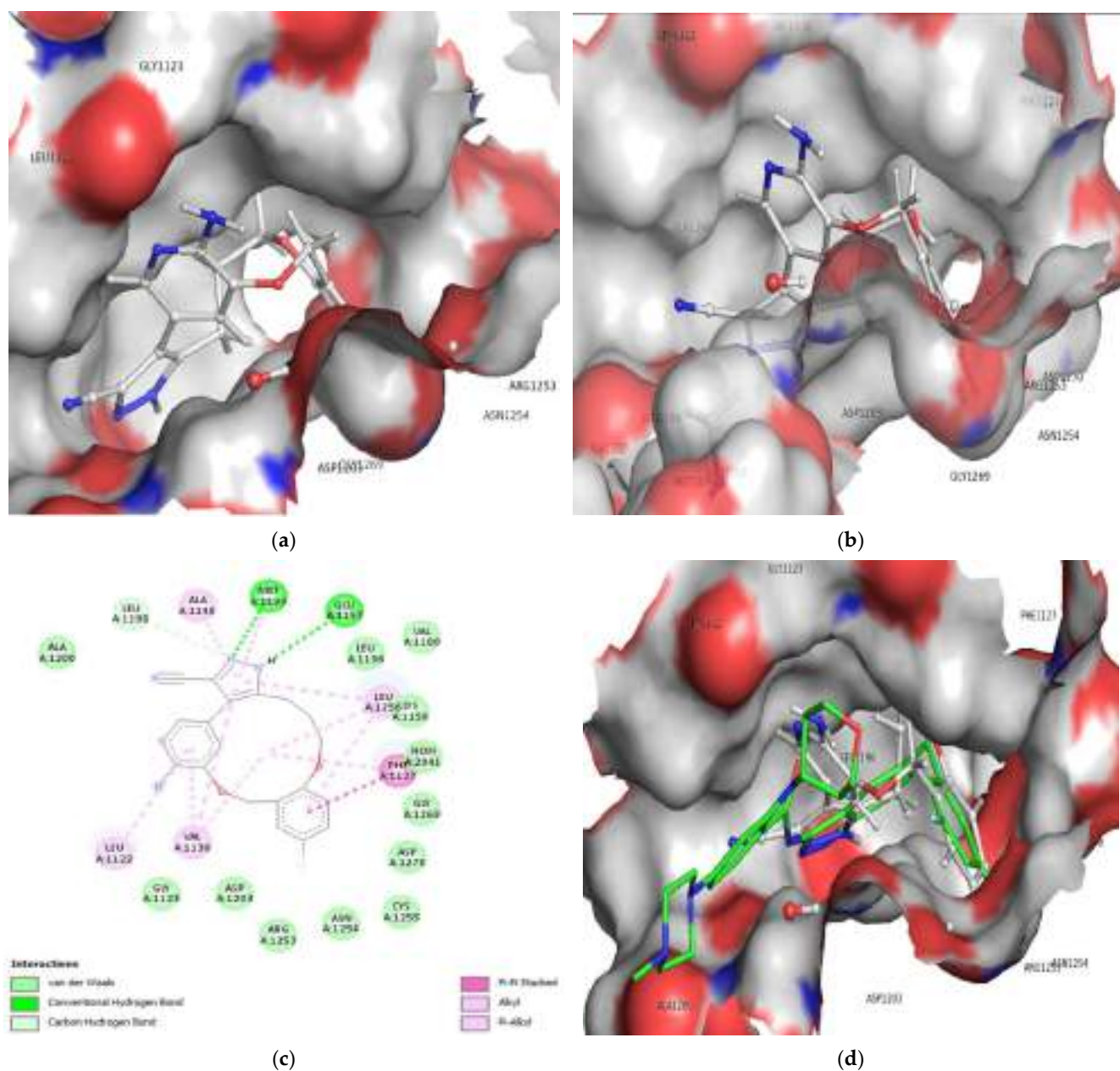


Figure 7. Depiction of alignment of compound 172 within the active site of ALK tyrosine kinase (a,b), 2D interaction of compound 172 with ALK tyrosine kinase (c) and Superimposed conformation of compound 172 with the pdb 5fto ligand (green colored); Entrectinib (d).

Moreover, at an interatomic distance of 4.81 Å, the pi-pi stacking hydrophobic contact has been facilitated by the indulgence of the pi orbital of the benzene ring and the pi orbital of the PHE1127. Furthermore, in compound 172, the pyrimidine ring is anchored with VAL1130 (interatomic distance 4.28 Å) and LEU1256 (interatomic distance 4.56 Å) and is coupled to the benzene cyclononaphane ring by alkyl hydrophobic contact. VAL1130 (interatomic distance 4.48 Å) and LEU1256 (interatomic distance 5.07 Å) make pi-alkyl hydrophobic contact with the pyrimidine ring and the cyclononaphane ring at the same time. Furthermore, the presence of an ether linkage in the unsaturated cyclononaphane ring amplifies the hydrophobicity of compound 172 as compared to the saturated benzene ring and, to a lesser extent, pyrimidine ring carbons. These findings validate the significance of the cyclononaphane ring in the molecule 172, which is primarily responsible for the compound's efficacy as mediated by the hydrophobic contact with the ALK tyrosine kinase.

Table 2. Presentation of docking interactions of the compound 172.

Residue	Distance in Å	Type of Interaction	Type of Bonding	From	Nature	To	Nature	Angle DHA	Angle HAY
MET1199:H	2.17	Hydrogen Bond	Conventional Hydrogen Bond	MET1199	H-Donor	O:N	H-Acceptor	142.364	93.83
GLU1197	2.87	Hydrogen Bond	Conventional Hydrogen Bond	O:H	H-Donor	GLU1197	H-Acceptor	111.868	157.456
LEU1198	2.82	Hydrogen Bond	Carbon Hydrogen Bond	LEU1198	H-Donor	O:N	H-Acceptor	140.175	119.444
PHE1127	4.81	Hydrophobic	Pi-Pi Stacked	PHE1127	Pi-Orbitals	0	Pi-Orbitals		
VAL1130	4.28	Hydrophobic	Alkyl	VAL1130	Alkyl	0	Alkyl		
LEU1256	4.59	Hydrophobic	Alkyl	LEU1256	Alkyl	0	Alkyl		
PHE1127	4.48	Hydrophobic	Pi-Alkyl	PHE1127	Pi-Orbitals	0	Alkyl		
VAL1130	5.07	Hydrophobic	Pi-Alkyl	0	Pi-Orbitals	VAL1130	Alkyl		
ALA1148	3.54	Hydrophobic	Pi-Alkyl	0	Pi-Orbitals	ALA1148	Alkyl		
MET1199	5.43	Hydrophobic	Pi-Alkyl	0	Pi-Orbitals	MET1199	Alkyl		
LEU1256	4.62	Hydrophobic	Pi-Alkyl	0	Pi-Orbitals	LEU1256	Alkyl		
LEU1122	4.93	Hydrophobic	Pi-Alkyl	0	Pi-Orbitals	LEU1122	Alkyl		
VAL1130	3.94	Hydrophobic	Pi-Alkyl	0	Pi-Orbitals	VAL1130	Alkyl		
LEU1256	4.17	Hydrophobic	Pi-Alkyl	0	Pi-Orbitals	LEU1256	Alkyl		

As a result, it can be concluded that the molecule 172 that binds to the ALK tyrosine kinase and drug receptor complex was mostly stabilized via conventional hydrogen, carbon hydrogen, pi-pi cation contact, alkyl hydrophobic and pi-alkyl hydrophobic interactions (See Figure 7).

Moreover, the residues: ALA1148 (interatomic distance 3.54 Å), MET1199 (interatomic distance 5.43 Å), LEU1256 (interatomic distance 4.62 Å), LEU1122 (interatomic distance 4.93 Å), VAL1130 (interatomic distance 3.94 Å) and LEU1256 forming part of a glycine rich loop (interatomic distance 4.17 Å) establishes a pi-alkyl hydrophobic interaction with the pi electrons of the pyrimidine and benzene rings, strengthening the molecule 172-ALK tyrosine receptor complex.

Furthermore, the docking analyses for compound 178 reveal the stability of the drug receptor complex through the formation of water-mediated hydrogen bonds, carbon hydrogen bonds, pi-pi stacking hydrophobic contacts, alkyl and pi-alkyl interactions, and a binding energy of -87.50 kcal/mol (docking score -7.84 kcal/mol, RMSD: 1.06 Å). The HOH2080 water molecules display hydrogen bonding contact with the N7 nitrogen atom of the cyclonaphane ring with an interatomic distance of 2.87 Å (see Figure 8a,b). At the same time, hydrogen of the N7 nitrogen produced by keto-enol tautomerism binds to the ASP1203 residue (interatomic distance 2.58 Å). This interaction is mediated by the presence of N7 nitrogen as a hydrogen bond donor and the oxygen atom of the ASP1203 residue. In addition, ASP1203 and the N1 nitrogen atom of the pyrazine ring in compound 178 formed another carbon hydrogen bond (see Figure 8a,b) (see Table 3). The superimposed conformation of compound 178 with the pdb-5fto ligand Entrectinib into the binding pocket of ALK TK is shown in Figure 8c,d.

Interestingly, the two pi-pi stacking hydrophobic contact into the drug receptor complex has been sustained and facilitated by the involvement of pi electrons from the pyrazine ring, pyrimidine ring, as well as pi electrons from the saturated benzene ring in residue A: PHE1127 (Interatomic distances 5.03 and 3.82 Å resp.). As a result, the residues A: ALA1148 (Interatomic distance 3.88 Å), A: LEU1122 (Interatomic distance 5.49 Å), A: LEU1198 (Interatomic distance 5.22 Å), and A: MET1199 (Interatomic distance 5.47 Å) establish an alkyl hydrophobic contact with the cyclonaphane ring's C4 substituted methyl moiety. Furthermore, the pi electrons of the benzene ring at residue A: PHE1127 (interatomic distance 4.51 Å) establish pi-alkyl hydrophobic contact with the alky moieties of the pyrazine and pyrimidine rings. The pyrazine ring and pyrazine then form a two pi-alkyl contact with the residue A: LEU1256, with interatomic distances of 5.43 and 4.60 Å, respectively.

Similarly, the residues A: VAL1130 and A: LEU1122 were linked with the pyrimidine ring and benzene ring via pi-alkyl hydrophobic interactions (Interatomic distances 4.14 Å and 3.84 Å, respectively). The Figure 9A,B displays the 2D interaction and surface view for the compound 172 and 178.

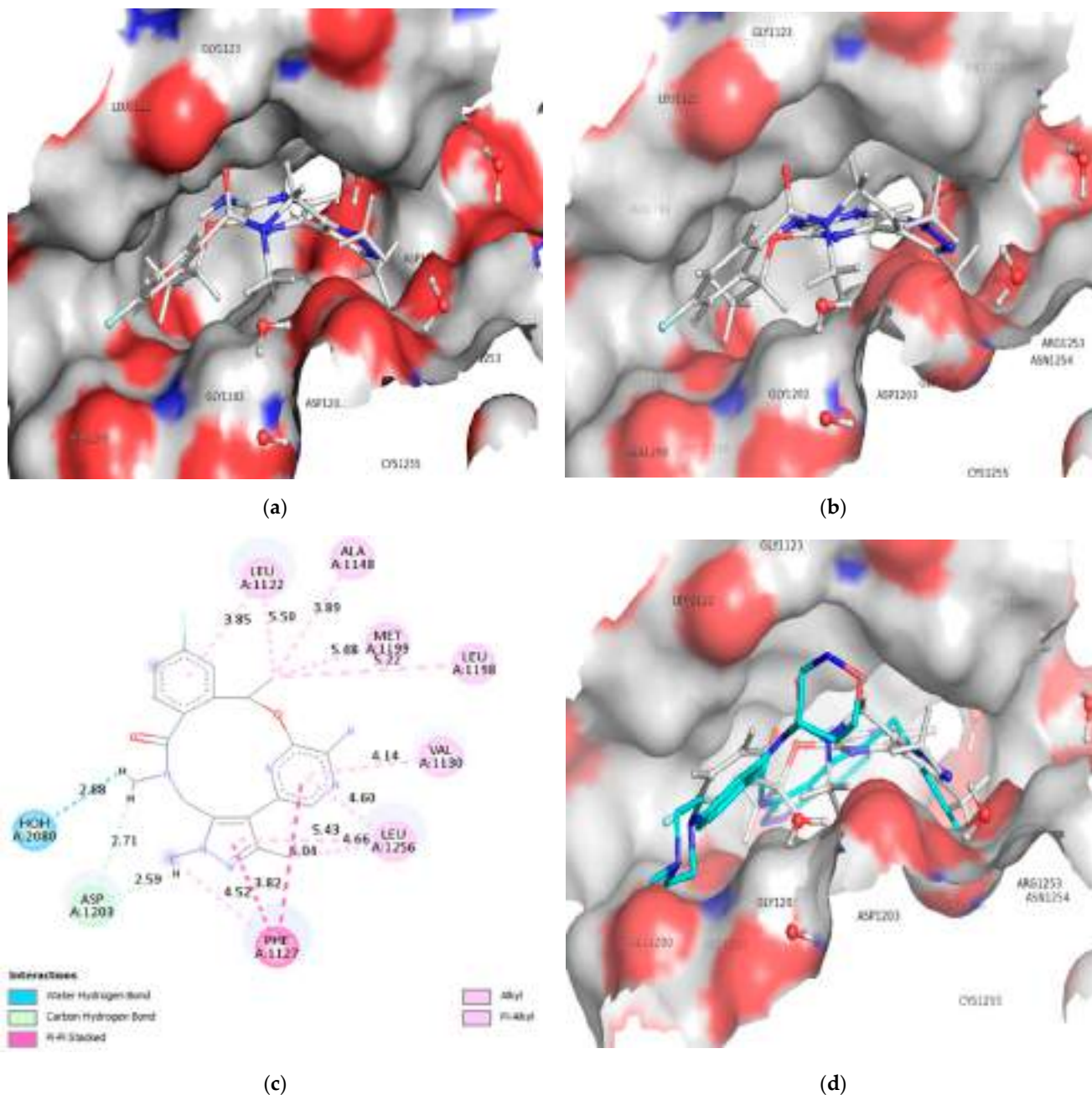


Figure 8. Depiction of alignment of compound 178 within the active site of ALK tyrosine kinase (a,b), 2D interaction of compound 178 with ALK tyrosine kinase (c), and Superimposed conformation of compound 178 with the pdb-5fto ligand (cyan colored): Entrectinib (d).

Table 3. Presentation of docking interactions of the compound 178.

Residue	Distance in Å	Type of Interaction	Types of Bonding	From	Nature	To	Nature	Angle DHA	Angle HAY
HOH2080	2.87	Hydrogen Bond	Water Hydrogen Bond; Carbon Hydrogen Bond	O:H3	H-Donor	HOH2080	H-Acceptor	114.83	91.8
ASP1203	2.58	Hydrogen Bond	Carbon Hydrogen Bond	O:H1	H-Donor	ASP1203	H-Acceptor	119.9	102.5
ASP1203	2.70	Hydrogen Bond	Carbon Hydrogen Bond	O:H1	H-Donor	ASP1203	H-Acceptor	147.9	110.3
PHE1127	5.03	Hydrophobic	Pi-Pi Stacked	PHE1127	Pi-Orbitals	0	Pi-Orbitals		
PHE1127	3.82	Hydrophobic	Pi-Pi Stacked	0	Pi-Orbitals	PHE1127	Pi-Orbitals		
ALA1148	3.88	Hydrophobic	Alkyl	ALA1148	Alkyl	O:C	Alkyl		
LEU1256	4.66	Hydrophobic	Alkyl	O:C	Alkyl	LEU1256	Alkyl		
LEU1122	5.49	Hydrophobic	Alkyl	O:C	Alkyl	LEU1122	Alkyl		
LEU1198	5.22	Hydrophobic	Alkyl	O:C	Alkyl	LEU1198	Alkyl		
MET1199	5.47	Hydrophobic	Alkyl	O:C	Alkyl	MET1199	Alkyl		
PHE1127	4.51	Hydrophobic	Pi-Alkyl	PHE1127	Pi-Orbitals	O:C	Alkyl		
LEU1256	5.43	Hydrophobic	Pi-Alkyl	0	Pi-Orbitals	LEU1256	Alkyl		
VAL1130	4.14	Hydrophobic	Pi-Alkyl	0	Pi-Orbitals	VAL1130	Alkyl		
LEU1256	4.60	Hydrophobic	Pi-Alkyl	0	Pi-Orbitals	LEU1256	Alkyl		
LEU1122	3.84	Hydrophobic	Pi-Alkyl	0	Pi-Orbitals	LEU1122	Alkyl		

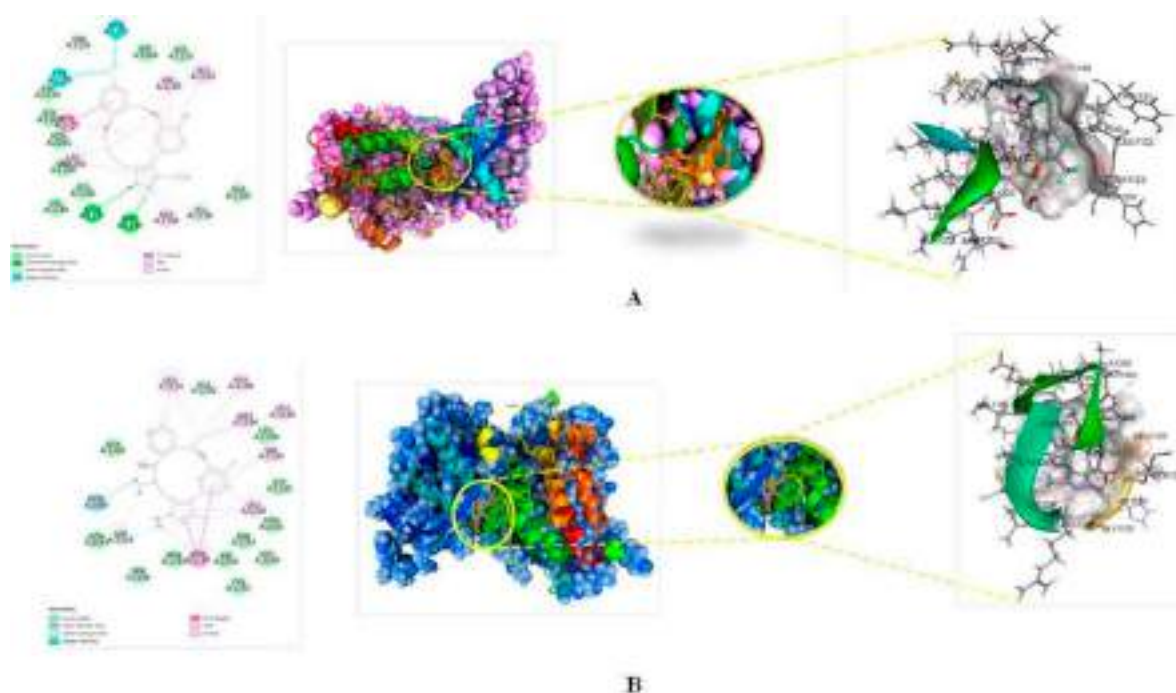


Figure 9. (A) Best docked pose of 172 with ALK displaying 2D interaction plot on the left panel. Pink dashed lines indicating the Pi-Alkyl bond and residues embedded in light green sphere indicating to involve in Van der Waals interactions. On the center panel, surface view of ALK displaying binding cavity of 172 and right panel displaying the zoomed out binding pocket having amino acid residues surrounding the 172 molecule; (B) Best docked pose of 178 with ALK displaying 2D interaction plot on the left panel. Pink dashed lines indicating the Pi-Alkyl bond and residues embedded in light green sphere indicate involvement in Van der Waals interactions. On the center panel, surface view of ALK displaying binding cavity of 178 and right panel displaying the zoomed out binding pocket having amino acid residues surrounding the 178 molecules.

In case of the molecule 178, there is a complete reversal of the conformation in comparison with the reported pdb 4CMU for the molecule 178. In the present work, we have used pdb 5fto for performing docking for the most active molecules 172 and the 178. The co-crystallized native ligand was docked along the 172 and the molecule 178. The RMSD values for the molecule 178 was found to be 1.06 Å, while co-crystallized ligand (pdb-5fto) displayed a RMSD of 1.19 Å, which was less than the molecule 178. This observation revealed the good fit of the molecule 178 into the binding pocket of the ALK tyrosine kinase due to the relatively higher flexibility of the ligand. The success rates in binding mode prediction for different docking programs such as, AutoDock 4 version (v4.2.6), FlexX 1.8, FRED (OEDocking 4.1.2.1), Glide 6.7, CCDC GOLD Suite 5.3, and ICM-Pro docking on the numerous known ligands when the RMSD cutoff ranges from 1.0 to 3.0 Å [29]. The same docked complex of the molecule 178 with ALK TK was analysed for the stability by MD simulation and MMGBSA. Although most of the existing docking programs were developed as a general methodology for different systems, they do have their own strengths and limitations and may show different performances on specific applications.

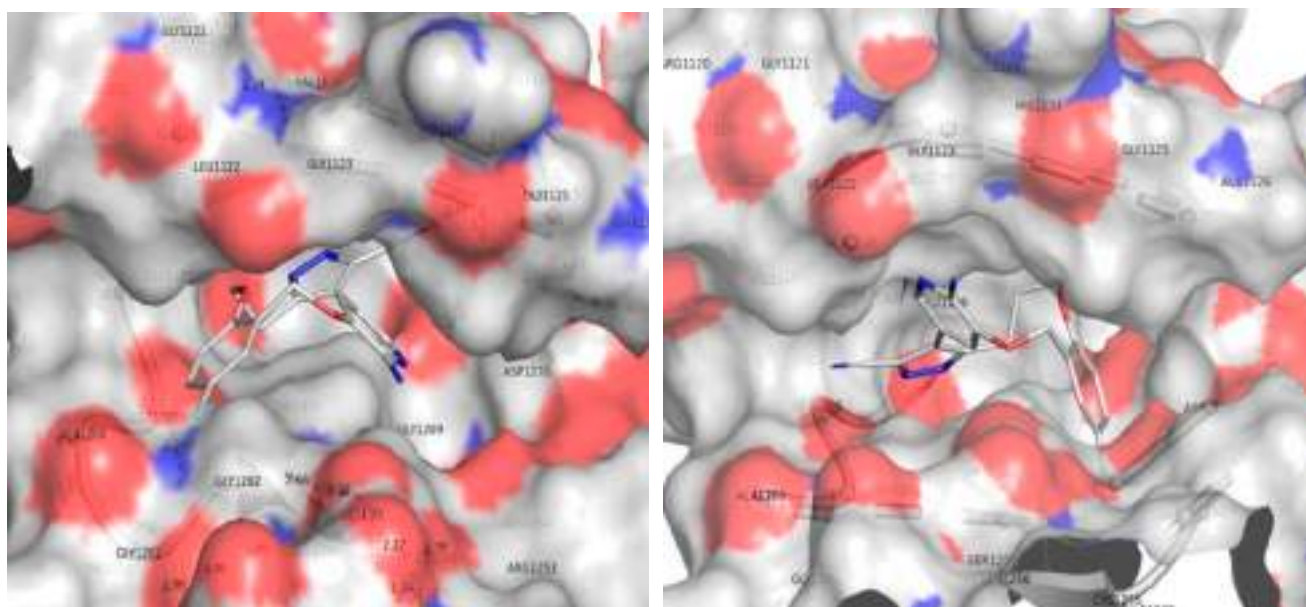
The simulation studies revealed the stability of 178 into the binding pocket of ALK TK, although it displayed reverse conformation. Moreover, the reversal of the conformation could also be attributed to the large size of active site of ALK TK (See Figure 8a,b), which allows adoption of different conformations for molecule 178. Additionally, recent studies point out that current docking software like AutoDock 4 version (v4.2.6), Dock (version 3 and 6), NRG Suite (PyMOL versions 1.2 and above) etc. and respective algorithms for docking scores are inclined toward flexibility of ligands which in turn is associated with loss of ligand conformational entropy on binding. Various factors such as binding site characteristics, one-dimensional properties of the compound library, the type of the binding pocket, ligand and protein flexibility and input differences apparently decide the docking performance [30]. Therefore, the molecule 178 (docking score -7.84 kcal/mol) with lower binding affinity for ALK TK has displayed a high degree of flexibility. Thus, all these combined factors resulted in an artificially more favorable binding score for more flexible decoys than for actives.

In addition to this, when we have redock the molecule 178 again into the binding pocket of ALK TK for comparing the docking results with the QSAR findings, it has attained similar conformation as that of the co-crystallized ligand (pdb-5fto) and the molecule 172. This observation supports the reported finding related to the loss of bioactive conformation due to the high degree of flexibility.

3.3. Comparison of Molecular Docking Results with the Reported X-ray Evidences

For comparing the docking results with the QSAR findings, it has attained similar conformation to that of the co-crystallized ligand (pdb-5fto) and the molecule 172. The docking position of molecule 172 shows that the phenyl with fluorine as a substituent is within the cavity formed by GLY1269 and ASN1254. Maria Menichincheri et al. [31] reported a similar observation. The molecule 171 has a comparable benzene ring with a fluorine substituent; however, the docking position shows that the fluorine carrying ring is unable to occupy the cavity produced by GLY1269 and ASN1254. Furthermore, conformation is completely reversed for molecules 171 and 172. One probable explanation is the existence of an extra carbon atom in molecule 172, which has resulted in increased flexibility and rsa (ratio of surface areas = ALL_MSA/ALL-SASA). As a result, QSAR and docking led to consistent and complementary results (see Figure 10).

Similarly, comparing molecules 176 and 178 indicates an intriguing impact of the N_Plan_4B and ringC_lipo_2B on docking position and activity profile. When compared to molecule 176, molecule 178 has a larger number of N_Plan_4B and a lower value of ringC_lipo_2B. This might be the explanation for the docking conformation reversal and variances in binding affinities. The added planer nitrogen appears to be boosting the polarity of the molecule (see Figure 11a,b).

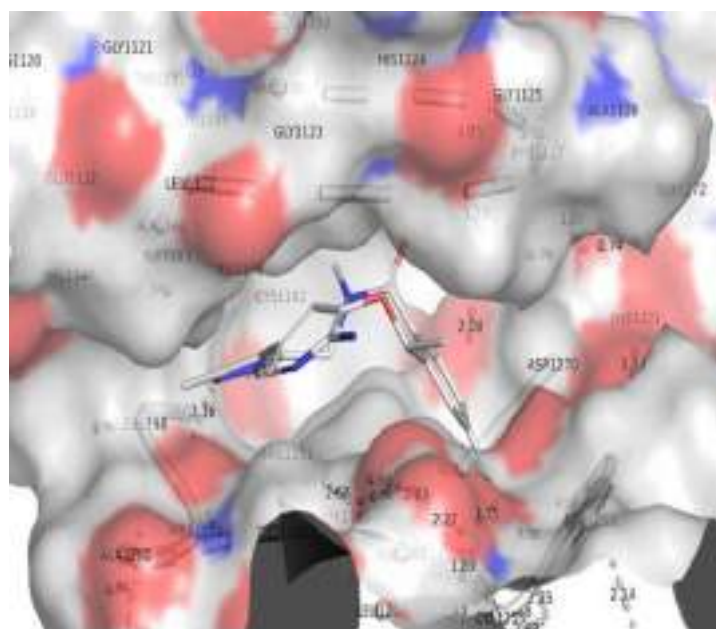


(a)

(b)

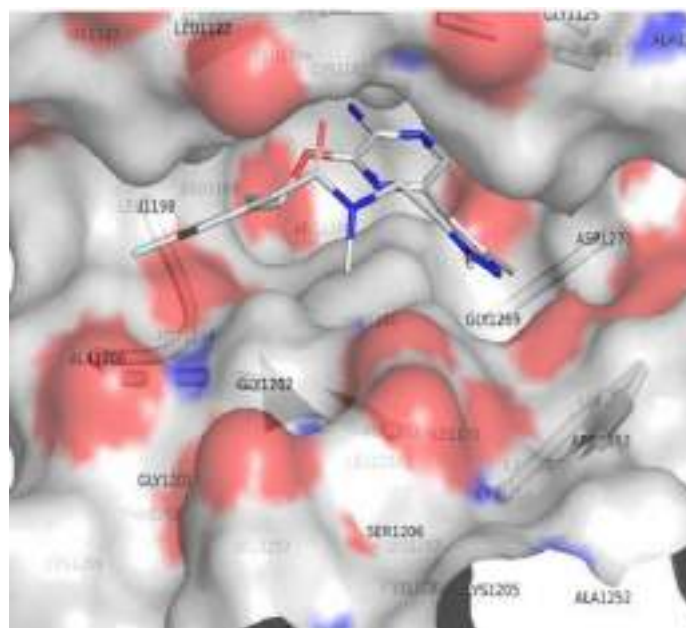
Figure 10. Depiction of the docking pose for the molecules 171 (a) and 172 (b) within the binding pocket of the ALK tyrosine kinase.

The combined effect of increase *rsa* and the presence of *notringO_F_5B* have resulted in increased potency for 218 as compared to 214. Another example is the pair of molecules 191 verses 205. This observation again divulges that the QSAR results and docking outcomes are in complete agreement with each other.



(a)

Figure 11. *Cont.*



(b)

Figure 11. Depiction of the docking pose for the molecule 178 (a) and 176 (b) within the binding pocket of the ALK tyrosine kinase.

3.4. Molecular Dynamics Simulation (MD)

Molecular dynamics and simulation (MD) experiments were performed to investigate the stability or convergence of the most active compounds 172 and 178 bound ALK complex. Based on the activity profile and molecular docking results, we have used dock complexes of the compound 172 and 178 for MD simulation analysis. When the root mean square deviation (RMSD) data were compared, each simulation including 100 ns revealed stable conformation. The C-backbone of ALK bound to 172 exhibited a deviation of about 2.2 Å (see Figure 12) while the C-backbone of ALK bound to 178 exhibited a deviation of about 1.8 Å (Figure 12); RMSD plots are within the acceptable range signifying the stability of proteins in the 172 and 178 bound state earlier than or after simulation; however, it can also be suggested that the two ligands, 172 and 178 bound to ALK is quite stable within complex.

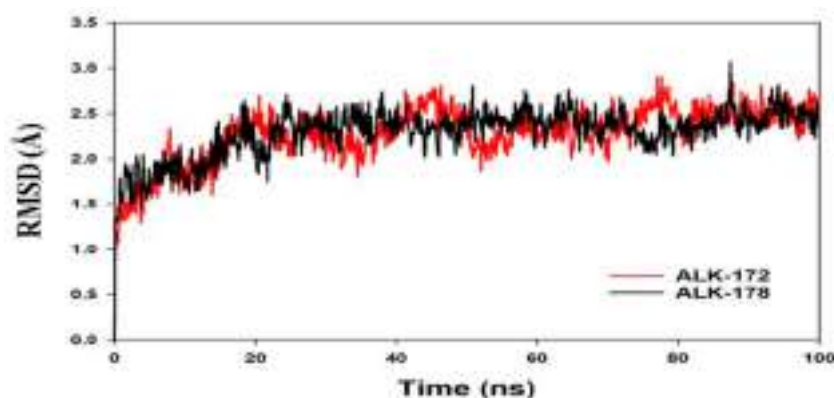


Figure 12. MD simulation trajectory analysis of Root Mean Square Deviations (RMSD) of 172 and 178 bound with ALK at 100 ns time frame displayed H-Bond plot of 172 bound ALK (red), 178 bound ALK (black).

The radius of gyration is a measure of the protein's compactness. The Radius of Gyration was reduced in 172 and 178 bound proteins, respectively (see Figure 13). According to

the overall quality analysis based on RMSD and Rg, 172 or 178 bound to the protein targets subsequently in the binding cavities and plays a significant role in the protein stability.

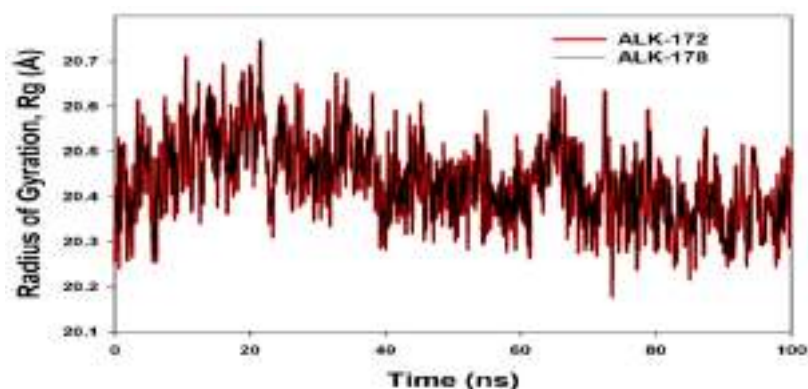


Figure 13. MD simulation trajectory analysis of Radius of gyration (Rg) of 172 and 178 bound with ALK at 100 ns time frame displayed H-Bond plot of 172 bound ALK (red), 178 bound ALK (black).

Plots for root mean square fluctuations (RMSF) of amino acid residues are shown at a time function of 100 ns. From the 100 ns simulation runs on ALK shown in Figure 14, ligand 172 has few variations peaks at residue indices 1145, 1220, and 1290, but ligand 178 has fluctuations at residues 1139, 1220, 1275, and 1345, although it was subsequently stabilized. As a result of the RMSF plots, it is reasonable to conclude that the protein structures were stable throughout the simulation within the 172, 178 bound conformation.

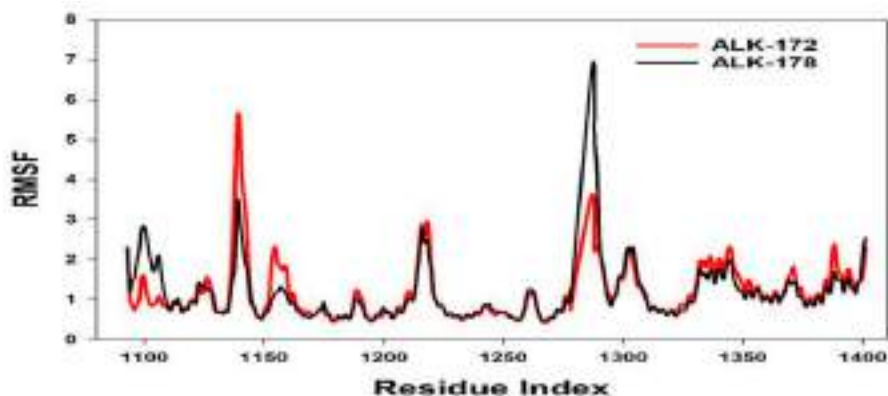


Figure 14. MD simulation trajectory analysis of Root Mean Square Fluctuations (RMSF) of 172 and 178 bound with ALK at 100 ns time frame displayed H-Bond plot of 172 bound ALK (red), 178 bound ALK (black).

The average hydrogen bonds established in 172 and 178 or the corresponding proteins throughout the 100 ns simulation were also noticed and recorded in Figure 15. From 0 ns to a 100 ns, an average of one hydrogen bonding is seen throughout the simulation or the same for MD simulations on 172 and 178 including ALK (Figure 15). Overall, three hydrogen bonds were generated during the simulation, as determined by a 2D ligand binding plot of 172 bound ALK protein, whereas in 178 bound along ALK, an average of one hydrogen bonding was produced. The quantity of hydrogen bonding over ALK along 172 and 178 hold strengthened the binding, assisting in making it more stable during the simulation (See Figure 15). In molecular docking studies, we have observed from the 2D interaction diagram where in ALK-172, we can see two hydrogen bonds were formed, on the other hand for ALK-178 we have observed single hydrogen bonding, therefore the same pattern for Molecular Dynamics.

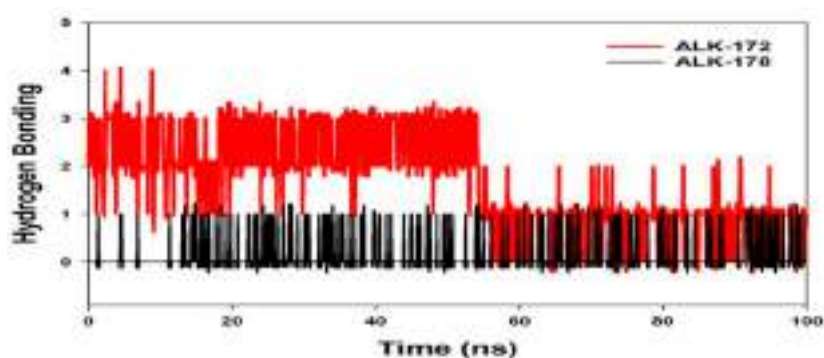


Figure 15. MD simulation trajectory analysis of Hydrogen Bonding (H-Bonds) of 172 and 178 bound with ALK at 100 ns time frame displayed H-Bond plot of 172 bound ALK (red), 178 bound ALK (black).

The step wise analysis of the stimulation trajectory of every 25 ns from beginning to end is depicted in Figure 16. The simulation trajectories exhibited the ligand 172 and 178 having no significant conformational changes throughout the 100 ns simulation. This signifies that the simulation complexes of ALK-172 and ALK-178 are stable and the ligand conformations at the active binding pocket of the ALK remains significantly unaltered (See Figure 16).

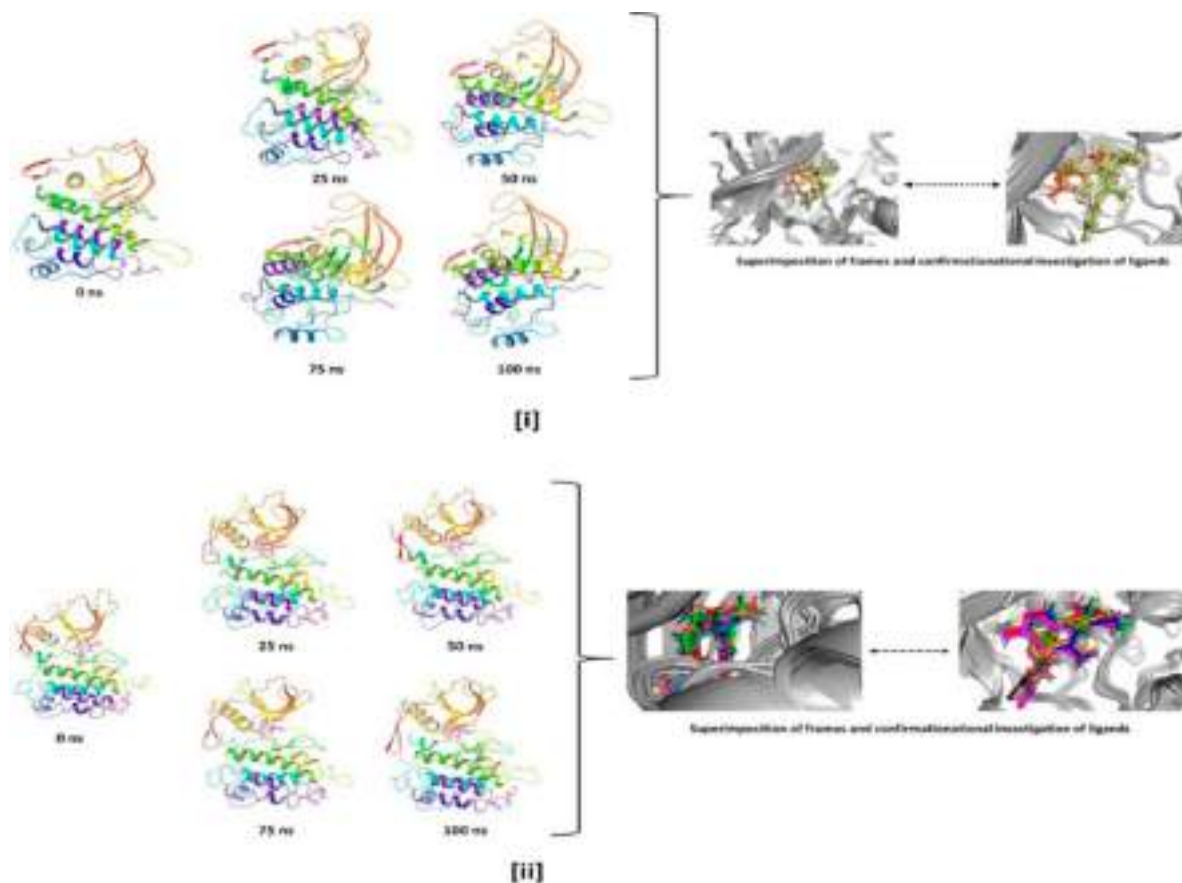


Figure 16. Stepwise trajectory analysis for every 25 ns displaying the protein and ligand conformation during 100 ns of simulation of [i] ALK-172 and [ii] ALK-178.

3.5. Molecular Mechanics Generalized Born and Surface Area (MMGBSA) Calculations and Energy Calculations

The MMGBSA technique is widely used to calculate the binding energy of ligands to protein molecules. With ALK, ligand 172 has the lowest binding energy of -49.3 kcal/mol, whereas 178 has a binding energy of -52.5 kcal/mol. The G_{bindvdW} , G_{bindLipo} , and $G_{\text{bindCoulomb}}$ energies contributed the most to the common binding energy of all kinds of interactions. G_{bind} is governed by non-bonded interactions such as $G_{\text{bindCoulomb}}$, $G_{\text{bindCovalent}}$, $G_{\text{bindHbond}}$, G_{bindLipo} , $G_{\text{bindSolvGB}}$, and G_{bindvdW} . Across all interactions, the G_{bindvdW} , G_{bindLipo} , and $G_{\text{bindCoulomb}}$ energies contributed the most to the average binding energy. On the other hand, the $G_{\text{bindSolvGB}}$ and $G_{\text{bindCovalent}}$ energies contributed the least to the final average binding energies. Furthermore, the $G_{\text{bindHbond}}$ interaction values of the 172-ALK and 178-ALK complexes indicated stable hydrogen bonds with amino acid residues. $G_{\text{bindSolvGB}}$ and $G_{\text{bindCovalent}}$ had negative energy contributions in all of the compounds, and so opposed binding. When coupled, $G_{\text{bindSolvGB}}$ and $G_{\text{bindCovalent}}$ verified adverse energy contributions and hence resisted binding. Figure 17 (left panel) shows that 172 and 178 at the ALK binding pocket experienced an angular shift of the angle (curved to straight) after post simulation at pre-simulation (0 ns) (100 ns) (see Figure 17). These conformational alterations result in improved binding pocket acquisition and engagement with residues, resulting in increased stability and binding energy (see Table 4).

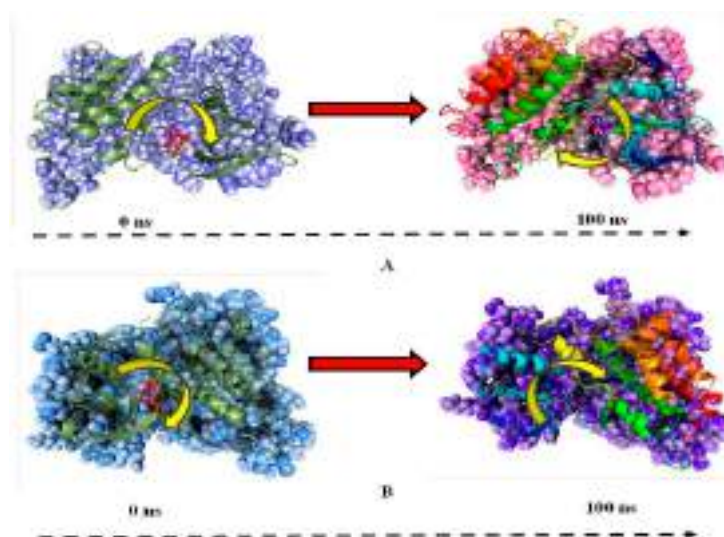


Figure 17. MMGBSA trajectory (0 ns, before simulation and 100 ns, after simulation) exhibited conformational changes upon binding the ligands with the protein, (A) ALK-172; (B) ALK-178. The arrows indicating the overall positional variation (movement and pose) of 172 and 178 at the binding site cavity.

Table 4. Binding energy calculation of 172 and 178 with ALK and non-bonded interaction energies from MMGBSA trajectories. (* indicates mean value of energy parameters).

Energies (kcal/mol) *	ALK-172	ALK-178
ΔG_{bind}	-49.4 ± 4.2	-52.6 ± 3.0
$\Delta G_{\text{bindLipo}}$	-17.4 ± 0.6	-19.5 ± 1.5
$\Delta G_{\text{bindvdW}}$	-41.1 ± 3.2	-44.8 ± 3.1
$\Delta G_{\text{bindCoulomb}}$	-9.1 ± 3.5	-5.7 ± 2.2
$\Delta G_{\text{bindHbond}}$	-1.4 ± 0.6	-0.3 ± 0.2
$\Delta G_{\text{bindSolvGB}}$	19.7 ± 3.1	18.2 ± 1.8
$\Delta G_{\text{bindCovalent}}$	1.0 ± 0.9	1.3 ± 1.2

As a result, it is possible that the 172 (See Figure 18A) and 178 (See Figure 18B) molecules have a high affinity for the primary target ALK. In ALK bounded 172 complex systems, the average total energy was -130.00 kcal/mol (green), while van der Waal's

energy (vdW) seemed to be merged over the total energy with an average energy of -30.00 kcal/mol, which was seen as the primary contributor to the stability of the ALK172 complex (cyan). Furthermore, coulombic interactions had a little impact on system stability, contributing an average energy of -101.00 kcal/mol (red) (See Figure 18). The energy profiles of the protein, ALK, and 178 complex systems were chosen to demonstrate the overall system's stability. In this regard, the Total Energy of the ALK-178 system has demonstrated to be completely stable, with an average total energy of -55.00 kcal/mol (dark green). However, van der Waal's energy (vdW) remained merged up-on the total energy with an average energy of -40 kcal/mol, taking into account as the primary contributor to the ALK-178 complex's stability (cyan). Furthermore, coulombic interactions performed a minimal influence in system stability, supplying an average energy of -10.00 kcal/mol (red) as seen in Figure 18.

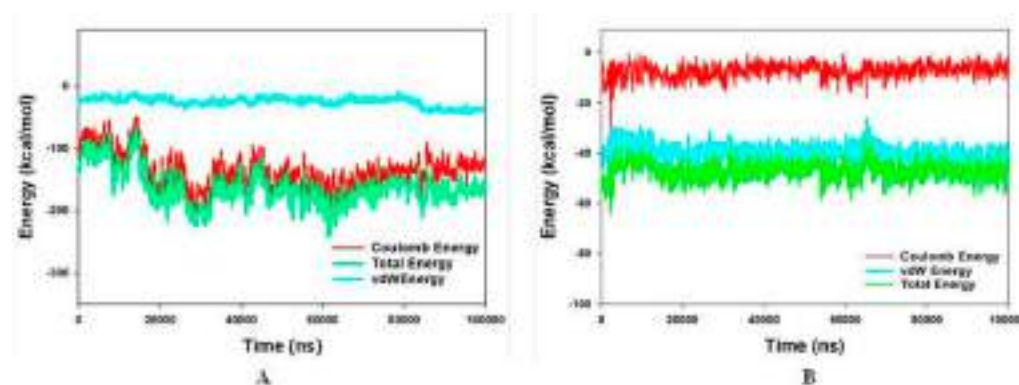


Figure 18. Energy plot of protein ALK with; (A) 172 complex system during the entire simulation event of 100 ns; (B) 178 complex system during the entire simulation event of 100 ns. The total energy (light green), van der Waal's energy (cyan) and coulomb energy (red) of the entire system indicating the stability of the individual systems bound to molecule.

Thus, MM-GBSA calculations resulted from MD simulation trajectories that were well justified with the binding energy obtained from docking results. Furthermore, the last frame (100 ns) of MMGBSA displayed the positional change of the 172 and 178 as compared to the 0 ns trajectory, indicating the better binding pose for best fitting in the protein's binding cavity (See Figure 19).

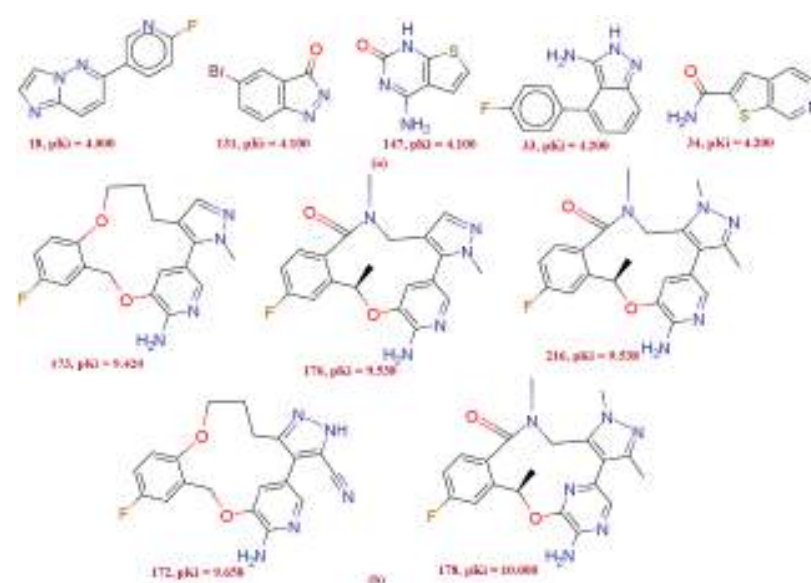


Figure 19. Variations in the K_i and chemical structure in the present dataset of ALK tyrosine kinase inhibitors: Five most active compounds (b), and five least active compounds (a) from the present series.

4. Materials and Methods

4.1. Selection of Data-Set

For the current study, 224 molecules with diverse structural features were selected due to the presence of different scaffolds and the substantial variation in the activity profile with an experimentally determined inhibition coefficient (K_i) for ALK tyrosine kinase [32–35]. K_i values ranging from 0.1 to 100,000 nM were changed to pK_i ($K_i = -\log K_i$) before actual QSAR evaluation for the ease of handling the data. The Figure 19 depicts five most active molecules followed by five least active molecules, indicating the heterogeneity of bio-activity and chemical properties. The Table 5 displays SMILES notations alongside ChEMBL id [36] and reported K_i and pK_i values for several sample compounds. (See Table S1 in Supplementary Materials displaying Sr no, ChEMBL id, smiles notation, K_i value, and pK_i values).

Table 5. Presentation of Serial number, ChEMBL ID, Smiles, pK_i and K_i value of 10 most active and 10 least active molecules in the dataset as representative examples only.

Sn	CHEMBL ID	Smiles	pK_i	K_i in nM
178	CHEMBL3286823	<chem>Cc1nn(C)c2c1-c1cnc(N)c(n1)O[C@H](C)c1cc(F)ccc1C(=O)N(C)C2</chem>	10	0.1
172	CHEMBL3286815	<chem>N#Cc1[nH]nc2c1-c1cnc(N)c(c1)OCc1cc(F)ccc1OCCC2</chem>	9.68	0.22
176	CHEMBL3286820	<chem>Cc1nn(C)c2c1-c1cnc(N)c(c1)O[C@H](C)c1cc(F)ccc1C(=O)N(C)C2</chem>	9.58	0.29
216	CHEMBL4286522	<chem>Cc1[nH][n+](C)c2c1-c1cnc(N)c(c1)O[C@H](C)c1cc(F)ccc1C(=O)N(C)C2</chem>	9.53	0.29
173	CHEMBL3286816	<chem>Cn1ncc2c1-c1cnc(N)c(c1)OCc1cc(F)ccc1OCCC2</chem>	9.42	0.38
161	CHEMBL3128064	<chem>Cc1nc(C(C)(C)O)sc1-c1cnc(N)c(O[C@H](C)c2cc(F)ccc2-n2nccn2)c1</chem>	9.39	0.4
181	CHEMBL3286832	<chem>C[C@H]1Oc2nc(cnc2N)-c2c(nc3ccc(C#N)cn23)CN(C)C(=O)c2ccc(F)cc21</chem>	9.25	0.56
174	CHEMBL3286818	<chem>C[C@H]1Oc2cc(cnc2N)-c2c(nn(C)c2C#N)CCOc2ccc(F)cc21</chem>	9.24	0.57
175	CHEMBL3286819	<chem>C[C@H]1Oc2cc(cnc2N)-c2c(nn(C)c2C#N)CCCOc2ccc(F)cc21</chem>	9.20	0.61
179	CHEMBL3286830	<chem>C[C@H]1Oc2cc(cnc2N)-c2c(nn(C)c2C#N)CN(C)C(=O)c2ccc(F)cc21</chem>	9.1	0.7
110	CHEMBL1995765	<chem>Nc1cc(=O)[nH]n1-c1ccccn1</chem>	4.5	31,622.7
13	CHEMBL1972934	<chem>Nc1ncnc2sccc12</chem>	4.4	39,810.7
39	CHEMBL1975212	<chem>Nc1ncnc2scc(-c3ccccc3)c12</chem>	4.4	39,810.7
48	CHEMBL1949855	<chem>O=c1[nH]cnc2c(Cl)cccc12</chem>	4.4	39,810.7
107	CHEMBL1994159	<chem>CC(=O)c1cccc(-c2ccc3nccn3n2)c1</chem>	4.3	50,118.7
129	CHEMBL2000879	<chem>c1ccc(C2CCc3[nH]nc3C2)cc1</chem>	4.3	50,118.7
33	CHEMBL1971519	<chem>Nc1n[nH]c2cccc(-c3ccc(F)cc3)c12</chem>	4.2	63,095.7
34	CHEMBL1971534	<chem>NC(=O)c1cc2ccncc2s1</chem>	4.2	63,095.7
50	CHEMBL1975921	<chem>O=c1[nH]c2cc(Br)cnc2[nH]1</chem>	4.2	63,095.7
131	CHEMBL2007097	<chem>Nc1nc(=O)[nH]c2sccc12</chem>	4.1	79,432.8

4.2. Molecular Structure Drawing and Optimization

The complete 224 molecules' 2D structures were drawn using free and open source software's ChemSketch 12 Freeware (<https://www.acdlabs.com/resources/free-chemistry->

[software-apps/chemsketch-freeware/](#) accessed on 2 March 2022, version 2021), while their 3D structures were generated using Open Babel 2.4, respectively. Furthermore, optimization of the full dataset molecules was achieved using the MMFF94 force field provided in TINKER (default settings), whilst Open3DAlign was used for molecular alignment, respectively [37].

4.3. Molecular Descriptor Calculation and Objective Feature Selection (OFS)

PyDescriptor, which is available as a plugin in the PyMOL 2.5 software application, was used to calculate descriptors for each molecule [38]. Molecular descriptors with almost constant values (>95 percent) and co-linearity ($|R|$) greater than 0.95 were eliminated using objective feature selection (OFS) stability among QSARINS v2.2.4 [39]. This approach removed unnecessary molecular descriptors that impact multi-collinear and mock variables in the GA-MLR model. As a result, following OFS treatment, about 3339 molecular descriptors were separated in order to develop QSAR models.

4.4. Subjective Feature Selection, QSAR Model—Development and Validation

The condensed pool of computed molecular descriptors includes 1D- and 3D-descriptors, as well as molecular properties or value descriptors, and so on. Coinciding with a huge molecule, followed by a vivid gap. After developing robust QSAR models, the Subjective Feature Selection (SFS) function in QSARINS v2.2.4 is used to run Genetic Algorithm (GA) based multi linear regression (MLR). QSAR models were developed in accordance with OECD guidelines and were then subjected to extensive internal or external statistical validation, Y-scrambling, or Applicability domain examination. The following steps are included in the QSAR model development practice. The whole dataset was utilised to create QSAR models, which were mostly based on the undivided (training set) dataset, however in this study, we are presenting one univariate divided set and another divided set multivariate QSAR models

1. The QSAR techniques have been anticipated to use a loosely split operation in QSARINSv2 software v2.2.4 based on a divided dataset. It divided a given dataset into 80% training (180 molecules in the training set) and 20% prediction (44 molecules in prediction set). The 180 molecules from the training set were used to generate the QSAR model, and external validation was completed on 44 compounds from the prediction set.
2. The QSARINS software v2.2.4 program was used to construct GA-MLR mainly based QSAR models, incorporating default parameters. Q^2_{LOO} is utilised as a fitness parameter to accomplish subjective feature selection. While doing SFS, the Q^2_{LOO} value was extraordinarily prolonged up to the four variables, but an insignificant uplift was seen after that. Thus, in order to keep the QSAR model from over-fitting, SFS was previously limited to a set of four descriptors. This resulted in the creation of simple and predictive QSAR models. (See Supplementary Materials Table S2 values for the selected four molecular descriptors present in QSAR models).

An important aspect of developing a good QSAR model with minimal over-fitting and appropriate interpretability is to have an enough number of molecular descriptors in QSAR the model. In the present study, a plan (see Figure 20) was projected in the large range of molecular descriptors included among the model yet R^2_{tr} and Q^2_{LOO} values to get the so-called breaking point. As a result, the variety of chemical descriptors related to the breakdown point used to be prioritized for model construction. Figure 20 shows that the breakage point correlates with four different factors. As a result, QSAR models with more than four descriptors were rejected.

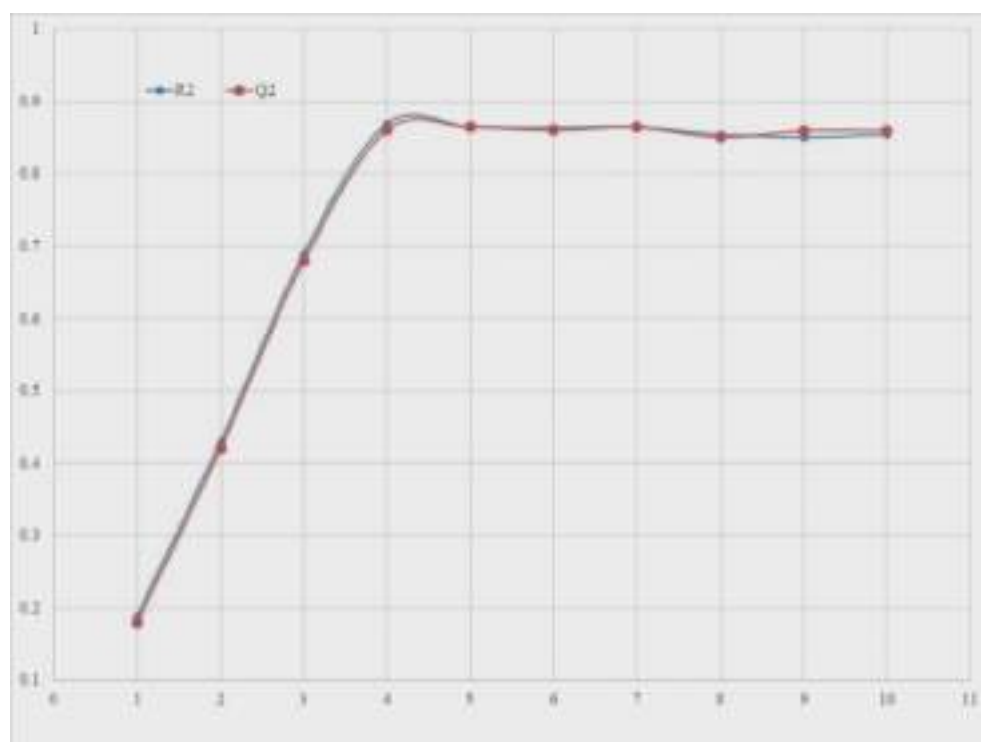


Figure 20. Plot of number of descriptors against Coefficient of Determination R^2 and Leave-One-Out Coefficient of Determination Q^2 to identify the optimum number of descriptors.

To perform acceptable validation, QSARINS v2.2.4 was used to do (a) leave-one-out (LOO) or leave-many-out (LMO) parameter-based internal validation; (b) external validation; (c) Y-scrambling or model applicability domain (AD) analysis in accordance with OECD requirements. The robustness of the GA-MLR-based QSAR model was previously assessed on the basis of how well the various statistical parameters perform on the respective starting value. Two QSAR models (1.1 and 1.2) consisting of univariate divided set and multivariate divided set models with excellent values on these parameters and best predictive capability were chosen for the analysis, but the rest of the QSAR models failed to fulfil some of these factors above-mention values and were omitted [35].

4.5. Molecular Docking

The protein data bank provided the ALK tyrosine kinase pdb file (pdb id-5fto, Resolution 1.7). For its X-ray resolution or sequel completeness, the pdb 5fto was carefully chosen. The optimized protein is appropriate for docking analysis. The protein preparation was carried by UCF chimera-1.16-win64 software (See Supplementary Materials for the detail procedure of protein preparation by chimera-1.16 software). Before the docking investigation, the natural ligand (Entrectinib) was removed. In the present study, the binding site for native ligand, namely the active site, has been studied. As a result, the compounds were docked between the active site, where the native ligand was originally bound along ALK tyrosine kinase, and the docking posture for the most active molecules 172 and 178 is shown below for convenience. The NRGSuite programme (PyMOL versions 1.2 and above) was used to do the molecular docking investigation. Because this is a free and open source software program, it may also be utilised as a PyMOL plugin [35]. It detects surface holes in a protein and uses them as target binding sites for docking simulations with the help of FlexAID [40]. It employs a genetic algorithm for function conformational search, model ligand and side-chain flexibility, and allows for covalent binding simulation. To achieve substantial performance with NRGSuite, the flexible-rigid docking approach was employed in conjunction with the following default settings: Because of the binding sites, the input technique is spherical (diameter: 17); the spacing on the three-dimensional grid is

0.385; facet band flexibility is no; ligand flexibility is yes; ligand posture is no; restrictions are no. Hetero groups-cloud molecules included; van der Waals permeability -0.1 ; solvent types-none; variation on chromosomes—1000; variation on generations—1000; fitness model-share; copy model-population boom; and variation on top complexes—5. After the docking process was validated, the molecule Entrectinib, a discovered tyrosine kinase inhibitor, was employed for validation.

4.6. MD Simulation Analysis

The virtual screening findings are utilised to evaluate the most active Molecule 178 with a docking score of -7.8 kcal/mol and Molecule 172 (-8.0 Kcal/mol) in molecular dynamics and simulation using the Schrodinger Desmond versus 2020.1 (MD simulation). The SPC (Simple factor charge) model was utilised to bind protein ligands using the docking complexes Molecule 178 and Molecule 172. In this system, the OPLS-2005 pressure subject and explicit solvent model with SPC water molecules were applied. To neutralise the charge, Na⁺ ions were added [41]. To imitate the physiological environment, 0.15 M NaCl alternatives are provided to the computer [42]. The Nose–Hoover chain coupling approach was employed to build up the NPT ensemble with temperature 300 K, leisure time of 1.0 ps, and pressure 1 bar, which was once as soon as maintained in all simulations. A 2 fs time step will be employed. The barostat approach with the Martyna–Tuckerman–Klein chain coupling scheme [43] was originally utilised for pressure control with a leisure time of 2 ps. The particle mesh Ewald technique [44] was used to calculate long-range electrostatic interactions with a radius of 9 for Coulomb interactions. The non-bonded forces were estimated using the RESPA integrator. The root mean square deviation (RMSD), root mean square fluctuation (RMSF), radius of gyration (Rg), and protein ligand interactions were assessed to have a check at the stability of the complex in MD simulations.

4.7. Molecular Mechanics Generalized Born and Surface Area (MMGBSA) Calculations

The binding free energy (G_{bind}) of docked complexes was determined using the molecular mechanics generalized born surface region (MM-GBSA) module in MD simulations comprising 5fto bonded with the most active molecule 178 and the molecule 172. (Schrodinger suite, LLC, New York, NY, USA, 2017-4). At around the same time, the binding free energy was estimated using the OPLS 2005 force field, the VSGB solvent model and rotamer search techniques [45]. Following the MD run, the MD trajectories frames were chosen at 10 ns intervals. The total free energy binding used to be calculated the usage of Equation (1):

$$\Delta G_{\text{bind}} = G_{\text{complex}} - (G_{\text{protein}} + G_{\text{ligand}}) \quad (3)$$

where,

ΔG_{bind} = binding free energy,

G_{complex} = free energy of the complex,

G_{protein} = free energy of the target protein, and

G_{ligand} = free energy of the ligand.

5. Conclusions

A cheminformatics technique was used effectively in the current investigation to predict ALK Tyrosine kinase inhibitory activity in order to uncover fundamental structural aspects important for anticancer activity. Two statistically robust univariate and four parametric QSAR models with exceptional external predictive capability were built, and the right number of molecular features were accurately positioned. The QSAR analysis effectively identified a combination based on previously unknown Pharmacophoric properties. The existence of fluorine atoms on the phenyl ring, as well as the presence of planar nitrogen atoms, must be retained in future drug design, coupled with some novel Pharmacophoric qualities such as *rsa*. The molecular descriptors identified in the developed QSAR models, such as the ratio of surface area (*rsa*), planar nitrogen within four bonds from the nitrogen

atom, fluorine atom within five bonds from the non-ring oxygen atom, lipophilic atoms within two bonds from the ring carbon atoms, and so on, can potentially enhance the ALK Tyrosine kinase inhibition potency. QSAR and molecular docking studies have successfully identified certain significant Pharmacophoric traits, such as the presence of an extra carbon atom in molecule 172, which results in increased flexibility and *rsa*; comparison of molecule 176 with 178 reveals an interesting influence of the N_PlanN_4B and ringC_lipo_2Bon docking pose and activity profile; and reversal in docking conformation and differences in the binding affinities of compounds 176 and 178. Moreover, identification of the additional polar nitrogen in QSAR analysis responsible for increasing the polarity of the molecule. This observation reveals that the QSAR and docking results are completely consistent with one another. The molecular docking studies on the 172 and 178 with the ALK tyrosine kinase receptor revealed that these compounds anchored to the ALK tyrosine kinase along the orientation or position extremely close to co-crystallized ligand; Entrectinib that resulted from crystallographic analysis of the ALK tyrosine kinase protein including its actual ligand. As a result, the created QSAR models meet the threshold values for several statistical parameters required to get the accuracy and applicability of a QSAR model. As a result, the obtained QSAR models include an appropriate mix of quantitative and qualitative characteristics. The pharmacophoric properties found in QSAR models show tremendous potential for optimizing dataset compounds in accordance with more potent ALK tyrosine kinase inhibitors as anticancer leads. Furthermore, MD simulation and binding free energy analyses support the findings of the QSAR and molecular docking studies.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/molecules27154951/s1>, Table S1: The SMILES notation for two hundred twenty four (224) ALK tyrosine kinase leads, along with their reported K_i and pK_i values. Table S2: The values for selected molecular descriptors present in QSAR models. Table S3: Details regarding performance of Univariate Divided set model 1.1. containing Different graphs associated with model, Table S4: Details regarding performance of Multivariate divided set model comprising Different graphs associated with model 1.2. Statistical parameters for used for validation of QSAR models. File S5: Description of the performance parameters in QSARINS.

Author Contributions: R.D.J.: Conceptualization, Data Curation and Draft Review and Editing, P.S.: Formal Analysis, Editing, N.J.: Formal Analysis, A.G. (Ajaykumar Gandhi): Review and editing, N.M.: Discussion, Editing, Formal Analysis, A.A.A.-M.: Review, Discussion and Editing, M.E.A.Z.: Review, Discussion and Editing, S.A.A.-H.: Formal Analysis, A.S.: Formal Analysis, V.H.M.: Discussion and Formal Analysis, A.G. (Arabinda Gosh): Discussion, Review and Editing, R.L.B.: Formal Analysis, Editing. All authors have read and agreed to the published version of the manuscript.

Funding: The authors Magdi E.A. Zaki, Sami A. Al-Hussain, and Aamal A. Al-Mutairi extend their sincere thanks the Deanship of Scientific Research, Imam Mohammad Ibn Saud Islamic University, Saudi Arabia, for its support of this research through the Research Group No RG-21-09-76.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data is available in the Supplementary Materials section.

Acknowledgments: The authors are thankful to Paola Gramatica and her team for providing QSARINS-v2.2.4 and developers of TINKER, ChemSketch 12 Freeware (ACD labs), and PyDescriptor for providing the free versions of their software. Rahul Jawarkar are grateful to the Rajendra Gode Institute of Pharmacy, Amravati, Maharashtra, India. Magdi E.A. Zaki and Sami A. Al-Hussain are thankful to the Deanship of Scientific Research at Imam Mohammad Ibn Saud Islamic University, Riyadh, KSA, for its support of this research through the Research group number 21-09-77.

Conflicts of Interest: The authors declare no conflict of interest.

Sample Availability: Samples of the compounds are available from the authors.

References

1. Hallberg, B.; Palmer, R.H. The role of the ALK receptor in cancer biology. *Ann. Oncol.* **2016**, *27*, iii4–iii15. [[CrossRef](#)]
2. Morris, S.W.; Kirstein, M.N.; Valentine, M.B.; Dittmer, K.G.; Shapiro, D.N.; Saltman, D.L.; Look, A.T. Fusion of a kinase gene, ALK, to a nucleolar protein gene, NPM, in non-Hodgkin's lymphoma. *Science* **1994**, *263*, 1281–1284. [[CrossRef](#)] [[PubMed](#)]
3. Holla, V.R.; Elamin, Y.Y.; Bailey, A.M.; Johnson, A.M.; Litztenburger, B.C.; Khotskaya, Y.B.; Sanchez, N.S.; Zeng, J.; Shufean, M.A.; Shaw, K.R.; et al. ALK: A tyrosine kinase target for cancer therapy. *Cold Spring Harb. Mol. Case Stud.* **2017**, *3*, a001115. [[CrossRef](#)]
4. Della Corte, C.M.; Viscardi, G.; Di Liello, R.; Fasano, M.; Martinelli, E.; Troiani, T.; Ciardiello, F.; Morgillo, F. Role and targeting of anaplastic lymphoma kinase in cancer. *Mol. Cancer* **2018**, *17*, 30. [[CrossRef](#)]
5. Kumar, N.M.; Mathew, M.; Anila, K.N.; Priyanka, S. A review on newer tyrosine kinase inhibitors and their uses. *J. Clin. Diagn. Res.* **2018**, *12*, XE01–XE06. [[CrossRef](#)]
6. Lin, J.J.; Riely, G.J.; Shaw, A.T. Targeting ALK: Precision Medicine Takes on Drug Resistance. *Cancer Discov.* **2017**, *7*, 137–155. [[CrossRef](#)] [[PubMed](#)]
7. Shaw, A.T.; Kim, D.-W.; Nakagawa, K.; Seto, T.; Crinó, L.; Ahn, M.-J.; De Pas, T.; Besse, B.; Solomon, B.J.; Blackhall, F.; et al. Crizotinib versus Chemotherapy in Advanced ALK-Positive Lung Cancer. *N. Engl. J. Med.* **2013**, *368*, 2385–2394. [[CrossRef](#)] [[PubMed](#)]
8. Solomon, B.J.; Mok, T.; Kim, D.-W.; Wu, Y.-L.; Nakagawa, K.; Mekhail, T.; Felip, E.; Cappuzzo, F.; Paolini, J.; Usari, T.; et al. First-line crizotinib versus chemotherapy in ALK-positive lung cancer. *N. Engl. J. Med.* **2014**, *371*, 2167–2177. [[CrossRef](#)]
9. Shaw, A.T.; Kim, D.W.; Mehra, R.; Tan, D.S.; Felip, E.; Chow, L.Q.; Camidge, D.R.; Vansteenkiste, J.; Sharma, S.; De Pas, T.; et al. Ceritinib in ALK-rearranged non-small-cell lung cancer. *N. Engl. J. Med.* **2014**, *370*, 1189–1197. [[CrossRef](#)] [[PubMed](#)]
10. Kim, D.-W.; Mehra, R.; Tan, D.S.W.; Felip, E.; Chow, L.Q.M.; Camidge, D.R.; Vansteenkiste, J.; Sharma, S.; De Pas, T.; Riely, G.J.; et al. Activity and safety of ceritinib in patients with ALK-rearranged non-small-cell lung cancer (ASCEND-1): Updated results from the multicentre, open-label, phase 1 trial. *Lancet Oncol.* **2016**, *17*, 452–463. [[CrossRef](#)]
11. Ou, S.-H.I.; Ahn, J.S.; De Petris, L.; Govindan, R.; Yang, J.C.-H.; Hughes, B.; Lena, H.; Moro-Sibilot, D.; Bearz, A.; Ramirez, S.V.; et al. Alectinib in crizotinib-refractory ALK-rearranged non-small-cell lung cancer: A phase II global study. *J. Clin. Oncol.* **2015**, *34*, 661–668. [[CrossRef](#)] [[PubMed](#)]
12. Shaw, A.T.; Gandhi, L.; Gadgeel, S.; Riely, G.J.; Cetnar, J.; West, H.; Camidge, D.R.; Socinski, M.A.; Chiappori, A.; Mekhail, T.; et al. Alectinib in ALK-positive, crizotinib-resistant, non-small-cell lung cancer: A single-group, multicentre, phase 2 trial. *Lancet Oncol.* **2016**, *17*, 234–242. [[CrossRef](#)]
13. Katayama, R.; Khan, T.M.; Benes, C.; Lifshits, E.; Ebi, H.; Rivera, V.M.; Shakespeare, W.C.; Iafrate, A.J.; Engelman, J.A.; Shaw, A.T. Therapeutic strategies to overcome crizotinib resistance in non-small cell lung cancers harboring the fusion oncogene EML4-ALK. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 7535–7540. [[CrossRef](#)] [[PubMed](#)]
14. Johnson, T.W.; Richardson, P.F.; Bailey, S.; Brooun, A.; Burke, B.J.; Collins, M.R.; Cui, J.J.; Deal, J.G.; Deng, Y.L.; Dinh, D.; et al. Discovery of (10R)-7-Amino-12-fluoro-2,10,16-trimethyl-15-oxo-10,15,16,17-tetrahydro-2H-8,4-(metheno)pyrazolo[4,3-h][2,5,11]-benzoxadiazacyclotetradecine-3-carbonitrile (PF-06463922), a macrocyclic inhibitor of anaplastic lymphoma kinase (ALK) and c-ros oncogene 1 (ROS1) with preclinical brain exposure and broad-spectrum potency against ALK-resistant mutations. *J. Med. Chem.* **2014**, *57*, 4720–4744. [[PubMed](#)]
15. Shaw, A.T.; Yeap, B.Y.; Solomon, B.J.; Riely, G.J.; Gainor, J.; Engelman, J.A. Effect of crizotinib on overall survival in patients with advanced non-small-cell lung cancer harbouring ALK gene rearrangement: A retrospective analysis. *Lancet Oncol.* **2011**, *12*, 4–12. [[CrossRef](#)]
16. Bellacasa, R.P.; Karachaliou, N.; Estrada-Tejedor, R.; Teixidó, J.; Costa, C.; Borrell, J.I. ALK and ROS1 as a joint target for the treatment of lung cancer: A review. *Transl. Lung Cancer Res.* **2013**, *2*, 72–86.
17. Roskoski, R., Jr. Anaplastic lymphoma kinase (ALK): Structure, oncogenic activation, and pharmacological inhibition. *Pharmacol. Res.* **2013**, *68*, 68–94. [[CrossRef](#)] [[PubMed](#)]
18. Cherkasov, A.; Muratov, E.N.; Fourches, D.; Varnek, A.; Baskin, I.I.; Cronin, M.; Dearden, J.; Gramatica, P.; Martin, Y.C.; Todeschini, R.; et al. QSAR modeling: Where have you been? Where are you going to? *J. Med. Chem.* **2014**, *57*, 4977–5010. [[CrossRef](#)] [[PubMed](#)]
19. Fujita, T.; Winkler, D.A. Understanding the Roles of the “two QSARs”. *J. Chem. Inf. Model.* **2016**, *56*, 269–274. [[CrossRef](#)] [[PubMed](#)]
20. Huang, J.; Fan, X. Why QSAR fails: An empirical evaluation using conventional computational approach. *Mol. Pharm.* **2011**, *8*, 600–608. [[CrossRef](#)] [[PubMed](#)]
21. Chirico, N.; Gramatica, P. Real external predictivity of QSAR models. Part 2. New intercomparable thresholds for different validation criteria and the need for scatter plot inspection. *J. Chem. Inf. Model.* **2012**, *52*, 2044–2058. [[CrossRef](#)] [[PubMed](#)]
22. Gramatica, P.; Cassani, S.; Roy, P.P.; Kovarich, S.; Yap, C.W.; Papa, E. QSAR modeling is not “Push a button and find a correlation”: A case study of toxicity of (Benzo-)triazoles on Algae. *Mol. Inform.* **2012**, *31*, 817–835. [[CrossRef](#)] [[PubMed](#)]
23. Martin, T.M.; Harten, P.; Young, D.M.; Muratov, E.N.; Golbraikh, A.; Zhu, H.; Tropsha, A. Does rational selection of training and test sets improve the outcome of QSAR modeling? *J. Chem. Inf. Model.* **2012**, *52*, 2570–2578. [[CrossRef](#)]
24. Masand, V.H.; Mahajan, D.T.; Nazeruddin, G.M.; Hadda, T.B.; Rastija, V.; Alfeefy, A.M. Effect of information leakage and method of splitting (rational and random) on external predictive ability and behavior of different statistical parameters of QSAR model. *Med. Chem. Res.* **2015**, *24*, 1241–1264. [[CrossRef](#)]
25. Gramatica, P. On the development and validation of QSAR models. *Methods Mol. Biol.* **2013**, *930*, 499–526.

26. Iwahara, T.; Fujimoto, J.; Wen, D.; Cupples, R.; Bucay, N.; Arakawa, T. Molecular characterization of ALK, a receptor tyrosine kinase expressed specifically in the nervous system. *Oncogene* **1997**, *14*, 439–449. [[CrossRef](#)]
27. Morris, S.W.; Naeve, C.; Mathew, P.; James, P.L.; Kirstein, M.N.; Cui, X.; Witte, D.P. ALK, the chromosome 2 gene locus altered by the t(2;5) in non-Hodgkin's lymphoma, encodes a novel neural receptor tyrosine kinase that is highly related to leukocyte tyrosine kinase (LTK). *Oncogene* **1997**, *14*, 2175–2188. [[CrossRef](#)] [[PubMed](#)]
28. Loren, C.E.; Scully, A.; Grabbe, C.; Edeen, P.T.; Thomas, J.; McKeown, M. Identification and characterization of DAlk: A novel *Drosophila melanogaster* RTK which drives ERK activation in vivo. *Genes Cells* **2001**, *6*, 531–544. [[CrossRef](#)] [[PubMed](#)]
29. Huang, S.Y. Comprehensive assessment of flexible-ligand docking algorithms: Current effectiveness and challenges. *Brief. Bioinform.* **2018**, *19*, 982–994. [[CrossRef](#)]
30. Menichincheri, M.; Ardini, E.; Magnaghi, P.; Avanzi, N.; Banfi, P.; Bossi, R.; Buffa, L.; Canevari, G.; Ceriani, L.; Colombo, M.; et al. Discovery of Entrectinib: A new 3-aminoindazole as a potent anaplastic lymphoma kinase (ALK), c-ros oncogene 1 kinase (ROS1), and pan-tropomyosin receptor kinases (Pan-TRKs) inhibitor. *J. Med. Chem.* **2016**, *59*, 3392–3408. [[CrossRef](#)]
31. Gramatica, P. Principles of QSAR Modeling. *Int. J. Quant. Struct. Prop. Relatsh.* **2020**, *5*, 61–97. [[CrossRef](#)]
32. Zaki, M.E.A.; Al-Hussain, S.A.; Masand, V.H.; Akasapu, S.; Lewaa, I. QSAR and Pharmacophore Modeling of Nitrogen Heterocycles as Potent Human N-Myristoyltransferase (Hs-NMT) Inhibitors. *Molecules* **2021**, *26*, 1834. [[CrossRef](#)]
33. Consonni, V.; Ballabio, D.; Todeschini, R. Comments on the definition of the Q2 parameter for QSAR validation. *J. Chem. Inf. Model.* **2009**, *49*, 1669–1678. [[CrossRef](#)]
34. Consonni, V.; Todeschini, R.; Ballabio, D.; Grisoni, F. On the Misleading Use of QF32 for QSAR Model Comparison. *Mol. Inform.* **2019**, *38*, 1800029. [[CrossRef](#)]
35. Chirico, N.; Gramatica, P. Real external predictivity of QSAR models: How to evaluate It? Comparison of different validation criteria and proposal of using the concordance correlation coefficient. *J. Chem. Inf. Model.* **2011**, *51*, 2320–2335. [[CrossRef](#)] [[PubMed](#)]
36. Gaulton, A.; Hersey, A.; Nowotka, M.; Bento, A.P.; Chambers, J.; Mendez, D.; Motow, P.; Atkinson, F.; Bellis, L.J.; Cibrián-Uhalte, E.; et al. The ChEMBL database in 2017. *Nucleic Acids Res.* **2017**, *45*, D945–D954. [[CrossRef](#)] [[PubMed](#)]
37. Tosco, P.; Balle, T.; Shiri, F. Open3DALIGN: An open-source software aimed at unsupervised ligand alignment. *J. Comput. Aided Mol. Des.* **2011**, *25*, 777–783. [[CrossRef](#)]
38. Masand, V.H.; Rastija, V. PyDescriptor: A new PyMOL plugin for calculating thousands of easily understandable molecular descriptors. *Chemom. Intell. Lab. Syst.* **2017**, *169*, 12–18. [[CrossRef](#)]
39. Gramatica, P.; Chirico, N.; Papa, E.; Cassani, S.; Kovarich, S. QSARINS: A new software for the development, analysis, and validation of QSAR MLR models. *J. Comput. Chem.* **2013**, *34*, 2121–2132. [[CrossRef](#)]
40. Gaudreault, F.; Morency, L.P.; Najmanovich, R.J. NRGsuite: A PyMOL plugin to perform docking simulations in real time using FlexAID. *Bioinformatics* **2015**, *31*, 3856–3858. [[CrossRef](#)]
41. Bowers, K.J.; Chow, D.E.; Xu, H.; Dror, R.O.; Eastwood, M.P.; Gregersen, B.A.; Klepeis, J.L.; Kolossvary, I.; Moraes, M.A.; Sacerdoti, F.D.; et al. Scalable algorithms for molecular dynamics simulations on commodity clusters. In Proceedings of the SC '06: 2006 ACM/IEEE Conference on Supercomputing, Tampa, FL, USA, 11–17 November 2006; IEEE: New York, NY, USA, 2006; p. 43. [[CrossRef](#)]
42. Jorgensen, W.L.; Maxwell, D.S.; Tirado-Rives, J. Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *J. Am. Chem. Soc.* **1996**, *118*, 11225–11236. [[CrossRef](#)]
43. Shivakumar, D.; Williams, J.; Wu, Y.; Damm, W.; Shelley, J.; Sherman, W. Prediction of Absolute Solvation Free Energies using Molecular Dynamics Free Energy Perturbation and the OPLS Force Field. *J. Chem. Theory Comput.* **2010**, *6*, 1509–1519. [[CrossRef](#)] [[PubMed](#)]
44. Martyna, G.J.; Klein, M.L.; Tuckerman, M. Nosé–Hoover chains: The canonical ensemble via continuous dynamics. *J. Chem. Phys.* **1992**, *97*, 2635–2643. [[CrossRef](#)]
45. Ylilauri, M.; Pentikäinen, O.T. MMGBSA as a Tool to Understand the Binding Affinities of Filamin–Peptide Interactions. *J. Chem. Inf. Model.* **2013**, *53*, 2626–2633. [[CrossRef](#)] [[PubMed](#)]

Article

QSAR Evaluations to Unravel the Structural Features in Lysine-Specific Histone Demethylase 1A Inhibitors for Novel Anticancer Lead Development Supported by Molecular Docking, MD Simulation and MMGBSA

Rahul D. Jawarkar ^{1,*}, Ravindra L. Bakal ^{1,†}, Nobendu Mukherjee ^{2,3,†}, Arabinda Ghosh ^{4,†}, Magdi E. A. Zaki ^{5,*}, Sami A. AL-Hussain ⁵, Aamal A. Al-Mutairi ⁵, Abdul Samad ⁶, Ajaykumar Gandhi ^{7,†} and Vijay H. Masand ^{8,†}

- ¹ Department of Medicinal Chemistry and Drug Discovery, Dr Rajendra Gode Institute of Pharmacy, University Mardi Road, Amravati 444603, India; rlbakal@gmail.com
- ² Department of Microbiology, Ramakrishna Mission Vivekananda Centenary College, Kolkata 700118, India; nabendu21@rkmvccrahara.org
- ³ Department of Health Sciences, Novel Global Community Educational Foundation, Hebersham, NSW 2770, Australia
- ⁴ Microbiology Division, Department of Botany, Gauhati University, Guwahati 781014, India; dra.ghosh@gauhati.ac.in
- ⁵ Department of Chemistry, Faculty of Science, Imam Mohammad Ibn Saud Islamic University, Riyadh 13318, Saudi Arabia; sahusain@imamu.edu.sa (S.A.A.-H.); aamutairi@imamu.edu.sa (A.A.A.-M.)
- ⁶ Department of Pharmaceutical Chemistry, Faculty of Pharmacy, Tishk International University, Erbil 44001, Iraq; abdul.samad@tiu.edu.iq
- ⁷ Department of Chemistry, Government Arts and Science College, Karur 639005, India; gacjay18@gmail.com
- ⁸ Department of Chemistry, Vidyabharati Mahavidyalaya, Camp, Amravati 444602, India; vijaymasand@gmail.com
- * Correspondence: rahuljawarkar@gmail.com (R.D.J.); mezaki@imamu.edu.sa (M.E.A.Z.)
- † These authors contributed equally to this work.



Citation: Jawarkar, R.D.; Bakal, R.L.; Mukherjee, N.; Ghosh, A.; Zaki, M.E.A.; AL-Hussain, S.A.; Al-Mutairi, A.A.; Samad, A.; Gandhi, A.; Masand, V.H. QSAR Evaluations to Unravel the Structural Features in Lysine-Specific Histone Demethylase 1A Inhibitors for Novel Anticancer Lead Development Supported by Molecular Docking, MD Simulation and MMGBSA. *Molecules* **2022**, *27*, 4758. <https://doi.org/10.3390/molecules27154758>

Academic Editor: Athanassios C. Tsipis

Received: 14 June 2022

Accepted: 19 July 2022

Published: 25 July 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Abstract: Using 84 structurally diverse and experimentally validated LSD1/KDM1A inhibitors, quantitative structure–activity relationship (QSAR) models were built by OECD requirements. In the QSAR analysis, certainly significant and understated pharmacophoric features were identified as critical for LSD1 inhibition, such as a ring Carbon atom with exactly six bonds from a Nitrogen atom, partial charges of lipophilic atoms within eight bonds from a ring Sulphur atom, a non-ring Oxygen atom exactly nine bonds from the amide Nitrogen, etc. The genetic algorithm–multi-linear regression (GA-MLR) and double cross-validation criteria were used to create robust QSAR models with high predictability. In this study, two QSAR models were developed, with fitting parameters like $R^2 = 0.83–0.81$, $F = 61.22–67.96$, internal validation parameters such as $Q^2_{LOO} = 0.79–0.77$, $Q^2_{LMO} = 0.78–0.76$, $CCC_{cv} = 0.89–0.88$, and external validation parameters such as, $R2_{ext} = 0.82$ and $CCC_{ex} = 0.90$. In terms of mechanistic interpretation and statistical analysis, both QSAR models are well-balanced. Furthermore, utilizing the pharmacophoric features revealed by QSAR modelling, molecular docking experiments corroborated with the most active compound's binding to the LSD1 receptor. The docking results are then refined using Molecular dynamic simulation and MMGBSA analysis. As a consequence, the findings of the study can be used to produce LSD1/KDM1A inhibitors as anticancer leads.

Keywords: LSD1; KDM1A; QSAR; anticancer; molecular docking; MD simulation; genetic algorithm–multi linear regression; MMGBSA

1. Introduction

Lysine-specific histone demethylase 1A (LSD1), also known as lysine (K)-specific demethylase 1A (KDM1A), is a crucial member of the monoamine oxidases family. LSD1

catalyzes two important and completely opposing enzymatic reactions with flavin adenine dinucleotide (FAD) as a cofactor: transcription repression via de-methylation at histone 3 lysine 4 methyl 1/2 (H3K4me1/2) and transcription activation via de-methylation at histone 3 lysine 9 methyl 1/2 (H3K9me1/2) [1]. LSD1 is also involved in the de-methylation of TP53, E2F1, and DNMT1 [2]. The typical healthy physiological condition is characterized by regimented epigenetic control of cyclic cellular processes, including rejuvenation, differentiation, and proliferation. LSD1 modulates differentiation and proliferation pathways in highly proliferative and widely metastatic small-cell lung cancer (SCLC) [3]. LSD1 overexpression is also found in other cancer types, including prostate cancer, breast cancer, colorectal cancer, and neuroblastoma [1–3].

According to reports, if LSD1 overexpression is reduced in several forms of blood malignancies (haematological sarcomas) such as leukaemia, multiple myeloma, and solid tumours, cell differentiation is also reduced. This makes LSD1 an appealing target for anticancer medication development [1].

There have been several reports of reversible LSD1 inhibitors to date [4]. Sorna et al. used high-throughput virtual screening to identify reversible LSD1 inhibitors; however, these drugs appeared to have considerable off-target and nonspecific effects [4,5]. Ma et al. found several pyrimidine–thiourea hybrids that showed a high sensitivity to LSD1 inhibition in vitro and in tumour xenografts. [4]. Furthermore, Li et al. developed a series of [1–3] triazolo [4,5-d] pyrimidine derivatives as selective LSD1 inhibitors, which were reported to block tumour cell migration [5]. The activity of documented reversible inhibitors, on the other hand, did not meet the advances of covalent inhibitors, due in part to the huge size and polarity of the LSD1 substrate binding pocket [6]. Furthermore, the reported compounds' erroneous binding techniques and the lack of theoretical research motivated us to hunt for hidden and buried structural features that are required for developing effective, efficient, and reversible LSD1 inhibitors. This convergence of circumstances encouraged us to do a computational study on the proven reversible LSD1 inhibitors from the appropriate database, as well as examine the underlying structural elements influencing the design and inhibition of potent and effective LSD1 inhibitors.

QSAR is a statistically-based intersectional plan of activities and standardized technique for identifying the mathematical relationship between the structural property of a molecule and its biological activity. General QSAR modelling protocol involves: (I) selecting a sufficiently abundant, admissible molecular dataset with accurate biological activity; (II) 3D-structure creation and optimization; (III) molecular descriptor calculation and constrained trimming using appropriate statistical methods; (IV) QSAR model development using an algorithm that fits favourable molecular descriptors; and (V) sufficient validation of the existing QSAR model (s) [5]. Circumstantial QSAR analysis quantifies the relationship between conspicuous but seemingly confusing molecule structural features and their experimentally studied biological activity. Statistical QSAR analysis predicts the biological activity of a drug prior to wet lab manufacturing and experimental in vivo testing. A QSAR that is conceptually neutral, illustrative, and statistically enhances pharmacokinetics knowledge [6,7]. This emphasizes the value of the QSAR study in promoting lead optimization.

QSAR models for LSD1 inhibitors have been discussed by a number of researchers. Rahman Abdizadeh et al. developed a 3D QSAR model for the set of tranylcypromine derivatives as a data set that performed similarly to the CoMFA ($q^2 = 0.67$; $r^2_{ncv} = 0.93$; $r^2_{pred} = 0.97$), CoMFA-RF ($q^2 = 0.69$; $r^2_{ncr} = 0.93$; $r^2_{pred} = 0.93$), CoMSIA ($q^2 = 0.83$; $r^2_{ncv} = 0.96$; $r^2_{pred} = 0.96$), and HQSAR models ($q^2 = 0.85$; $r^2_{ncv} = 0.90$; $r^2_{pred} = 0.73$) for training as well as the test set of LSD1 inhibition. Moreover, the significant gap between the q^2 and the r^2 values indicate the overfitting in both the CoMFA and CoMSIA models. However, because of the absence of mechanistic interpretation and atom-by-atom pharmacophoric features in CoMFA and CoMSIA investigations, their use has been limited to the optimization of a few pharmacological classes [8]. To date, several LSD1 inhibitors have been approved, and some of them, including ORY-1001, GSK-2879552, IMG-7289,

INCB059872, CC-90011, and ORY-2001 (See Figure 1), are currently being studied in clinical trials for cancer treatment, particularly in small lung cancer cells (SCLC) and acute myeloid leukaemia (AML) [9].

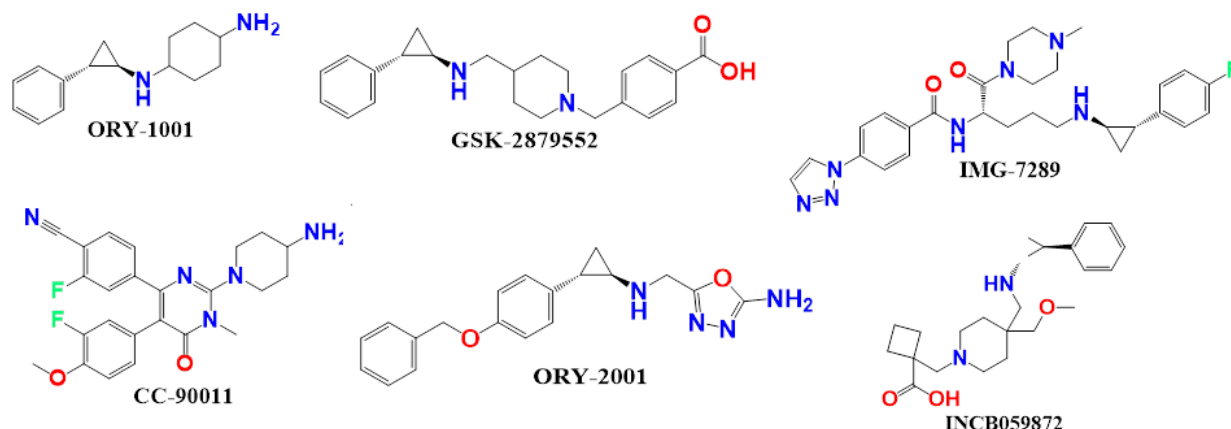


Figure 1. Presentation of the Structures of some clinical trial molecules.

A moderate-sized dataset-based QSAR with enough predictive capability and mechanistic understanding is clearly useful for boosting lead potency. In this study, we used molecular docking, MD simulation, and MMGBSA to create robust QSAR models for 84 structurally varied molecules with empirically established LSD1 inhibitory efficacy.

2. Results

Despite the fact that the current study is based on a moderate size dataset of 85 molecules, the existence of multiple molecular scaffolds, functional groups, substituents, diverse rings viz. non-aromatic, homoaromatic, heteroaromatic, fused rings; spiro compounds, etc., has significantly covered a vast chemical space. The QSAR models developed are based on a split and entire data set. R^2 , R^2_{adj} , CCC_{tr} , and other fitting metrics have values far above the allowed threshold values, indicating that the QSAR models are statistically tolerable with the required number of chemical descriptors. Internal validation parameters include Q^2_{LOO} , Q^2_{LMO} , and others with values that condescend to give the statistical robustness of the QSAR models. The external predictability of both models can be seen in the high values of external validation aspects like R^2_{ex} and Q^2_{Fn} . Model applicability domain is supported by William's plots (See Figure 2) (Applicability Domain). Fulfillment of allowed threshold values for numerous parameters, as well as poor correlation among molecular descriptors, rule out the possibility of serendipitous QSAR model construction [10–14] (see Table S2, Supplementary Information). These grounds validate these models' statistical robustness and strong external prediction.

2.1. Outlier Behavior of the Dataset Molecules

The third type of outlier, outliers toward the model, can only be identified after the regression model has been established. They indicate an X-Y link. Because of the variety of chemical structures explored in the study, model outliers are a specific form of outlier that may be found in high numbers in the QSAR/QSPR data set.

Based on the Williams plot, molecule 60 was identified as the third type of outlier in the divided set model, molecule 79 as an X outlier, and molecule 82 as a Y outlier. Figure 3 illustrates the core plot and loading plot of the Descriptor in a split-set QSAR model. The descriptor ring, CH3B, has a significant impact on molecule 60's outlier characteristics, but the descriptors lipo_ringS_8Bc and com_sp2O_4A have a substantial impact on molecule 82. The descriptor **fringCH3B**, on the other hand, had a major impact on molecule 79. The aforementioned conclusion explained the impact of particular molecular descriptors on the cluster of molecules in the dataset (See Figure 3).

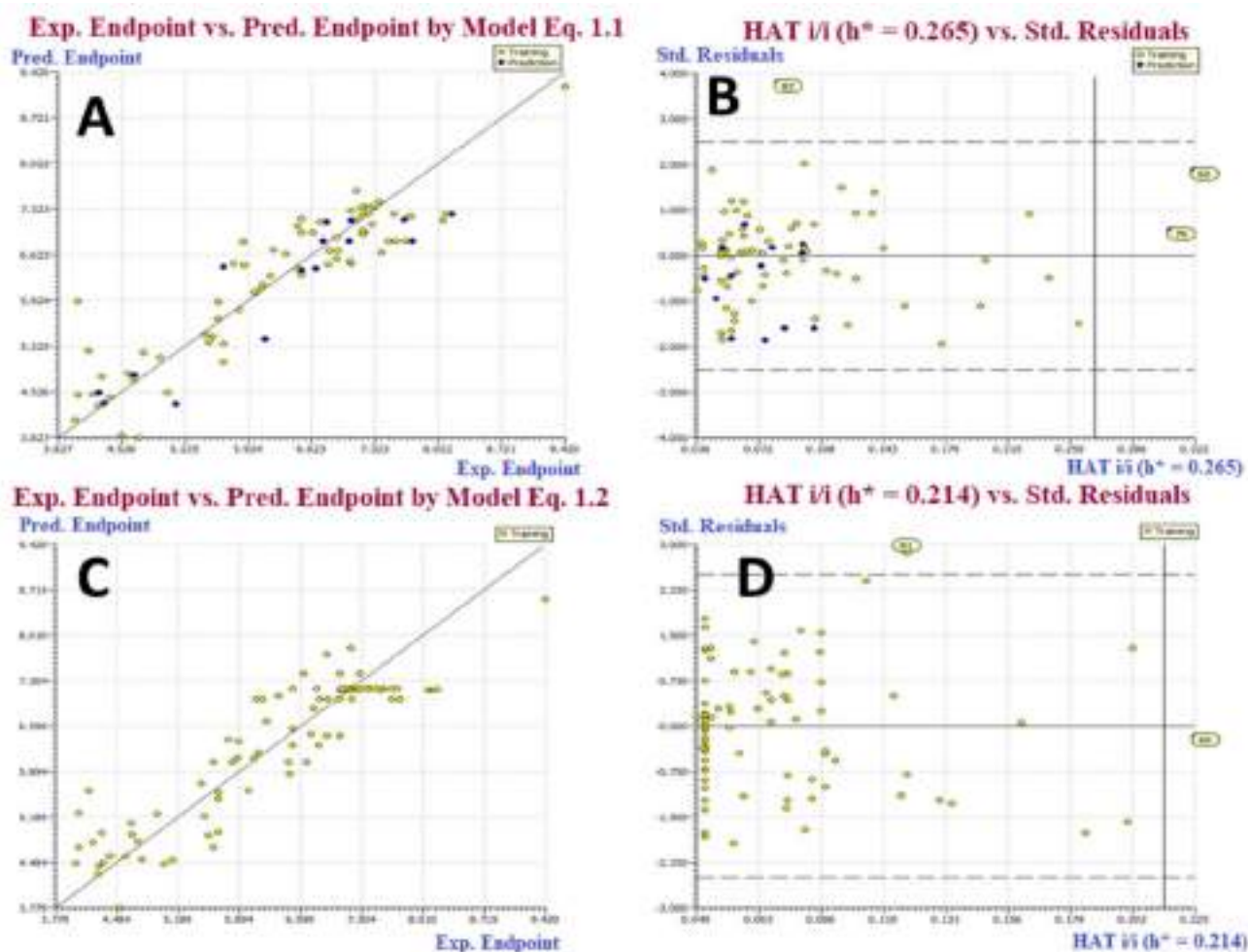


Figure 2. (A) Graph of experimental vs. Predicted pEC₅₀ values for model 1.1. (B) Williams plot for model 1.1. (C) Graph of experimental vs. Predicted pEC₅₀ values for model 1.2. (D) Williams plot for model 1.2.

2.2. GA-MLR QSAR Models

Model-1.1 (Divided Set: Training Set-80% (67 molecules) and Prediction Set-20% (17 molecules)):

$$\text{pEC}_{50} = 13.856 (\pm 1.734) - 0.832 (\pm 0.226) \text{ avg_molweight} - 0.211 (\pm 0.087) \text{ fringCH3B} + 0.263 (\pm 0.092) \text{ fNringC6B} + 4.482 (\pm 1.892) \text{ lipo_ringS_8Bc} - 0.639 (\pm 0.279) \text{ com_sp2O_4A}.$$

[R² = 0.83, R²_{adj} = 0.82, Q²_{LOO} = 0.79, Q²_{LMO} = 0.78, RMSE_{tr} = 0.49, MAE_{tr} = 0.37, RSS_{tr} = 16.37, CCC_{tr} = 0.91, RMSE_{cv} = 0.54, MAE_{cv} = 0.40, PRESS_{cv} = 20.07, CCC_{cv} = 0.89, R²_{ext} = 0.82, Q²_{F1} = 0.81, Q²_{F2} = 0.81, Q²_{F3} = 0.81, CCC_{ex} = 0.90].

Model-1.2 (Full Set: Training Set-100%, (84 molecules)):

$$\text{pEC}_{50} = 6.488 (\pm 0.315) - 0.151 (\pm 0.078) \text{ fringCH3B} + 2.921 (\pm 1.496) \text{ lipo_ringS_8Bc} + 0.972 (\pm 0.349) \text{ famdNnotringO9B} + 0.347 (\pm 0.104) \text{ fdonsp3C2B} - 0.775 (\pm 0.302) \text{ fsp3CamdN4B}.$$

[R² = 0.81, R²_{adj} = 0.80, Q²_{LOO} = 0.78, Q²_{LMO} = 0.77, RMSE_{tr} = 0.51, MAE_{tr} = 0.41, RSS_{tr} = 22.24, CCC_{tr} = 0.90, RMSE_{cv} = 0.56, MAE_{cv} = 0.45, PRESS_{cv} = 26.57, CCC_{cv} = 0.88].

In this QSAR investigation, model 1.1 was constructed using the extended dataset, whereas model 1.2 was created using the entire dataset. The developed models are distinct in three of the five descriptors out of a total of five. The effects of variation in each molecular descriptor on the biological activity of the associated molecule are demonstrated with examples in the next section, even if the permutation in the bioactivity of each molecule in the dataset is the total of all five molecular descriptors.

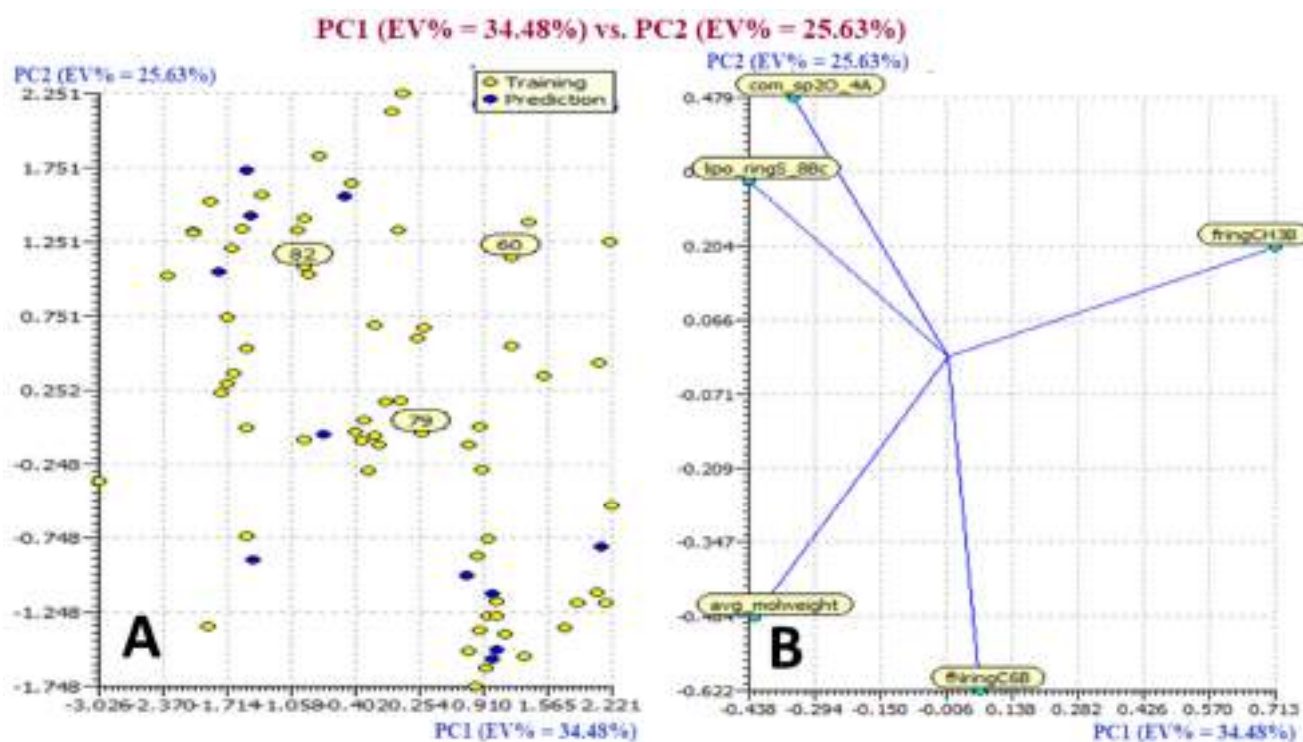


Figure 3. Presentation of Score Plot (A) and Loading Plot (B) for the Descriptor in divided set QSAR Model.

3. Discussion

3.1. Mechanistic Interpretation of Descriptors

fNringC6B, lipo_rings_8Bc, famdNnotringO9B, and fdonsp3C2B: These four molecular descriptors had positive coefficient values in both the divided and full set models, showing that amplification in the values of these molecular descriptors improves the anticancer potential of LSD1 inhibitors. The importance of some molecular descriptors is demonstrated by comparing variations in the pEC_{50} or EC_{50} values with transformations in the values of molecular descriptors.

fNringC6B (frequency of occurrence of ring carbon atom exactly at 6 bonds from nitrogen atom). This observation is supported by comparing compound **1** ($fNringC6B = 1$; $pEC_{50} = 9.42$) with compound **6** ($fNringC6B = 0$; $pEC_{50} = 7.71$). Possibly, an increase in the value of $fNringC6B$ to 1 for compound **6** enhanced its LSD1 inhibitory potency by about two hundred and twenty-two times ($\Delta pEC_{50} = 2.22$) (See Figure 4).

This observation was also seen by comparing the subsequent pair of molecules: **17** ($pIC_{50} = 7.25$, $fNringC6B = 2$) with **18** ($pIC_{50} = 7.21$, $fNringC6B = 0$), **31** ($pIC_{50} = 7.04$, $fNringC6B = 4$) with **32** ($pIC_{50} = 6.91$, $fNringC6B = 0$), **52** ($pIC_{50} = 6.11$, $fNringC6B = 1$) with **56** ($pIC_{50} = 5.88$, $fNringC6B = 0$), etc.

Vianello Paola et al. and colleagues also reported the synthesis of chemical **2** (4-ethyl-N-[3-(methoxymethyl)-2-[(4-[(3R)-pyrrolidin-3-yl] methoxyphenoxy) methyl] phenyl]-4H-thieno [3,2-b] pyrrole-5-carboxamide) from the dataset (see Figure 5). The most efficient basic moiety was compound **2** with pyrrolidin-3-yl-methanol substituent, which had potent inhibitory activity against LSD1 ($IC_{50} = 0.08570.02$ M) according to structure–activity relationship studies. He went on to say that the polar interaction with two negatively charged regions of the LSD1 catalytic site is responsible for the compound **2**'s increased potency [10] (See Figure 5).

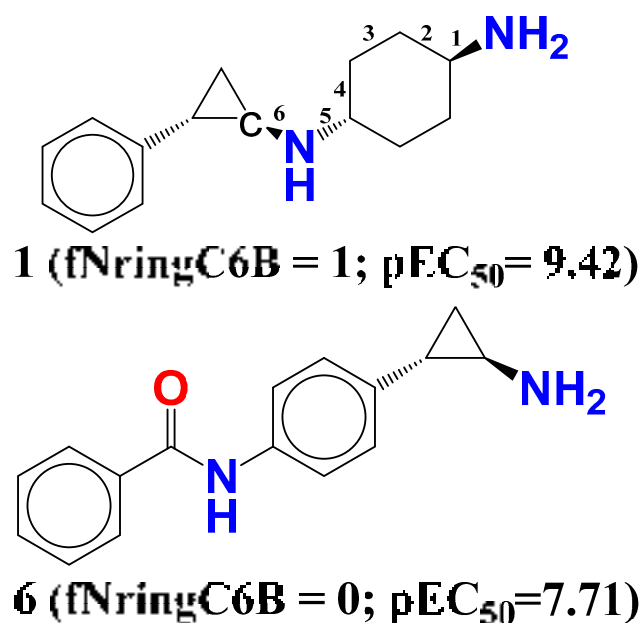


Figure 4. Illustration of molecular descriptor $fNringC6B$ for the molecules 1 and 6 only.

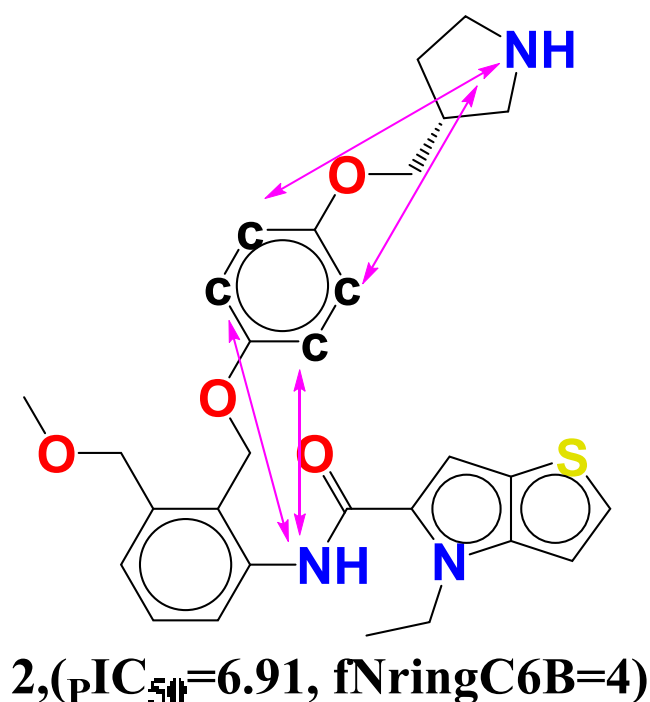


Figure 5. Depiction of the molecular descriptor $fNringC6B$ in the compound 2.

Thus, the present observation supports that the pyrrolidine ring enhances the polarity of the compound 2 that occurred exactly at 6 bonds. As a whole, the same feature has been captured in the QSAR model through the descriptor $fNringC6B$; therefore, QSAR results are complimentary with the reported findings. At the end, the QSAR model not only identified the polar nitrogen, but it also recognized the lipophilic carbon atom important for LSD1 inhibitory activity.

$lipo_ringS_8Bc$ (Sum of partial charges of lipophilic atoms within 8 bonds from ring sulfur atom). The molecule with the better LSD1 inhibition might be obtained by enhancing the number of lipophilic atoms that accounted within 8 bonds from the sulfur atom. Just a four-fold amplification in the value of $lipo_ringS_8Bc$ sufficed about 2×10^3

fold more potent ($\Delta pEC_{50} = 3.32$) LSD1 inhibitor compound **4** (**lipo_ringS_8Bc** = 0.19; $pEC_{50} = 8.04$) than compound **72** (**lipo_ringS_8Bc** = 0.05; $pEC_{50} = 4.72$). Several other pairs of compounds also support this observation: **31** (**lipo_ringS_8Bc** = -0.23 ; $pEC_{50} = 7.046$) with **32** (**lipo_ringS_8Bc** = 0; $pEC_{50} = 6.917$), **35** (**lipo_ringS_8Bc** = 0; $pEC_{50} = 6.827$) with **36** (**lipo_ringS_8Bc** = -0.28 ; $pEC_{50} = 6.81$), and **45** (**lipo_ringS_8Bc** = -0.207 ; $pEC_{50} = 6.511$) with **46** (**lipo_ringS_8Bc** = -0.231 ; $pEC_{50} = 6.509$).

Whence merely adding the number of carbon atoms is restricted (here average_molweight, i.e., molecular property average molecular weight, is with negative correlation) or inadequate, it is advisable to add electronegative atoms to the carbon atoms within 8 bonds from the ring sulfur to intensify the partial positive charge on the lipophilic atoms that boost up the LSD1 potency of the compound, respectively (see Figure 6).

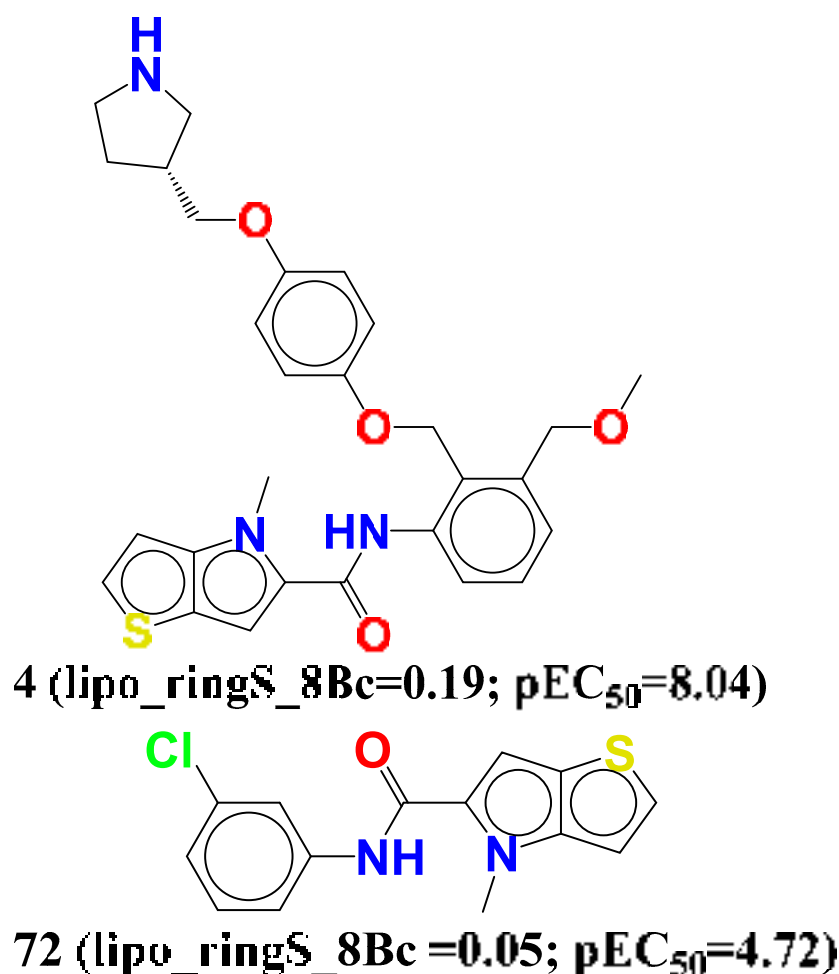


Figure 6. Depiction of molecular descriptor **lipo_ringS_8Bc** for the molecules **4**, **72**, and reported molecule **28186757** only.

Furthermore, a comparison of compound **4** to the previously reported molecule **28186757** suggests that increasing the amount of carbon atoms at the 8th position, specifically in the ether-containing carbon atom, will enhance the LSD1 inhibitory activity even more [10].

famdNnotringO9B (Frequency of occurrence of non-ring oxygen atom exactly at 9 bonds from the amide nitrogen) with a positive coefficient exhibit a direct correlation with LSD1 inhibitory potency. The four displayed compounds, **10**, **65**, **13**, and **14**, in Figure 7 illustrate the influence of the present molecular descriptor on the LSD1 inhibitory potency of the compound. It can be noted that, if the same non-ring carbon atom simultaneously

occurred at one to eight bonds or more than 9 bonds from the amide nitrogen, then it is eluded during the calculation of famdNnotringO9B (see Figure 7).

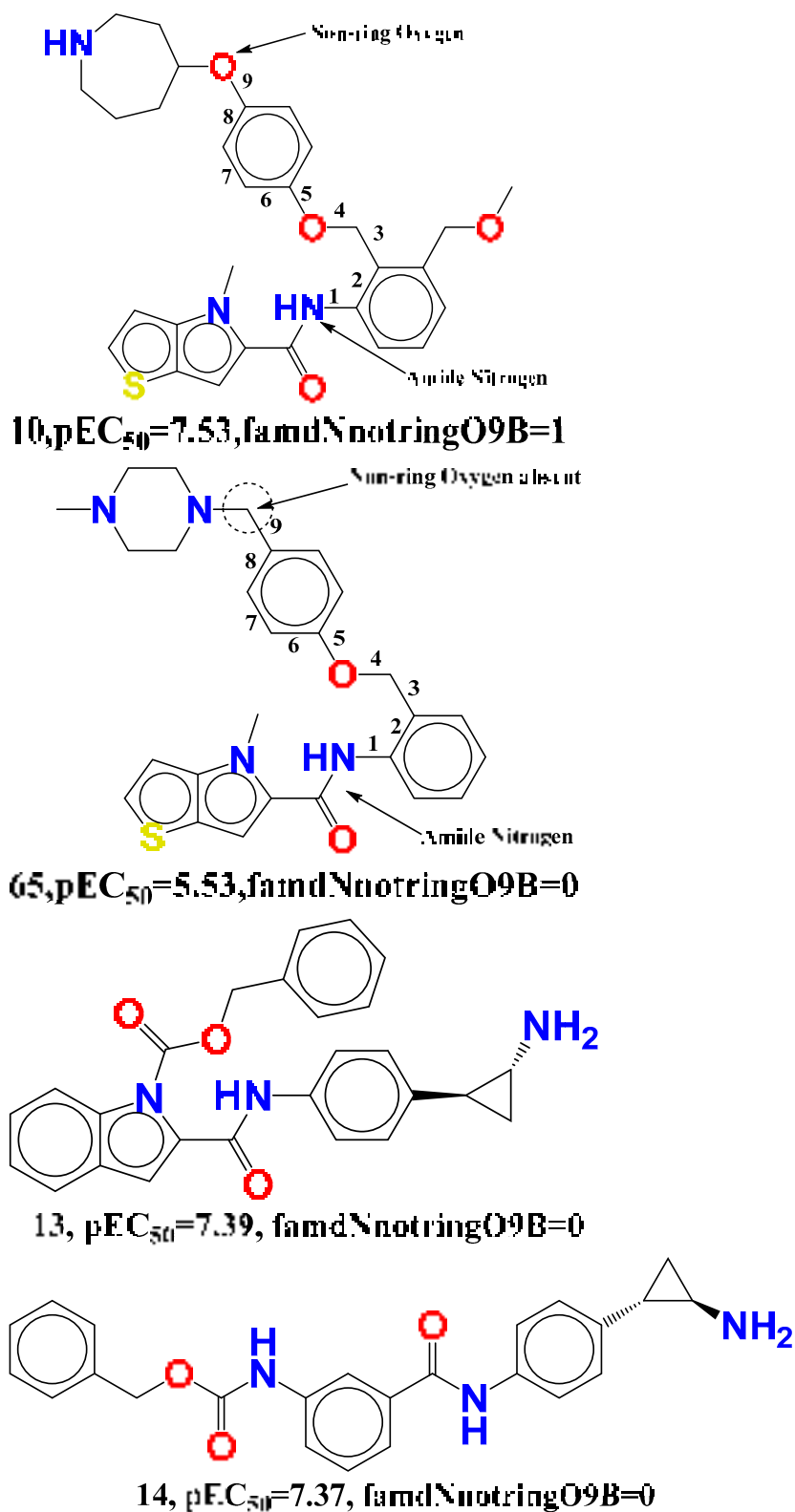


Figure 7. Pictorial depiction of molecular descriptor famdNnotringO9B for the molecules 10, 65, 13, and 14 only.

Non-ring oxygen was detected exactly 9 links from the amide nitrogen in compound **10**, but the same oxygen was missing in compounds **65**, **13**, and **14**. This finding further supports the idea that the appropriate distance between the amide nitrogen and the non-ring oxygen is important for LSD1 inhibition. This helps to explain why molecules **10** and **65**, **13**, and **14** have different LSD1 inhibitory action. Instead, Vianello, Paola, and colleagues found that removing oxygen had only a little effect on the LSD1 inhibitory function. This new discovery backs up the QSAR concept, emphasizing the significance of the oxygen atom in the 9th position from the amide nitrogen. In addition, Vianello Paola emphasized the importance of thieno [3,2-b]pyrrole-5-carboxamides as novel reversible inhibitors of the LSD1 receptor, noting that the same amide nitrogen was successfully detected as `famdNnotringO9B` in QSAR modelling. As a result, the QSAR results are consistent with the stated findings.

Another key chemical characteristic, `fdonsp3C2B` (frequency of occurrence of sp³ hybridised carbon atom exactly at 2 bonds from donor atom), is strongly linked with the reported bioactivity of LSD1 inhibitors. When comparing compound **1** to compound **57**, it can be shown that increasing the number of sp³ hybridised carbon atoms exactly at 2 bonds enhances the LSD1 inhibitory potency (see Figure 8).

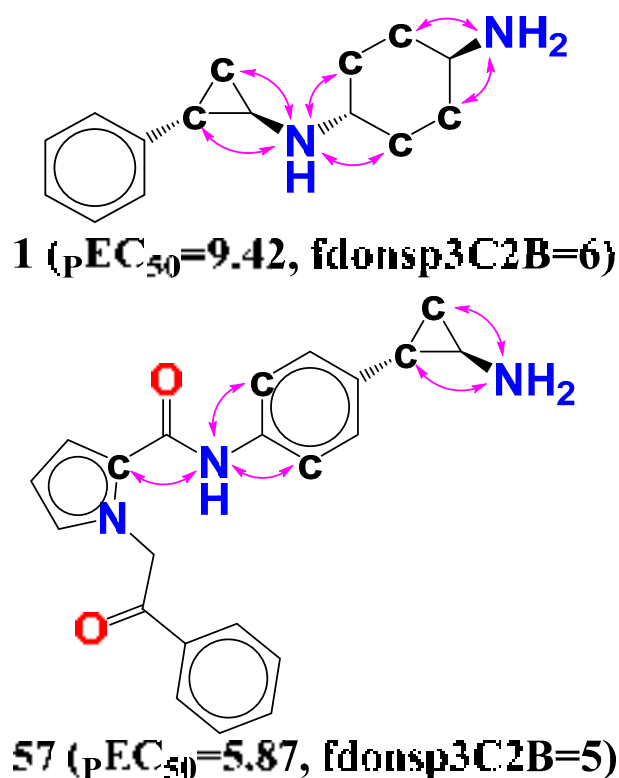


Figure 8. Pictorial display of molecular descriptor `fdonsp3C2B` for the molecules **1** and **57** only.

Furthermore, the same result holds true for a few other compounds: the most active compound **1** ($pEC_{50} = 9.42$, $fdonsp3C2B = 6$), as well as the compounds **2** ($pEC_{50} = 8.17$, $fdonsp3C2B = 2$), **3** ($pEC_{50} = 8.10$, $fdonsp3C2B = 2$), **4** ($pEC_{50} = 8.07$, $fdonsp3C2B = 2$), and **5** ($pEC_{50} = 7.49$, $fdonsp3C2B = 2$). The LSD1 inhibitory activity will be increased by 3.55 units if the value of the molecular descriptor `fdonsp3C2B` for the molecule **57** is increased from 2 to 6 (about a 35-fold increase in LSD1 inhibitory potency). Furthermore, sp³ hybridised carbon atoms should be added to boost LSD1 inhibitory activity, according to the current findings. Furthermore, increasing the amount of such sp³ hybridised carbons along the donor increases the electrical and hydrophobic interaction with the LSD1 receptor, showing lipophilicity.

Following that, it was discovered that during the MD modeling of compound **1** that the NH moiety, which acts as a donor with two bonds from the sp³ hybridized carbon atom (fdonsp3C2B), demonstrated significant hydrogen bonding with GLU308 (86 percent) and thus plays an important role in the stability of the LSD1–compound **1** complex. Furthermore, by including a water molecule, the same NH moiety created hydrogen bonds with a similar residue (GLU308), increasing the stability of the drug receptor complex. Furthermore, another NH₂ substituent (91 percent) developed hydrogen bonding connections with the Glu801 residue, increasing the stability of the drug receptor complex (see Figure 9). This implies that the QSAR modelling has effectively identified certain important pharmacophores involved in the stability of the drug receptor complex, in addition to finding the many hidden structural elements crucial for LSD1 inhibition. As a consequence, the QSAR findings are entirely consistent with the molecular docking and MD simulation experiments.

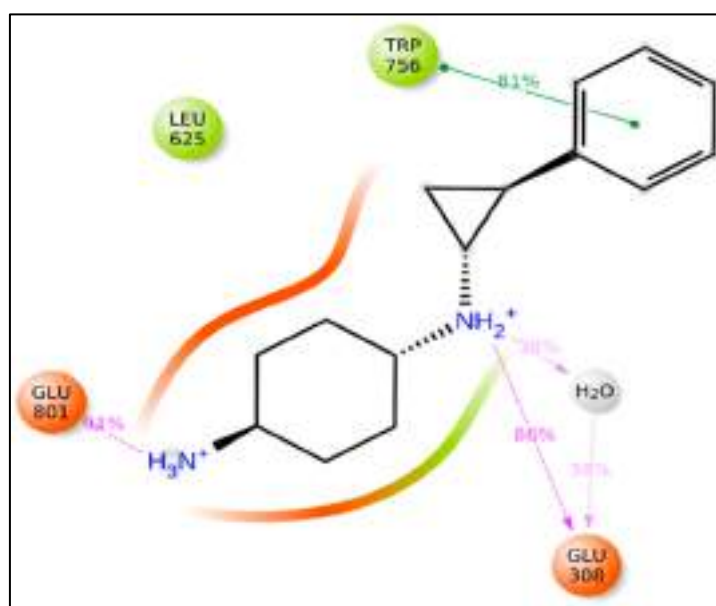


Figure 9. Depiction of the involvement of the molecular descriptor fdonsp3C2B in the LSD1–Compound **1** interactions during MD simulations.

No one chemical descriptor can explain the variation in inhibitory effectiveness of medicines in a dataset. The performance of the QSAR model is impacted by the synchronous effect of many molecular descriptors, some of which are not included in the QSAR models.

3.2. Molecular Docking

The molecular interaction of the five most active molecules with the LSD1 protein at the active site was investigated using molecular docking. The crystal structure of LSD1 was obtained using the RCSB protein data repository (<https://www.rcsb.org/structure/2dw4>, accessed on 24 May 2022) (PDB code: 2DW4). The full length of LSD1 comprises 852 amino acids with three key structure domains [11–14]: N-terminal Swi3-Rsc8-Moira domain (SWIRM domain, residues 172–270); C-terminal amine oxidase-like domain (AOL domain, residues 271–417 and 523–833); and central tower-like domain (Tower domain, residues 418–522). The SWIRM domain of LSD1 consists of six long α -helices (SW α 1–6) and two stranded β -sheets (SW β 1–2), which regulates the chromatin remodeling and histone modification by taking part in protein–protein interactions.

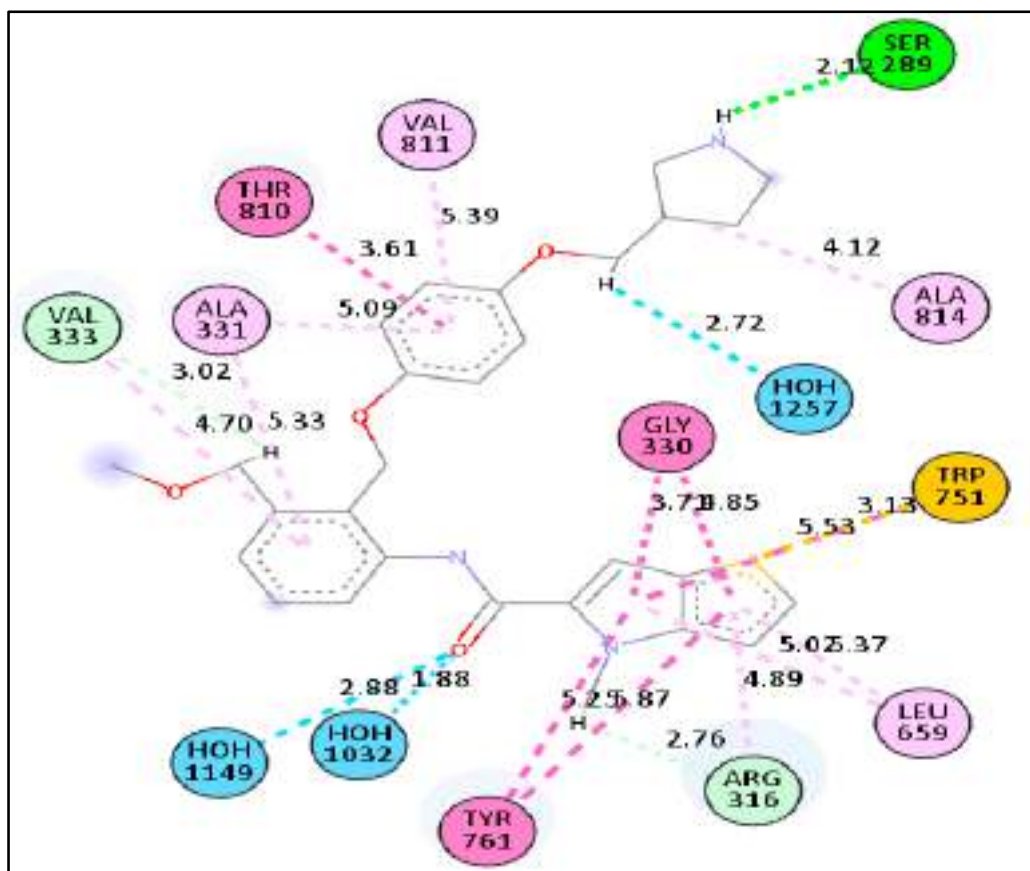
We investigated the probable interactions of inhibitors inside the active site of LSD1 to better understand the SAR and QSAR models of the five most active drugs. With an RMSD of 1.3618 Å, the 2DW4 ligand was redocked into the LSD1 binding pocket. Because the

accuracy of the docking results was determined by RMSD, this indicates that NRG Suite docking was able to effectively recognise the correct binding configuration (2.0). The docking scores for the five compounds, **1** ($EC_{50} = 0.38$ nm), **2** ($EC_{50} = 6.7$ nm), **3** ($EC_{50} = 7.8$ nm), **4** ($EC_{50} = 8.4$), **5** ($EC_{50} = 18$), and pdb-2dw4 ligand, were found to be -8.33 (RMSD -1.38 Å), -10.47 (RMSD 1.82 Å), -11.16 (RMSD 1.32 Å), -11.10 (RMSD 1.58 Å), -10.96 (RMSD 1.13 Å), and -11.31 (RMSD 1.36 Å) Kcal/mol, respectively, demonstrating that docking scores could predict ligand EC_{50} values. The 2D interactions for the five compounds were displayed in Figures 10–14.

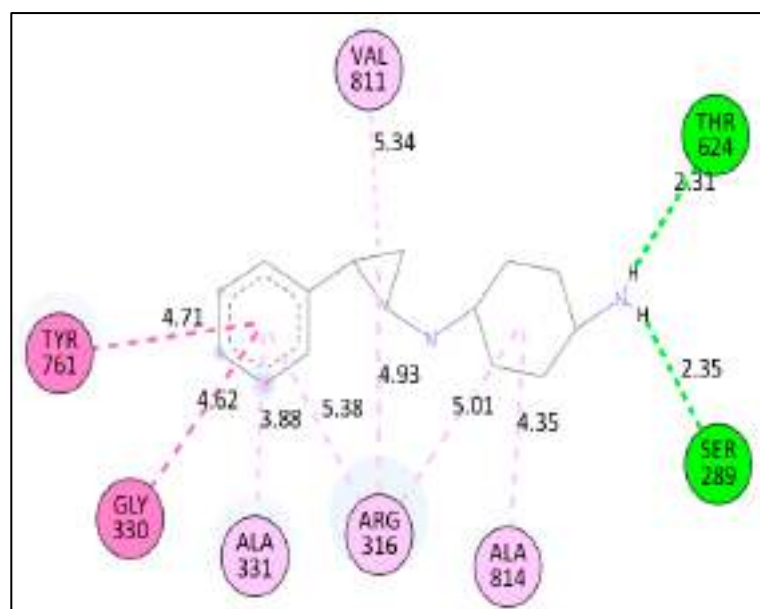
In terms of compound **5**'s low activity, the amide nitrogen forms a conventional hydrogen bond with the neutral non-polar amino acid residue MET332, a water–hydrogen bond with HOH1032, and a neutral polar amino acid residue with the pyrrolidine ring. THR624 forms a second hydrogen bond. With HOH1251, it creates a third water–hydrogen bond. TRP751, GLY330, LEU859, ALA331, TYR761, VAL811, ARG316, and ALA814, on the other hand, have been shown to form hydrophobic bonds with a thiene-pyrrole ring, a benzamide ring, a phenoxy ring, or a pyrrolidine ring (pi-pi T-shaped, amide-pi stacked, alky and pi-alkyl interactions). Despite the wide and flexible structure of compound five, the active conformation and compound-**5**–LSD1 complex were maintained via a variety of hydrophobic interactions and hydrogen bonding.

Compound **5** and compound **4** have similar interactions, although compound **4** is three times more powerful than compound **5**. In the structure, compound **5** has a folded shape, whereas compound **4** has an extended conformation akin to the pdb-2dw4 ligand. Within the active area of the LSD1 receptor, chemicals **5** and **4** are diametrically opposed. Except for one hydrophobic interaction with a TYR761 amino acid residue, the thien-pyrrole ring orientation was different. The QSAR models demonstrate the importance of the thiene-pyrrole ring for the reversible inhibition of the LSD1 receptor. The molecular descriptor **lipo_ringS_8Bc** indicates the importance of the Sum of partial charges of lipophilic atoms within 8 bonds from ring sulfur atoms. With TRP751, compound **4** (**lipo_ringS_8Bc** = -0.1869) made more than eight types of hydrophobic connections and one pi-sulphure interaction, whereas compound **5** (**lipo_ringS_8Bc** = -0.2319) made seven hydrophobic contacts (See Figure 11A,B). The difference in the reactivity of these compounds was linked with the occurrence of positively charged lipophilic atoms. The present observation indicates that the decrease in the negative charge promotes more hydrophobic contacts in the compound **4**. Furthermore, in compound **1** (**lipo_ringS_8Bc** = 0), partial positive charges are zero, underlining the observation of declining negative charges and intensifying positive charges within the thiene-pyrrole ring, which promotes better hydrophobic contact with the LSD1 receptor. The compounds **3** (**lipo_ringS_8Bc** = -0.1869) and **2** (**lipo_ringS_8Bc** = -0.2339) showed the same behavior. Finally, QSAR analysis was successful in uncovering latent pharmacophoric characteristics that determine not only the LSD1 inhibitory action of these compounds, but also their binding pattern. As a result, the molecular docking analysis results are entirely congruent with the QSAR findings.

Moreover, compound **2** ($EC_{50} = 6.7$ nm), was marginally more potent than compound **3** ($EC_{50} = 7.8$ nm). The 2D interactions for compounds **2** and **3** show that compound **2** produced three standard hydrogen bonding contacts with SER760, LYS661, ARG316, and GLU801, but compound **3** did not form any conventional hydrogen bonding interactions with SER760, ALA809, THR810, or HOH1257. Moreover, compound **2** executed more than 11 different hydrophobic contacts with the HIS564, ALA539, VAL333, GLY330, TRP751, VAL811, VAL317, ALA814, etc. Moreover, the thiene-pyrrole ring in the compound didn't contribute in any of the hydrophobic contact, but it aligned over the solvent accessible surface area of the LSD1 receptor (See Figure 12A,B). Furthermore, when the conformations of compounds **2** and **3** are compared to the pdb-2dw4 ligand, it is clear that compound **2** aligns and superimposes entirely along the docked conformation of the pdb ligand. Following that, in compound **3**, the thiene-pyrrole ring aligns vertically in the receptor (LSD1) binding pocket, which is completely different from the bioactive conformation of the pdb ligand and could explain the difference in potency between these compounds

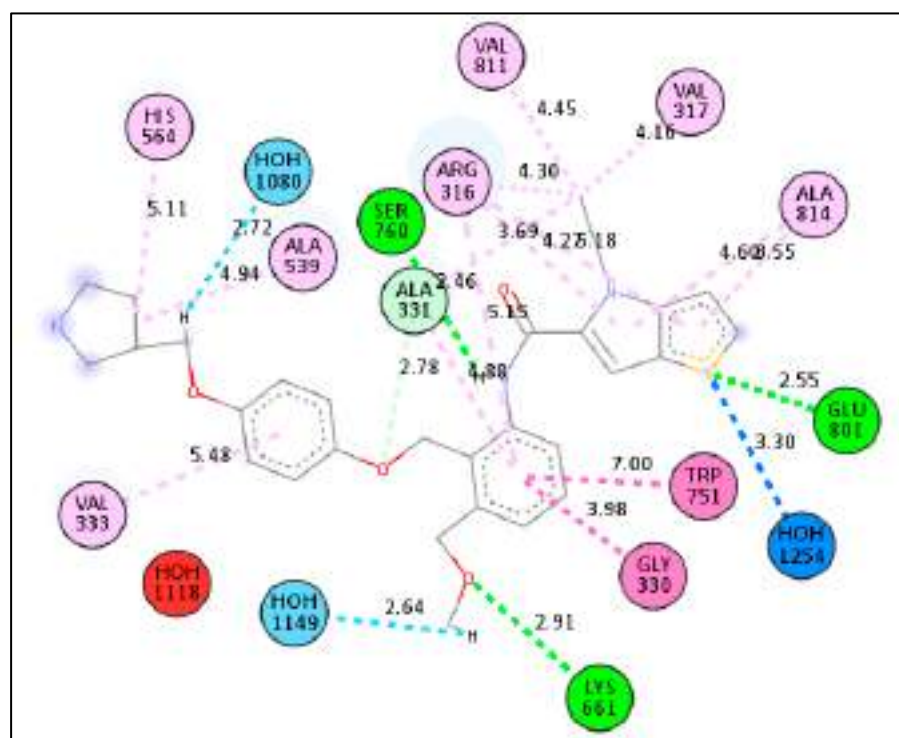


(A)

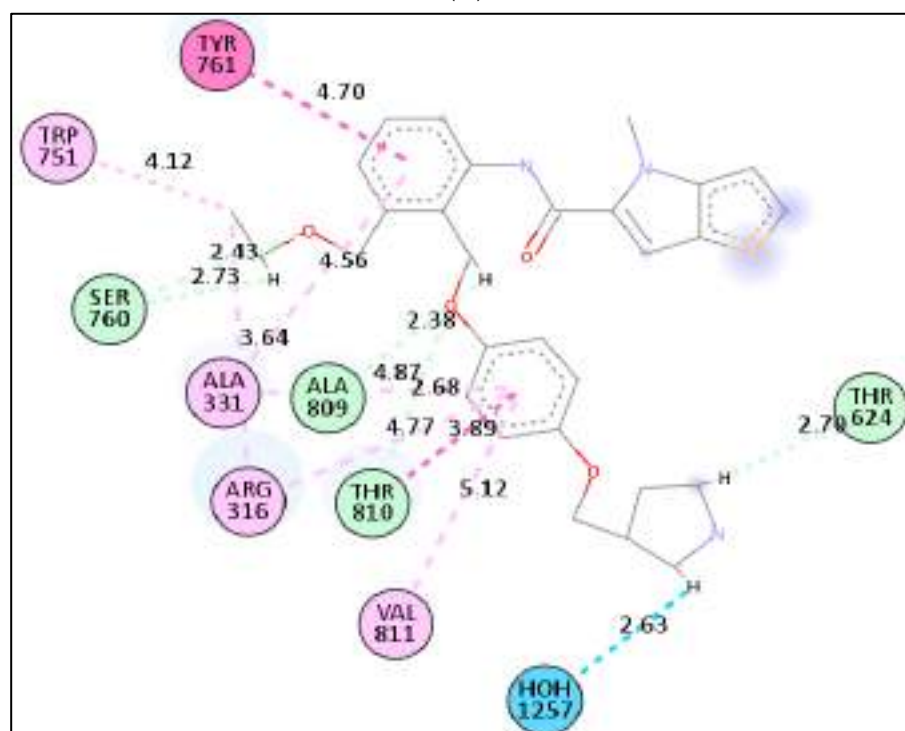


(B)

Figure 11. Presentation of the 2D and 3D interactions of compound 4 (A) and compound 1 (B) with LSD1 receptor.



(A)



(B)

Figure 12. Presentation of the 2D and 3D interactions of compound 4 (A) and compound 1 (B) with LSD1 receptor.

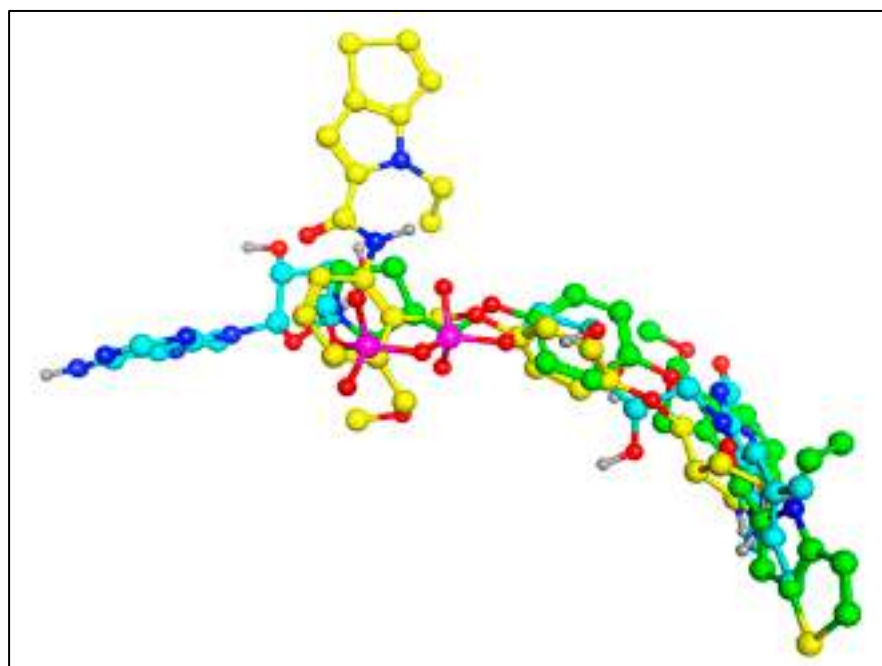


Figure 13. Comparison of the docked conformation of the compound 2 and 3 with the pdb-2dw4 ligand. (Green colored—Comp 2; yellow colored—Comp 3; s and cyan colored—pdb-2dw4 ligand).

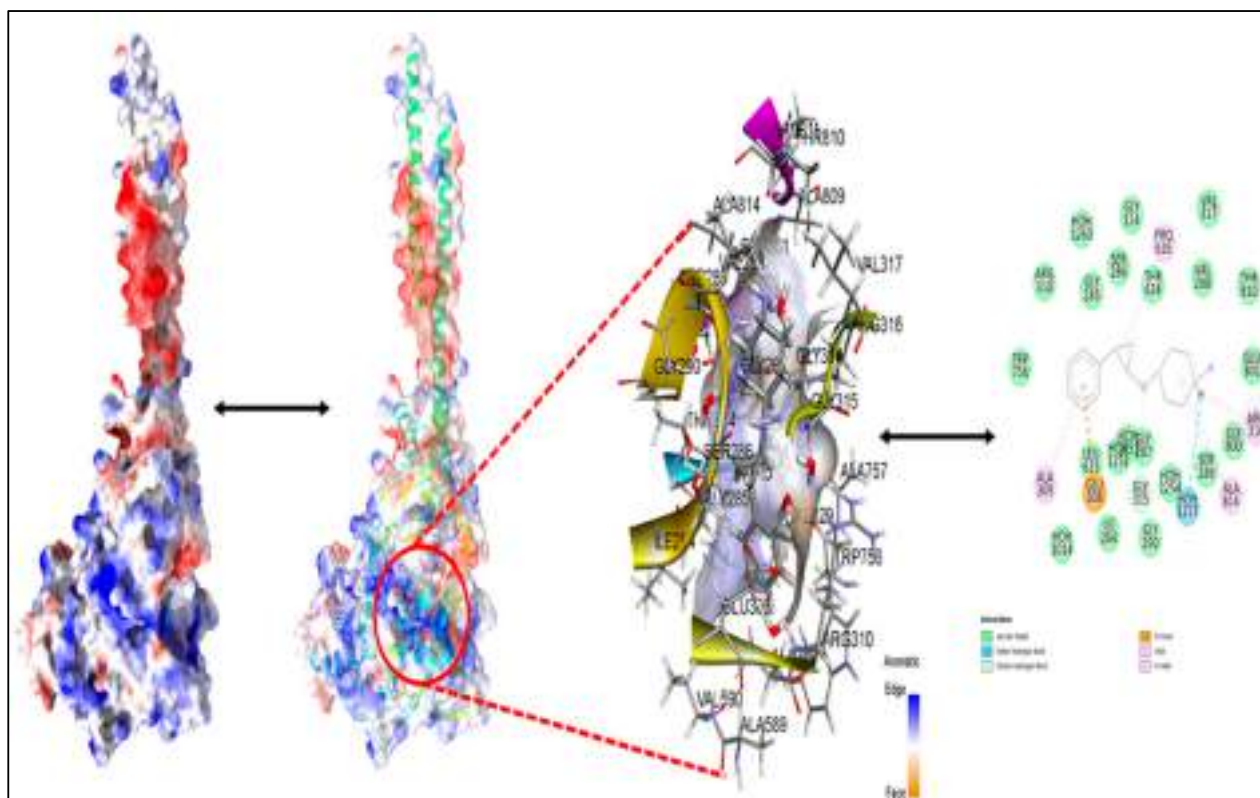


Figure 14. Best docked pose of compound 1 with LSD displaying 2D interaction plot on the left panel. Pink dashed lines indicate the Pi-Alkyl bond, and residues embedded in light green sphere indicate the involvement in Van der Waals interactions. On the center panel, surface view of LSD displaying binding cavity of Compound 1 and right panel displaying the zoomed out binding pocket having amino acid residues surrounding the Compound 1.

The compound **2** ($pIC_{50} = 8.174$, $fdonsp3C2B = 2$, $lipo_ringS\ 8Bc = -0.2339$) and **3** ($pIC_{50} = 8.174$, $fdonsp3C2B = 2$, $lipo_ringS\ 8Bc = -0.2339$) differed from the compound **1** ($pIC_{50} = 9.42$, $fdonsp3C2B = 6$, $lipo_ringS\ 8Bc = 0$) in terms of two descriptors: **fdonsp3C2B** and **lipo_ringS 8Bc**. Compound **3** can't form hydrogen bonds or hydrophobic interactions with the receptor because of its altered orientation. The amide donor produced hydrogen bonds with the SER760 residue in compound **2**, whereas another donor nitrogen of the terminal pyrrolidine ring aligned over the solvent accessible surface region, and the sp^3 hybridised carbon atom made hydrophobic interactions with the receptor. In the QSAR model, the same feature was captured. In addition, the thiene-pyrrole ring sulphure atom formed conventional hydrogen bonds with the ARG316 and GLU 801 residues, as well as a water-hydrogen link with the HOH1254 residue. The relevance of the thiene-pyrrole sulphure atom, which was captured in the QSAR model as **lipo ringS 8Bc** descriptors, is highlighted by this observation. Furthermore, the lipophilic carbon atoms in the benzene ring of compound **2** connected to the thiene-pyrrole ring via amide linkage generate distinct hydrophobic interactions with the receptor. This finding emphasises the significance of positively charged lipophilic carbon atoms in drug receptor interactions. Thus, QSAR modelling was successful in identifying the features required to improve binding affinity, and the results were in perfect agreement with the molecular docking data. In addition, comparison with the most active compound **1** ($PIC_{50} = 9.42$, $fdonsp3C2B = 6$, $lipo\ ringS\ 8Bc = 0$) indicated the importance of the lipophilic, as well as the electronic properties required for binding affinity and, ultimately, LSD1 receptor inhibition.

The docking results revealed that the descriptors, **fdonsp3C2B** and **lipo_ringS_8Bc**, played important roles in the inhibition of the LSD1 receptor, which was consistent with the QSAR findings.

3.3. Molecular Dynamic Simulations

During the simulation, monitoring the protein's RMSD can provide insight into its structural conformation. The RMSD analysis can identify if the fluctuations at the end of the simulation are centred on some thermal average structure if the simulation has equilibrated. For tiny, spherical proteins, changes on the order of 1–3 are perfectly acceptable. Larger changes, on the other hand, imply that the protein is significantly changing form during simulation. It's also crucial that your simulation converges, which means the RMSD values settle around a fixed number. If the average RMSD of the protein is still increasing or dropping at the end of the simulation, your system has not equilibrated, and your simulation may not be lengthy enough to do a thorough analysis. Ligand RMSD (right Y-axis): the ligand RMSD (right Y-axis) shows how stable the ligand is in relation to the protein and its binding pocket.

When the protein–ligand complex is aligned on the reference protein backbone first, and then the RMSD of the ligand-heavy atoms is measured, the RMSD of the ligand is plotted. If the observed values are significantly greater than the RMSD of the protein, the ligand has most likely diffused away from its initial binding site.

The above-mentioned diagram is the triple run result of Root Mean Square Deviations (RMSD) MD simulation trajectory analysis. The RMSD plot of the LSD–compound **1** complex (Figure 3) indicates that the complex stabilizes at about 20 ns. After that, for the length of the simulation, swings in RMSD values for target remain within 0.5, which is absolutely acceptable. The ligand fit-to-protein RMSD values fluctuate within 0.7 Angstrom after they have been equilibrated. These findings indicate that the ligands stayed firmly connected to the receptor's binding site throughout the simulation period. The RMSD values for ligand fit to protein do not change much during the simulation duration, showing that the ligands remain securely attached to the receptor's binding site, as shown in Figure 15.

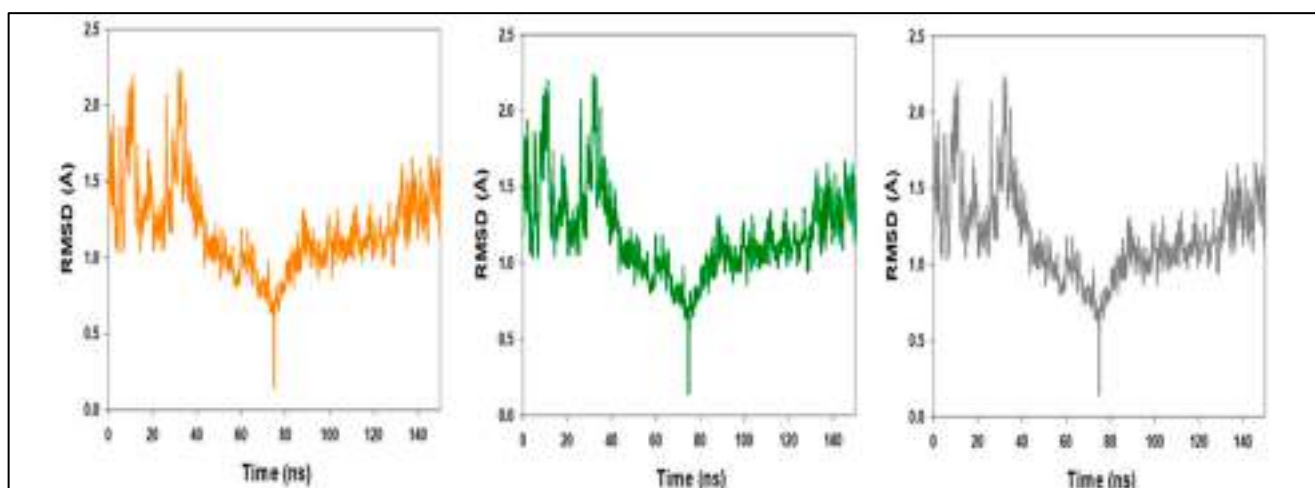


Figure 15. MD simulation trajectory analysis of Root Mean Square Divisions (RMSD) of compound 1 bound with LSD; 150 ns time frame in triplicate displayed.

Figure 16 shows the average hydrogen bonds established throughout the 150 ns triple simulation between compound 1 and the various proteins. From 0 to 150 ns, an average of four hydrogen bonds are observed for LSD, and the same is true for triple MD simulations of compound 1 and LSD (Figure 16). Throughout the simulation, two hydrogen bonds were established, as shown by the 2D ligand binding figure. The number of hydrogen bonds between LSD and compound 1 has increased, making the binding stronger and more robust over simulation.

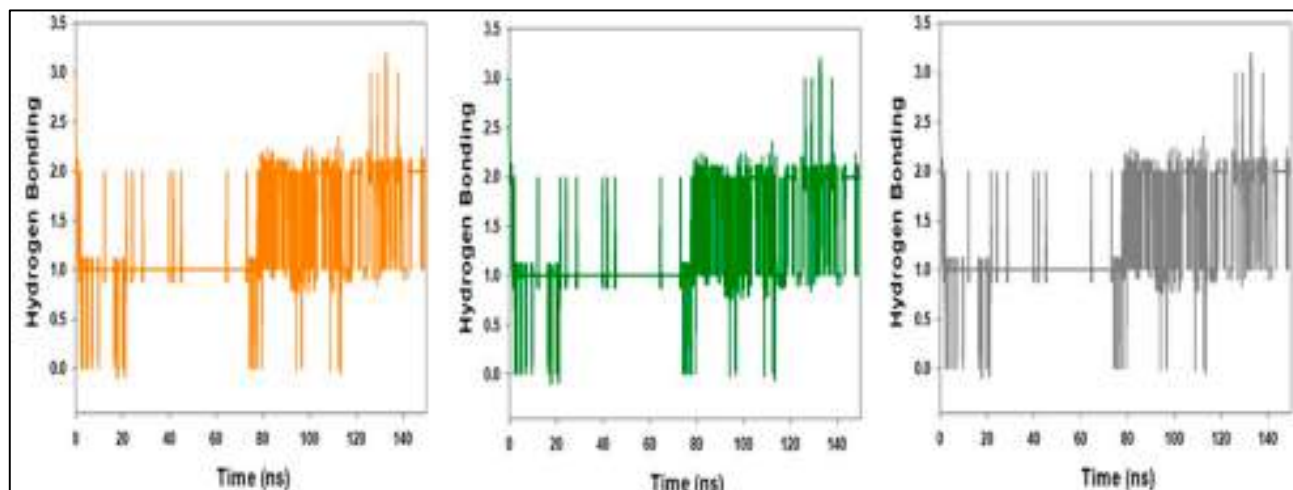


Figure 16. MD simulation trajectory analysis of Hydrogen-bonding (H-bonding) of compound 1 bound with LSD 150 ns time frame in triplicate displayed.

On the RMSF plot, peaks represent portions of the protein that fluctuate the most during the simulation. Protein tails (both N- and C-terminal) typically change more than any other part of the protein. Alpha helices and beta strands, for example, are usually stiffer than the unstructured component of the protein and fluctuate less than loop sections. According to MD trajectories, the residues with greater peaks belong to loop areas or N- and C-terminal zones (Figure 17). Although there is some instability between 400 and 600 residues, the stability of ligand binding to the protein is demonstrated by stable RMSF values of binding site residues.

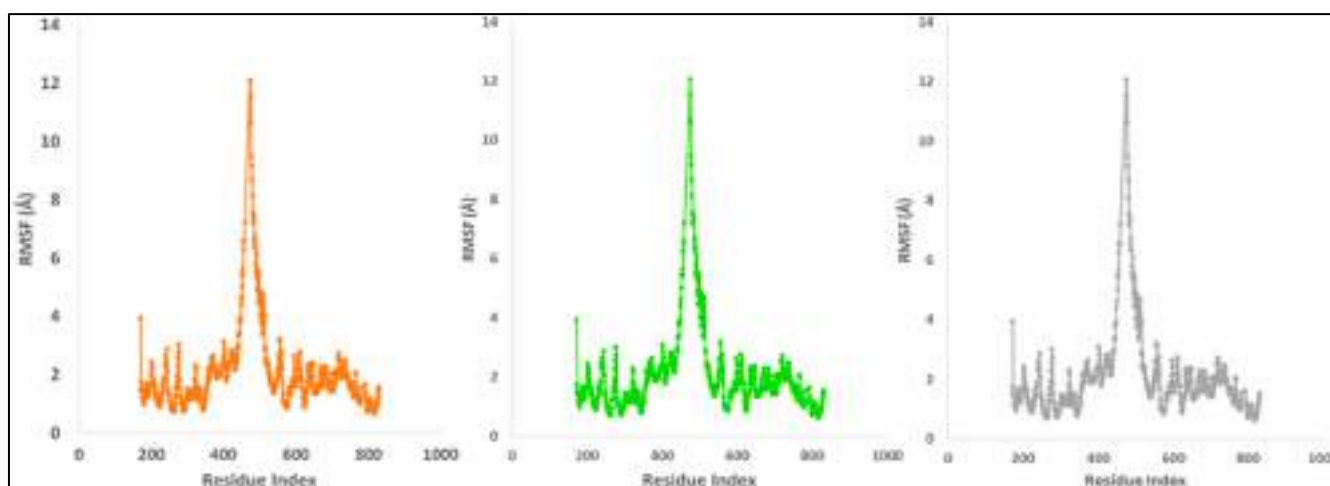


Figure 17. MD simulation trajectory analysis of Root Mean Square Fluctuations (RMSF) of compound 1 bound with LSD with their triplicate runs.

The compactness of proteins is measured by the radius of gyration. The Radius of Gyration of LSD proteins bound to compound 1 was reduced (Figure 18). Compound 1 bonded to the protein targets posthumously in the binding cavities and plays a substantial role in the stability of the proteins, according to the overall quality analysis using RMSD and Rg.

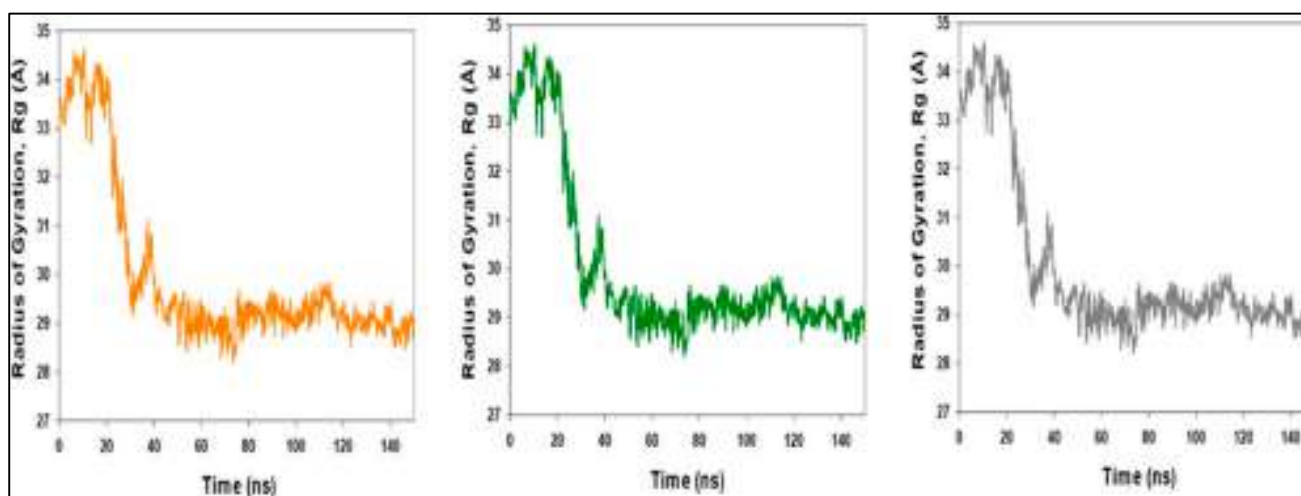


Figure 18. MD simulation trajectory analysis of Radius of Gyration (RoG) of compound 1 bound with LSD with their triplicate runs.

Protein interactions with the ligand can be detected throughout the simulation. These interactions can be categorized and summarized by type, as shown in the graphs below. The four types of protein–ligand interactions (or ‘contacts’) are hydrogen bonds, hydrophobic, ionic, and water bridges. The ‘Simulation Interactions Diagram’ panel in Maestro can be used to analyse the subtypes of each interaction type. Over the course of the trajectory, the stacked bar charts are standardised; for example, a value of 0.7 indicates that the specific interaction is maintained for 70% of the simulation duration. Values exceeding 1.0 are possible because some protein residues may have several interactions with the same subtype of ligand. The majority of the important ligand–protein interactions found by MD are hydrogen bonds and hydrophobic interactions, as shown in Figure 19. In terms of H-bonds, the LSD–compound 14 complex residues VAL 4288, GLY 290, TYR 571, ASP 754, and SER 760 are the most essential. Over the course of the trajectory, the stacked bar charts

were standardised; for example, a value of 1.0 signifies that the specific interaction was maintained for 100% of the simulation duration. Values exceeding 1.0 are possible because some protein residues may have several interactions with the same subtype of ligand.

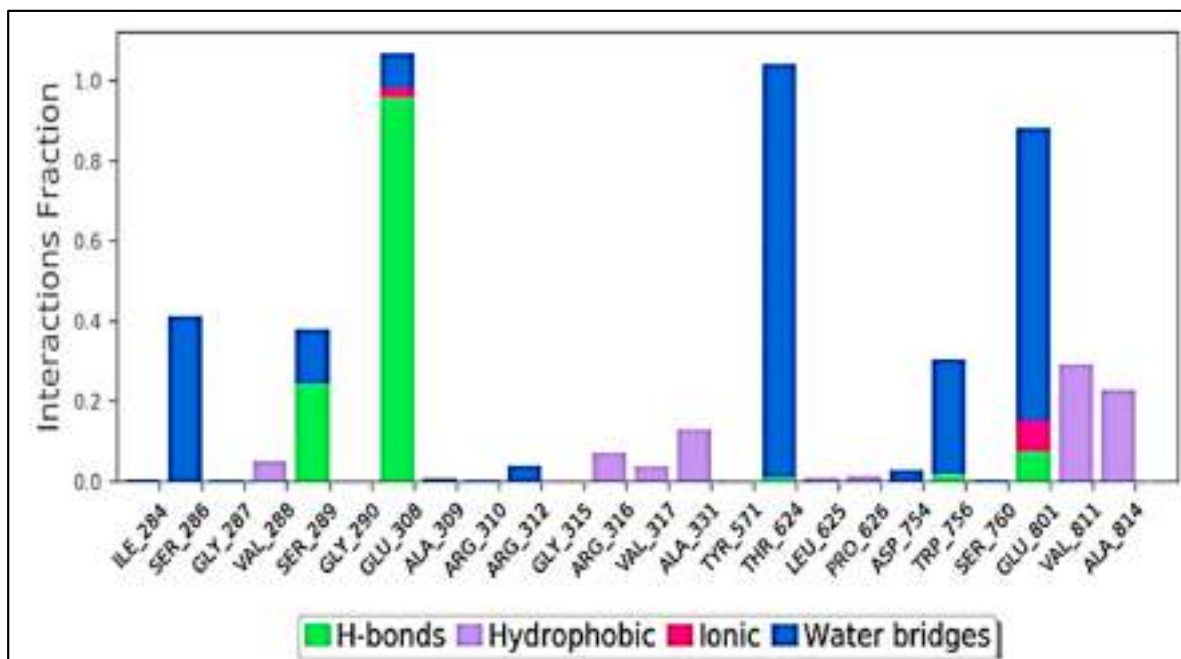


Figure 19. Protein-ligand contact histogram (H-bonds, Hydrophobic, Ionic, Water bridges) of LSD and compound 1.

Figure 20 depicts individual ligand atom interactions with protein residues. Interactions that last more than 30.0 percent of the simulation time (0.00 through 150.0 ns) in the chosen trajectory are shown.

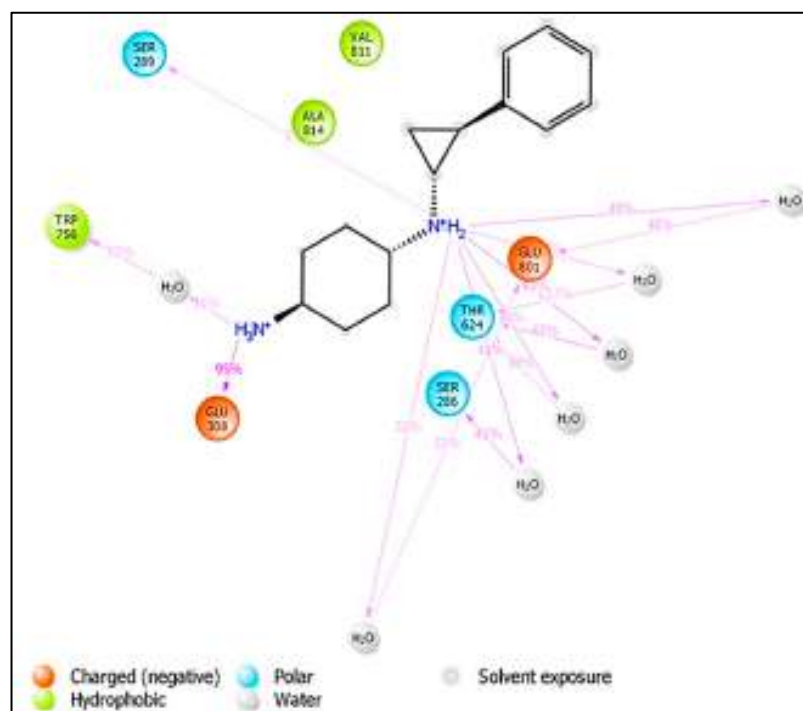


Figure 20. Ligand atom interactions with the protein residues LSD-compound 1.

The presence of protein secondary structural elements (SSE) such as alpha helices and beta strands is checked throughout the simulation to guarantee that they are not present. The plot above shows the distribution of SSE by residue index over the entire protein structure, and it includes all residues. The graphs at the bottom illustrate the evolution of each residue and its SSE assignment throughout the experiment, in contrast to the charts below, which show a summary of the SSE composition for each trajectory frame during the simulation (as shown in Figure 21).

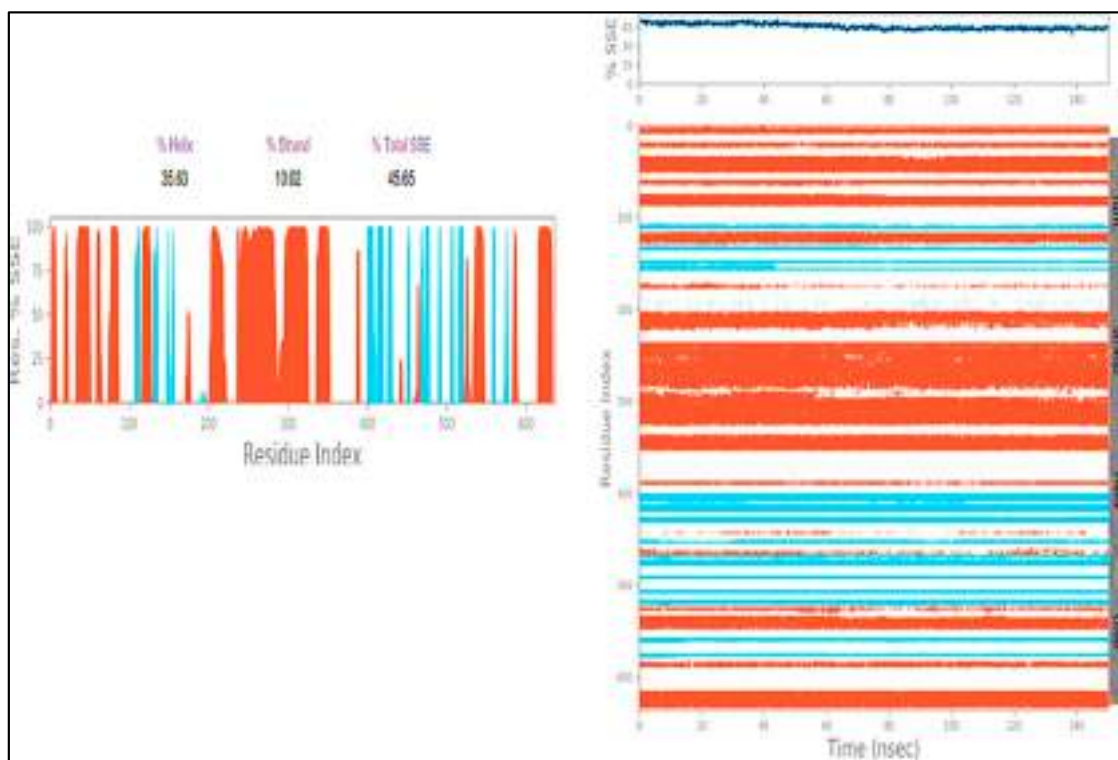


Figure 21. Secondary Structure element distribution by residue index throughout the protein structure. Red indicates alpha helices, and blue indicate beta-strands of; LSD-compound 1.

In comparison to the 0 ns structure, the positional change was obvious in the stepwise trajectory analysis of every 25 ns of compound 1 simulation with LSD (Figure 22). In order to achieve conformational stability and convergence, the ligand, compound 1, was discovered to exhibit structural angular mobility at the end frame.

The ligand torsions graphic depicts the conformational evolution of each rotatable bond (RB) in the ligand throughout the simulation trajectory (0.00 through 150.00 ns). The top panel shows a two-dimensional schematic of a ligand with color-coded rotatable bonds. Each rotatable bond torsion is accompanied with a dial plot and a bar plot of the same colour. The structure of the torsion during the simulation is depicted by dial (or radial) charts. The simulation begins in the radial display's centre, and the time evolution is plotted radially outwards.

In the bar charts, which summarize the data from the dial plots, the probability density of the torsion is shown. If torsional potential data is provided, the graphic also displays the potential of the rotatable bond (by summing the potential of the related torsions). The potential values are given in kcal/mol and are displayed on the chart's left Y-axis. The correlations between the histogram and torsion potential can reflect the conformational strain that the ligand undergoes in order to maintain a protein-bound shape (See Figure 23).

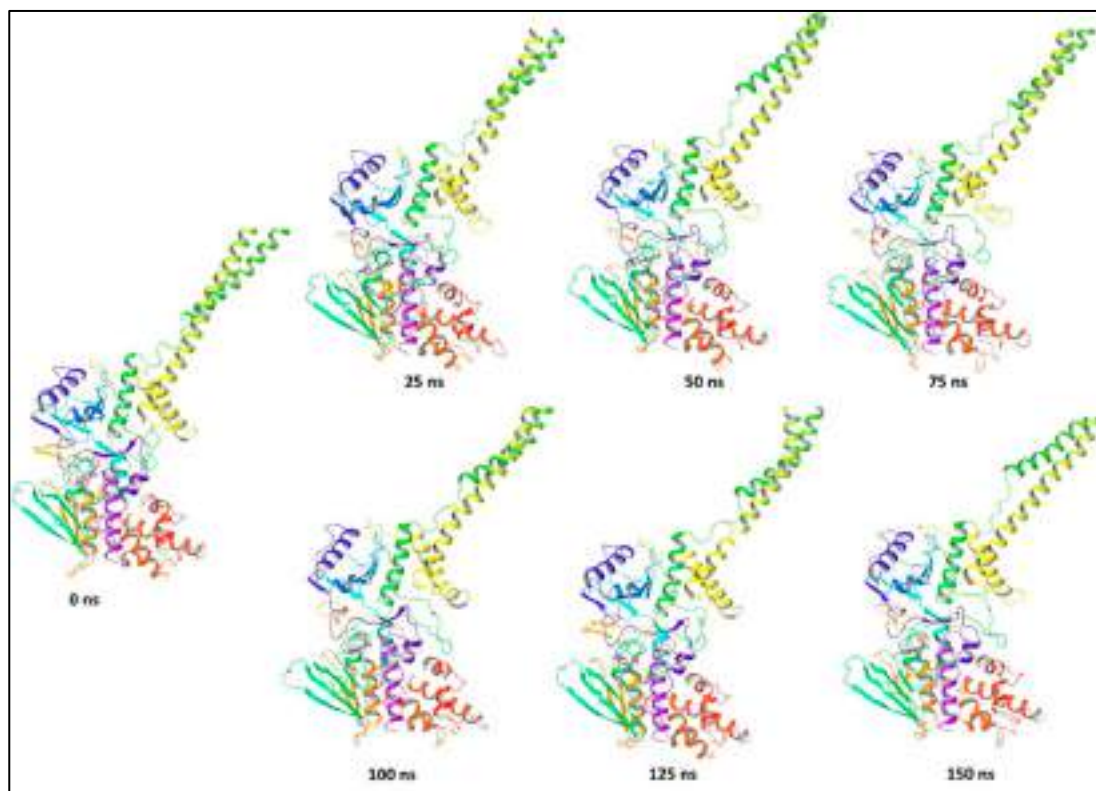


Figure 22. Stepwise trajectory analysis for every 25 ns displaying the protein and ligand conformation during 150 ns of simulation of LSD-compound 1.

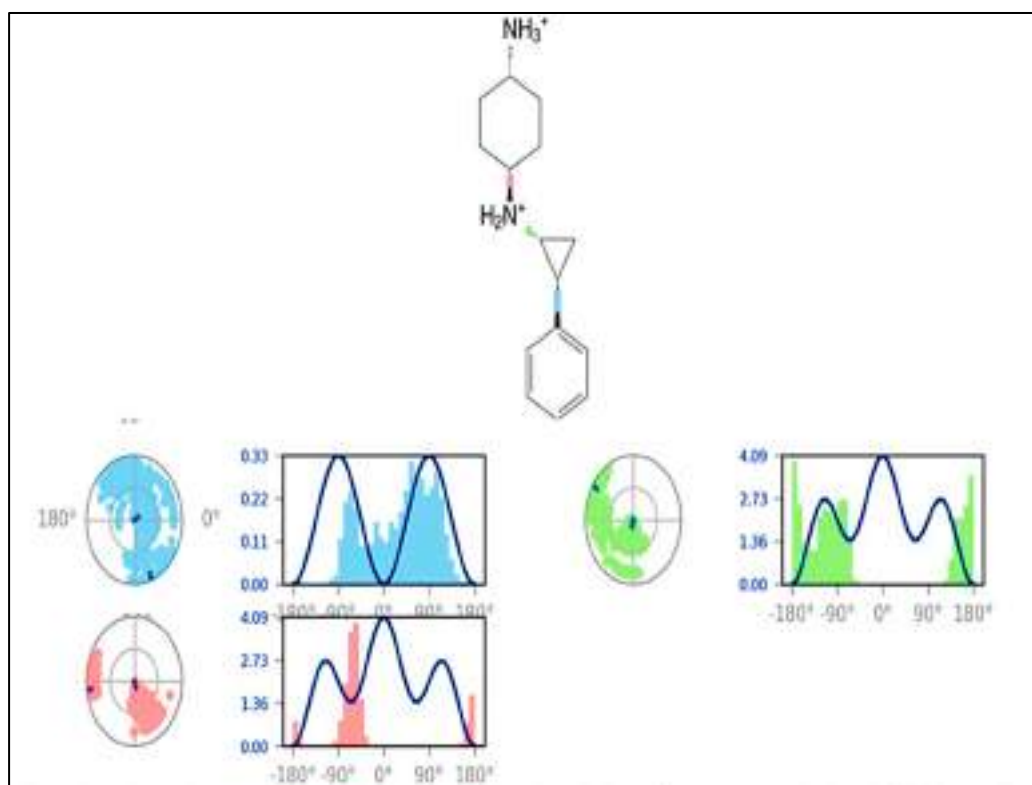


Figure 23. Ligand torsion profile of LSD-compound 1 displayed during 150 ns of simulation 1.

Molecular Mechanics Generalized Born and Surface Area (MMGBSA) calculations. The MMGBSA method is often used to determine the binding energy of ligands to protein molecules. The binding free energy of each protein–compound **1** complex was calculated, as well as the influence of the other non-bonded interactions energies (Table 1). The binding energy of ligand compound **1** with LSD is -59.78 kcal/mol. G_{bind} is governed by non-bonded interactions such as $G_{\text{bindCoulomb}}$, $G_{\text{bindCovalent}}$, $G_{\text{bindHbond}}$, G_{bindLipo} , $G_{\text{bindSolvGB}}$, and G_{bindvdW} . The G_{bindvdW} , G_{bindLipo} , and $G_{\text{bindCoulomb}}$ energies contributed the most to the average binding energy across all types of interactions. These conformational alterations result in improved binding pocket acquisition and engagement with residues, resulting in increased binding energy and stability. Thus, the binding energy obtained from docking results was well justified by MM-GBSA calculations. Furthermore, the last frame (150 ns) of MMGBSA displayed the positional change of compound **1** as compared to the 0 ns trajectory, indicating a better binding pose for best fitting in the protein's binding cavity (see Figure 24).

Table 1. Binding energy calculation of compound **1** with LSD and non-bonded interaction energies from MMGBSA trajectories. (star indicates mean of all the energy value).

Energies (kcal/mol) Mean *	LSD1 + Comp-1
ΔG_{bind}	-42.18 ± 7.60
$\Delta G_{\text{bindLipo}}$	-07.62 ± 4.78
$\Delta G_{\text{bindvdW}}$	-15.63 ± 7.77
$\Delta G_{\text{bindCoulomb}}$	-12.54 ± 4.07
$\Delta G_{\text{bindHbond}}$	-11.68 ± 2.00
$\Delta G_{\text{bindSolvGB}}$	31.53 ± 9.70
$\Delta G_{\text{bindCovalent}}$	10.22 ± 4.00

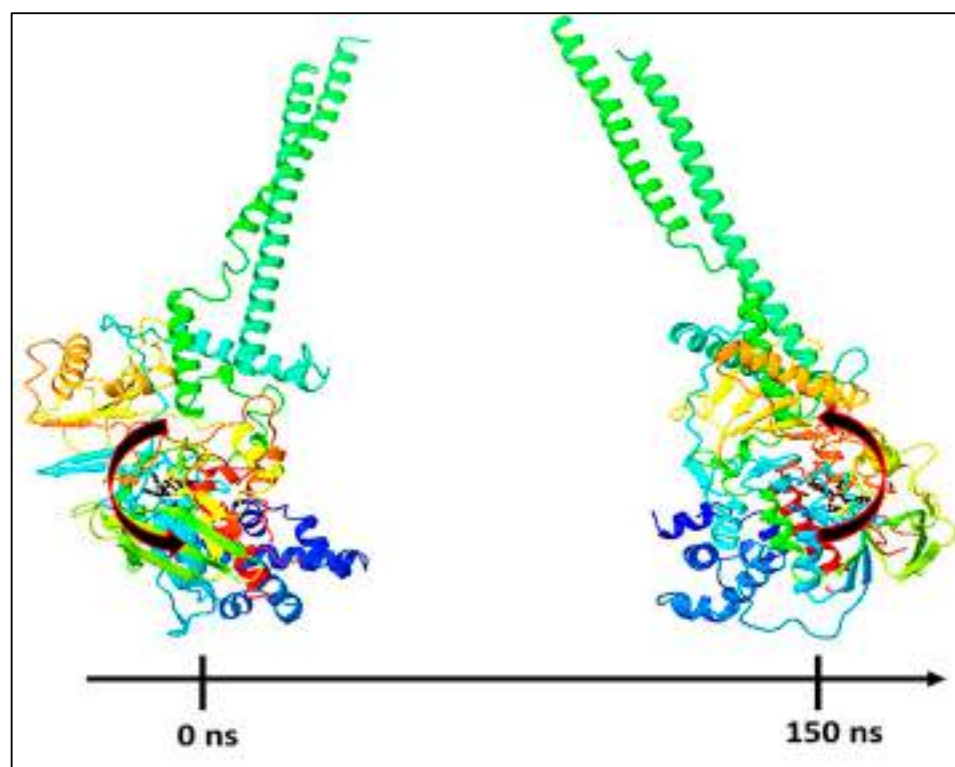


Figure 24. MMGBSA trajectory (0 ns, before simulation and 150 ns, after simulation) exhibited conformational changes of compound **1** upon binding with the proteins, LSD, the arrows indicating the overall positional variation (movement and pose) of compound **1** at the binding site cavity.

4. Materials and Methods

4.1. Preparation of Data Sets/Modeling Set Preparation from ChEMBL Data

Only compounds having experimental LSD1 inhibitory potency tested against a range of LSD1 inhibition assays were used in the ChEMBL [9] database. A limited data set of 84 LSD1 inhibitors with accurate EC_{50} values (0.38–89500 nM) was created from a crude dataset of 191 compounds with experimental EC_{50} values after removing structural duplicates, multi-component compounds or salts, and compounds with imprecise EC_{50} values. The EC_{50} values in nanomolar (nM) units were converted to molar units first (M). For the sake of data set handling, EC_{50} (M) values for each molecule were transformed to pEC_{50} ($pEC_{50} = -\log EC_{50}$). SMILES notations for all 84 substances with experimental EC_{50} and pEC_{50} values are listed in Table S1 in Supplementary. Figure 25 shows a representative example of the five least active and five most active LSD1 inhibitors.

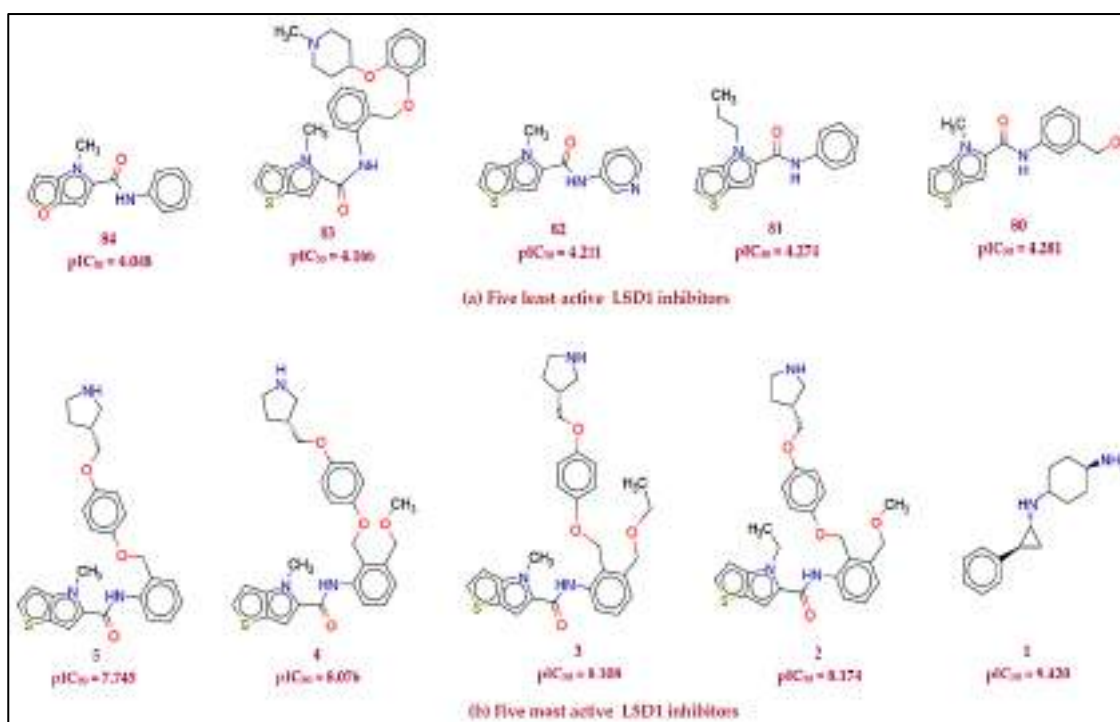


Figure 25. Representative (a) five least active and (b) five most active LSD1 inhibitors from the selected data set.

4.2. Calculation of Molecular Descriptors and Objective Feature Selection (OFS)

Using Open Babel 3.1, the SMILES notations were translated to 3D structures [15]. The most stable conformation is found in the geometry optimized molecule. As a result, calculating molecular descriptors on a dataset of optimized molecules assures that all physico-chemical attributes for all molecules in the dataset are uniform. Prior to calculating molecular descriptors, all of the compounds in the current dataset were optimized using TINKER (force field MMFF94). An appropriate calculation of many molecular descriptors is required in QSAR analysis to improve mechanistic understanding. A huge collection of more than 30,000 unique 1D- to 3D-molecular descriptors may be found in PyDescriptor, a PyMOL plugin [16]. Data trimming was performed to prevent the risk of overfitting due to noisy duplicated descriptors. Then, using QSARINS-2.2.4 [17], objective feature selection (OFS) was used to exclude near-constant, constant, and significantly inter-correlated ($|R| > 0.90$) molecular descriptors. Despite the fact that only 1733 molecular descriptors were accepted into the contracted molecular descriptor pool, it nevertheless has a wide range of descriptors that cover a wide chemical spectrum.

4.3. Splitting of the Data Set Molecules into Training and External Sets and Subjective Feature Selection

To avoid information leaking, it is critical to divide the entire data set into training and prediction sets with correct configuration and sizes prior to rigorous subjective feature selection [18]. To avoid bias, the entire data set was arbitrarily divided into two sets: training (an 80%, or 67 molecules) and prediction (20%, or 17 molecules). The sole objective of a training set is to select an acceptable number of molecular descriptors for developing QSAR models, whereas the prediction set is used to validate these models externally (Predictive QSAR). The genetic algorithm-reinforced multilinear regression (GA-MLR) method, as implemented in QSARINS-2.2.4, was used to pick acceptable descriptors using Q^2_{LOO} as a fitness parameter for subjective feature selection.

To construct a good QSAR model, it is critical to avoid overfitting and to choose an appropriate number of molecular descriptors in order to provide satisfactory interpretability. As a result, a graph of the number of molecular descriptors (X -axis) involved in the models against R^2_{tr} and Q^2_{LOO} values (Y -axis) has been plotted in the current communication to achieve breaking point, with the number of molecular descriptors corresponding to the breaking point being an optimum number of descriptors in QSAR model building. Because the graph in Figure 3 shows a breaking point at five variables, QSAR models with more than five descriptors were eliminated (See Figure 26).

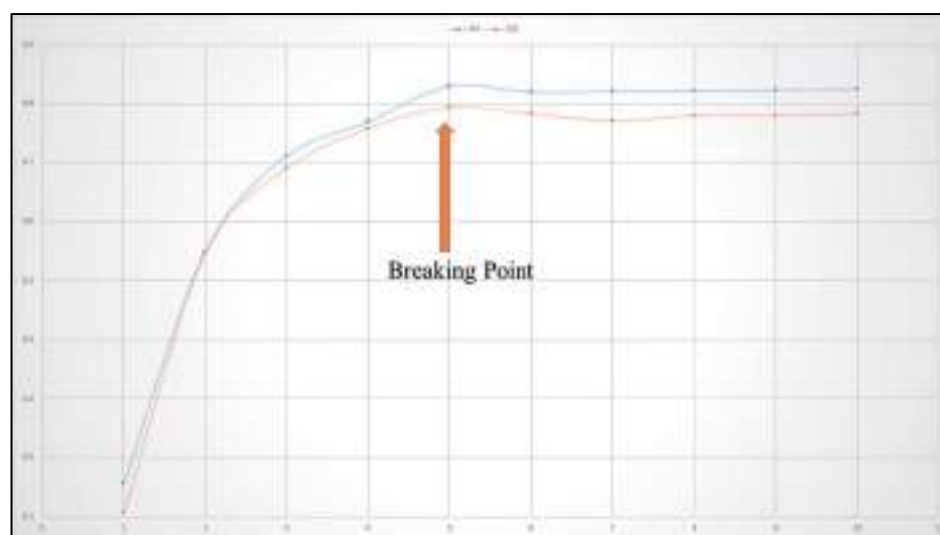


Figure 26. Depiction of Plot for the number of descriptors against the Coefficient of Determination R^2 and Leave-One-Out Coefficient of Determination Q^2 to identify the optimum number of descriptors.

4.4. Model Development and Validation

The robustness of the created models was determined using a variety of validation criteria reported in the literature. Internal predictability and statistical quality of the developed model were tested using parameters such as the coefficient of determination (r^2), leave-one-out cross-validation (Q^2_{LOO}), and leave-many-out cross-validation to achieve this (Q^2_{LMO}). In addition, for each developed model, the standard error of estimate(s) was defined. For the given QSAR models for the stated dataset, RMSE (Root Mean Squared of Errors) for the training (RMSE_{TR}) and external prediction sets (RMSE_{ext}) that denote the complete error of the model that was predicted as an extra portion of the accuracy [5,18] were used.

The QUIK rule (Q Under the Influence of K) was used to examine the inter-correlation between descriptors. To reduce inter-correlation among descriptors, the QUICK rule was set to 0.05. The fit of the randomly reordered Y -data was checked using Y -randomization with 2000 iterations to ensure the trustworthiness of the created QSAR model. The dependent variables (pEC50 value) of the training set were shuffled, and new coefficients of

determination were produced for the randomization of the constructed QSAR model. The new models' coefficients of determination are significantly low, indicating that the reported model in this QSAR research was not acquired by chance correlation [19].

All models were externally validated using the following validation criteria: r^2_{ext} (external determination coefficient), Q^2_{F1} , Q^2_{F2} , Q^2_{F3} , Concordance Correlation Coefficient (CCC), CCC_{ex} , r^2_m , and r^2_o . The R^2_m (overall) parameter penalizes a model when there are big disparities between observed and predicted values of all the compounds in the collection (considering both training and test sets). The difference between the values of the expected and the resultant experimental activity data was assessed using the r^2_m (pEC_{50} value). It has been suggested that the observed value for the r^2_m should be lower than 0.2 if the r^2_o value is more than 0.5. To validate model reliability and robustness, all QSAR models were examined for validation parameters such as Golbraikh and Tropsha's criterion.

In general, the created QSAR model's predictive ability is determined by how well the anticipated value matches the observed (experimental biological activity) value. Even the presence of a single outlier reduces the generated QSAR model's prediction ability. Following that, we attempted to identify the outliers based on compound with a considerably high residual value in GA-MLR QSAR models. Furthermore, by comparing the predicted value to the standardized residual values, we were able to identify the outlier compounds. Similarly, the leverage effect in Williams' plot revealed structural variation in database compounds. The created QSAR model's applicability domain is determined by combining the leverage and standard residuals [20–23].

4.5. Molecular Docking Analysis

The protein data bank (<https://www.rcsb.org/structure/2DW4>, accessed on 24 May 2022) was used to obtain the pdb file for the LSD1 receptor. The pdb 2dw4 [24] was chosen because of its X-ray resolution and sequence completion. The health of the protein was evaluated before actual docking simulations by plotting Ramchandran's plot [25] (See Figure 27). For docking analysis, the optimized protein is acceptable. Although all of the compounds were docked into the active site, the docking pose for the most active compound 1 as a representative has been shown below for convenience.

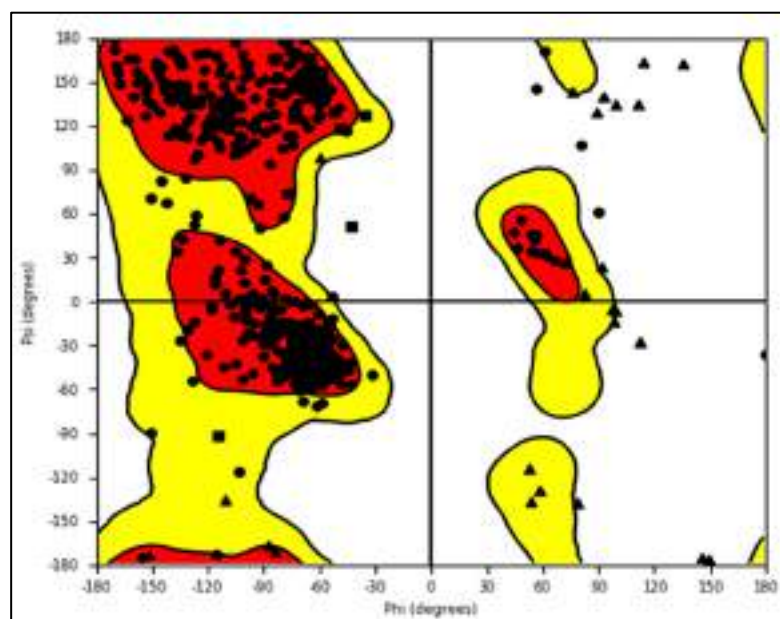


Figure 27. Ramchandran plot for LSD1 receptor.

The software NRGSuite [26] was utilized for molecular docking analysis. This open-source software is accessible as a PyMOL plugin (www.pymol.org, accessed on 7 July 2022). With the help of FlexAID [27], it can detect the surface cavities in a protein and use them as

target-binding sites for docking simulations. It models ligand and side-chain flexibility, as well as covalent docking, and employs a genetic algorithm for conformational search. To gain the best performance using NRGsuite, we used a flexible-rigid docking technique with the following default settings: Side chain flexibility—no; ligand flexibility—yes; ligand pose as reference—no; constraints- no; HET groups- included water molecules; van der Waals permeability—0.1; solvent types—no type; number of chromosomes—1000; number of generations—1000; fitness model—share; reproduction model—population boom; number of TOP complexes—5. The native ligand, a known tranylcypromine inhibitor of LSD1 [24], was used to validate the docking technique for molecular docking.

4.6. Molecular Dynamic Simulation

Desmond, a package from Schrödinger LLC [28], was used to simulate molecular dynamics for 150 nanoseconds. Docking experiments provided the earliest step of protein and ligand complexes for molecular dynamics simulation. In static settings, Molecular Docking Studies can predict the ligand binding state. Because docking provides a static view of a molecule's binding pose in a protein's active site [29], MD simulations tend to compute atom movements over time by integrating Newton's classical equation of motion. The ligand binding status in the physiological milieu was predicted using simulations [30,31].

Protein Preparation Wizard or Maestro was used to preprocess the protein–ligand complex, which included complex optimization and minimization. The System Builder tool was used to prepare all of the systems. TIP3P was chosen as a solvent model with an orthorhombic box (Transferable Intermolecular Interaction Potential 3 Points). In the simulation, the OPLS 2005 force field was used [32]. Counter ions were added to the models to make them neutral. A total of 0.15 M salt (NaCl) was added to replicate physiological circumstances. For the entire simulation, the NPT ensemble with 300 K temperature and 1 atm pressure was chosen. Before the simulation, the models were loosened. After every 100 ps, the trajectories were saved for analysis, and the simulation's stability was determined by measuring the root mean square deviation (RMSD) of the protein and ligand over time.

4.7. Molecular Mechanics Generalized Born and Surface Area (MMGBSA) Calculations

During MD simulations of LSD complexed with complex 1, the binding free energy (G_{bind}) of docked complexes was calculated using the premier molecular mechanics generalized born surface area (MM-GBSA) module (Schrodinger suite, LLC, New York, NY, USA, 2017-4). The binding free energy was calculated using the OPLS 2005 force field, VSGB solvent model, and rotamer search methods [16–18]. After the MD run, 10 ns intervals were used to choose the MD trajectories frames. The total free energy binding was calculated using Equation (1):

$$\Delta G_{\text{bind}} = G_{\text{complex}} - (G_{\text{protein}} + G_{\text{ligand}}) \quad (1)$$

where, ΔG_{bind} = binding free energy, G_{complex} = free energy of the complex, G_{protein} = free energy of the target protein, and G_{ligand} = free energy of the ligand. The MMGBSA outcome trajectories were analyzed further for post-dynamics structural modifications.

5. Conclusions

Pharmacophoric traits responsible for improved LSD1 inhibition unraveled by present QSAR evaluation are interconnected and thus easy to incorporate to optimize present LSD1 inhibitors towards more potent analogues; for example, a higher number of Nitrogen atoms precisely at six bonds and a lower number of Hydrogen atoms at three bonds from the ring Carbon atom can be introduced at the same time to optimize the LSD1 inhibitors towards better activity, and a higher number of non-ring Oxygen atoms precisely at nine bonds from the amide Nitrogen and a less frequent occurrence of sp^2 Oxygen within 4Å boosts the LSD1 inhibitory activity. Likewise, the hydrogen bond donor atom at two bonds and amide Nitrogen at four bonds from sp^3 hybridized Carbon atoms enhances the desired

activity. Two of the five descriptors in the split-set model emphasize the relevance of the ring carbon atom, whereas one descriptor represents the importance of the ring Sulphur atom, indicating that there is room for modification of dataset compounds for greater LSD1 inhibition. On the other hand, two out of five descriptors emphasize the relevance of amide nitrogen, suggesting that the current dataset compounds might be optimized for improved LSD1 inhibition. Lipophilic atoms, such as ring carbon atoms, were identified as a possible center for the optimization of LSD1 inhibitors for anticancer efficacy by certain chemical descriptors. As a result, the created QSAR model may be used to improve compounds for better LSD1 inhibition and cancer prevention. The docking results revealed that the descriptors, **fdonsp3C2B** and **lipo_ringS_8Bc**, played important roles in the inhibition of the LSD1 receptor, which was consistent with the QSAR findings. The MD simulation results display that the ligands were tightly bound to the binding site of the receptor during the simulation. The ligands are still firmly connected to the receptor's binding site, as evidenced by the fact that the RMSD values for the ligand fit-to-protein does not significantly vary over the course of the simulation. Compound 1's position was altered in the last 150-ns frame of the MMGBSA simulation, compared to the 0-ns trajectory, indicating a more advantageous binding pose for the binding cavity of the protein. Therefore, the MD simulation and MMGBSA analysis strengthens the outcome of the QSAR and Molecular docking studies.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/molecules27154758/s1>. Table S1: The SMILES notation for 84 leads, along with their reported EC₅₀ and pEC₅₀ values. Table S2: The values for selected molecular descriptors present in QSAR models.

Author Contributions: Conceptualization, R.D.J., R.L.B. and N.M.—formal analysis; data curation, N.M., A.G. (Arabinda Ghosh), M.E.A.Z. and S.A.A.-H.; formal analysis, N.M., A.A.A.-M.—formal analysis, A.G. (Ajaykumar Gandhi), A.S. and V.H.M.; methodology, N.M., A.G. (Ajaykumar Gandhi) and V.H.M.; writing—original draft, N.M., A.G. (Ajaykumar Gandhi) and S.A.A.-H.; writing—review and editing. All authors have read and agreed to the published version of the manuscript.

Funding: The authors extend their sincere appreciation to the Deanship of Scientific Research, Imam Mohammad Ibn Saud Islamic University, Saudi Arabia, for its support of this research through the Research Group No RG-21-09-77.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data is available in the Supplementary section.

Acknowledgments: The authors are thankful to Paola Gramatica and her team for providing QSARINS-v2.2.4, and the developers of TINKER, ChemSketch 12 Freeware (ACD labs), and Py-Descriptor for providing the free versions of their software.

Conflicts of Interest: The authors declare no conflict of interest.

Sample Availability: Not applicable.

Abbreviations

CADD	Computer Aided Drug Designing
SMILES	Simplified Molecular-Input Line-Entry System
GA	Genetic Algorithm
MLR	Multiple Linear Regression
QSAR	Quantitative Structure-Activity Relationship
QSARINS	QSAR Insubria
OECD	Organization for Economic Co-operation and Development
OFS	Objective Feature Selection
SFS	Subjective Feature Selection
LSD1	Lysine-specific demethylase 1

References

1. Maiques-Diaz, A.; Somervaille, T.C.P. LSD1: Biologic roles and therapeutic targeting. *Epigenomics* **2016**, *8*, 1103–1116. [[CrossRef](#)]
2. McGrath, J.P.; Williamson, K.E.; Balasubramanian, S.; Odate, S.; Arora, S.; Hatton, C.; Edwards, T.M.; O'Brien, T.; Magnuson, S.; Stokoe, D.; et al. Pharmacological Inhibition of the Histone Lysine Demethylase KDM1A Suppresses the Growth of Multiple Acute Myeloid Leukemia Subtypes. *Cancer Res.* **2016**, *76*, 1975–1988. [[CrossRef](#)]
3. Mohammad Helai, P.; Smitheman Kimberly, N.; Kamat Chandrashekar, D.; Soong, D.; Federowicz Kelly, E.; Van Aller Glenn, S.; Schneck Jess, L.; Carson Jeffrey, D.; Liu, Y.; Butticello, M.; et al. A DNA Hypomethylation Signature Predicts Antitumor Activity of LSD1 Inhibitors in SCLC. *Cancer Cell* **2015**, *28*, 57–69. [[CrossRef](#)]
4. Ma, L.-Y.; Zheng, Y.-C.; Wang, S.-Q.; Wang, B.; Wang, Z.-R.; Pang, L.-P.; Zhang, M.; Wang, J.-W.; Ding, L.; Li, J.; et al. Design, Synthesis, and Structure–Activity Relationship of Novel LSD1 Inhibitors Based on Pyrimidine–Thiourea Hybrids As Potent, Orally Active Antitumor Agents. *J. Med. Chem.* **2015**, *58*, 1705–1716. [[CrossRef](#)]
5. Cherkasov, A.; Muratov, E.N.; Fourches, D.; Varnek, A.; Baskin, I.I.; Cronin, M.; Dearden, J.; Gramatica, P.; Martin, Y.C.; Todeschini, R.; et al. QSAR Modeling: Where Have You Been? Where Are You Going To? *J. Med. Chem.* **2014**, *57*, 4977–5010. [[CrossRef](#)]
6. Fujita, T.; Winkler, D.A. Understanding the Roles of the “Two QSARs”. *J. Chem. Inf. Model.* **2016**, *56*, 269–274. [[CrossRef](#)]
7. Abdizadeh, R.; Heidarian, E.; Hadizadeh, F.; Abdizadeh, T. QSAR Modeling, Molecular Docking and Molecular Dynamics Simulations Studies of Lysine-Specific Demethylase 1 (LSD1) Inhibitors as Anticancer Agents. *Anti-Cancer Agents Med. Chem.* **2021**, *21*, 987–1018. [[CrossRef](#)]
8. Fang, Y.; Liao, G.; Yu, B. LSD1/KDM1A inhibitors in clinical trials: Advances and prospects. *J. Hematol. Oncol.* **2019**, *12*, 129. [[CrossRef](#)]
9. Gaulton, A.; Hersey, A.; Nowotka, M.; Bento, A.P.; Chambers, J.; Mendez, D.; Mutowo, P.; Atkinson, F.; Bellis, L.J.; Cibrián-Uhalte, E.; et al. The ChEMBL database in 2017. *Nucleic Acids Res.* **2017**, *45*, D945–D954. [[CrossRef](#)]
10. Vianello, P.; Sartori, L.; Amigoni, F.; Cappa, A.; Fagá, G.; Fattori, R.; Legnaghi, E.; Ciossani, G.; Mattevi, A.; Meroni, G.; et al. Thieno[3,2-b]pyrrole-5-carboxamides as New Reversible Inhibitors of Histone Lysine Demethylase KDM1A/LSD1. Part 2: Structure-Based Drug Design and Structure–Activity Relationship. *J. Med. Chem.* **2017**, *60*, 1693–1715. [[CrossRef](#)]
11. Zheng, Y.-C.; Ma, J.; Wang, Z.; Li, J.; Jiang, B.; Zhou, W.; Shi, X.; Wang, X.; Zhao, W.; Liu, H.-M. A Systematic Review of Histone Lysine-Specific Demethylase 1 and Its Inhibitors. *Med. Res. Rev.* **2015**, *35*, 1032–1071. [[CrossRef](#)]
12. Niwa, H.; Sato, S.; Hashimoto, T.; Matsuno, K.; Umehara, T. Crystal Structure of LSD1 in Complex with 4-[5-(Piperidin-4-ylmethoxy)-2-(p-tolyl)pyridin-3-yl]benzotrile. *Molecules* **2018**, *23*, 1538. [[CrossRef](#)]
13. Stazi, G.; Zwergel, C.; Valente, S.; Mai, A. LSD1 inhibitors: A patent review (2010–2015). *Expert Opin. Ther. Pat.* **2016**, *26*, 565–580. [[CrossRef](#)]
14. Magliulo, D.; Bernardi, R.; Messina, S. Lysine-Specific Demethylase 1A as a Promising Target in Acute Myeloid Leukemia. *Front. Oncol.* **2018**, *8*, 255. [[CrossRef](#)]
15. O'Boyle, N.M.; Banck, M.; James, C.A.; Morley, C.; Vandermeersch, T.; Hutchison, G.R. Open Babel: An open chemical toolbox. *J. Cheminform.* **2011**, *3*, 33. [[CrossRef](#)]
16. Masand, V.H.; Rastija, V. PyDescriptor: A new PyMOL plugin for calculating thousands of easily understandable molecular descriptors. *Chemom. Intell. Lab. Syst.* **2017**, *169*, 12–18. [[CrossRef](#)]
17. Gramatica, P.; Chirico, N.; Papa, E.; Cassani, S.; Kovarich, S. QSARINS: A new software for the development, analysis, and validation of QSAR MLR models. *J. Comput. Chem.* **2013**, *34*, 2121–2132. [[CrossRef](#)]
18. Masand, V.H.; Mahajan, D.T.; Nazeruddin, G.M.; Hadda, T.B.; Rastija, V.; Alfeefy, A.M. Effect of information leakage and method of splitting (rational and random) on external predictive ability and behavior of different statistical parameters of QSAR model. *Med. Chem. Res.* **2014**, *24*, 1241–1264. [[CrossRef](#)]
19. Consonni, V.; Todeschini, R.; Ballabio, D.; Grisoni, F. On the Misleading Use of QF32 for QSAR Model Comparison. *Mol. Inform.* **2019**, *38*, 1800029. [[CrossRef](#)]
20. Krstajic, D.; Buturovic, L.J.; Leahy, D.E.; Thomas, S. Cross-validation pitfalls when selecting and assessing regression and classification models. *J. Cheminform.* **2014**, *6*, 10. [[CrossRef](#)]
21. Martin, T.M.; Harten, P.; Young, D.M.; Muratov, E.N.; Golbraikh, A.; Zhu, H.; Tropsha, A. Does Rational Selection of Training and Test Sets Improve the Outcome of QSAR Modeling? *J. Chem. Inf. Model.* **2012**, *52*, 2570–2578. [[CrossRef](#)]
22. Chirico, N.; Gramatica, P. Real External Predictivity of QSAR Models. Part 2. New Intercomparable Thresholds for Different Validation Criteria and the Need for Scatter Plot Inspection. *J. Chem. Inf. Model.* **2012**, *52*, 2044–2058. [[CrossRef](#)]
23. Roy, P.P.; Kovarich, S.; Gramatica, P. QSAR model reproducibility and applicability: A case study of rate constants of hydroxyl radical reaction models applied to polybrominated diphenyl ethers and (benzo-)triazoles. *J. Comput. Chem.* **2011**, *32*, 2386–2396. [[CrossRef](#)]
24. Mimasu, S.; Sengoku, T.; Fukuzawa, S.; Umehara, T.; Yokoyama, S. Crystal structure of histone demethylase LSD1 and tranylcypromine at 2.25 Å. *Biochem. Biophys. Res. Commun.* **2008**, *366*, 15–22. [[CrossRef](#)]
25. Hollingsworth, S.A.; Karplus, P.A. A fresh look at the Ramachandran plot and the occurrence of standard structures in proteins. *BioMol. Concepts* **2010**, *1*, 271–283. [[CrossRef](#)]
26. Gaudreault, F.; Morency, L.-P.; Najmanovich, R.J. NRGsuite: A PyMOL plugin to perform docking simulations in real time using FlexAID. *Bioinformatics* **2015**, *31*, 3856–3858. [[CrossRef](#)]

27. Gaudreault, F.; Najmanovich, R.J. FlexAID: Revisiting Docking on Non-Native-Complex Structures. *J. Chem. Inf. Model.* **2015**, *55*, 1323–1336. [[CrossRef](#)]
28. Bowers, K.J.A.C.; David, E.; Xu, H.; Dror Ron, O.; Eastwood Michael, P.; Gregersen Brent, A.; Klepeis John, L.; Kolossvary, I.; Moraes, M.A.; Sacerdoti Federico, D.; et al. Scalable Algorithms for Molecular Dynamics Simulations on Commodity Clusters. In Proceedings of the 2006 ACM/IEEE Conference on Supercomputing, Tampa, FL, USA, 11–17 November 2006. [[CrossRef](#)]
29. Ferreira, L.G.; Dos Santos, R.N.; Oliva, G.; Andricopulo, A.D. Molecular docking and structure-based drug design strategies. *Molecules* **2015**, *20*, 13384–13421. [[CrossRef](#)]
30. Hildebrand, P.W.; Rose, A.S.; Tiemann, J.K.S. Bringing Molecular Dynamics Simulation Data into View. *Trends Biochem. Sci.* **2019**, *44*, 902–913. [[CrossRef](#)]
31. Rasheed, M.A.; Iqbal, M.N.; Saddick, S.; Ali, I.; Khan, F.S.; Kanwal, S.; Ahmed, D.; Ibrahim, M.; Afzal, U.; Awais, M. Identification of Lead Compounds against Scm (fms10) in *Enterococcus faecium* Using Computer Aided Drug Designing. *Life* **2021**, *11*, 77. [[CrossRef](#)]
32. Shivakumar, D.; Williams, J.; Wu, Y.; Damm, W.; Shelley, J.; Sherman, W. Prediction of Absolute Solvation Free Energies using Molecular Dynamics Free Energy Perturbation and the OPLS Force Field. *J. Chem. Theory Comput.* **2010**, *6*, 1509–1519. [[CrossRef](#)] [[PubMed](#)]

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/363036509>

OPEN ACCESS EDITED BY Repurposing food molecules as a potential BACE1 inhibitor for Alzheimer's disease

Article · August 2022

CITATIONS

0

READS

55

31 authors, including:



Badrah Alghamdi

King Abdulaziz University

73 PUBLICATIONS 563 CITATIONS

SEE PROFILE



Filia Alexiou

4 PUBLICATIONS 4 CITATIONS

SEE PROFILE



Rahul Jawarkar

Dr Rajendra Gode Institute of Pharmacy Amravati Maharashtra

64 PUBLICATIONS 406 CITATIONS

SEE PROFILE



Swastika Maitra

Adamas knowledge city

8 PUBLICATIONS 4 CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:



Hodgkin disease [View project](#)



Qualitative QSAR: Easy Interpretation of QSAR Models [View project](#)



OPEN ACCESS

EDITED BY
Mercè Pallàs,
University of Barcelona, SpainREVIEWED BY
Hoon Kim,
Sunchon National University,
South Korea
Rehan Khan,
Institute of Nano Science
and Technology (INST), India
Anas Shamsi,
Jamia Millia Islamia, India*CORRESPONDENCE
Nobendu Mukerjee
nabendu21@arkmvccrahara.org
Ghulam Md Ashraf
ashraf.gm@gmail.com
Arabinda Ghosh
dra.ghosh@gauhati.ac.in†These authors have contributed
equally to this workSPECIALTY SECTION
This article was submitted to
Alzheimer's Disease and Related
Dementias,
a section of the journal
Frontiers in Aging Neuroscience

RECEIVED 17 February 2022

ACCEPTED 07 July 2022

PUBLISHED 22 August 2022

COPYRIGHT

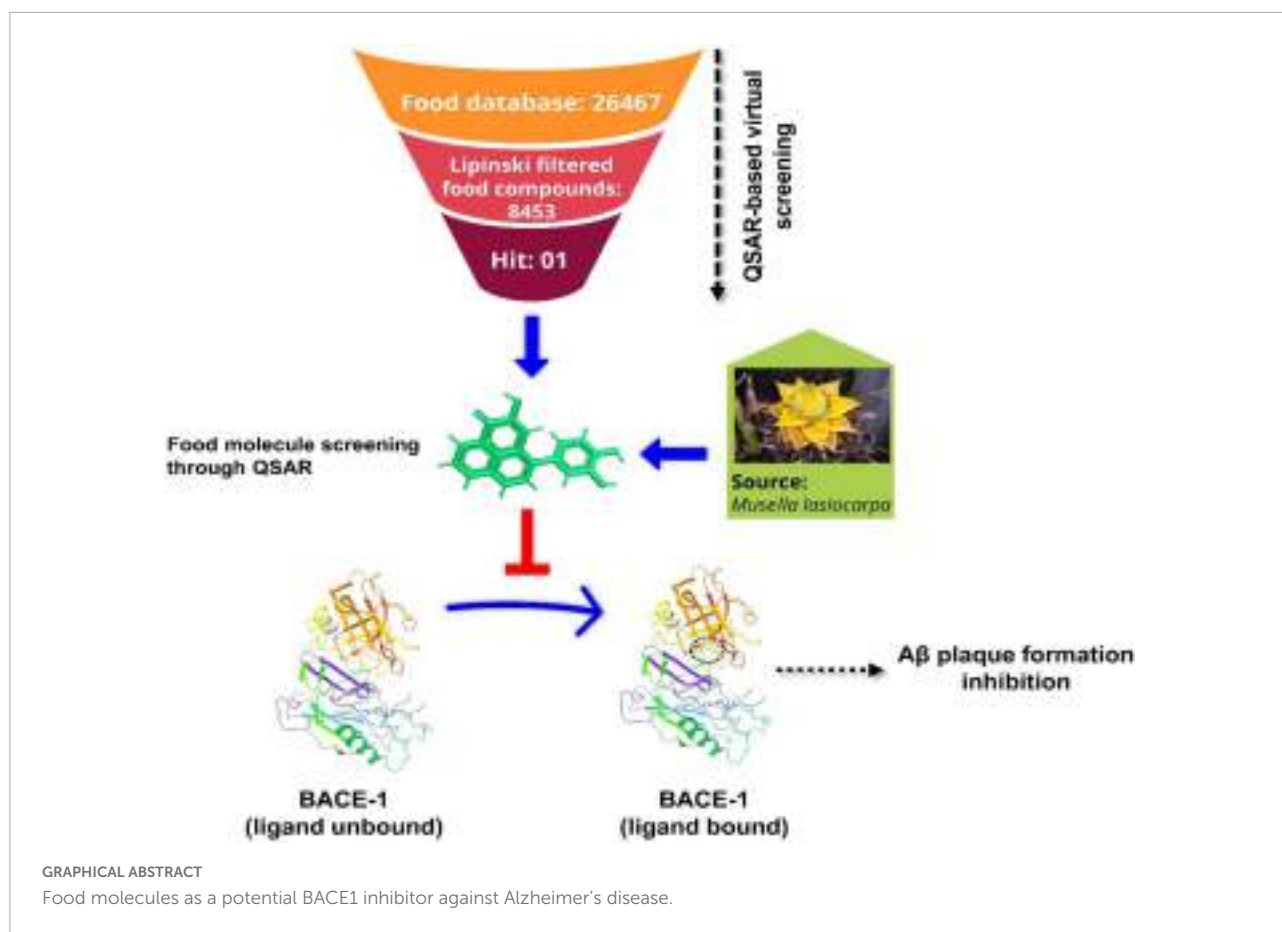
© 2022 Mukerjee, Das, Jawarkar,
Maitra, Das, Castrosanto, Paul, Samad,
Zaki, Al-Hussain, Masand, Hasan,
Bukhari, Perveen, Alghamdi, Alexiou,
Kamal, Dey, Malik, Bakal, Abuzenadah,
Ghosh and Md Ashraf. This is an
open-access article distributed under
the terms of the [Creative Commons
Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use,
distribution or reproduction in other
forums is permitted, provided the
original author(s) and the copyright
owner(s) are credited and that the
original publication in this journal is
cited, in accordance with accepted
academic practice. No use, distribution
or reproduction is permitted which
does not comply with these terms.

Repurposing food molecules as a potential BACE1 inhibitor for Alzheimer's disease

Nobendu Mukerjee^{1,2*†}, Anubhab Das^{3†}, Rahul D. Jawarkar^{4†},
Swastika Maitra^{5†}, Padmashree Das⁶, Melvin A. Castrosanto⁷,
Soumyadip Paul¹, Abdul Samad⁸, Magdi E. A. Zaki⁹,
Sami A. Al-Hussain⁹, Vijay H. Masand¹⁰,
Mohammad Mehedi Hasan¹¹, Syed Nasir Abbas Bukhari¹²,
Asma Perveen¹³, Badrah S. Alghamdi^{14,15,16},
Athanasios Alexiou^{17,18}, Mohammad Amjad Kamal^{19,20,21,22},
Abhijit Dey²³, Sumira Malik²⁴, Ravindra L. Bakal⁴,
Adel Mohammad Abuzenadah^{20,25}, Arabinda Ghosh^{26*} and
Ghulam Md Ashraf^{15,25*}

¹Department of Microbiology, Ramakrishna Mission Vivekananda Centenary College, Khardaha, India, ²Department of Health Sciences, Novel Global Community Educational Foundation, Hebersham, NSW, Australia, ³Institute of Health Sciences, Presidency University, Kolkata, India, ⁴Department of Medicinal Chemistry, Dr. Rajendra Gode Institute of Pharmacy, Amravati, India, ⁵Department of Microbiology, Adamas University, Kolkata, India, ⁶Central Silk Board, Guwahati, India, ⁷Institute of Chemistry, University of the Philippines Los Baños, Los Baños, Philippines, ⁸Department of Pharmaceutical Chemistry, Faculty of Pharmacy, Tishk International University, Erbil, Iraq, ⁹Department of Chemistry, Faculty of Science, Imam Mohammad Ibn Saud Islamic University, Riyadh, Saudi Arabia, ¹⁰Department of Chemistry, Vidya Bharati Mahavidyalaya, Amravati, India, ¹¹Department of Biochemistry and Molecular Biology, Faculty of Life Sciences, Mawlana Bhashani Science and Technology University, Tangail, Bangladesh, ¹²Department of Pharmaceutical Chemistry, College of Pharmacy, Jouf University, Sakaka, Saudi Arabia, ¹³Glocal School of Life Sciences, Glocal University, Saharanpur, India, ¹⁴Department of Physiology, Faculty of Medicine, King Abdulaziz University, Jeddah, Saudi Arabia, ¹⁵Pre-Clinical Research Unit, King Fahd Medical Research Center, King Abdulaziz University, Jeddah, Saudi Arabia, ¹⁶The Neuroscience Research Unit, Faculty of Medicine, King Abdulaziz University, Jeddah, Saudi Arabia, ¹⁷Department of Science and Engineering, Novel Global Community Educational Foundation, Hebersham, NSW, Australia, ¹⁸AFNP Med, Vienna, Austria, ¹⁹Institutes for Systems Genetics, Frontiers Science Center for Disease-Related Molecular Network, West China Hospital, Sichuan University, Chengdu, China, ²⁰King Fahd Medical Research Center, King Abdulaziz University, Jeddah, Saudi Arabia, ²¹Department of Pharmacy, Faculty of Allied Health Sciences, Daffodil International University, Dhaka, Bangladesh, ²²Enzymoics, Novel Global Community Educational Foundation, Hebersham, NSW, Australia, ²³Department of Life Sciences, Presidency University, Kolkata, India, ²⁴Amity Institute of Biotechnology, Amity University, Jharkhand, Ranchi, India, ²⁵Department of Medical Laboratory Sciences, Faculty of Applied Medical Sciences, King Abdulaziz University, Jeddah, Saudi Arabia, ²⁶Microbiology Division, Department of Botany, Gauhati University, Guwahati, India

Alzheimer's disease (AD) is a severe neurodegenerative disorder of the brain that manifests as dementia, disorientation, difficulty in speech, and progressive cognitive and behavioral impairment. The emerging therapeutic approach to AD management is the inhibition of β -site APP cleaving enzyme-1 (BACE1), known to be one of the two aspartyl proteases that cleave β -amyloid precursor protein (APP). Studies confirmed the association of high BACE1 activity with the proficiency in the formation of β -amyloid-containing neurotoxic plaques, the characteristics of AD. Only a few FDA-approved BACE1 inhibitors



are available in the market, but their adverse off-target effects limit their usage. In this paper, we have used both ligand-based and target-based approaches for drug design. The QSAR study entails creating a multivariate GA-MLR (Genetic Algorithm-Multilinear Regression) model using 552 molecules with acceptable statistical performance ($R^2 = 0.82$, $Q^2_{loo} = 0.81$). According to the QSAR study, the activity has a strong link with various atoms such as aromatic carbons and ring Sulfur, acceptor atoms, sp^2 -hybridized oxygen, etc. Following that, a database of 26,467 food compounds was primarily used for QSAR-based virtual screening accompanied by the application of the Lipinski rule of five; the elimination of duplicates, salts, and metal derivatives resulted in a truncated dataset of 8,453 molecules. The molecular descriptor was calculated and a well-validated 6-parametric version of the QSAR model was used to predict the bioactivity of the 8,453 food compounds. Following this, the food compounds whose predicted activity (pKi) was observed above 7.0 M were further docked into the BACE1 receptor which gave rise to the Identification of 4-(3,4-Dihydroxyphenyl)-2-hydroxy-1H-phenalen-1-one (PubChem I.D: 4468; Food I.D: FDB017657) as a hit molecule (Binding Affinity = -8.9 kcal/mol, $pKi = 7.97$ nM, $Ki = 10.715$ M). Furthermore, molecular dynamics simulation for 150 ns and molecular mechanics generalized born and surface area (MMGBSA) study aided in identifying structural motifs involved in interactions with the BACE1 enzyme. Molecular docking and QSAR yielded complementary and congruent results. The validated analyses

can be used to improve a drug/lead candidate's inhibitory efficacy against the BACE1. Thus, our approach is expected to widen the field of study of repurposing nutraceuticals into neuroprotective as well as anti-cancer and anti-viral therapeutic interventions.

KEYWORDS

beta-site APP cleaving enzyme 1, BACE1, Alzheimer's disease, glioblastoma, QSAR, molecular docking, MD simulations, golden lotus banana

Introduction

Alzheimer's disease (AD) is a devastating mental illness, which leads to an irreversible, progressive brain disorder that slowly destroys memory skills and learning abilities (De Strooper and Karran, 2016). Though the disease progression and risk factors of AD are not completely understood, a large number of evidence suggest the formation of amyloid-beta ($A\beta$) is central to the pathophysiology of AD (Vassar et al., 1999). AD progression stages vary from mild to severe in middle-aged people to older persons detected with cognitive tests (Hall et al., 2019). During the preclinical stage of AD, patients seem to be symptom-free, but neurodegenerative changes occur in the brain. Abnormal accumulation of $A\beta$ -containing plaques and hyperphosphorylated tau throughout the brain causes healthy neurons to exhibit loss of synaptic connections, ion-channel dysfunctions, and severe deterioration in neuronal health. This ultimately leads to neuronal cell death and cognitive decline in elderly persons (Musi et al., 2018; Sebastián-Serrano et al., 2018). This progressive accumulation of $A\beta$ is caused by imbalances in the levels of $A\beta$ production, aggregation, and clearance (Murphy and LeVine, 2010; Jabir et al., 2021). Moreover, alterations in synaptic plasticity and integral neuronal circuitries severely hamper neurogenesis.

The Special Report examines MCI, including Alzheimer's, from the perspectives of consumers and primary care providers. AD affects 6.5 million Americans who are 65 years old or older. By 2060, this figure may increase to 13.8 million unless medical advances prevent, slow down, or cure AD. A total of 121,499 AD deaths were reported on official death certificates in 2019. According to one COVID-19 report, AD is the sixth most common cause of death in the United States in 2019 and will move up to the seventh most common cause in 2020 and 2021. Among Americans 65 and older, AD is the seventh most common cause of death. Between 2000 and 2019, mortality from HIV, heart disease, and stroke all dropped, whereas AD deaths rose by 145 percent. In 2021, 16 billion hours of care were supplied by over 11 million family members and other unpaid caregivers. These statistics demonstrate a decrease in carers and an increase in the amount of care that each caregiver who is still working gives. \$271.6 billion was

spent on unpaid dementia care in 2021. Family caregivers now have a higher risk of emotional discomfort as well as poorer mental and physical health due to COVID-19. Caregivers of dementia are also impacted by COVID-19. To safeguard their own health and the health of their families, several caregivers have resigned. However, there is a need for more dementia carers. Medicaid payments are more than 22 times greater than Medicare payments, and payments to beneficiaries 65 and older with AD or other dementias are approximately three times more than payments to beneficiaries without these illnesses. Hospice, long-term care, and healthcare for dementia patients will cost \$321 billion in 2022. A poll by the Alzheimer's Association indicates MCI obstacles. The study revealed low MCI knowledge, a lack of desire to seek medical attention for symptoms, and challenges with MCI diagnosis. According to survey findings, clinical trial participation should be increased together with MCI awareness and diagnosis, particularly in impoverished communities (Alzheimer's Association, 2012).

$A\beta$ is a neurotoxic aggregate produced by the consecutive proteolysis of β -amyloid precursor protein (APP) by two aspartyl proteases, beta-site APP cleaving enzyme, Beta-Secretase (BACE1), and finally by γ secretase (Sinha et al., 1999; Vassar et al., 1999; Yan et al., 1999; Cui et al., 2011; Zhang et al., 2011). BACE1 is a novel target, a type 1 transmembrane aspartic protease, related to the pepsin and retroviral aspartic protease families (Moussa-Pacha et al., 2020), having 501 amino acids and is predominantly expressed in the human brain (Zacchetti et al., 2007; Hassan et al., 2018).

As there is a strong association between $A\beta$ accumulation and AD, the primary therapeutic strategy for the treatment of AD targets lowering the concentration of $A\beta$. One such strategy that has come up in recent findings involves inhibiting the enzymes that generate $A\beta$ in the first place. BACE1 antisense oligonucleotide treatment to APP overexpressing cells is reportedly responsible for a decreased production of β -secretase cleaved APP fragments. Recent studies suggest that the levels of BACE1 protein and their activity were raised to approximately double in patients with AD where BACE1 might initiate or enhance AD pathogenesis. Several studies proved that BACE1 is a vital drug target and active site of the enzyme (Sinha et al., 1999; Vassar et al., 1999; Dingwall, 2001;

Fukumoto et al., 2002; Yang et al., 2003; Li and Südhof, 2004). It is covered by a flexible antiparallel β -hairpin, called a flap (Hussain et al., 2000). It is elucidated that the flap can control substrate access to the receptor site and set the substrate into the perfect orientation for the catalytic process (Lin et al., 2000). Hence, the inhibition of proteases such as BACE1 may represent modifying treatment for AD by controlling the production of A β (Tresadern et al., 2011; Arif et al., 2020). Therefore, BACE1 inhibitors may be used to treat AD.

Natural compounds have been investigated for many years for their ability to target a range of *trans*-acetyl peptidases in the search for new medications with fewer side effects due to their high biocompatibility. Apart from their benefits, the majority of natural compounds have a variety of disadvantages, including a high molecular weight, low stability, and in many cases, insufficient solubility. These difficulties can be resolved by using computer-assisted research to generate more effective and selective inhibitors. They are less harmful and more readily absorbed into the body than manufactured drugs. Computational techniques enable the modeling and screening of such compounds in a cost-effective and informative manner from a vast library of choices. The binding sites for various bioactive substances in BACE1 protein were predicted, and their interactions were explored utilizing the Food database and molecular docking. *Musella lasiocarpa*'s bioactive component, 4-(3,4-Dihydroxyphenyl)-2-hydroxy-1H-phenalen-1-one, is found to be the potential inhibitor of BACE1. *M. lasiocarpa*, also known as Chinese Dwarf Banana, is the only species in the genus *Musella*, which belongs to the family of Musaceae. Some novel compounds have also been isolated from this plant, which showed remarkable *in vitro* anticancer activities and some degree of antimicrobial activity (Bo et al., 2000; Dong et al., 2011).

These results demonstrate the potential of the *M. lasiocarpa* as a source of novel drugs, nutraceuticals, and functional foods. The result from this study will inspire the perspective of natural compounds and computer-aided drug discovery and, in the long run, significantly reduce the time required to research and build new lead compounds with specified biological activity and structural variety.

Methodology

QSAR modeling

Organization for economic corporation and development guidelines and a standard protocol recommended by different researchers, which involved sequential execution of (1) data collection and its curation, (2) structure generation and calculation of molecular descriptors, (3) objective feature selection (OFS), (4) splitting the dataset into training and external validation sets, and (5) subjective feature selection

involving building a regression model and validation of the developed model, has been followed to build a widely applicable QSAR model for the BACE1 inhibitory activity. This also ensures thorough validation and successful application of the model.

Selection of data set

The data set of BACE1 inhibitory activity used for building, training, and validating the QSAR model in the present work was downloaded from the ChEMBL database (accessed on 22nd December 2021)¹, which is a free and publicly accessible database. Initially, the data set comprised of 552 molecules (Davies et al., 2015). Then, as a part of data curation, entries with ambiguous K_i values, duplicates, salts, metal-based inhibitors, etc., were omitted. The final data set comprises structurally diverse 371 molecules with remarkable variation in structural scaffolds, which were tested experimentally for potency in terms of K_i (nM). A collection of 371 molecules with various structural configurations with mentioned enzyme inhibitory concentration (K_i) was decided on for the existing work (O'Boyle et al., 2011; Tosco et al., 2011; Fujita and Winkler, 2016). The K_i values ranging from 0.47 to 1,400 nM have been transformed to pK_i ($pK_i = -\log K_i$) earlier than real QSAR evaluation for ease of dealing with the data. Five least and five most active molecules are depicted in Figure 1 to illustrate the version in bio-activity with chemical features. The SMILES strings with mentioned K_i and pK_i values for all of the molecules are depicted inside Supplementary Table 1 with the Supplementary material.

Molecular structure drawing and optimization

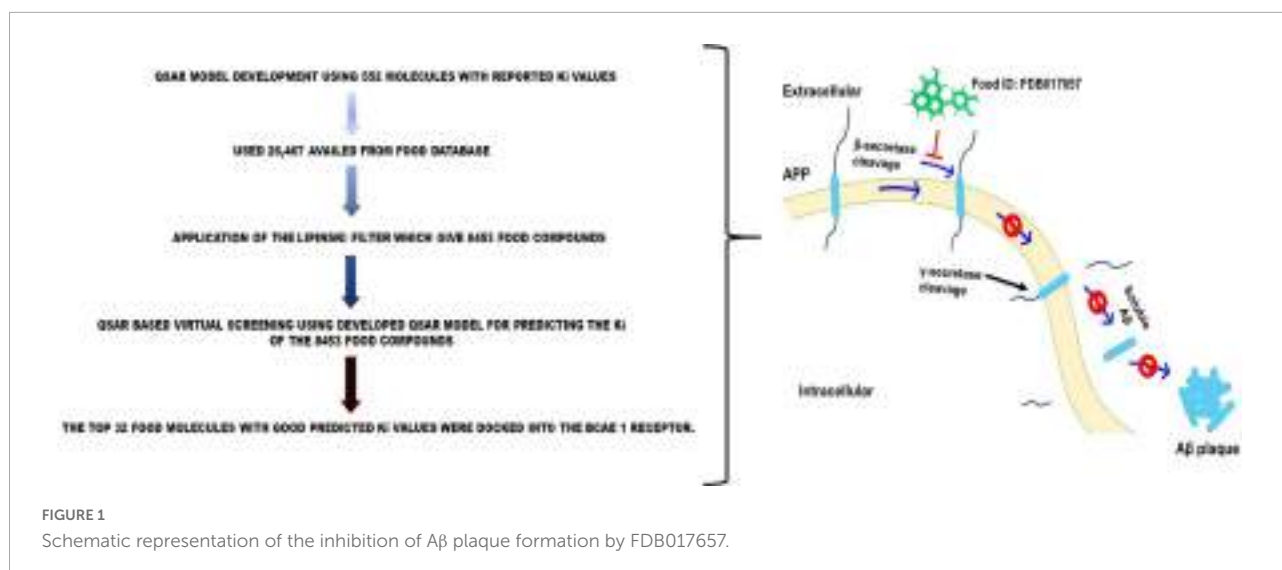
Drawing the 2D structures of all 371 and converting them to their corresponding 3D structures was done using ChemSketch 12 Freeware and Open Babel 2.4, (O'Boyle et al., 2011) both free and open-source software, respectively.² After that, the MMFF94 force field available in TINKER (default setting) and Open3DAlign is used for optimization and molecular alignment, respectively.

Molecular descriptor calculation and objective feature selection

PyDescriptor has over 30,000 molecular descriptors for each molecule (Masand and Rastija, 2017). Molecular descriptors with near constant values (>95%) and resonance ($|R| > 0.95$) were eliminated using OFS in QSARINS v2.2.4 (Gramatica et al., 2013, 2014). This work removed duplicate molecular descriptors and prevented collisions of multicollinearity and hypothetical variables in the Genetic Algorithm multiple linear regression (GA-MLR) model. Only 3,281 molecular descriptors

¹ [https://www.ebi.ac.uk/chembl/g/#search_results/all/query bace 1](https://www.ebi.ac.uk/chembl/g/#search_results/all/query%20bace%201)

² www.acdlabs.com



were found in the reduced pool of molecular descriptors generated after OFS processing.

Subjective feature selection, QSAR model development, and validation

A reduced set of molecular descriptors consisting of 1D to 3D descriptors, molecular properties and charge descriptors, etc., covers a fairly complete descriptor space. A statistically robust genetic algorithm (GA) based on multiple linear regression (MLR) was deployed to run the QSAR model using the subjective feature selection (SFS) task in QSARINS v2.2.4. Models derived according to the organization for economic corporation and development (OECD) principles have been subjected to rigorous internal and external statistical validation, scrambling, and range analysis. The QSAR deployment process goes through the following steps:

To generate a QSAR model from the split data set, a random split operation is performed in QSARINS v2.2.4 to split the given data set into 80% training set (297 molecules) and 20% predictions (74 molecules set). A total of 297 molecules in the training set were used to develop the QSAR model, and external validation was performed on 74 molecules in the prediction set (Gramatica et al., 2013).

The GA-MLR-based QSAR model was built with the default setting QSARINS v2.2.4. Subjective feature selection was performed by setting Q^2_{LOO} as a fitness feature. The Q^2_{LOO} score increased significantly to six variables but then increased slightly. Therefore, to avoid overfitting the model, SFS was limited to a set of seven descriptors. This helped to get a simple and informative QSAR model (see **Supplementary Table 2** in **Supplementary information**) for additional information on the six selected molecular descriptors present in the QSAR).

A good QSAR model which has been properly validated using various methods such as cross-validation, external

validation, Y-randomization, and applicability domain (Williams plot) is useful for future utilization in virtual screening, molecular optimization, and decision making, etc. The following statistical parameters and their recommended threshold values are routinely used to validate a model (Bellacasa et al., 2013; Roskoski, 2013; Fujita and Winkler, 2016; Gramatica, 2020): $R^2_{tr} \geq 0.6$, $Q^2_{loo} \geq 0.5$, $Q^2_{LMO} \geq 0.6$, $R^2 > Q^2$, $R^2_{ex} \geq 0.6$, $RMSE_{tr} < RMSE_{cv}$, $\Delta K \geq 0.05$, $CCC \geq 0.80$, $Q^2_{-Fn} \geq 0.60$, $r^2_m \geq 0.5$, $(1-r^2/r_o^2) < 0.1$, $0.9 \leq k \leq 1.1$ or $(1-r^2/r_o^2) < 0.1$, $0.9 \leq k' \leq 1.1$, $(r_o^2 - r_o'^2) < 0.3$, $RMSE_{ex}$, MAE_{ex} , R^2_{ex} , Q^2_{F1} , Q^2_{F2} , Q^2_{F3} , and low R^2_{Yscr} , $RMSE$ and MAE . The formulae for calculating these statistical parameters are available in **Supplementary Table 3** in the **Supplementary material**. In addition, the Williams plot was plotted to evaluate the applicability domain of the QSAR model.

QSAR-based virtual screening

A database of 26,467 food compounds was downloaded from FoodDB (accessed December 28, 2021)³, primarily for QSAR-based virtual screening (VS) accompanied through the utility of Lipinski rule of five; the elimination of duplicates, salts, and metal derivatives resulted in a truncated dataset of 8,453 molecules. Therefore, the 8,453-food molecule was mainly used for VS-based QSAR. Prior to the calculation of the molecular descriptor, the 3D molecular system was organized within the same Method as a modeling set. Next, the molecular descriptor was calculated and a well-validated 6-parametric version of the QSAR model was used to predict the bioactivity of the new compound (Gramatica, 2013; Neves et al., 2018; Zaki et al., 2021; Jawarkar et al., 2022; Mukerjee et al., 2022). (The calculated

³ <http://fooddb.ca/>

molecular descriptors along with the predicted pKi and Ki values with smiles strings are given in **Supplementary Table 4** in the **Supplementary material**.

Virtual screening of natural compounds using molecular docking

The structure-based virtual screening of the compounds was performed using AutoDock vina Version 1.1.2 (Trott and Olson, 2010). Binding sites of BACE1 for screening were predicted using DoGSiteScorer (Volkamer et al., 2012) and information about the binding site of the native ligand. The size of the grid box was set to be $96 \times 52 \times 56$ Å for BACE1, centered around the identified binding site. The compounds with the best binding affinity (kcal/mol) corroborating ligand-based screening in the QSAR analysis were selected for the studies.

Preparation of protein and ligand molecule

The target of interest BACE1 (PDB I.D: 2ZHV) crystallographic structure was discovered in the Protein Data Bank's structural database, and a molecular editor with an open-source license was used to import the structure (Discovery studio visualizer 4.0)⁴ (Figure 2). To get the structure optimized, the UCSF Chimera utilized the steepest descent to find 1,000 steps, followed by the conjugate gradient of energy minimization approach to optimize the structure of 4-(3,4-dihydroxyphenyl)-2-hydroxy-1H-phenalen-1-one (PubChem I.D: 4468; Food I.D: FDB017657) was acquired from Food Database after screening and QSAR. This dataset was imported into the DS visualizer and saved as PDB files.

The FOOD database was used to obtain the library of ligands by screening the metabolites of *M. lasiocarpa* (Bonvino et al., 2018). The smiles notation and the three-dimensional structures of the selected ligands were downloaded in SDF format from the PubChem database (Kim et al., 2019), and further ligand structure files were converted to PDB format using Open Babel software (O'Boyle et al., 2011). The energy minimization of the ligands was performed in UCSF Chimera software (Pettersen et al., 2004) using the Amber ff 14 sb force field. The receptor used in the study is receptor tyrosine kinase; RCSB Protein Data Bank (PDB) (Berman et al., 2000) was used to download receptors BACE1 with PDB I.D: 2ZHV, Resolution: 1.85 Å. The protein structure was prepared by removing ligand, water molecules, and metal ions. Polar hydrogens were added, and non-polar hydrogens were merged. Finally, Kollman charges were added to the protein molecule before converting to PDBQT format by AutoDock Tools (v.1.5.6) of the MGL software package (Forli et al., 2016).

⁴ <https://www.rcsb.org/structure/2ZHV>

Molecular docking for validation of docking score

The best hits from the QSAR modeling and virtual screening were re-docked against BACE1 (PDB I.D: 2ZHV). Protein and ligand preparations were done using AutoDock Tools (v.1.5.6) (Forli et al., 2016). Gasteiger charges were added to the ligand molecules prior converting to the PDBQT format. Online server DoGSiteScorer and the information about the binding site residues of native ligand were used to construct the grid box. The grid box of dimensions $96 \times 52 \times 56$ Å for BACE1 with 0.375 Å grid spacing was constructed using auto grid 4.2. Semi-flexible docking was done keeping the receptor molecule rigid and ligands flexible. Molecular docking was done via AutoDock 4.2 (Morris et al., 2009) using the Lamarckian Genetic Algorithm (LGA) scoring function with the number of GA runs = 100, population size = 550, and a maximum number of evaluations = 24,000,000.

Molecular dynamics simulation (MD-simulation) and free energy landscape (FEL) analysis

The MD simulations studies were carried out in triplicate on dock complexes for BACE1 (PDB I.D: 2ZHV) with FDB017657 using the Desmond 2020.1 from Schrödinger, LLC. The triplicate samplings were made using the same parameters for each MD run to obtain the reproducibility of the results. The OPLS-2005 force field (Bowers et al., 2006; Chow et al., 2008; Shivakumar et al., 2010) and explicit solvent model with the stocktickerSPC water molecules were used in this system (Jorgensen et al., 1983). Na⁺ ions were added to neutralize the charge of 0.15 M, and NaCl solutions were added to the system to simulate the physiological environment. Initially, the system was equilibrated using an NVT ensemble for 150 ns to retrain over the protein- FDB017657 complex. Following the previous step, a short run of equilibration and minimization was carried out using an NPT ensemble for 12 ns. The NPT ensemble was set up using the Nose-Hoover chain coupling scheme (Martyna et al., 1994) with the temperature at 27°C, the relaxation time of 1.0 ps, and pressure of 1 bar maintained in all the simulations. A time step of 2fs was used. The Martyna-Tuckerman-Klein chain coupling scheme (Martyna et al., 1992) barostat method was used for pressure control with a relaxation time of 2 ps. The particle mesh Ewald method (Toukmaji and Board, 1996) was used to calculate long-range electrostatic interactions, and the Radius for the Coulomb interactions was fixed at 9Å. RESPA integrator was used for a time step of 2 fs for each trajectory to calculate the bonded forces. The root means square deviation (RMSD), radius of gyration (Rg), root mean square fluctuation (RMSF),

and the number of hydrogen (H-bonds) were calculated to monitor the stability of the MD simulations. The free energy landscape of protein folding on the FDB017657 bound complex was measured using `geo_measures v 0.8` (Kagami et al., 2020). `Geo_measures` include a powerful library of `g_sham` and form the MD trajectory against RMSD and Radius of gyration (Rg) energy profile of folding recorded in a 3D plot using `matplotlib` python package.

Molecular mechanics generalized born and surface area (MMGBSA) calculations

During MD simulations of BACE1 complexed with FDB017657, the binding free energy (G_{bind}) of docked complexes was calculated using the premier molecular mechanics generalized Born surface area (MM-GBSA) module (Schrodinger suite, LLC, New York, NY, United States, 2017-4). The binding free energy was calculated using the OPLS 2005 force field, VSGB solvent model, and rotamer search methods (Piao et al., 2019). After the MD run, 10 ns intervals were used to choose the MD trajectories frames. The total free energy binding was calculated using equation 1:

$$\Delta G_{\text{bind}} = G_{\text{complex}} - (G_{\text{protein}} + G_{\text{ligand}}) \quad (1)$$

Where, ΔG_{bind} = binding free energy, G_{complex} = free energy of the complex, G_{protein} = free energy of the target protein, and G_{ligand} = free energy of the ligand. The MMGBSA outcome trajectories were analyzed further for post-dynamics structure modifications.

Dynamic cross-correlation and principal component (PCA) analysis

During a 150 ns MD simulation, a dynamic cross-correlation matrix (DCCM) was constructed across all C-atoms for all complexes to examine domain correlations. During a 150 ns simulation of the BACE1 (PDB I.D- 2ZHV) complexed with FDB017657, PCA analysis was used to recover the global movements of the trajectories. To calculate the PCA, a covariance matrix was created as stated. For conformational analysis of the FDB017657 in bound complex, 20 alternative conformational modes of the main component as movements of trajectories were calculated, and a comparison of the first highest mode (PC2) with PC10 was investigated. `Geo measures v 0.8` was used to calculate the free energy landscape of protein folding on an FDB017657-bound complex (Kagami et al., 2020). The MD trajectory versus PC2 and PC10 energy folding profiles were recorded in a 3D plot using the `matplotlib` python package using `Geo measures`, which includes a comprehensive library of `g_sham`.

Results

The present QSAR analysis is performed using a moderate-size data set comprising structurally diverse compounds with experimentally measured K_i in the range from 0.47 to 1,400 nM. Thus, it covers a sufficiently broad chemical and data range. This helped to derive a properly validated (Martin and Muchmore, 2012; Gramatica et al., 2013; Masand et al., 2015; Fujita and Winkler, 2016) genetic algorithm unified with multilinear regression (GA-MLR) model to collect or extend thorough information about the pharmacophoric features that control the desired bio-activity (Descriptive QSAR) and having adequate external predictive capability (Predictive QSAR). The six variable-based GA-MLR QSAR model along with selected internal and external validation parameters (see **Supplementary material** for additional parameters) is as follows:

QSAR Model (Divided Set: Training Set-80% and Prediction Set-20%):

$$pKi = 4.415 (\pm 0.236) + 0.046 (\pm 0.017) * \text{com_lipohyd_5A} + 0.12 (\pm 0.027) * \text{faccC3B} + 0.353 (\pm 0.105) * \text{aroC_sumpc} + 0.273 (\pm 0.034) * \text{N_acc_5B} + 0.109 (\pm 0.02) * \text{aroC_ringS_6B} + 0.269 (\pm 0.04) * \text{fsp3OringC8B} +$$

Statistical parameters related to fitting, double cross-validation, and Y-scrambling for the *de novo* QSAR model with thresholds for some parameters (bottom of table) are shown in **Table 1**.

Thresholds for some important statistical parameters: $R^2 \geq 0.6$, $Q^2_{\text{LOO}} \geq 0.5$, $Q^2_{\text{LMO}} \geq 0.6$, $R^2 > Q^2$, $R^2_{\text{ex}} \geq 0.6$, $\text{RMSE}_{\text{tr}} < \text{RMSE}_{\text{cv}}$, $\Delta K \geq 0.05$, $\text{CCC} \geq 0.80$, $Q^2_{\text{Fn}} \geq 0.60$, $r^2_{\text{m}} \geq 0.6$, $0.9 \leq k \leq 1.1$ and $0.9 \leq k' \leq 1.1$ at $\text{RMSE} \approx 2$, $\text{MAE} \approx 0.4 R^2_{\text{adj}}$ of fitting parameter R^2 , etc.) The threshold is significantly exceeded, confirming the adequacy of the number of molecular descriptors in the model and the statistical acceptability of the QSAR model. The values of Q^2_{LOO} , Q^2_{LMO} , etc., (internal validation parameters) confirm the statistical robustness of the QSAR model. The high values of the external test parameters R^2_{ex} , Q^2_{F1} , Q^2_{F2} , Q^2_{F3} , etc., emphasize the external predictability of the QSAR model. The coverage area of the model (Applicability Domain) is determined from Williams plots for the QSAR model (see **Figure 3**). Almost all statistical parameters reached values well above the accepted threshold, and the minimal correlation between the selected molecular descriptors precluded the possibility of random development of the QSAR model. This data confirms the statistical reliability and high external predictability of the developed QSAR models.

A nicely proven correlation among salient capabilities of the molecules represented through molecular descriptors, and their bioactivity expands statistics approximately mechanistic elements of molecules, specificity, and quantity (presence or even absence) of various structural developments for preferred bioactivity. Although, within the QSAR analysis, we've as compared the K_i values of various molecules in correlation and as an impact of a specific molecular descriptor, a similar

TABLE 1 Statistical parameters for the developed QSAR model.

Statistical parameters	Model
Fitting	
R ²	0.8120
R ² _{adj}	0.8168
R ² -R ² _{adj}	0.0037
LOF	0.1807
K _{xx}	0.2382
Delta K	0.0928
RMSE _{tr}	0.4079
MAE _{tr}	0.3276
RSS _{tr}	49.5888
CCC _{tr}	0.9014
s	0.4128
F	221.7565
Internal validation	
Q ² _{LOO}	0.8120
R ² -Q ² _{LOO}	0.0085
RMSE _{cv}	0.4175
MAE _{cv}	0.3352
PRESS _{cv}	51.9415
CCC _{cv}	0.8968
Q ² _{LMO}	0.8123
R ² _{Yscr}	0.0197
RMSE AV _{Yscr}	0.9534
Q ² _{Yscr}	-0.0283
External validation	
RMSE _{ext}	0.4604
MAE _{ext}	0.3803
PRESS _{ext}	15.4724
R ² _{ext}	0.7829
Q ² _{F1}	0.7832
Q ² _{F2}	0.7819
Q ² _{F3}	0.7714
CCC _{ext}	0.8742
r ² _{m aver.}	0.6607
r ² _{m delta}	0.1977
K'	0.9967
K	0.9999
Clos'	0.0496
Clos	0.0000

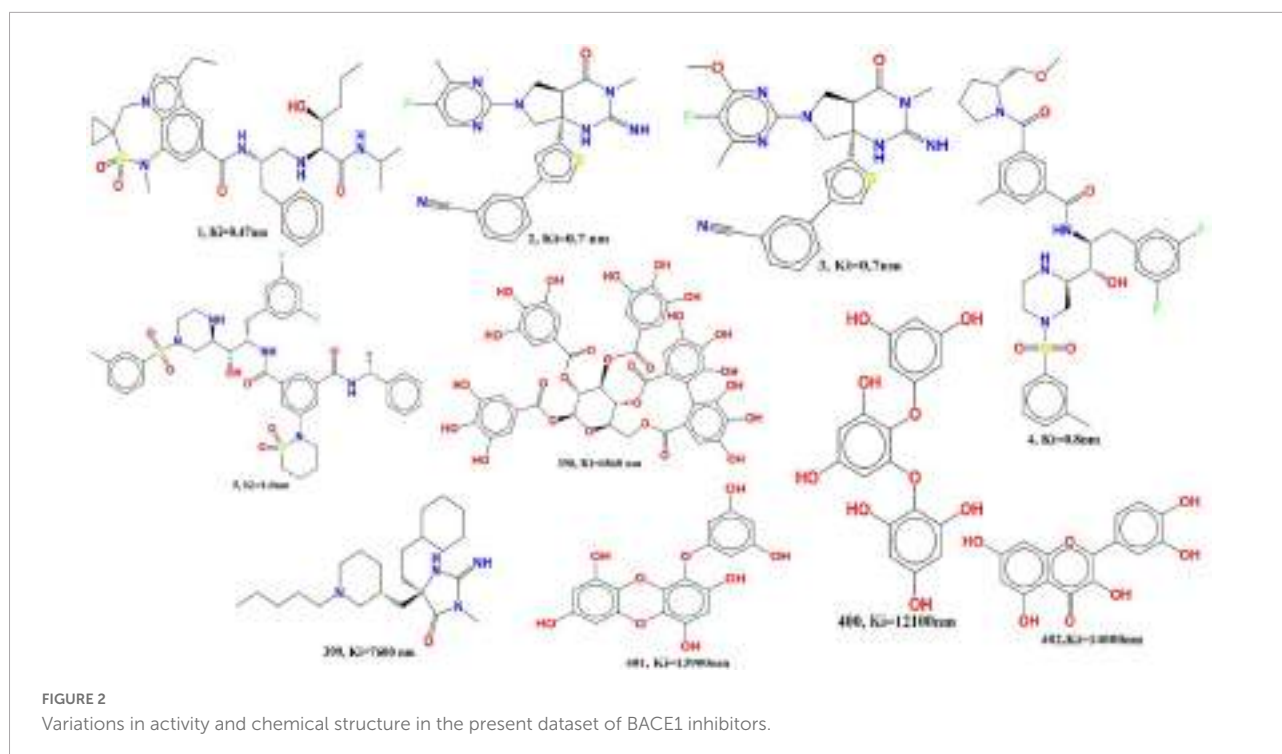
or contrary impact of different molecular descriptors or unknown descriptors has a dominant impact in figuring out the general Ki value of a molecule cannot be neglected. In different words, a single molecular descriptor is incapable of absolutely explaining the experimental Ki value for this sort of numerous set of molecules. That is, the successful usage of the advanced QSAR model is based on the concomitant usage of molecular descriptors.

Mechanistic interpretation

com_lipohyd_5A (occurrence of the lipo-hydrophobic atoms within 5A from the center of mass of the molecule) Since this descriptor received a positive coefficient in the developed QSAR model, increasing the value of the molecular descriptor com_lipohyd_5A was observed to increase BACE1 inhibitory activity. This can be observed by comparing the compound 368 (pKi = 7.75, com_lipohyd_5A = 14) with 48 (pKi = 6.37, com_lipohyd_5A = 13) for which an increase in the value of the molecular descriptor com_lipohyd_5A from 13 for the molecule 48 to the 14 will give rise to increase in pKi value by about 1 unit (about ten-fold increase in BACE1 inhibitory activity). This observation is further reinforced by the following pair of molecules; 243 (pKi = 7.35, com_lipohyd_5A = 8) and 134 (pKi = 6.93, com_lipohyd_5A = 7), 211 (pKi = 7.92, com_lipohyd_5A = 12) and 207 (pKi = 7.58, com_lipohyd_5A = 11), 189 (pKi = 8.4, com_lipohyd_5A = 15) and 255 (pKi = 8.19, com_lipohyd_5A = 10), 262 (pKi = 5.52, com_lipohyd_5A = 15) and 115 (pKi = 4.97, com_lipohyd_5A = 13), etc. Shifting the molecular descriptor com_lipohyd_5A with com_lipohyd_6A statistically improves the performance (R: 0.84) of the developed QSAR model, while replacing com_lipohyd_5A with com_lipohyd_4A statistically drops the performance of the QSAR model (R: 0.81). This observation underscores the importance of the molecular descriptor com_lipohyd_5A. In addition, the optimum distance between lipo-hydrophobic atoms from the center of mass of the molecule should be maintained at 6 Å to show better inhibitory activity against BACE1.

faccC3B (occurrence of the carbon atom exactly at three bonds from the acceptor atom). If the same descriptor exists in two or four bonding at the same time, it will be removed during the calculation of faccC3B. This descriptor received a positive sign in the developed QSAR model. Therefore, further increases in faccC3B value increase the inhibitory efficacy of BACE1 inhibitors. Comparison of the compound 331 (pKi = 7.6, faccC3B = 7) with 142 (pKi = 6.87, faccC3B = 6) illustrate the influence of the molecular descriptor faccC3B. If the value of the molecular descriptor faccC3B is enhanced from 6 for the molecule 142 to 7 will upsurge the pki value by about 1 unit (about 10-fold amplification in the BACE1 inhibitory activity). Following pair from **Figures 3C,D** of molecules support this observation; 295 (pKi = 8.72, faccC3B = 6) and 255 (pKi = 8.19, faccC3B = 4), 262 (pKi = 5.52, faccC3B = 3) and 115 (pKi = 4.97, faccC3B = 2), 258 (pKi = 8.7, faccC3B = 9) and 274 (pKi = 7.58, faccC3B = 8), 128 (pKi = 6.28, faccC3B = 3) and 60 (pKi = 6.1, faccC3B = 2), 169 (pKi = 8.7, faccC3B = 9) and 286 (pKi = 7.55, faccC3B = 7), etc.

This shows that simple carbon atoms along the 3-bond acceptor atom are essential for inhibitory efficacy, but the molecular descriptors faccC3B and fdonC3B (frequency of carbon atoms exactly 3 bonds away from the donor atom). Or



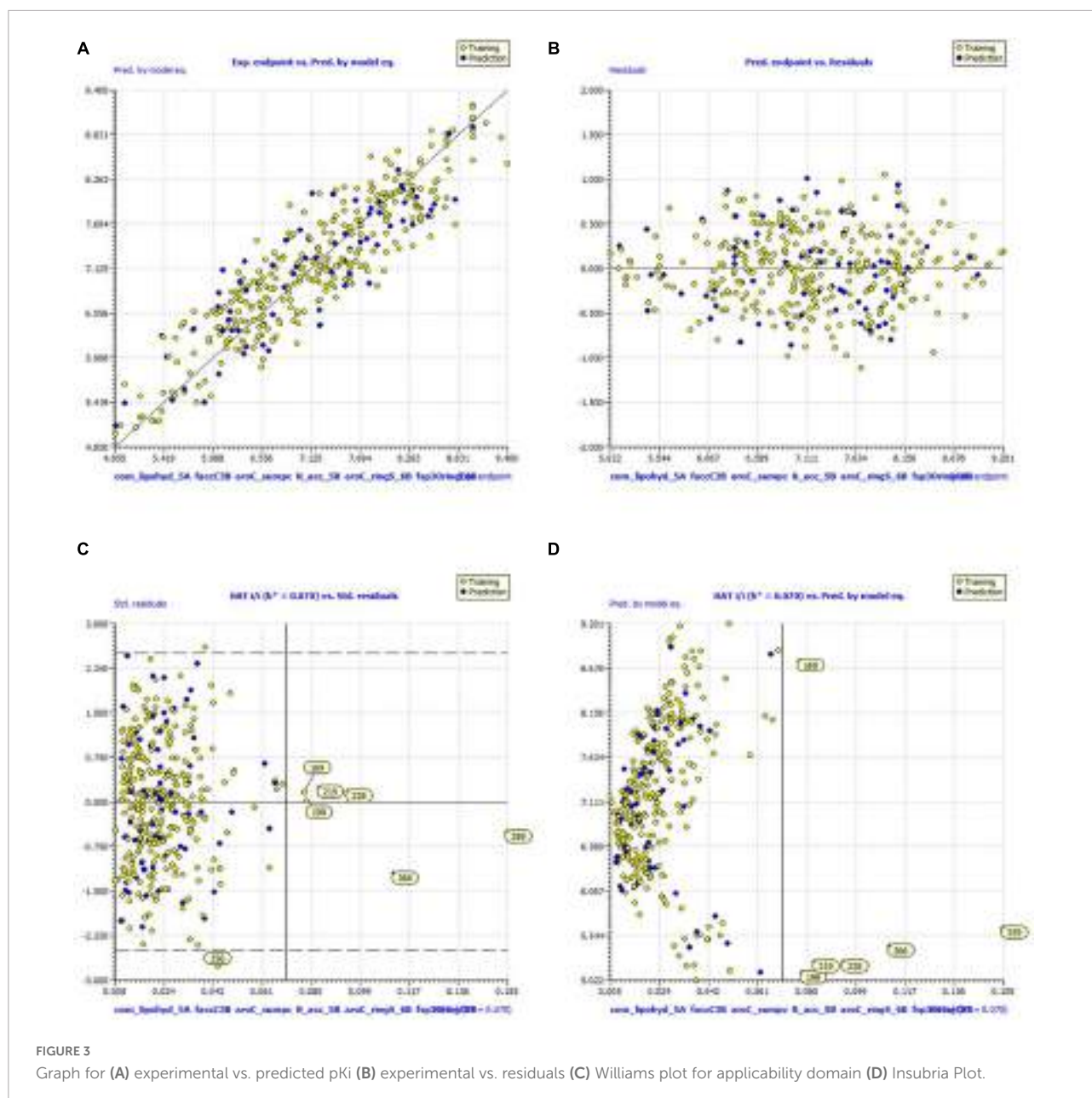
fdonlipo5B (an acceptor atom with exactly 5 bonds from the donor atom) greatly improves the statistical detection power of the developed QSAR model (fdonC3B, R: 0.89) (fdonlipo5B, R: 0.88). This observation shows that the inhibitory effect can be enhanced by replacing the acceptor atom with a donor atom with 3 bonds of the same topology distance, or by moving the carbon atom to any lipophilic atom along the donor by a distance of five bonds. This observation shows that the carbon atom is more important along the donor atom located at the optimum distance of the three bonds. Therefore, since most H-bond donors or acceptors are nitrogen or oxygen, the presence of donor atoms close to the carbon atom may help to enhance the interaction with the polar residues of the receptor (BACE1). In addition, the descriptor also shows the important role that carbon atoms definitely play in lipophilicity.

aroC_sumpc (sum of partial charges on aromatic carbon atoms in the range of + 0.2 and -0.2). A positive coefficient for aroC_sumpc indicates that the higher the value of this descriptor, the better the activity profile. The presence of a large number of carbon atoms makes the molecule lipophilic, while the presence of positively or negatively charged carbon atoms causes various types of hydrophobic interactions with the receptor (BACE1), such as pi-alkyl, pi-cation, and pi-pi stacking. This observation underscores the importance of the molecular descriptor aroC_sumpc. This observation is supported by comparing the following pairs of molecules: 316 (pKi = 8.5, aroC_sumpc = 0.224000013) and 23 (pKi = 7.4, aroC_sumpc = -0.088000003),

159 (pKi = 9, aroC_sumpc = 0.013999997) and 298 (pKi = 8.5, aroC_sumpc = -0.197999999), 369 (pKi = 9, aroC_sumpc = 0.356000006) and 109 (pKi = 8.14, aroC_sumpc = -0.322999999), etc. This observation pointed out that negatively charged carbons are not favorable for BACE1 inhibitory activity, hence, ring carbon atom with neutral or positively charged possesses a better BACE1 inhibitory activity.

Subsequently, shifting the molecular descriptor aroC_sumpc to the molecular descriptor ringC_sumpc (sum of partial charges of the ring carbon atom) for future drug optimization increases the statistical power (R: 0.88) of the developed QSAR model. Furthermore, assuming that the molecular descriptor aroC_sumpc is replaced with the molecular descriptor aroCminus_sumpc (sum of partial charges of negatively charged aromatic carbon atoms), this reduces the statistical power (R: 0.75) of the QSAR model. Therefore, incorporating the ring carbon atom (ringC_sumpc) is a better choice for future optimization of hits to the lead molecule for better inhibitory activity of BACE1.

N_acc_5B (occurrence of acceptor atom within 5 bonds from the nitrogen atom). The molecular descriptor N_acc_5B indicates that the acceptor atom is within the four bonds of the nitrogen atom. This molecular descriptor has a positive coefficient in the developed QSAR model, so increasing the number of such combinations can enhance BACE1 inhibition. The effect of N_acc_5B can be explained by comparing the molecule 258 (pKi = 8.7, N_acc_5B = 5) and 57 (pKi = 6.4, N_acc_5B = 3). For molecule 57, if the value of the molecule descriptor increase from 3 to 5 will further amplify the pKi value



by about 2 units, therefore, enhancing the BACE1 inhibitory activity by about 20 folds. These descriptors pointed out the importance of the nitrogen atom in BACE1 inhibitory activity. Moreover, additional molecular pair also illustrate the effect of N_acc_5B on BACE1 inhibitory activity include; 250 (pKi = 7.64, N_acc_5B = 5) and 55 (pKi = 6.52, N_acc_5B = 3), 189 (pKi = 8.4, N_acc_5B = 5) and 252 (pKi = 7.57, N_acc_5B = 4), 297 (pKi = 8.0, N_acc_5B = 8) and 102 (pKi = 7.17, N_acc_5B = 7), 312 (pKi = 8.5, N_acc_5B = 4) and 213 (pKi = 6.2, N_acc_5B = 3), etc.

The shift of the molecular descriptor N_acc_5B by the molecular descriptor N_acc_2B (the occurrence of acceptor

atoms within two bonds from the nitrogen atom) strongly affects the statistical performance (R: 0.86) of the developed QSAR model (N_don_5B). This observation shows that the subsequent descriptor is N_acc_2B, which is useful as a better alternative to future drug optimization and to enhance the BACE1 inhibitor activity. Again, it is suggested that the optimum distance between the acceptor atom and the nitrogen atom may be two bonds.

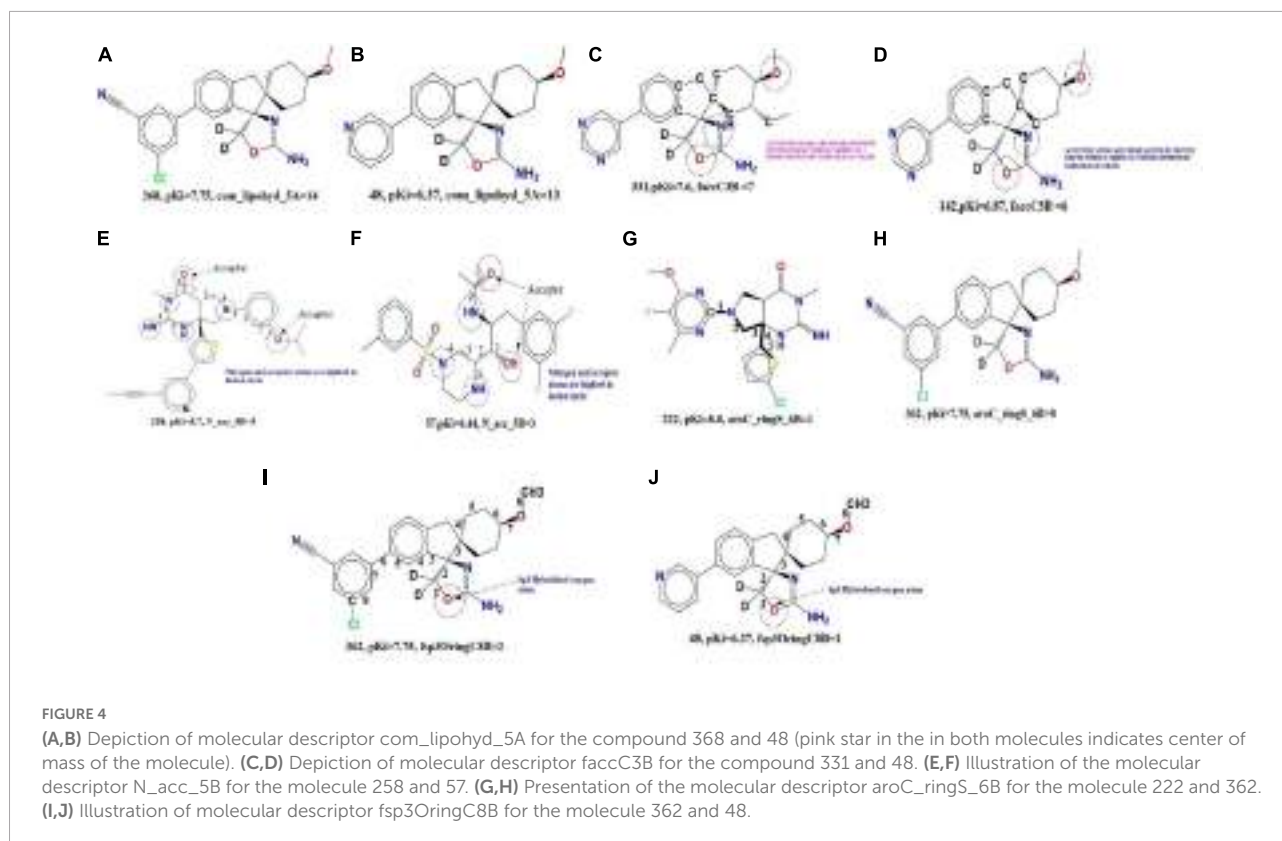
On the other hand, replacing the molecular descriptor N_acc_5B with the molecular descriptor N_don_5B (the appearance of a donor atom within 5 bonds from the nitrogen atom) significantly increases the statistical power (R: 0.85)

of the developed QSAR model. This remark reveals two strategies for future optimization, consisting of reducing the topological distance from five bonds to two bonds according to the N_{acc_2B} descriptor, and replacement of acceptors with donors by maintaining the optimal distance of five bonds. This observation underscores the importance of nitrogen atoms and acceptor/donor properties for better inhibitory activity of BACE1.

$aroC_ringS_6B$ (occurrence of the ring sulfur atoms within six bonds from the aromatic carbon atoms). The positive coefficient of the developed QSAR model descriptor justifies the increase in the value of the $aroC_ringS_6B$ descriptor and further enhances the BACE1 inhibitory activity. Shifting the molecular descriptor $aroC_ringS_6B$ by $fringSringC3B$ (frequency of ring carbon atoms in exactly three bonds from the ring sulfur atom) significantly increases the statistical power ($R:0.91$) of the QSAR model. By comparing the molecule 222 ($pK_i = 8.0$, $aroC_ringS_6B = 1$) and 362 ($pK_i = 7.75$, $aroC_ringS_6B = 0$), the influence of $aroC_ringS_6B$ as illustrated in the **Figures 4A–J**. Moreover, amplification in the value of the molecular descriptor $aroC_ringS_6B$ from 0 to 1 for the molecule will uplift the pK_i value by about 0.75 units (about 7 fold increase in the BACE1 inhibitory activity). Another pair of molecules also illustrate the effect of $aroC_ringS_6B$ on BACE1 inhibitory activity include; 256 ($pK_i = 8.5$, $aroC_ringS_6B = 2$) and 2 ($pK_i = 7.5$,

$aroC_ringS_6B = 1$), 295 ($pK_i = 8.7$, $aroC_ringS_6B = 2$) and 324 ($pK_i = 7.3$, $aroC_ringS_6B = 0$), 189 ($pK_i = 8.4$, $aroC_ringS_6B = 2$) and 254 ($pK_i = 8.1$, $aroC_ringS_6B = 1$), 144 ($pK_i = 7.8$, $aroC_ringS_6B = 2$) and 140 ($pK_i = 7.0$, $aroC_ringS_6B = 0$), etc.

Furthermore, replacing the molecular descriptor $aroC_ringS_6B$ with $fringSringC5B$ (frequency of ring carbon atoms in exactly five bonds from the ring sulfur atom) shows an increase in the statistical power ($r: 0.94$) of the developed QSAR model. On the other hand, replacing the molecular descriptor $aroC_ringS_6B$ with $fringSC6B$ (frequency of carbon atoms appearing in exactly 6 bonds from the ring sulfur atom) significantly reduces the statistical power ($R: 0.80$) of the QSAR model. This observation emphasizes the importance of the ring carbon atom and the ring sulfur atom. Furthermore, it can be concluded that reducing the topology bond distance from six to five bonds strongly affects the inhibitory activity of BACE1. Therefore, in future drug designs, it is recommended to keep the optimal distance between the ring carbon and the ring sprue atom at five bonds in order to increase the inhibitory activity of base 1. Aromatic and heterocyclic rings are very powerful motifs in drug discovery with target proteins such as the classical arene-arene interactions (π stacking), π -sulfur interactions, and arene-cation interaction. It offers many unique and powerful interactions such as bonds (end-face interactions) and recently identified π cations.



fsp3OringC8B (frequency of occurrence of the ring carbon atom exactly at eight bonds from the sp³ oxygen atom) If the same ring carbon atom occurs simultaneously in seven or nine bonds, it will be bypassed in the fsp3OringC8B calculation. This molecular descriptor has a positive coefficient in the developed QSAR model, so increasing its value can improve the Ki value.

The effect of fsp3OringC8B can be assessed by comparing the molecule 362 9 (pKi = 7.75, fsp3OringC8B = 2) with 48 (pKi = 6.37, fsp3OringC8B = 1). Increasing the value of fsp3OringC8B for molecule 1 to 2 for molecule 48 will upsurge the pKi value by 1.38unit (about 13-fold amplification in the BACE1 inhibitory potential). Following pair of molecules also explain the effect of fsp3OringC8B on BACE1 inhibitory activity; 211 (pKi = 7.9, fsp3OringC8B = 3) and 207 (pKi = 7.5, fsp3OringC8B = 1), 299 (pKi = 8.7, fsp3OringC8B = 2) and 255 (pKi = 8.1, fsp3OringC8B = 1), 352 (pKi = 8.4, fsp3OringC8B = 2) and 30 (pKi = 6.0, fsp3OringC8B = 0), 110 (pKi = 7.5, fsp3OringC8B = 4) and 66 (pKi = 6.5, fsp3OringC8B = 0), etc. Replacing the fsp3OringC8B molecular descriptor with the fsp3OringC9B descriptor significantly reduces the statistical power (R: 0.78) of the developed QSAR model.

This observation underscores the importance of the molecular descriptor fsp3OringC8B. Therefore, future drug designs need to maintain the optimal distance between sp³ hybridized oxygen and ring carbon atoms at eight bonds in order to achieve better BACE1 inhibitory activity. Most oxygen atoms act as either acceptors or ether linkage, which can confer lipophilicity on the molecule. This may help to enhance both polar and hydrophobic interactions with the BACE1 receptor.

Molecular docking for validation of docking score

In molecular docking analysis of the BACE1 with FDB017657 in Autodock output, a dock complex displayed the best conformation. Receptors and ligands were saved in the.pdbqt format for subsequent usage using the MGL 1.5.6 suite. Vina was launched from a command prompt using the command line. In the setup, the default grid point spacing was 0.525 and the exhaustiveness was set to 8. The output files were in.pdbqt format, and they were analyzed using PyMol and the Discovery studio visualizer 2021. The ligand-binding was validated and optimized using the co-crystal ligand. Both the receptor and ligands were made by combining 48 polar hydrogen bonds and detecting 1 rotatable bond and adding Kollman and Gasteiger charges. Finally, both receptor and ligand molecules were stored in the.pdbqt format. With the values $X = -1.655$, $Y = 57.005$, and $Z = 133.83$, a grid box was produced with a spacing of 0.375. Docking experiments of the protein-ligand complex were carried out using Genetic Algorithm (GA) parameters were set with 100, population

TABLE 2 Screening of phytochemicals based on their best binding energy.

Protein-ligand	Binding affinity (kcal/mol)
2zhv_8265	-4.8
2zhv_8263	-4.7
2zhv_8262	-7.3
2zhv_8079	-4.7
2zhv_7888	-5
2zhv_7701	-6.4
2zhv_7594	-6.3
2zhv_7334	-2.9
2zhv_7032	-4.9
2zhv_6574	-5.1
2zhv_5179	-5.6
2zhv_4844	-7.6
2zhv_4817	-5.6
2zhv_4693	-4.9
2zhv_4688	-5.8
2zhv_4605	-5.5
2zhv_4468	-8.9
2zhv_4340	-7.6
2zhv_4009	-5.7
2zhv_3981	-4.6
2zhv_3805	-4.7
2zhv_3207	-5.7
2zhv_2839	-4.8
2zhv_1976	-5.4
2zhv_1749	-5.7
2zhv_734	-7.3
2zhv_686	-6.7
2zhv_673	-8.3
2zhv_603	-6.5
2zhv_442	-4.5
2zhv_41	-3.8
2zhv_4	-5.6

size was made 300 with a maximum number of evaluates was set to low at 2,500,000 and maximum generations of 27,000. Further docking experiments of the protein-ligand complex were carried out using the Lamarckian Genetic Algorithm (LGA) to obtain the lowest free energy of binding (G). The 2ZHV-FDB017657 complex showed free energy of binding (ΔG) -8.9 kcal/mol, inhibitory concentration (Ki) 990.57 μ M, ligand efficiency -1.26, total internal energy -1.45 kJ/mol, and torsional energy 0.3 kJ/mol. The docking scores are mentioned in **Table 2**.

The principal residues making the binding pocket around 4-(3,4-dihydroxyphenyl)-2-hydroxy-1H-phenalen-1-one (Food I.D: FDB017657) are comprised of THR72, THR231, GLY230, ASP228, GLY34, ILE118, SER35, PHE108, ASP106, LYS107, GLY74, and GLN73 by Van der Waals interaction forces; ASP32

is involved in conventional hydrogen bonding, while TYR71 is involved in forming a conventional Pi-Pi bond (Figure 5, left).

Molecular dynamics simulation (MD) and free energy landscape analysis

Molecular dynamics and simulation (MD) studies were carried out to determine the stability and convergence of the 4-(3,4-Dihydroxyphenyl)-2-hydroxy-1H-phenalen-1-one (PubChem I.D: 4468; Food I.D: FDB017657) bound BACE1 (PDB I.D: 2ZHV) complex. Each simulation of 150 ns displayed stable conformation while comparing the root mean square deviation (RMSD) values.

The Root Mean Square Deviation (RMSD) is a metric for calculating the average change in the displacement of a group of atoms in relation to a reference frame. It is calculated for each and every frame of the trajectory. The RMSD for frame x is:

$$RMSD_x = \sqrt{\frac{1}{N} \sum_{i=1}^N (r'_i(t_x) - r_i(t_{ref}))^2}$$

where N is the number of atoms in the atom selection; t_{ref} is the reference time (the first frame is usually used as the reference and is treated as time $t = 0$); and where r' is the position of the selected atoms after superimposing on the reference frame in frame x , where frame x is recorded at time t_x . Every frame in the simulation trajectory is subjected to the same technique (Maiorov and Crippen, 1994).

For characterizing local changes along the protein chain, the Root Mean Square Fluctuation (RMSF) is useful. The RMSF for residue i is:

$$RMSF_i = \sqrt{\frac{1}{T} \sum_{t=1}^T \langle (r'_i(t) - r_i(t_{ref}))^2 \rangle}$$

The angle brackets indicate that the average of square distance is taken over the selection of atoms in the residue. where T is the trajectory time over which the RMSF is calculated, t_{ref} is the reference time, r_i is the position of residue i r' is the position of atoms in residue i after superposition on the reference. Its simulation paths of Desmond were examined. MD trajectory analysis was used to calculate the root mean square deviation (RMSD), root mean square fluctuation (RMSF), and protein–ligand interactions. Protein RMSD: The graphs depict the evolution of a protein's RMSD (left Y-axis). The RMSD is estimated based on the atom selection once all protein frames are aligned on the reference frame backbone.

The α -backbone of BACE1 bound to 4-(3,4-dihydroxyphenyl)-2-hydroxy-1H-phenalen-1-one (PubChem I.D: 4468; Food I.D: FDB017657) exhibited a deviation of 0.4 Å (Figure 6A). RMSD plots are within the acceptable range signifying the stability of proteins in the FDB017657 bound state before and after simulation and it can also be suggested

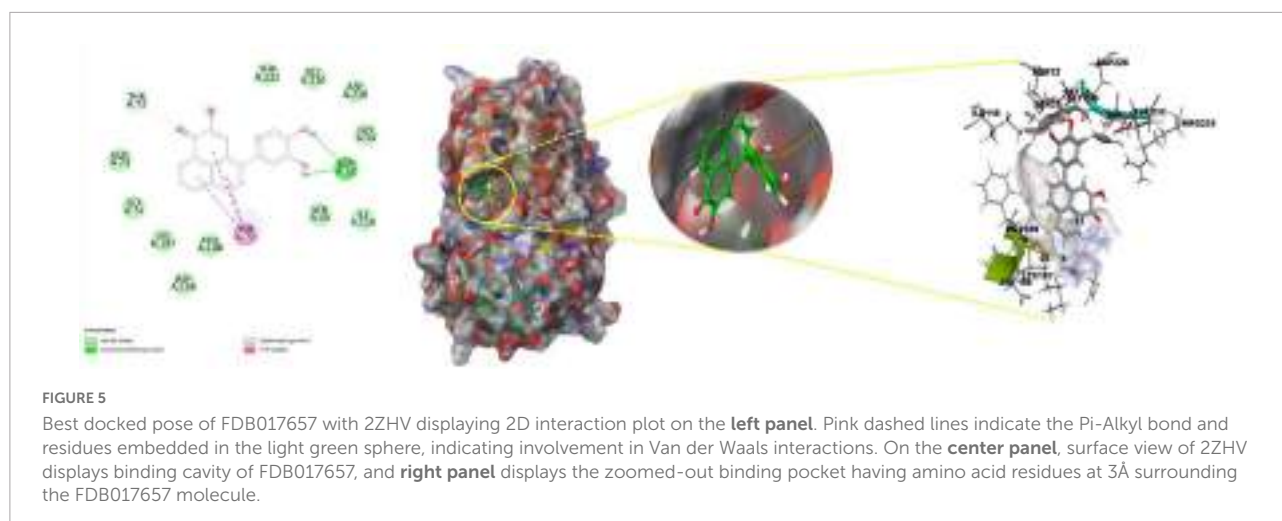
that FDB017657 bound BACE1 (PDB I.D: 2ZHV) is quite stable in the complex might be due to significant binding of the ligand.

The radius of gyration is the measure of the compactness of the protein. FDB017657 bound proteins displayed a lowering of radius of gyration (R_g) (Figure 6B; R1, R2, R3). The lowering of R_g indicates the compactness of the protein–ligand complex. From the overall quality analysis from RMSD and R_g , it can be suggested that FDB017657 bound to the protein targets posthumously in the binding cavities and plays a significant role in the stability of the proteins.

The plots for root mean square fluctuations (RMSF) displayed a significant RMSF in BACE1 protein at a few residues at the specific time function of 150 ns. Peaks show sections of the protein that fluctuate the greatest during the simulation on the RMSF plot. Typically, the tails (N- and C-terminal) of proteins change more than any other portion of the protein. Secondary structural parts such as alpha helices and beta strands are usually more rigid than the unstructured portion of the protein and fluctuate less than loop areas. The residues with higher peaks belong to loop areas or N and C-terminal zones, as determined by MD trajectories (Figure 6C). The stability of ligand binding to the protein is shown by low RMSF values of binding site residues. From the triplicate runs of BACE1, as shown in Figure 6C, a few fluctuating peaks can be seen although mostly the complex is found to be stabilized as shown in Figure 6C. The RMSF values are acceptable for stabilizing the protein–ligand complex. Therefore, in RMSF plots, it can be suggested that the protein structures were stable during simulation in FDB017657 bound conformation.

The average hydrogen bonds formed between FDB017657 and the respective protein, BACE1 (PDB I.D: 2ZHV), during the 150 ns simulation were also recorded (Figure 6D). From 0 ns to 150 ns a formation of hydrogen bonding was found throughout the simulation and the same for triplicate MD simulation of FDB017657 with BACE1 (Figure 6D). Moreover, the pattern of two hydrogen bond formation with BACE1 (PDB I.D: 2ZHV), in docking was corroborated by the number of hydrogen plot analyses after 150 ns molecular dynamics (Figure 6D). The amount of hydrogen bonds between BACE1 with FDB017657 has strengthened the binding and facilitated to conform to a more stable complex during the simulation.

Throughout the simulation, protein interactions with the ligand can be observed. As seen in the graph above, these interactions can be classified and summarized by type. Hydrogen bonds, hydrophobic, ionic, and water bridges are the four forms of protein–ligand interactions (or “contacts”). Each interaction type has a number of subtypes that can be examined using Maestro's “Simulation Interactions Diagram” panel (see Figure 7A). The stacked bar charts are standardized over the course of the trajectory. Some protein residues may make several interactions of the same subtype with the ligand, values above 1.0 are feasible. As shown in Figure 7A, the majority of



the significant ligand–protein interactions discovered by MD are hydrogen bonds and hydrophobic interactions. For 2ZHV–FDB017657, complex residues VAL_31, ASP_32, TYR_71, and THR_72 are the most important ones in terms of H-bonds.

Individual ligand atom interactions with protein residues are depicted in **Figure 7B**. Interactions that occur for more than 30.0% of the simulation period in the chosen trajectory (0.00 through 150.0 ns) are displayed. From **Figure 7B**, it can be concluded that the amino acid residues: PHE108, TRP76, TYR71, and VAL69 involve a hydrophobic interaction, LYS107, ARG128 possess a positive charge bonding with the ligand, GLN73, SER35, and ASN37 are involved in polar interactions, and ASP32 and ASP228 are involved in negatively charged interaction with the ligand, FDB017657 in 150 ns simulation time scale.

Throughout the simulation, the existence of protein secondary structural elements (SSE) such as alpha helices and beta strands is examined to ensure that they are not present. The plot shown in **Figure 7C** depicts the distribution of SSE by residue index over the complete protein structure, and it encompasses the full protein structure. In contrast to the charts, which show the summary of the SSE composition for each trajectory frame during the course of the simulation, the graphs at the bottom show the evolution of each residue and its SSE assignment throughout the experiment. Throughout the simulation, alpha-helices and beta-strands are monitored as secondary structure elements (SSE). The left graph shows the distribution of SSE across the protein structure by the residue index. The top image highlights the SSE composition for each trajectory frame throughout the simulation, while the bottom plot tracks each residue's SSE assignment through time.

It can be observed from **Figure 7D**, how each rotatable bond (RB) in the ligand alters its conformation throughout the simulation on the ligand torsions map (0.00 through 150.15 ns). The top panel shows a two-dimensional schematic of a ligand with color-coded rotatable bonds. There includes a dial plot

as well as bar plots in the same color for each rotatable bond torsion. The evolution of the torsion's conformation during the simulation is depicted using dial (or radial) graphs. The simulation's time evolution is depicted radially outwards from the simulation's start point at the center of the radial plot. The data from the dial plots are summarized in the bar plots, which show the torsion probability density in the data. Alternatively, if torsional potential data is available, the graphic will also indicate the rotatable bond's potential (by summing the potential of the related torsions) kcal/mol. The potential values are given as kcal/mol and plotted on the graph's left Y-axis. The histogram and torsion potential correlations can reveal the conformational strain that the ligand is under in order to maintain a protein-bound conformational state.

The stepwise trajectory analysis of every 25 ns of simulation of FDB017657 with BACE1 displayed the positional alteration with reference to the 0 ns structure (**Figure 8**). It has been observed that the ligand, FDB017657 has possessed a structural angular movement at the end frame to achieve its conformational stability and convergence.

The free energy landscape (FEL) of achieving global minima of C α backbone atoms of proteins with respect to RMSD and radius of gyration (Rg) is displayed in **Figure 9**, BACE1 bound to the ligand, FDB017657 achieved the global minima (lowest free energy state) at 1.1 Å and Rg 20.9 Å (**Figure 9**). The FEL envisaged a deterministic behavior of BACE1 to the lowest energy state owing to its high stability and best conformation at FDB017657 bound state.

Molecular mechanics generalized born and surface area calculations

To assess the binding energy of ligands to protein molecules, the MMGBSA technique is commonly employed. The binding free energy of each BACE1– FDB017657 complex, as well

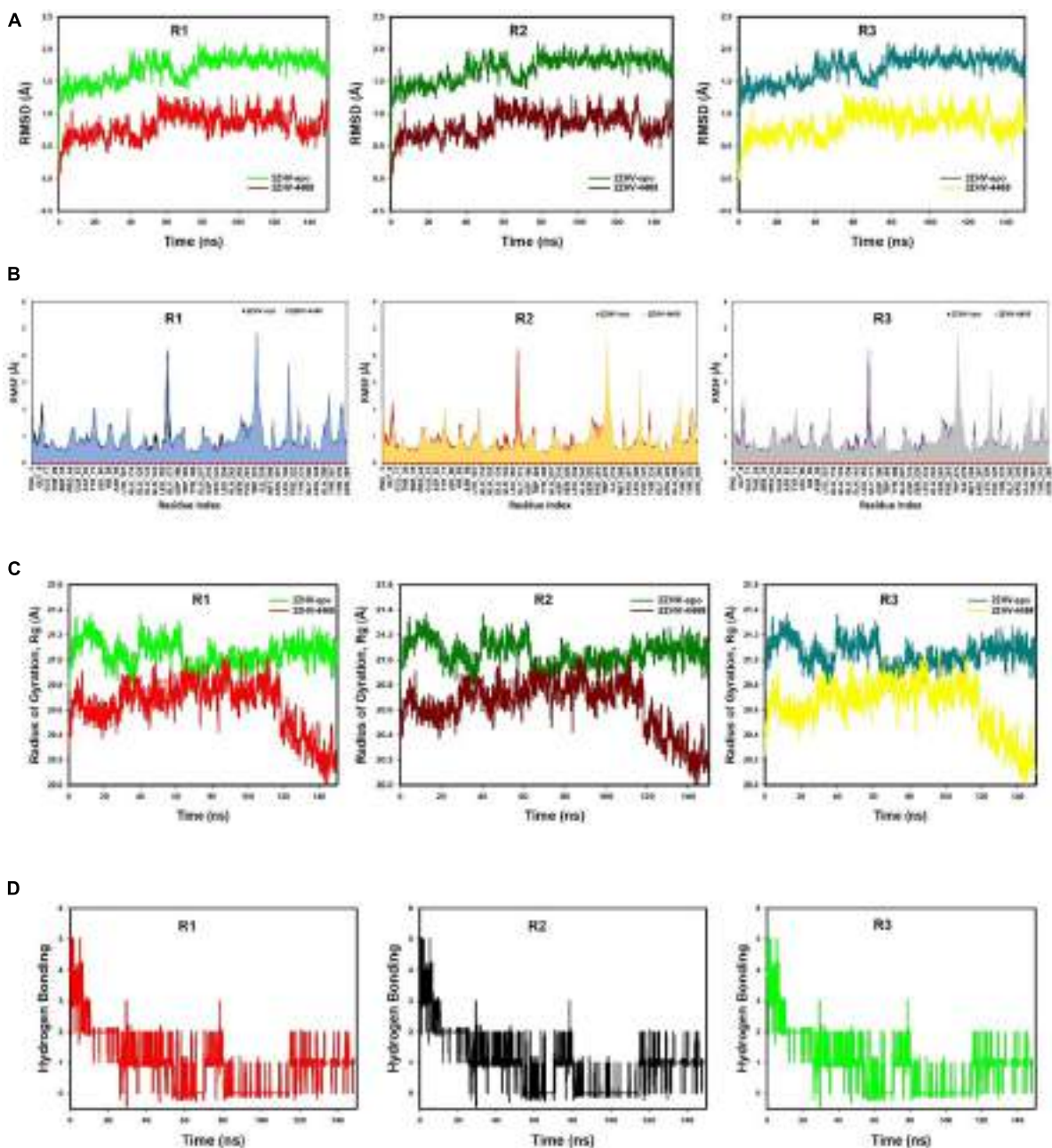


FIGURE 6

(A) MD simulation trajectory analysis of Root Mean Square Divisions (RMSD) of FDB017657 bound with 2ZHV, i.e., BACE1 150 ns time frame in triplicate displayed: R1 (replicate 1) RMSD plot of FDB017657 bound BACE1 (PDB I.D.: 2ZHV) (red) with control protein BACE1 (PDB I.D.: 2ZHV) (light green); R2 (replicate 2) RMSD plot of FDB017657 bound BACE1 (PDB I.D.: 2ZHV) (dark maroon) with control protein BACE1 (PDB I.D.: 2ZHV) (juniper green); R3 (replicate 3) RMSD plot of FDB017657 bound BACE1 (PDB I.D.: 2ZHV) (lemon yellow) with control protein BACE1 (PDB I.D.: 2ZHV) (cyan). (B) MD simulation trajectory analysis of Root Mean Square Fluctuations (RMSF) of FDB017657 bound with BACE1 (PDB I.D.: 2ZHV) at 150 ns time frame in triplicate displayed: R1 (replicate 1) RMSF plot of FDB017657 bound BACE1 (PDB I.D.: 2ZHV) (navy blue) with control protein BACE1 (PDB I.D.: 2ZHV) (black); R2 (replicate 2) RMSF plot of FDB017657 bound BACE1 (PDB I.D.: 2ZHV) (canary yellow) with control protein BACE1 (PDB I.D.: 2ZHV) (red); R3 (replicate 3) RMSF plot of FDB017657 bound BACE1 (PDB I.D.: 2ZHV) (gray) with control protein BACE1 (PDB I.D.: 2ZHV) (purple). (C) MD simulation trajectory analysis of Radius of gyration (Rg) of FDB017657 bound with BACE1 (PDB I.D.: 2ZHV) at 150 ns time frame in triplicate displayed: R1 (replicate 1) Rg plot of FDB017657 bound BACE1 (PDB I.D.: 2ZHV) (red) with control protein BACE1 (PDB I.D.: 2ZHV) (light green); R2 (replicate 2) Rg plot of FDB017657 bound BACE1 (PDB I.D.: 2ZHV) (dark maroon) with control protein BACE1 (PDB I.D.: 2ZHV) (juniper green); R3 (replicate 3) Rg plot of FDB017657 bound BACE1 (PDB I.D.: 2ZHV) (cyan) with control protein BACE1 (PDB I.D.: 2ZHV) (lemon yellow). (D) MD simulation trajectory analysis of Hydrogen Bonding (H-Bonds) of FDB017657 bound with BACE1 (PDB I.D.: 2ZHV) at 150 ns time frame in triplicate displayed: R1 (replicate 1) H-Bond plot of FDB017657 bound BACE1 (PDB I.D.: 2ZHV) (red); R2 (replicate 2) H-Bond plot of FDB017657 bound BACE1 (PDB I.D.: 2ZHV) (black); R3 (replicate 3) H-Bond plot of FDB017657 bound BACE1 (PDB I.D.: 2ZHV) (light green).

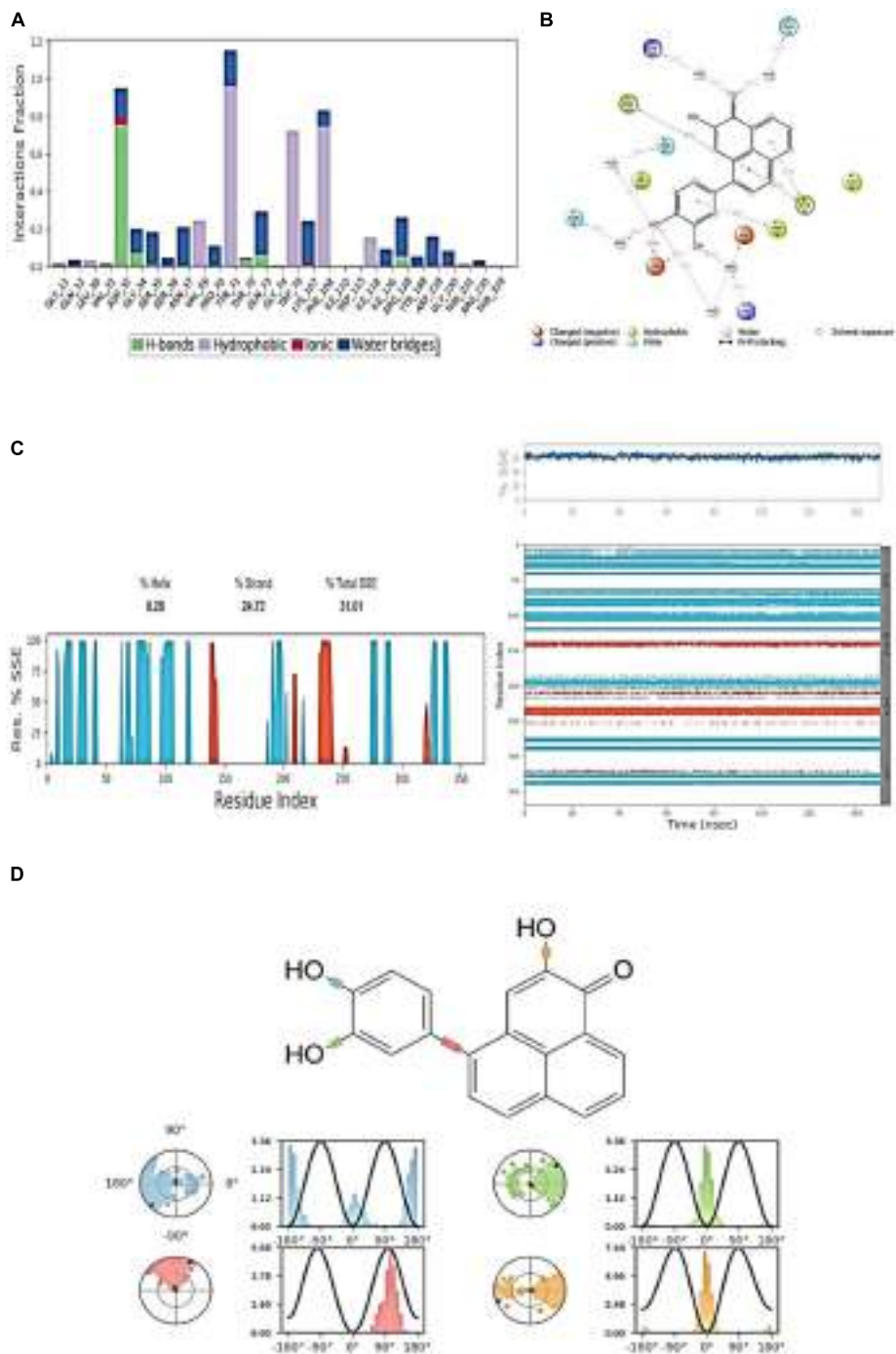


FIGURE 7

(A) Protein-ligand contact histogram (H-bonds, Hydrophobic, Ionic, Water bridges) of the ligand, FDB017657 bound with 2ZHV recorded in a 150 ns simulation interval. (B) Ligand atom interactions with the protein residues of 2ZHV bound with FDB017657. (C) Secondary Structure element distribution by residue index throughout the protein structure. Red indicates alpha helices, and blue indicate beta-strands of 2ZHV bound with FDB017657. (D) Ligand torsion profile.

as the impact of other non-bonded interaction energies, were estimated. With BACE1, the ligand FDB017657 has a binding energy of -53.4670 kcal/mol. Non-bonded interactions like $G_{\text{bindCoulomb}}$, $G_{\text{bindCovalent}}$, $G_{\text{bindHbond}}$, G_{bindLipo} ,

$G_{\text{bindSolvGB}}$, and G_{bindvdW} govern G_{bind} . Across all types of interactions, the G_{bindvdW} , G_{bindLipo} , and $G_{\text{bindCoulomb}}$ energies contributed the most to the average binding energy. On the other side, the $G_{\text{bindSolvGB}}$ and $G_{\text{bindCovalent}}$

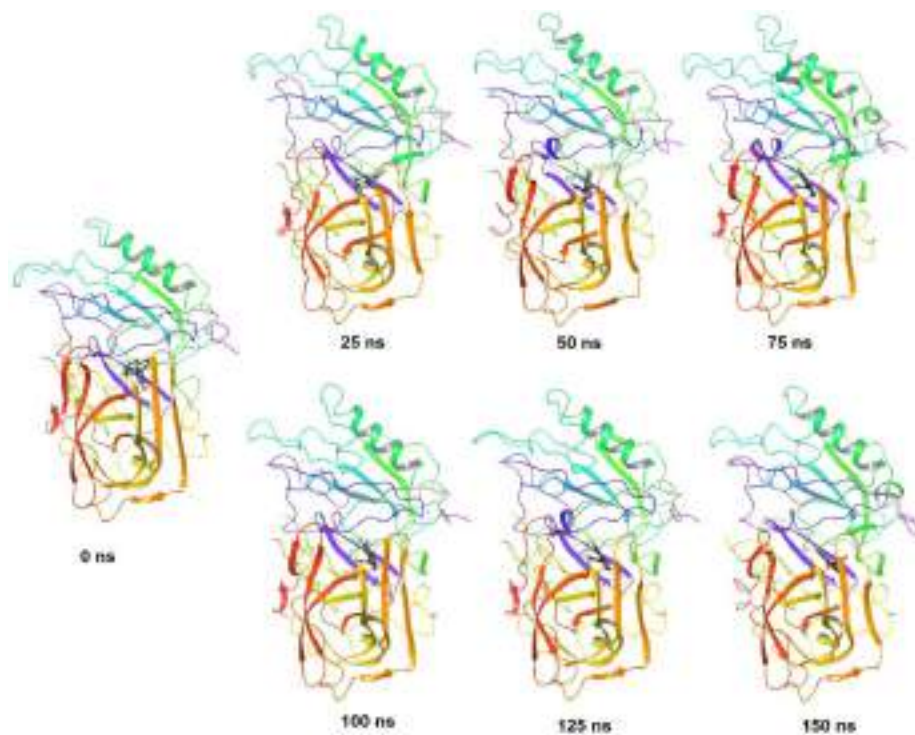


FIGURE 8

Stepwise trajectory analysis for every 25 ns displaying the protein, BACE1 (PDB I.D: 2ZHV) and ligand conformation during 150 ns of simulation of 4-(3,4-Dihydroxyphenyl)-2-hydroxy-1H-phenalen-1-one (PubChem I.D: 4468; Food I.D: FDB017657).

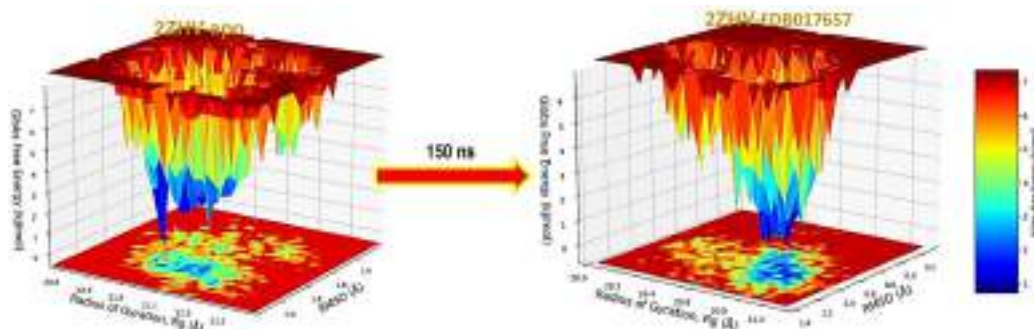


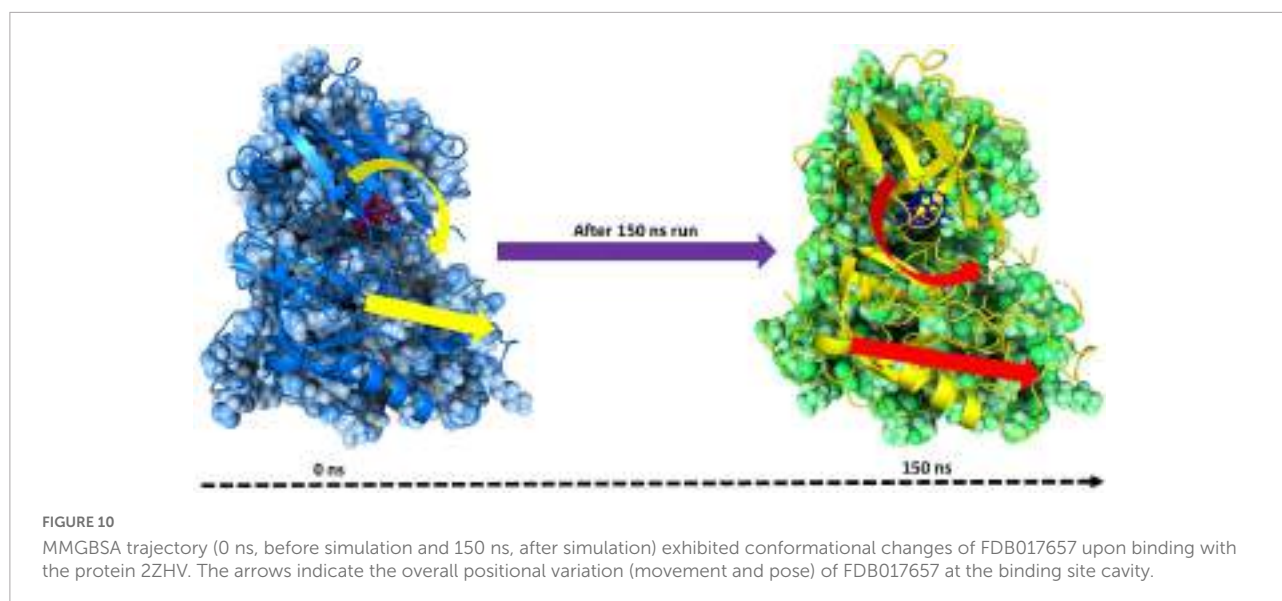
FIGURE 9

Free Energy Landscape displaying the achievement of global minima (ΔG , kJ/mol) of BACE1 in presence of FDB017657 with respect to their RMSD (nm) and Radius of gyration (R_g , nm).

energies contributed the least to the final average binding energies. Furthermore, the $G_{\text{bind}}^{\text{Hbond}}$ interaction values of BACE1–FDB017657 complexes demonstrated stable hydrogen bonds with amino acid residues. In all of the compounds, $G_{\text{bind}}^{\text{SolvGB}}$ and $G_{\text{bind}}^{\text{Covalent}}$ exhibited unfavorable energy contributions and so opposed binding. **Figure 10** (left panel) reveals that between pre-simulation (0 ns) and post-simulation (150 ns), FDB017657 in the binding pocket of BACE1 has undergone a large angular change in the pose (curved to

straight) (150 ns). These conformational changes lead to better binding pocket acquisition and interaction with residues, which leads to enhanced stability and binding energy (mentioned in **Table 3**).

Thus, MM-GBSA calculations resulted from MD simulation trajectories well justified with the binding energy obtained from docking results; moreover, the last frame (150 ns) of MMGBSA displayed the positional change of FDB017657 as compared to the 0 ns trajectory signifying the better binding



pose for best fitting in the binding cavity of the protein (see [Figure 10](#)).

Therefore, it can be suggested that the FDB017657 molecule has a good affinity for the major target BACE1.

Dynamic cross-correlation, principal component analysis (PCA), and energy calculations

Molecular dynamics simulation trajectories are analyzed for dynamic cross-correlation among the domains within protein chains bound with the FDB017657 molecule. For correlative dynamic motion, the cross-correlation matrices of BACE1 were generated and displayed in [Figure 11](#). The blue blocks displayed in the figure indicated the residues having high correlated movement and red having the least correlation. The amino acid residues of FDB017657 bound BACE1 showed the concerted movement of residues ([Figure 11](#)).

Principal component analysis (PCA) determines the relationship between statistically meaningful conformations

(major global motions) sampled during the trajectory. PCA of the MD simulation trajectories for BACE1 bound to the FDB017657 molecule was analyzed to interpret the randomized global motion of the atoms of amino acid residues. The internal coordinates mobility into three-dimensional space in the spatial time of 150 ns were recorded in a covariance matrix and the rational motion of each trajectory is interpreted in the form of orthogonal sets or Eigen vectors. In the BACE1 trajectory, PCA indicates statistically significant conformations. It is possible to identify the major motions within the trajectory as well as the critical motions required for conformational changes. In BACE1 bound to FDB017657, two different clusters along the PC1 and PC2 planes are exhibited that indicate a non-periodic conformational shift ([Figure 12A](#)). While these global motions are periodic because the groupings along the PC3 and PC4 planes do not totally cluster separately ([Figure 12B](#)). Moreover, a high periodic global motion was observed along the PC9 and PC10 planes due to the grouping of trajectories in a single cluster at the center of the PCA plot ([Figure 12C](#)). Centering of the trajectories in a single cluster indicates the periodic motion of MD trajectories due to stable conformational global motion.

The energy profiles of the protein, BACE1 and FDB017657 complex systems, were determined to display the stability of the entire system. In this regard, the total energy (ETOT) of the BACE1 bound FDB017657 system was shown to be very stable with an average total energy of -69.00 kcal/mol (green). However, van der Waal's energy (vdW) displayed to be merged over the total energy with an average energy of -40.00 kcal/mol and contemplated as a principal contributor to the stability of the BACE1-FDB017657 complex (cyan). In addition, Coulombic interactions played a minor role in the system stability and contributed to an average energy of -32.00 kcal/mol (red), (see [Figure 12D](#)).

TABLE 3 Binding energy calculation of FDB017657 with 2ZHV and non-bonded interaction energies from MMGBSA trajectories.

Energies (kcal/mol)	2ZHV
ΔG_{bind}	-53.467 ± 3.001
$\Delta G_{\text{bindLipo}}$	-22.124 ± 2.448
$\Delta G_{\text{bindvdW}}$	33.667 ± 0.0701
$\Delta G_{\text{bindCoulomb}}$	-9.827 ± 5.083
$\Delta G_{\text{bindHbond}}$	-1.465 ± 0.775
$\Delta G_{\text{bindSolvGB}}$	-8.989 ± 1.695
$\Delta G_{\text{bindCovalent}}$	-1.079 ± 1.049

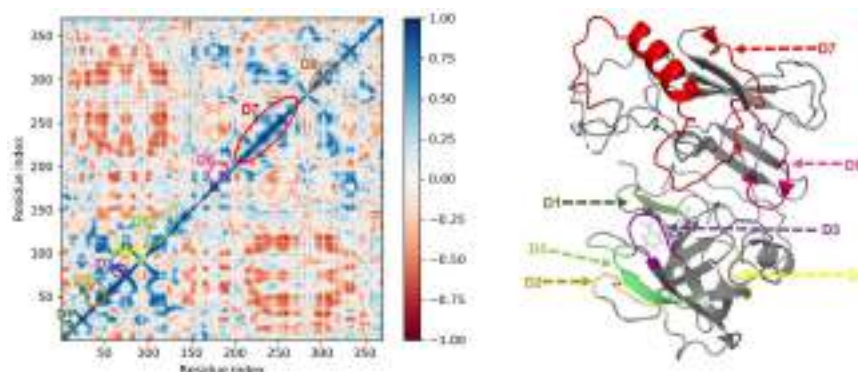


FIGURE 11

Dynamic Cross Correlation matrix (DCCM) of 2ZHV and correlated amino acids conformed into secondary structural domains (colored) and non-correlated domains (gray) of 2ZHV.

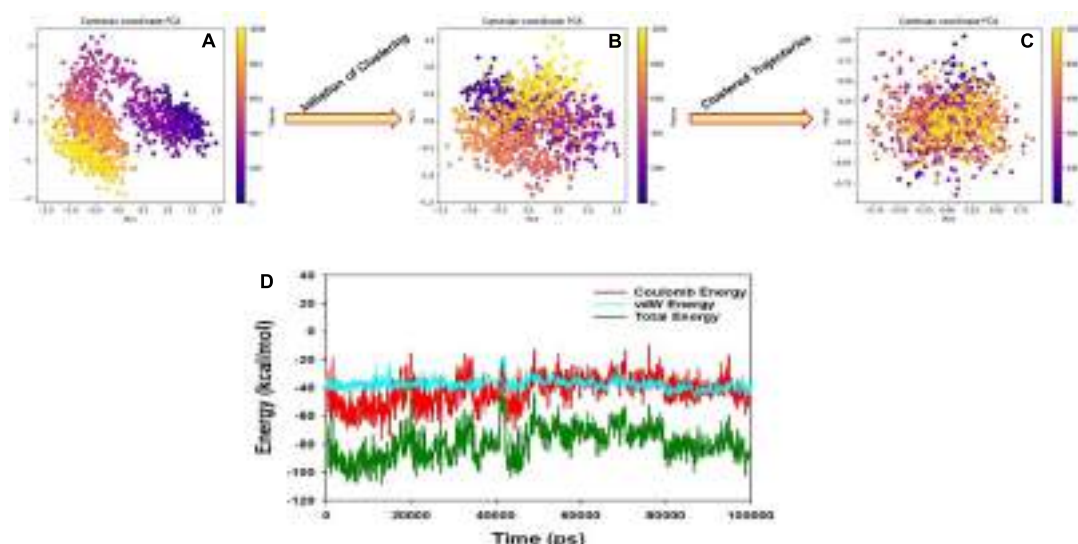


FIGURE 12

(A) PCA of 2ZHV- FDB017657 showing a stable configuration. (B) Energy plot of protein BACE1 and FDB017657 complex system during the entire simulation event of 150 ns. (C,D) The change in PCA movements. The total energy (dark green), van der Waal's energy (cyan) and Coulomb energy (red) of the entire system indicate the stability of the individual systems bound to FDB017657 molecule.

Discussion

Proteolytic processing of APP by BACE1 is the rate-determining step in A β production, hence BACE1 is employed as a therapeutic target for creating innovative lead compounds in AD in this study. According to the earlier reports, it was suggested that the enzyme BACE1 is also associated with different types of cancers and viruses in conjunction with AD. In this study, we have tried to reveal the potential of naturally available food molecules to bind the BACE1's active site in a highly specific binding pattern. The aim of our study is toward the development of a drug from food compounds with the help of computational biology as

it has the additional advantage regarding safety, and lesser chance of side effects. The low toxicity profile of natural products inspired by small-molecule inhibitors may prove to be a great asset during the frenetic development period of drug discovery when time is of the essence. Current state-of-the-art computational approaches can be used to identify structural and pharmacophoric properties of active natural compounds that can be used as drugs. Our results suggest that the selected 8,453 compounds from the Food database are majorly phenols and naphthol metabolites having a high potential of showing inhibitory activity against BACE1. The Food database is a recent database that proved the potential of food metabolites that we use in our daily life and found its

major application in developing different therapeutics treating depressive disorders and others.

A nicely proven correlation among salient capabilities of the molecules represented through molecular descriptors, and their bioactivity expands statistics approximately mechanistic elements of molecules, specificity, and quantity (presence or even absence) of various structural developments for preferred bioactivity. Although, with the QSAR analysis, we've compared the K_i values of various molecules in correlation and as an impact of a specific molecular descriptor, a similar or contrary impact of different molecular descriptors or unknown descriptors has a dominant impact in figuring out the general K_i value of a molecule can't be neglected. In other words, a single molecular descriptor is incapable of absolutely explaining the experimental K_i value for this sort of numerous sets of molecules. That is, the successful usage of the advanced QSAR model is based on the concomitant usage of molecular descriptors.

A QSAR model with multiple chemical descriptors is built using a dataset of 371 compounds. The resulting model was rigorously verified for fitting and internal validation to prove its strong external prediction capacity and resilience. In addition, virtual screening using QSAR yielded a novel food molecule with a better K_i value of 10.715 nM. Combined QSARs and molecular docking studies offered complimentary information and helped discover prodigious and under-privileged chemical characteristics that might be leveraged to change a molecule to produce better BACE1 inhibitors with higher K_i values. In the future, structural modifications that result in augmented values for the molecular descriptors with positive coefficients in the developed model for the anti-BACE1 activity will be performed to generate novel hits suitable for construction and *in vitro* evaluation as anti-Alzheimer's (AD) disease therapy. These reports already demonstrated the potential of this plant as a source of novel drugs, nutraceuticals, and functional foods. Our present study perhaps supports a further avenue for *in vivo* and clinical trial of the food molecule, 4-(3,4-dihydroxyphenyl)-2-hydroxy-1H-phenalen-1-one to target BACE1 for any future scope to treat the AD along with viruses and cancer.

Data availability statement

The original contributions presented in this study are included in the article/**Supplementary material**, further inquiries can be directed to the corresponding authors.

Author contributions

NM and ADa performed the concept design. NM, AG, SwM, and RJ designed and carried out the experimental procedures. NM, AG, RJ, RB, MC, and VM done the analysis. NM, ADa, PD, AP, GM, SB, BA, and SA-H done the manuscript preparation.

NM, ADe, GM, MK, AAl, ADa, SuM, AMA, and MZ edited the manuscript. All authors contributed to the article and approved the submitted version.

Funding

This research work was funded by the Institutional Fund Projects under grant no. IFPDP-85-22. This research work was funded by the Institutional Fund Projects under grant no. (IFPDP-85-22). Therefore, authors gratefully acknowledge technical and financial support from Ministry of Education and Deanship of Scientific Research (DSR), King Abdulaziz University, Jeddah, Saudi Arabia.

Conflict of interest

AA is honorary associated with the scientific board of the company AFNP Med in Austria.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fnagi.2022.878276/full#supplementary-material>

SUPPLEMENTARY TABLE 1

Smileskipki.

SUPPLEMENTARY TABLE 2

Descriptor used in QSAR.

SUPPLEMENTARY TABLE 3

Formulas for calculation of model parameters.

SUPPLEMENTARY TABLE 4

Food molecules predicted pki and descriptors.

References

- Alzheimer's Association (2012). 2012 Alzheimer's disease facts and figures. *Alzheimer's Dement.* 8, 131–168.
- Arif, N., Subhani, A., Hussain, W., and Rasool, N. (2020). In silico inhibition of BACE-1 by selective phytochemicals as novel potential inhibitors: Molecular docking and DFT studies. *Curr. Drug Disc. Technol.* 17, 397–411. doi: 10.2174/1570163816666190214161825
- Bellacasa, R. P., Karachaliou, N., Estrada-Tejedor, R., Teixidó, J., Costa, C., and Borrrell, J. I. A. L. K. (2013). and ROS1 as a joint target for the treatment of lung cancer: A review. *Transl. Lung Cancer Res.* 2, 72–86.
- Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., et al. (2000). The protein data bank. *Nucleic Acids Res.* 28, 235–242. doi: 10.1093/nar/28.1.235
- Bo, Q., Runhua, L., Hanqing, W., Min, W., Wenpeng, L., and Jinlun, X. (2000). Chemical constituents from *Musella lasiocarpa* (Franch.) CY Wu. *Nat. Product Res. Develop.* 12, 41–44.
- Bonvino, N. P., Liang, J., McCord, E. D., Zafiris, E., Benetti, N., Ray, N. B., et al. (2018). OliveNetTM: A comprehensive library of compounds from *Olea europaea*. *Database* 2018:bay016. doi: 10.1093/database/bay016
- Bowers, K. J., Chow, D. E., Xu, H., Dror, R. O., Eastwood, M. P., Gregersen, B. A., et al. (2006). "Scalable algorithms for molecular dynamics simulations on commodity clusters," in *SC'06: Proceedings of the 2006 ACM/IEEE Conference on Supercomputing*, (Tampa, FL: IEEE), 43–43. doi: 10.1109/SC.2006.54
- Chow, E., Rendleman, C. A., Bowers, K. J., Dror, R. O., Hughes, D. H., Gullingsrud, J., et al. (2008). *Desmond performance on a cluster of multicore processors*. New York, NY: DE Shaw Research Technical Report DESRES/TR-2008-01.
- Cui, H., Hung, A. C., Klaver, D. W., Suzuki, T., Freeman, C., Narkowicz, C., et al. (2011). Effects of heparin and enoxaparin on APP processing and a β production in primary cortical neurons from Tg2576 mice. *PLoS One* 6:e23007. doi: 10.1371/journal.pone.0023007
- Davies, M., Nowotka, M., Papadatos, G., Dedman, N., Gaulton, A., and Atkinson, F. (2015). ChEMBL web services: Streamlining access to drug discovery data and utilities. *Nucleic Acids Res.* 43, D945–D954. doi: 10.1093/nar/gkv352
- De Strooper, B., and Karran, E. (2016). The cellular phase of Alzheimer's disease. *Cell* 164, 603–615. doi: 10.1016/j.cell.2015.12.056
- Dingwall, C. (2001). Spotlight on BACE: The secretases as targets for treatment in Alzheimer disease. *J. Clin. Invest.* 108, 1243–1246. doi: 10.1172/JCI14402
- Dong, L. B., He, J., Li, X. Y., Wu, X. D., Deng, X., Xu, G., et al. (2011). Chemical constituents from the aerial parts of *Musella lasiocarpa*. *Nat. Products Bioprospect.* 1, 41–47. doi: 10.1007/s13659-011-0007-7
- Forli, S., Huey, R., Pique, M. E., Sanner, M. F., Goodsell, D. S., and Olson, A. J. (2016). Computational protein–ligand docking and virtual drug screening with the AutoDock suite. *Nat. Protocols* 11, 905–919. doi: 10.1038/nprot.2016.051
- Fujita, T., and Winkler, D. A. (2016). Understanding the Roles of the "Two QSARs". *J. Chem. Inf. Model* 56, 269–274. doi: 10.1021/acs.jcim.5b00229
- Fukumoto, H., Cheung, B. S., Hyman, B. T., and Irizarry, M. C. (2002). β -Secretase protein and activity are increased in the neocortex in Alzheimer disease. *Arch. Neurol.* 59, 1381–1389. doi: 10.1001/archneur.59.9.1381
- Gramatica, P. (2013). On the development and validation of QSAR models. *Methods Mol. Biol.* 930, 499–526. doi: 10.1007/978-1-62703-059-5_21
- Gramatica, P. (2020). Principles of QSAR Modeling. *Int. J. Quant. Struct. Property Relation.* 5, 61–97. doi: 10.4018/IJQSPR.2020.0701.oa1
- Gramatica, P., Cassani, S., and Chirico, N. (2014). QSARINS-Chem: Insubria Datasets and New QSAR/QSPR Models for Environmental Pollutants in QSARINS. *J. Comput. Chem. Softw. News Updates* 35, 1036–1044. doi: 10.1002/jcc.23576
- Gramatica, P., Chirico, N., Papa, E., Kovarich, S., and Cassani, S. (2013). QSARINS: A New Software for the Development, Analysis, and Validation of QSAR MLR Models. *J. Comp. Chem. Softw. News Updates* 34, 2121–2132. doi: 10.1002/jcc.23361
- Hall, A., Pekkala, T., Polvikoski, T., Van Gils, M., Kivipelto, M., Lötjönen, J., et al. (2019). Prediction models for dementia and neuropathology in the oldest old: The Vantaa 85+ cohort study. *Alzheimer's Res. Ther.* 11, 1–2. doi: 10.1186/s13195-018-0450-3
- Hassan, M., Shahzadi, S., Seo, S. Y., Alashwal, H., Zaki, N., and Moustafa, A. A. (2018). Molecular docking and dynamic simulation of AZD3293 and solanezumab effects against BACE1 to treat Alzheimer's disease. *Front. Comput. Neurosci.* 12:34. doi: 10.3389/fncom.2018.00034
- Hussain, I., Powell, D. J., Howlett, D. R., Chapman, G. A., Gilmour, L., Murdock, P. R., et al. (2000). ASP1 (BACE2) cleaves the amyloid precursor protein at the β -secretase site. *Mol. Cell. Neurosci.* 16, 609–619. doi: 10.1006/mcne.2000.0884
- Jabir, N. R., Rehman, M. T., Alsolami, K., Shakil, S., Zughabi, T. A., Alserihi, R. F., et al. (2021). Concatenation of molecular docking and molecular simulation of BACE-1, γ -secretase targeted ligands: In pursuit of Alzheimer's treatment. *Ann. Med.* 53, 2332–2344. doi: 10.1080/07853890.2021.2009124
- Jawarkar, R. D., Bakal, R. L., Zaki, M. E. A., Al-Hussain, S., Ghosh, A., Gandhi, A., et al. (2022). QSAR based virtual screening derived identification of a novel hit as a SARS CoV-229E 3CLpro Inhibitor: GA-MLR QSAR modeling supported by molecular Docking, molecular dynamics simulation and MMGBSA calculation approaches. *Arab. J. Chem.* 15:103499. doi: 10.1016/j.arabj.2021.103499
- Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., and Klein, M. L. (1983). Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* 79, 926–935. doi: 10.1063/1.445869
- Kagami, L. P., das Neves, G. M., Timmers, L. F. S. M., Caceres, R. A., and Eifler-Lima, V. L. (2020). Geo-Measures: A Pymol plugin for protein structure ensembles analysis. *Comp. Biol. Chem.* 87:107322. doi: 10.1016/j.compbiolchem.2020.107322
- Kim, S., Chen, J., Cheng, T., Gindulyte, A., He, J., He, S., et al. (2019). PubChem 2019 update: Improved access to chemical data. *Nucleic Acids Res.* 47, D1102–D1109. doi: 10.1093/nar/gky1033
- Li, Q., and Südhof, T. C. (2004). Cleavage of amyloid- β precursor protein and amyloid- β precursor-like protein by BACE 1. *J. Biol. Chem.* 279, 10542–10550. doi: 10.1074/jbc.M310001200
- Lin, X., Koelsch, G., Wu, S., Downs, D., Dashti, A., and Tang, J. (2000). Human aspartic protease memapsin 2 cleaves the β -secretase site of β -amyloid precursor protein. *Proc. Natl. Acad. Sci.* 97, 1456–1460. doi: 10.1073/pnas.97.4.1456
- Maiorov, V. N., and Crippen, G. M. (1994). Significance of root-mean-square deviation in comparing three-dimensional structures of globular proteins. *J. Mol. Biol.* 235, 625–634. doi: 10.1006/jmbi.1994.1017
- Martin, Y. C., and Muchmore, S. W. (2012). Frozen out: Molecular modeling in the age of cryocrystallography. *J. Comput. Aided Mol. Design* 26, 91–92.
- Martyna, G. J., Klein, M. L., and Tuckerman, M. (1992). Nosé–Hoover chains: The canonical ensemble via continuous dynamics. *J. Chem. Phys.* 97, 2635–2643. doi: 10.1063/1.463940
- Martyna, G. J., Tobias, D. J., and Klein, M. L. (1994). Constant pressure molecular dynamics algorithms. *J. Chem. Phys.* 101, 4177–4189. doi: 10.1063/1.467468
- Masand, V. H., Mahajan, D. T., Alafeefy, A. M., Bukhari, S. N., and Elsayed, N. N. (2015). Optimization of antiproliferative activity of substituted phenyl 4-(2-oxoimidazolidin-1-yl) benzenesulfonates: QSAR and CoMFA analyses. *Eur. J. Pharm. Sci.* 77, 230–237.
- Masand, V. H., and Rastija, V. (2017). PyDescriptor: A new PyMOL plugin for calculating thousands of easily understandable molecular descriptors. *Chemom. Intell. Labor. Syst.* 169, 12–18. doi: 10.1016/j.chemolab.2017.08.003
- Morris, G. M., Huey, R., Lindstrom, W., Sanner, M. F., Belew, R. K., Goodsell, D. S., et al. (2009). AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *J. Comp. Chem.* 30, 2785–2791. doi: 10.1002/jcc.21256
- Moussa-Pacha, N. M., Abdin, S. M., Omar, H. A., Alniss, H., and Al-Tel, T. H. (2020). BACE1 inhibitors: Current status and future directions in treating Alzheimer's disease. *Med. Res. Rev.* 40, 339–384. doi: 10.1002/med.21622
- Mukerjee, N., Das, A., Maitra, S., Ghosh, A., Khan, P., Alexiou, A., et al. (2022). Dynamics of natural product lupenone as a potential fusion inhibitor against the spike complex of novel semliki forest virus. *PLoS One* 17:e0263853. doi: 10.1371/journal.pone.0263853
- Murphy, M. P., and LeVine, III, H. (2010). Alzheimer's disease and the amyloid-beta peptide. *J. Alzheimers Dis.* 19, 311–323. doi: 10.3233/JAD-2010-1221
- Musi, N., Valentine, J. M., Sickora, K. R., Baeuerle, E., Thompson, C. S., Shen, Q., et al. (2018). Tau protein aggregation is associated with cellular senescence in the brain. *Aging cell* 17:e12840. doi: 10.1111/acer.12840
- Neves, B. J., Braga, R. C., Melo-Filho, C. C., Moreira-Filho, J. T., Muratov, E. N., and Andrade, C. H. (2018). QSAR-Based Virtual Screening: Advances and Applications in Drug Discovery. *Front. Pharmacol.* 9:1275. doi: 10.3389/fphar.2018.01275





- O'Boyle, N. M., Banck, M., James, C. A., Morley, C., Vandermeersch, T., and Hutchison, G. R. (2011). Open Babel: An open chemical toolbox. *J. Cheminform.* 3:33. doi: 10.1186/1758-2946-3-33
- Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., et al. (2004). UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comp. Chem.* 25, 1605–1612. doi: 10.1002/jcc.20084
- Piao, L., Chen, Z., Li, Q., Liu, R., Song, W., Kong, R., et al. (2019). Molecular dynamics simulations of wild type and mutants of SAPAP in complexed with Shank3. *Int. J. Mol. Sci.* 20:224. doi: 10.3390/ijms20010224
- Roskoski, R. Jr. (2013). Anaplastic lymphoma kinase (ALK): Structure, oncogenic activation, and pharmacological inhibition. *Pharmacol. Res.* 68, 68–94. doi: 10.1016/j.phrs.2012.11.007
- Sebastián-Serrano, Á, Diego-García, D., and Díaz-Hernández, M. (2018). The neurotoxic role of extracellular tau protein. *Int. J. Mol. Sci.* 19:998. doi: 10.3390/ijms19040998
- Shivakumar, D., Williams, J., Wu, Y., Damm, W., Shelley, J., and Sherman, W. (2010). Prediction of absolute solvation free energies using molecular dynamics free energy perturbation and the OPLS force field. *J. Chem. Theory Comp.* 6, 1509–1519. doi: 10.1021/ct900587b
- Sinha, S., Anderson, J. P., Barbour, R., Basí, G. S., Caccavello, R., Davis, D., et al. (1999). Purification and cloning of amyloid precursor protein β -secretase from human brain. *Nature* 402, 537–540. doi: 10.1038/990114
- Tosco, P., Balle, T., and Shiri, F. (2011). Open3DALIGN: An open-source software aimed at unsupervised ligand alignment. *J. Comp.Aided Mol. Design* 25, 777–783. doi: 10.1007/s10822-011-9462-9
- Toukmaji, A. Y., and Board, J. A. Jr. (1996). Ewald summation techniques in perspective: A survey. *Comp. Phys. Commun.* 95, 73–92. doi: 10.1016/0010-4655(96)00016-1
- Tresadern, G., Delgado, F., Delgado, O., Gijzen, H., Macdonald, G. J., Moechars, D., et al. (2011). Rational design and synthesis of aminopiperazinones as β -secretase (BACE) inhibitors. *Bioorgan. Med. Chem. Lett.* 21, 7255–7260. doi: 10.1016/j.bmcl.2011.10.050
- Trott, O., and Olson, A. J. (2010). AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.* 31, 455–461. doi: 10.1002/jcc.21334
- Vassar, R., Bennett, B. D., Babu-Khan, S., Kahn, S., Mendiaz, E. A., Denis, P., et al. (1999). β -Secretase cleavage of Alzheimer's amyloid precursor protein by the transmembrane aspartic protease BACE. *Science* 286, 735–741. doi: 10.1126/science.286.5440.735
- Volkamer, A., Kuhn, D., Grombacher, T., Rippmann, F., and Rarey, M. (2012). Combining global and local measures for structure-based druggability predictions. *J. Chem. Inf. Model* 52, 360–372. doi: 10.1021/ci200454v
- Yan, R., Bienkowski, M. J., Shuck, M. E., Miao, H., Tory, M. C., Pauley, A. M., et al. (1999). Membrane-anchored aspartyl protease with Alzheimer's disease β -secretase activity. *Nature* 402, 533–537. doi: 10.1038/990107
- Yang, L. B., Lindholm, K., Yan, R., Citron, M., Xia, W., Yang, X. L., et al. (2003). Elevated β -secretase expression and enzymatic activity detected in sporadic Alzheimer disease. *Nat. Med.* 9, 3–4. doi: 10.1038/nm0103-3
- Zacchetti, D., Chieregatti, E., Bettegazzi, B., Mihailovich, M., Sousa, V. L., Grohovaz, F., et al. (2007). BACE1 expression and activity: Relevance in Alzheimer's disease. *Neurodegener. Dis.* 4, 117–126. doi: 10.1159/000101836
- Zaki, M. E. A., Al-Hussain, S. A., Masand, V. H., Akasapu, S., Bajaj, S. O., El-Sayed, N. N. E., et al. (2021). Identification of Anti-SARS-CoV-2 Compounds from Food Using QSAR-Based Virtual Screening, Molecular Docking, and Molecular Dynamics Simulation Analysis. *Pharmaceuticals* 14:357. doi: 10.3390/ph14040357
- Zhang, Y. W., Thompson, R., Zhang, H., and Xu, H. (2011). APP processing in Alzheimer's disease. *Mol. Brain* 4:3. doi: 10.1186/1756-6606-4-3

CITATION

Mukerjee N, Das A, Jawarkar RD, Maitra S, Das P, Castrosanto MA, Paul S, Samad A, Zaki MEA, Al-Hussain SA, Masand VH, Hasan MM, Bukhari SNA, Perveen A, Alghamdi BS, Alexiou A, Kamal MA, Dey A, Malik S, Bakal RL, Abuzenadah AM, Ghosh A and Md Ashraf G (2022) Repurposing food molecules as a potential BACE1 inhibitor for Alzheimer's disease. *Front. Aging Neurosci.* 14:878276. doi: 10.3389/fnagi.2022.878276

Article

Quinoline Derivatives with Different Functional Groups: Evaluation of Their Catecholase Activity

Mohamed Moutaouakil ¹, Said Tighadouini ¹, Zainab M. Almarhoon ^{2,*}, Maha I. Al-Zaben ²,
Abir Ben Bacha ^{3,4}, Vijay H. Masand ⁵, Jamal Jamaledine ¹ and Rafik Saddik ^{1,*}

¹ Laboratory of Organic Synthesis, Extraction, and Valorization, Faculty of Sciences Ain Chock, Hassan II University, Casablanca 20000, Morocco

² Department of Chemistry, College of Science, King Saud University, P.O. Box 2455, Riyadh 11451, Saudi Arabia

³ Biochemistry Department, College of Science, King Saud University, P.O. Box 2245, Riyadh 11495, Saudi Arabia

⁴ Laboratory of Plant Biotechnology Applied to Crop Improvement, Faculty of Science of Sfax, University of Sfax, Sfax 3038, Tunisia

⁵ Department of Chemistry, Vidya Bharati Mahavidyalaya, Amravati 444 602, India

* Correspondence: zalmarhoon@ksu.edu.sa (Z.M.A.); rafik.saddik@gmail.com (R.S.)

Abstract: In this work, we are interested in finding new catalysts for catecholase, whose principle is based on the oxidation reaction of catechol to *o*-quinone. In this context, we have studied a series of seven quinoline-based compounds. The present work indicates that the complexes formed between seven selected quinoline compounds and the copper salts viz. Cu(OAc)₂, CuSO₄, Cu(NO₃)₂, and CuCl₂ elicit catalytic activities for the oxidation of catechol to *o*-quinone. The complexes formed with the Cu(OAc)₂ salt show a much higher catalytic activity than the others, whereas the Cu(NO₃)₂ and CuCl₂ salts formed complexes with low catalytic activity. This study also shows that the oxidation rate depends on two factors, namely the chemical structure of the ligands and the nature of the ions coordinated with the copper.

Keywords: catalytic activity; quinoline; catechol oxidase; *o*-quinone



Citation: Moutaouakil, M.; Tighadouini, S.; M. Almarhoon, Z.; I. Al-Zaben, M.; Ben Bacha, A.; H. Masand, V.; Jamaledine, J.; Saddik, R. Quinoline Derivatives with Different Functional Groups: Evaluation of Their Catecholase Activity. *Catalysts* **2022**, *12*, 1468. <https://doi.org/10.3390/catal12111468>

Academic Editor: László Poppe

Received: 16 October 2022

Accepted: 15 November 2022

Published: 18 November 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



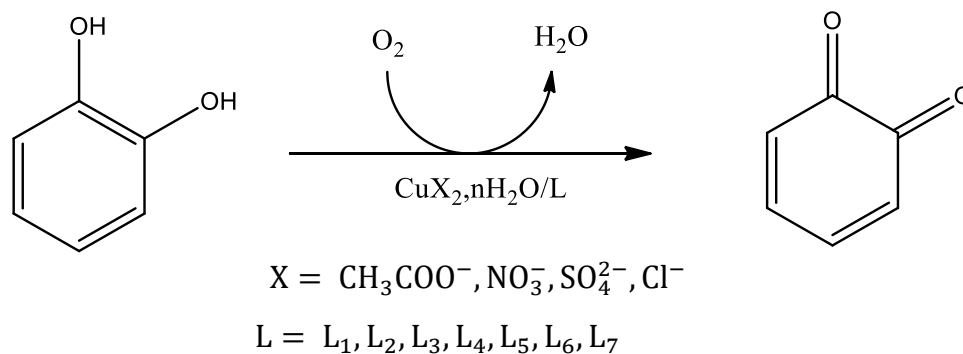
Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

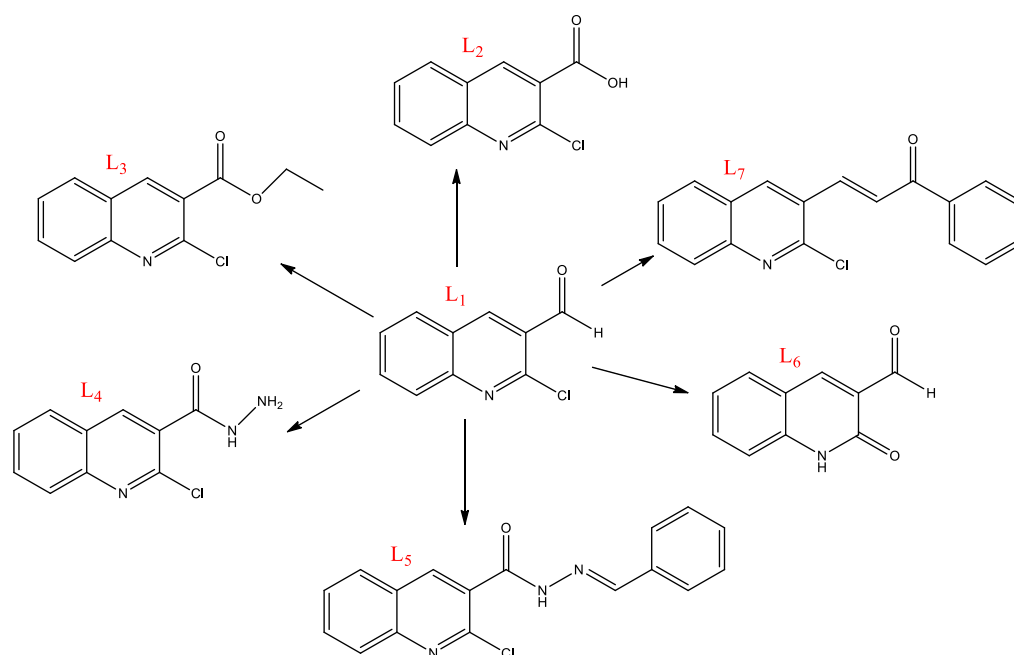
Copper is among the important metals in many catalytic processes and is characterized by its ability to combine with organic ligands to catalyze several diverse biological processes [1]. For example, catechol oxidase is a copper-based moiety that catalyzes the oxidation of phenols to quinones in the presence of oxygen. It is found for example in plants, wherein it plays an essential role in catalyzing the oxidation of catechol to *o*-quinone to produce after polymerization melanin, which gives a dark brown color to damaged fruits [2]. Dopamine is an important neurotransmitter, which after oxidation and polymerization gives polydopamine, an adhesive agent with many industrial applications. The oxidation of phenols to corresponding quinones is a very interesting process that has various applications in many fields [3].

Quinoline derivatives are among the compounds with great pharmacological powers [4], such as antimicrobial [5], anticancer [6–8], antifungal [9,10], antiviral [11], anti-inflammatory [12], antioxidants [13,14], antitumor [15], anti-SARS-CoV-2 [16], corrosion inhibitors [17,18], and antimalarial [19]. On the other hand, copper–quinoline complexes are similar systems that can catalyze many biological processes. For this reason, we are interested in this work to better understand this compelling mechanism and to discover the efficiency of quinoline derivatives in the oxidation of catechol.

This work aims to study the effect of ligands–copper(II) complexes on the oxidation of catechol to *o*-quinone in the presence of O₂ (Scheme 1). In this respect, the reaction was first performed without catalysts, then using the copper salts Cu(OAc)₂, CuSO₄, Cu(NO₃)₂, and CuCl₂, then using the synthesized ligands (L₁, L₂, L₃, L₄, L₅, L₆, and L₇—Scheme 2), and finally, the reaction was catalyzed by the ligands–copper (II) complexes. In order to be able to compare and discuss the obtained results, the oxidation rate for each catechol transformation was calculated.



Scheme 1. The oxidation reaction of catechol to *o*-quinone.



Scheme 2. Chemical structure of the studied quinoline ligands.

2. Results and Discussions

The results of this study are represented in Figures 1 and 2, which give the absorbance as a function of time for the different cases. Figure 1a presents the oxidation of catechol without any catalyst, while Figure 1b represents the reaction catalyzed by the synthesized ligands. In Figure 1c, the reaction is catalyzed by the metal salts and finally, Figure 2 provides the obtained results using the complexes formed between the synthesized ligands and the metal salts. Table 1 represents the oxidation rate for the different cases.

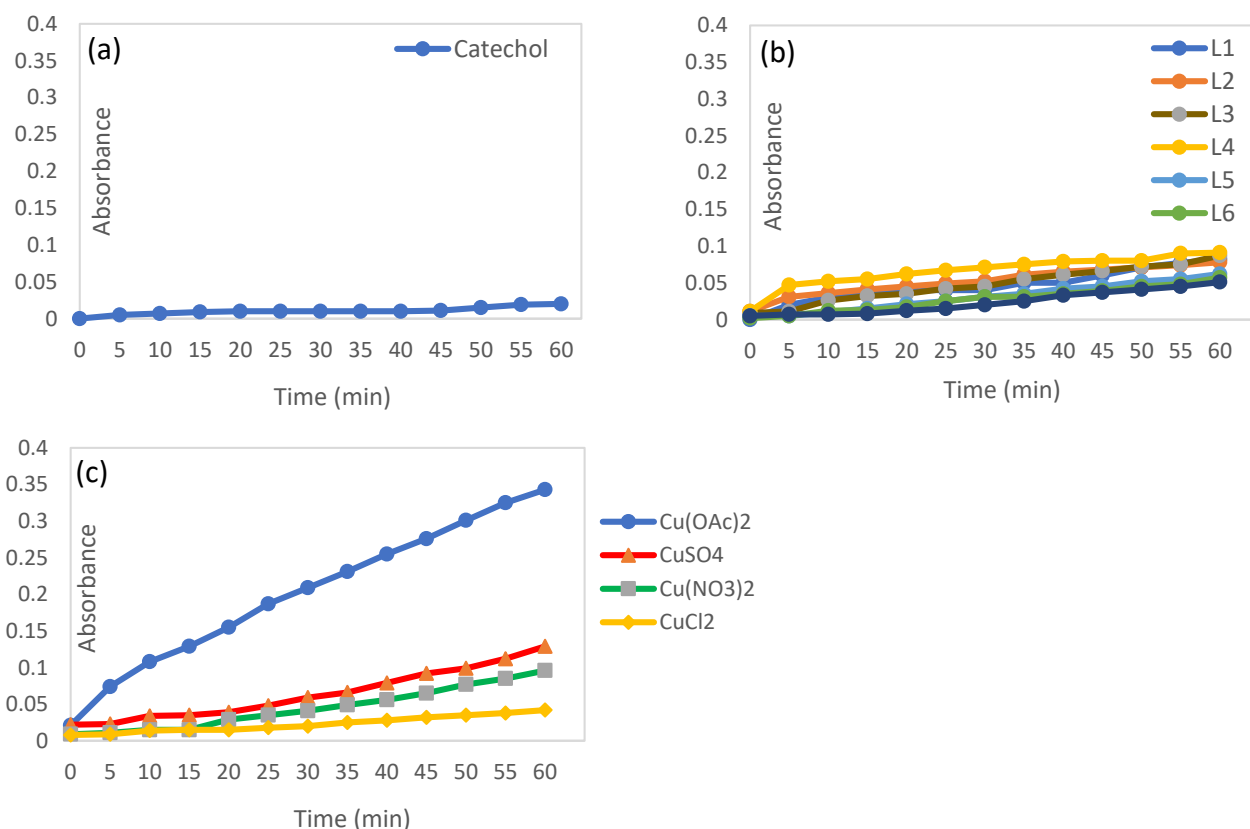


Figure 1. Oxidation of catechol to *o*-quinone in the absence of copper complexes: (a) Reaction without any catalyst, (b) reaction catalyzed by the synthesized ligands, and (c) reaction synthesized by the metal salts. Reaction conditions: methanol solutions, 0.15 mL metal salt at 2×10^{-3} mol/L, 0.15 mL ligand at 2×10^{-3} mol/L, 2 mL catechol at 10^{-1} mol/L are mixed, $T = 25^\circ\text{C}$, $\lambda = 390$ nm.

Table 1. Oxidation rate of catechol to *o*-quinone in ($\mu\text{mol L}^{-1} \text{s}^{-1}$).

	Cu(OAc) ₂	CuSO ₄	Cu(NO ₃) ₂	CuCl ₂	Ligands Only
L ₁	71.38	34.86	16.53	13.33	11.39
L ₂	94.30	26.25	18.61	11.25	10.83
L ₃	85.27	27.91	17.36	12.22	12.08
L ₄	126.80	65.13	31.25	15.55	12.64
L ₅	114.44	52.63	27.64	14.03	8.61
L ₆	69.30	26.53	14.03	11.39	7.92
L ₇	89.58	48.75	17.91	12.64	7.08
Salt only	47.63	17.91	13.33	5.83	
Without catalysis			2.78		

We notice from Figure 1a that the absorbance remains at almost zero over time and the oxidation rate is very low viz. $2.78 \mu\text{mol L}^{-1} \text{s}^{-1}$. Figure 1b also shows a very low absorbance and conversion rate between $7.08 \mu\text{mol L}^{-1} \text{s}^{-1}$ and $12.64 \mu\text{mol L}^{-1} \text{s}^{-1}$. From Figure 1c, we also notice that the absorbance and oxidation rate remain low for the metal salts CuSO₄, Cu(NO₃)₂, and CuCl₂; however, the salt Cu(OAc)₂ has a better oxidation rate of $47.22 \mu\text{mol L}^{-1} \text{s}^{-1}$. Therefore, it can be deduced that the oxidation reaction cannot take place without a catalyst and that neither ligands nor salts alone can catalyze this reaction, except Cu(OAc)₂, which has a slightly greater catalytic effect compared to the others.

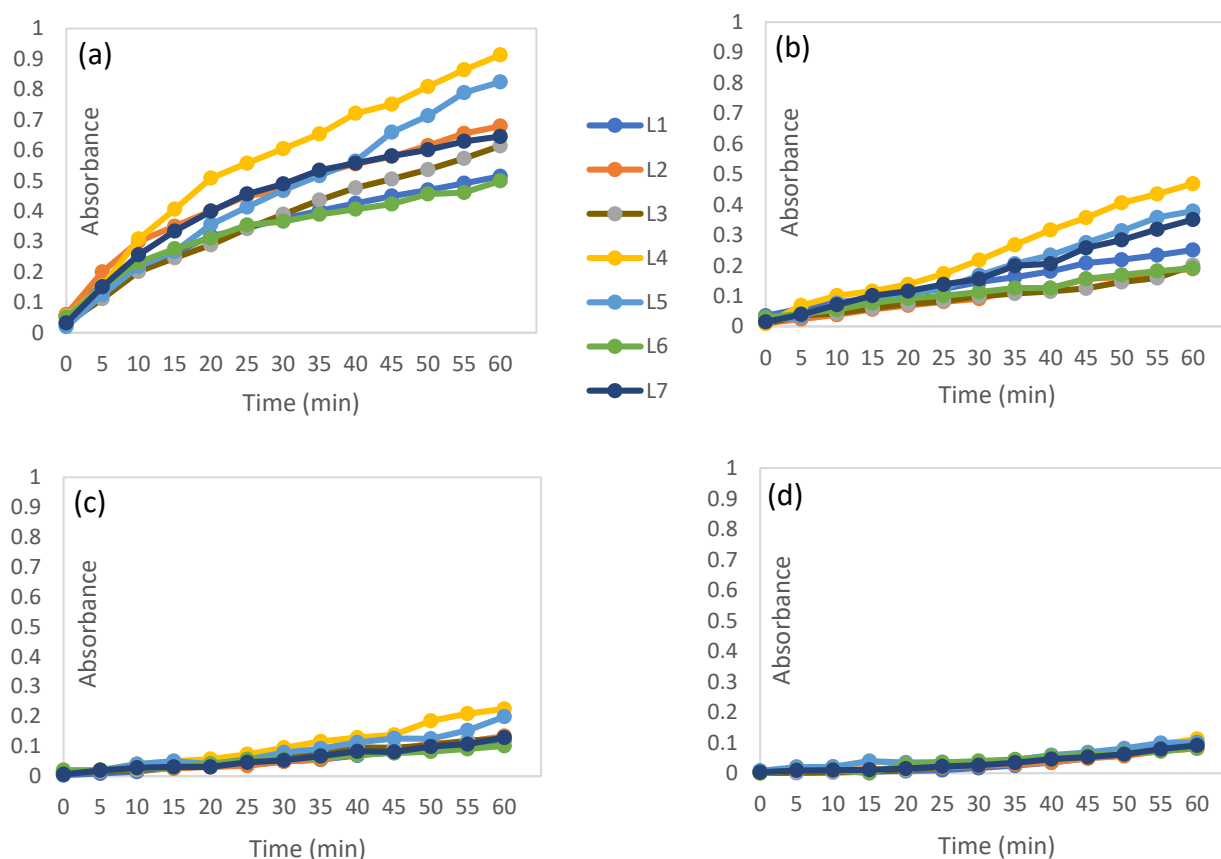


Figure 2. Oxidation of catechol to *o*-quinone in the presence of copper complexes: (a) Reaction in the presence of $\text{Cu}(\text{OAc})_2$ and ligands, (b) reaction in the presence of CuSO_4 and ligands, (c) reaction in the presence of $\text{Cu}(\text{NO}_3)_2$ and ligands, and (d) reaction in the presence of CuCl_2 and ligands. Reaction conditions: methanol solutions, 0.15 mL metal salt at 2×10^{-3} mol/L, 0.15 mL ligand at 2×10^{-3} mol/L, 2 mL catechol at 10^{-1} mol/L are mixed, $T = 25^\circ\text{C}$, $\lambda = 390$ nm.

The results obtained in Figure 2 show that the complexes formed between the $\text{Cu}(\text{OAc})_2$ salt and the ligands possess better catalytic activity for the oxidation of catechol, as the obtained oxidation rates are the highest (Figure 2a). The $\text{L}_4/\text{Cu}(\text{OAc})_2$ complex shows an oxidation rate of $126.80 \mu\text{mol L}^{-1} \text{s}^{-1}$, followed by the $\text{L}_5/\text{Cu}(\text{OAc})_2$ complex ($114.44 \mu\text{mol L}^{-1} \text{s}^{-1}$), then the $\text{L}_2/\text{Cu}(\text{OAc})_2$, $\text{L}_7/\text{Cu}(\text{OAc})_2$, and $\text{L}_3/\text{Cu}(\text{OAc})_2$ complexes with oxidation rates of 94.30 , 89.58 , and $85.27 \mu\text{mol L}^{-1} \text{s}^{-1}$, respectively, and finally, $\text{L}_1/\text{Cu}(\text{OAc})_2$ and $\text{L}_6/\text{Cu}(\text{OAc})_2$, which give the lowest oxidation rates (71.38 and $69.30 \mu\text{mol L}^{-1} \text{s}^{-1}$, respectively).

The complexes formed between the ligands and the CuSO_4 salt also give high oxidation rates but are generally lower than those obtained in the case of $\text{Cu}(\text{OAc})_2$, as the highest oxidation rate is $65.13 \mu\text{mol L}^{-1} \text{s}^{-1}$ for L_4/CuSO_4 , followed by L_5/CuSO_4 and L_7/CuSO_4 , with oxidation rates of 52.63 and $48.75 \mu\text{mol L}^{-1} \text{s}^{-1}$, respectively, and then the complexes formed between CuSO_4 and ligands L_1 , L_3 , L_6 , and L_2 come last with oxidation rates of 34.86 , 27.91 , 26.53 , and $26.25 \mu\text{mol L}^{-1} \text{s}^{-1}$, respectively (Figure 2b).

The complexes formed between the ligands and $\text{Cu}(\text{NO}_3)_2$ also catalyze the oxidation reaction, but with slightly lower oxidation rates compared to the previous ones, as the $\text{L}_3/\text{Cu}(\text{NO}_3)_2$ complex gives the largest oxidation rate of $31.25 \mu\text{mol L}^{-1} \text{s}^{-1}$, followed by the $\text{L}_5/\text{Cu}(\text{NO}_3)_2$ complex ($27.64 \mu\text{mol L}^{-1} \text{s}^{-1}$), and then the complexes formed between $\text{Cu}(\text{NO}_3)_2$ and the ligands L_2 , L_7 , L_3 , L_1 , and L_6 , with oxidation rates of 18.61 , 17.91 , 17.36 , 16.53 , and $14.03 \mu\text{mol L}^{-1} \text{s}^{-1}$, respectively (Figure 2c).

The complexes formed between the ligands and CuCl_2 do not show a significant catalytic effect, as there is no significant increase in absorbance in all cases, and the oxidation

rate remains quite low, the largest conversion rate being $15.55 \mu\text{mol L}^{-1} \text{s}^{-1}$ for L_4/CuCl_2 and the smallest being $11.25 \mu\text{mol L}^{-1} \text{s}^{-1}$ for L_2/CuCl_2 (Figure 2d). We deduce from these results that the chemical structure of the ligands and the nature of the copper salts play an important role in the catalytic activity of the studied complexes.

It can be seen from this study that all the complexes formed between the ligands and $\text{Cu}(\text{OAc})_2$ have very high catalytic activities compared to the complexes formed with the other salts, and this is due to the weak bonding between the OAc^- anions and the Cu^{2+} cations, which facilitates the coordination between the ligands and copper. On the other hand, the ligands L_4 and L_5 form complexes with a very high catalytic activity, because the presence of electron donor groups increases the electron density at the nitrogen atom, which favors the coordination with the metal and the formation of stable complexes. However, in the case of L_1 and L_6 , the catalytic activity decreases due to the presence of electron-withdrawing groups (Cl and CO) that weaken the electron density at the oxygen atom and thus disfavor the formation of the copper–metal bond. In the case of complexes formed between the ligands and the metal salt CuCl_2 , the absorbance remains low, and the catalytic activity decreases because the Cl^- anions are strongly bound to copper, and the coordination between the metal and the ligands thus becomes very difficult.

In summary, the catalytic activity of copper salts and quinoline ligands is very low, but their assembly results in complexes that efficiently catalyze the oxidation of catechol to *o*-quinone. The results show that the oxidation rate depends on two factors, namely the ions' nature, and the ligands' chemical structure. The ions strongly bound to copper reduce the coordination of the ligands, resulting in complexes of low catalytic activity, and the reverse is true for the ions weakly bound to copper, which facilitate the coordination of the ligands, giving stable complexes and high catalytic activities. On the other hand, the chemical structure of the ligands plays an essential role, and the presence of electron-donating groups enriches the coordination site in electron density, which increases the stability of the studied complex as well as its catalytic activity. However, the presence of electron-withdrawing groups decreases the electron density at the coordination site, which decreases the stability of the complex and its catalytic activity.

3. Materials and Methods

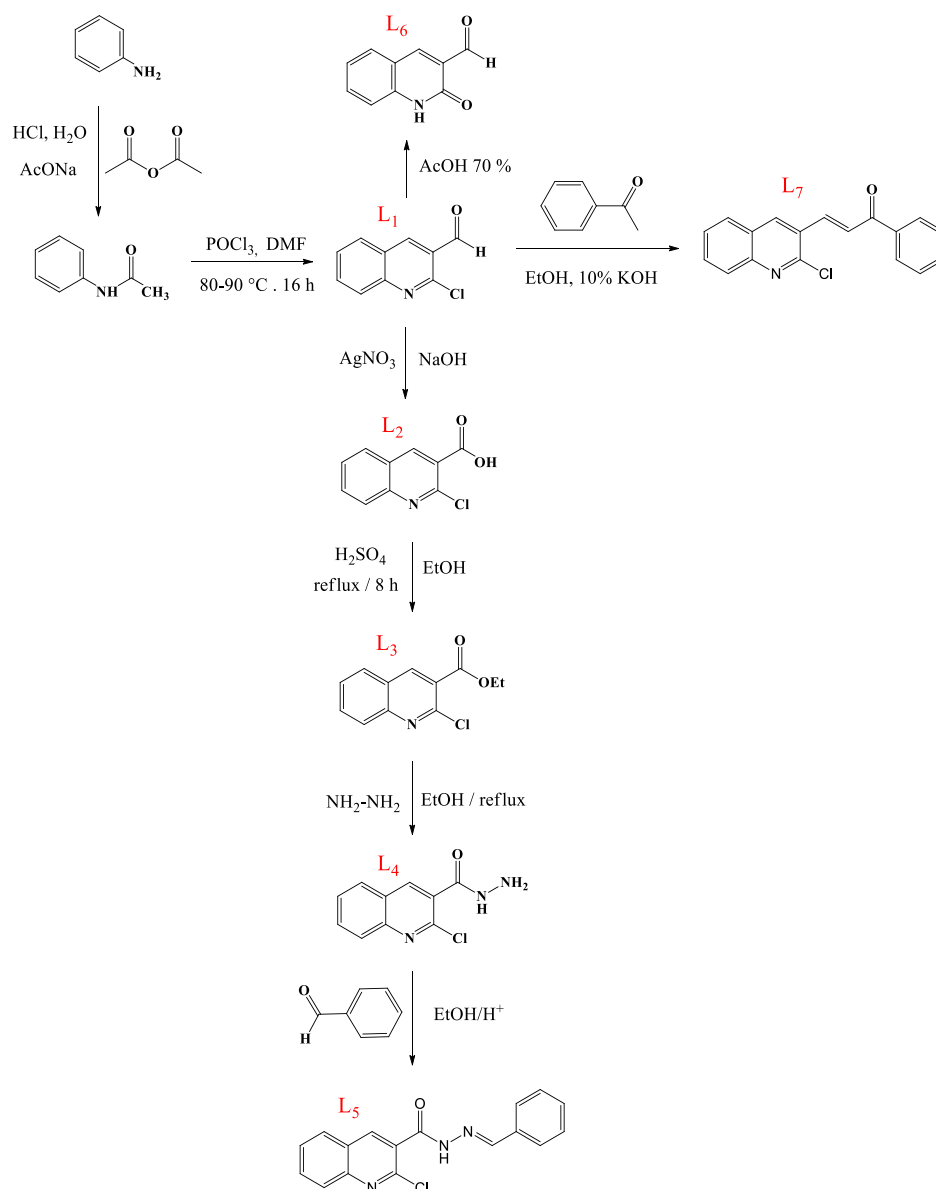
3.1. Reaction and Method

The reaction studied is schematized in Scheme 1, the kinetics of this reaction was followed by measuring the absorbance as a function of time by a UV–Vis spectrophotometer for one hour at 390 nm (absorption maximum of *o*-quinone), under the following conditions: $T = 25 \text{ }^\circ\text{C}$, $\epsilon = 1.6 \text{ L mol}^{-1} \text{ cm}^{-1}$. The solutions are prepared by dissolving in methanol, and the complexes are synthesized in situ by mixing 0.15 mL of a solution ($2 \times 10^{-3} \text{ mol L}^{-1}$) of CuX_2 , $n\text{H}_2\text{O}$, and 0.15 mL of a solution ($2 \times 10^{-3} \text{ mol L}^{-1}$) of the ligand, and then 2 mL of a solution of catechol at a concentration of $10^{-1} \text{ mol L}^{-1}$ is added [20].

3.2. Synthesis of Ligands

The studied quinoline derivatives were synthesized according to the procedures described in the literature (Scheme 3):

Synthesis of 2-chloroquinoline-3-carbaldehyde (compound L_1): DMF (3 eq) and POCl_3 (4.5 eq) were stirred for 30 min at $0 \text{ }^\circ\text{C}$, then acetanilide (1 eq) in CHCl_3 (15 mL) was added slowly, and after the addition, the reaction mixture was heated for 16 h ($80\text{--}90 \text{ }^\circ\text{C}$). When the reaction was complete, the mixture was poured into crushed ice and neutralized with saturated NaHCO_3 solution; the resulting precipitate was filtered, washed with water, and recrystallized in ethyl acetate [21].



Scheme 3. Protocol for the synthesis of quinoline derivatives.

Synthesis of 2-chloroquinoline-3-carboxylic acid (compound L_2): 89 mg of compound L_1 were dissolved in a minimum of ethanol, then an ethanolic solution of AgNO_3 (0.7 mmol) and NaOH (2.5 mmol) were added. The mixture was left under stirring at room temperature for 4 h; at the end of the reaction, the excess AgNO_3 was removed by filtration, then a few drops of concentrated sulfuric acid were added to neutralize the solution. The precipitate formed was filtered, washed with water, and dried to give a dark yellow product [22].

Synthesis of ethyl 2-chloroquinoline-3-carboxylate (compound L_3): In a minimum of ethanol, 100 mg of compound L_2 and 4 drops of concentrated sulfuric acid were added, and the mixture was refluxed for 8 h. After cooling, the solid formed was recovered by filtration, washed with water, dried, and recrystallized in ethanol [23].

Synthesis of 2-chloroquinoline-3-carbohydrazide (compound L_4): 0.5 mmol of compound L_3 and 0.5 mmol of hydrazine were heated at reflux for 4 h in a minimum of ethanol, and at the end of the reaction, 50 g of ice was added to the solution and the obtained precipitate was filtered, washed with water, dried, and recrystallized in ethanol [24].

Synthesis of *N'*-benzylidene-2-chloroquinoline-3-carbohydrazide (compound L₅): To 15 mL of ethanol, 56 mg of compound L₄ and 40 mg of benzaldehyde and a few drops of acetic acid were added and heated at reflux for 12 h. At the end of the reaction, the reaction mixture was allowed to reach room temperature and the resulting precipitate was filtered, washed with water, dried, and recrystallized in ethanol [25].

Synthesis of 2-oxo-1,2-dihydroquinoline-3-carbaldehyde (compound L₆): Compound L₁ (1 mmol) was heated to 110 °C for 12 h in acetic acid (70%), and at the end of the reaction, the reaction mixture was allowed to reach room temperature. The precipitate formed was filtered and washed with water, dried, and recrystallized in ethanol [26].

Synthesis of 3-(2-chloroquinolin-3-yl)-1-phenylprop-2-en-1-one (compound L₇): To 20 mL of ethanol, 1 mmol of acetophenone, 1 mmol of compound L₁, and 5 mL of 10% NaOH solution were added. After 8 h of stirring at room temperature, the precipitate formed was filtered, washed with water, and recrystallized in ethanol [27].

NMR data of all these ligands are available in the file Supplementary Materials.

4. Conclusions

The results of this study show that the studied quinoline-derived complexes possess catalytic activities, and in particular, the complexes formed between the ligands and the metal salt Cu(OAc)₂ efficiently catalyze the oxidation of catechol to *o*-quinone. The 2-chloroquinoline-3-carbohydrazide ligand (compound L₄) exhibits the highest catalytic activity, and 2-oxo-1,2-dihydroquinoline-3-carbaldehyde (compound L₆) exhibits the lowest catalytic activity. In general, the oxidation efficiency of the studied complexes depends on the ions' nature and the ligands' chemical structure. Ions weakly bound to the metal and electron-rich coordination sites yield stable complexes that strongly catalyze catechol oxidation. Further studies are still in progress in our laboratory to synthesize new quinoline derivatives, evaluate their biological and catalytic activities, as well as to obtain more details on this compelling catalytic process.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/catal12111468/s1>.

Author Contributions: R.S. and S.T.: conceptualization, supervision, project administration, methodology, resources, data curation, writing of the original draft, review, and editing; M.M.: conduct of experiment; V.H.M.: review and editing, methodology; J.J.: project administration; Z.M.A., M.I.A.-Z. and A.B.B.: project administration, methodology, writing of original draft and paying publication fees. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Acknowledgments: The authors extend their appreciation to the Deanship of Scientific Research at King Saud University for funding this work through research group No (RGP-070).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Trammell, R.; Rajabimoghadam, K.; Garcia-Bosch, I. Copper-promoted functionalization of organic molecules: From biologically relevant Cu/O₂ model systems to organometallic transformations. *Chem. Rev.* **2019**, *119*, 2954–3031. [CrossRef] [PubMed]
2. Tripathy, R.; Singha, S.; Sarkar, S. A review on bio-functional models of catechol oxidase probed by less explored first-row transition metals. *J. Coord. Chem.* **2022**, *8*, 1–51. [CrossRef]
3. Hauser, D.; Septiadi, D.; Turner, J.; Petri-Fink, A.; Rothen-Rutishauser, B. From bioinspired glue to medicine: Polydopamine as a biomedical material. *Materials* **2020**, *13*, 1730. [CrossRef] [PubMed]
4. Tabassum, R.; Ashfaq, M.; Oku, H. Current pharmaceutical aspects of synthetic quinoline derivatives. *Mini-Rev. Med. Chem.* **2021**, *21*, 1152–1172. [CrossRef] [PubMed]
5. Insuasty, D.; Vidal, O.; Bernal, A.; Marquez, E.; Guzman, J. Antimicrobial activity of quinoline-based hydroxyimidazolium hybrids. *Antibiotics* **2019**, *8*, 239. [CrossRef]

6. Guan, Y.F.; Liu, X.J.; Yuan, X.Y.; Liu, W.B.; Li, Y.R.; Yu, G.X.; Zhang, S.Y. Design, Synthesis, and Anticancer Activity Studies of Novel Quinoline-Chalcone Derivatives. *Molecules* **2021**, *26*, 4899. [[CrossRef](#)]
7. Diaconu, D.; Antoci, V.; Mangalagiu, V.; Amariuca-Mantu, D.; Mangalagiu, I.I. Quinoline-imidazole/benzimidazole derivatives as dual-/multi-targeting hybrids inhibitors with anticancer and antimicrobial activity. *Sci. Rep.* **2022**, *12*, 1–17. [[CrossRef](#)]
8. Zablotsky, D.; Segal, I.; Zablotskaya, A.; Maiorov, M.; Nguyen, T.A. 19—Antimicrobial Activity of Hybrid Organic-Inorganic Core-Shell Magnetic Nanocomposites. In *Magnetic Nanoparticle-Based Hybrid Materials*; Woodhead Publishing: Cambridge, UK, 2021; pp. 501–507. ISBN 978-0-12-823688-8. [[CrossRef](#)]
9. Ajani, O.O.; Iyaye, K.T.; Ademosun, O.T. Recent advances in chemistry and therapeutic potential of functionalized quinoline motifs—A review. *RSC Adv.* **2022**, *12*, 18594–18614. [[CrossRef](#)]
10. Qin, T.H.; Liu, J.C.; Zhang, J.Y.; Tang, L.X.; Ma, Y.N.; Yang, R. Synthesis and biological evaluation of new 2-substituted-4-amino-quinolines and-quinazoline as potential antifungal agents. *Bioorganic Med. Chem. Lett.* **2022**, *72*, 128877. [[CrossRef](#)]
11. Wyman, K.A.; Girgis, A.S.; Surapaneni, P.S.; Moore, J.M.; Abo Shama, N.M.; Mahmoud, S.H.; Panda, S.S. Synthesis of Potential Antiviral Agents for SARS-CoV-2 Using Molecular Hybridization Approach. *Molecules* **2022**, *27*, 5923. [[CrossRef](#)]
12. Ji, K.L.; Liu, W.; Yin, W.H.; Li, J.Y.; Yue, J.M. Quinoline alkaloids with anti-inflammatory activity from *Zanthoxylum avicennae*. *Org. Biomol. Chem.* **2022**, *20*, 4176–4182. [[CrossRef](#)] [[PubMed](#)]
13. Abdi, B.; Fekadu, M.; Zeleke, D.; Eswaramoorthy, R.; Melaku, Y. Synthesis and Evaluation of the Antibacterial and Antioxidant Activities of Some Novel Chloroquinoline Analogs. *J. Chem.* **2021**, *2021*, 2408006. [[CrossRef](#)]
14. El-Saghier, A.M.; El-Naggar, M.; Hussein, A.H.M.; El-Adasy, A.B.A.; Olish, M.; Abdelmonsef, A.H. Eco-Friendly Synthesis, Biological Evaluation, and In Silico Molecular Docking Approach of Some New Quinoline Derivatives as Potential Antioxidant and Antibacterial Agents. *Front. Chem.* **2021**, *397*, 679967. [[CrossRef](#)] [[PubMed](#)]
15. El Rhabori, S.; El Aissouq, A.; Chtita, S.; Khalil, F. Design of novel quinoline derivatives as anti-breast cancer using 3D-QSAR, molecular docking, and pharmacokinetic investigation. *Anti-Cancer Drugs* **2022**, *33*, 789–802. [[CrossRef](#)] [[PubMed](#)]
16. Gentile, D.; Fuochi, V.; Rescifina, A.; Furneri, P.M. New Anti SARS-CoV-2 Targets for Quinoline Derivatives Chloroquine and Hydroxychloroquine. *Int. J. Mol. Sci.* **2020**, *21*, 5856. [[CrossRef](#)]
17. Verma, C.; Quraishi, M.A.; Ebenso, E.E. Quinoline and its derivatives as corrosion inhibitors: A review. *Surf. Interfaces* **2020**, *21*, 100634. [[CrossRef](#)]
18. Alamiery, A. Investigations on corrosion inhibitory effect of newly quinoline derivative on mild steel in HCl solution complemented with antibacterial studies. *Biointerface Res. Appl. Chem.* **2022**, *2*, 1561–1568.
19. Da Silva Neto, G.J.; Silva, L.R.; Annunziato, Y. Dual quinoline-hybrid compounds with antimalarial activity against *Plasmodium falciparum* parasites. *New J. Chem.* **2022**, *46*, 6502–6518. [[CrossRef](#)]
20. Saddik, R.; Abridgach, F.; Benchat, N.; El Kadiri, S.; Hammouti, B.; Touzani, R. Catecholase Activity Investigation for Pyridazinone- and Thiopyridazinone-Based Ligands. *Res. Chem. Intermed.* **2012**, *38*, 1987–1998. [[CrossRef](#)]
21. Belferdi, F.; Merabet, N.; Belkhir, L.; Douara, B. Regioselective demethylation of quinoline derivatives. A DFT rationalization. *J. Mol. Struct.* **2016**, *1118*, 10–17. [[CrossRef](#)]
22. Syniugin, A.R.; Chekanov, M.O.; Volynets, G.P.; Starosyla, S.A.; Bdzhol, V.G.; Yarmoluk, S.M. Design, synthesis, and evaluation of 3-quinoline carboxylic acids as new inhibitors of protein kinase CK2. *J. Enzym. Inhib. Med. Chem.* **2016**, *31*, 160–169. [[CrossRef](#)] [[PubMed](#)]
23. Krishnakumar, V.; Khan, F.-R.N.; Mandal, B.K.; Mitta, S.; Dhasamandha, R.; Govindan, V.N. Quinoline-3-Carboxylates as Potential Antibacterial Agents. *Res. Chem. Intermed.* **2012**, *38*, 1819–1826. [[CrossRef](#)]
24. Abass, M.; Alzandi, A.R.A.; Hassan, M.M.; Mohamed, N. Recent Advances on Diversity Oriented Heterocycle Synthesis of Fused Quinolines and Its Biological Evaluation. *Polycycl. Aromat. Compd.* **2021**, *41*, 2120–2209. [[CrossRef](#)]
25. As, T.; Sal, S. Synthesis and Characterisation of Substituted Quinoline by Vilsmeier-Haack Reagent. *Int. J. Chem. Stud.* **2017**, *5*, 1–4.
26. Hamama, W.S.; Ibrahim, M.E.; Gooda, A.A.; Zoorob, H.H. Recent Advances in the Chemistry of 2-Chloroquinoline-3-Carbaldehyde and Related Analogs. *RSC Adv.* **2018**, *8*, 8484–8515. [[CrossRef](#)] [[PubMed](#)]
27. Li, W.; Xu, F.; Shuai, W.; Sun, H.; Yao, H.; Ma, C.; Xu, S.; Yao, H.; Zhu, Z.; Yang, D.-H.; et al. Discovery of Novel Quinoline-Chalcone Derivatives as Potent Antitumor Agents with Microtubule Polymerization Inhibitory Activity. *J. Med. Chem.* **2019**, *62*, 993–1013. [[CrossRef](#)]



Article

Pharmacophore Synergism in Diverse Scaffold Clinches in Aurora Kinase B

Vijay H. Masand^{1,*}, Sami A. Al-Hussain², Mithilesh M. Rathore¹, Sumer D. Thakur³, Siddhartha Akasapu⁴, Abdul Samad⁵ , Amal A. Al-Mutairi² and Magdi E. A. Zaki^{2,*}

¹ Department of Chemistry, Vidya Bharati Mahavidyalaya, Amravati 444602, Maharashtra, India

² Department of Chemistry, Faculty of Science, Imam Mohammad Ibn Saud Islamic University, Riyadh 11623, Saudi Arabia

³ Department of Chemistry, RDIK and NKD College, Badnera, Amravati 444701, Maharashtra, India

⁴ Curia Global, Springfield, MO 65807, USA

⁵ Department of Pharmaceutical Chemistry, Faculty of Pharmacy, Tishk International University, Erbil 44001, Iraq

* Correspondence: vijaymasand@gmail.com (V.H.M.); mezaki@imamu.edu.sa (M.E.A.Z.)

Abstract: Aurora kinase B (AKB) is a crucial signaling kinase with an important role in cell division. Therefore, inhibition of AKB is an attractive approach to the treatment of cancer. In the present work, extensive quantitative structure–activity relationships (QSAR) analysis has been performed using a set of 561 structurally diverse aurora kinase B inhibitors. The Organization for Economic Cooperation and Development (OECD) guidelines were used to develop a QSAR model that has high statistical performance ($R^2_{tr} = 0.815$, $Q^2_{LMO} = 0.808$, $R^2_{ex} = 0.814$, $CCC_{ex} = 0.899$). The seven-variable-based newly developed QSAR model has an excellent balance of external predictive ability (Predictive QSAR) and mechanistic interpretation (Mechanistic QSAR). The QSAR analysis successfully identifies not only the visible pharmacophoric features but also the hidden features. The analysis indicates that the lipophilic and polar groups—especially the H-bond capable groups—must be present at a specific distance from each other. Moreover, the ring nitrogen and ring carbon atoms play important roles in determining the inhibitory activity for AKB. The analysis effectively captures reported as well as unreported pharmacophoric features. The results of the present analysis are also supported by the reported crystal structures of inhibitors bound to AKB.

Keywords: aurora kinase B; QSAR; pharmacophore modeling



Citation: Masand, V.H.; Al-Hussain, S.A.; Rathore, M.M.; Thakur, S.D.; Akasapu, S.; Samad, A.; Al-Mutairi, A.A.; Zaki, M.E.A. Pharmacophore Synergism in Diverse Scaffold Clinches in Aurora Kinase B. *Int. J. Mol. Sci.* **2022**, *23*, 14527. <https://doi.org/10.3390/ijms232314527>

Academic Editors: Jesús Vicente de Julián-Ortiz, Gloria Castellano and Francisco Torrens

Received: 19 September 2022

Accepted: 11 November 2022

Published: 22 November 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The machinery for cell division, also known as mitosis, is completely regulated. Any irregularity or imperfect mitosis results in nondiploid DNA content, which ultimately causes cancer [1]. Researchers have therefore become interested in developing cancer chemotherapeutics that target centrosome maturation and separation, mitotic spindle assembly, chromosomal separation, and cytokinesis involving the participation of numerous important signaling kinases, including aurora, polo-like-kinase (Plk), and cyclin-dependent kinase (Cdk) [2,3]. The successful transition to mitosis depends on the aurora kinase family of serine/threonine kinases [4–7]. Since their discovery in 1995 and the initial detection of their expression in human cancer tissue in 1998 [2,5,7–9], these kinases have received a great deal of attention. This is due to their aberrant and excessive expression in a wide range of solid and liquid tumors, such as pancreatic, lung, liver, and breast tumors, as well as their oncogenic activity [2,4,5,7–11].

The aurora kinase family consists of three isoforms (A, B, and C), each of which differs in the length and amino acid composition of the N-terminal domain, but they share a common and conserved ATP binding site [2,12]. In order for the centrosome to mature, and for spindle assembly, meiosis, and metaphase spindle orientation to occur, aurora-A

is essential [2,12]. In order to achieve precise chromosomal segregation and cytokinesis, aurora kinase B (AKB) is required [2,12]. Massive polyploidization and failure to bio-orientate chromosomes result from AKB inhibition [2,12]. Since aurora kinase C (AKC), which complements the activity of AKB, has received less attention to date, we decided to focus only on AKB in this investigation, due to a lack of data for AKC [12].

Aurora kinases have been suggested as prospective targets for anticancer treatments due to their crucial function in controlling the cell cycle. At this time, none of the ATP-competitive inhibitors targeting AKB that are in clinical development (Figure 1) have been granted approval [4,5,13].



Figure 1. Structures of some known aurora inhibitors in different clinical trial stages.

In these conditions, a quick and effective strategy to find AKB inhibitors is still a key goal for medicinal chemists. To fulfill this goal, there is a need to use modern methods such as computer-aided drug design (CADD) to reduce time, costs, trial-and-error procedures, and other required resources [14,15]. The vibrant and developing field of CADD is successful due to the result-oriented performance of molecular docking, QSAR, and its other branches [14–16]. In QSAR, a mathematical model is created to connect chemical descriptors (structural features) to a desired bioactivity profile using a wide range of machine learning techniques [17,18]. In a more pragmatic sense, QSAR allows one to prioritize compounds with desirable attributes for a subsequent (and presumably successful) biological evaluation [17–19]. Traditional QSAR concentrates on producing statistically significant models [17–19]. Previously, different researchers have reported QSAR models for AKB using different techniques. For example, Neaz et al. [20] reported a 3D-QSAR model for a dataset of forty-eight quinazoline derivatives possessing other heterocyclic rings. The developed model had a leave-one-out cross-validated correlation coefficient (Q2LOO) of 0.56. Another 3D-QSAR and molecular docking study of azaindole derivatives as AKB inhibitors was accomplished by Lan and co-workers [21]. The best developed QSAR model based on forty-one molecules had Q2LOO = 0.575. Likewise, Ashraf et al. [22] used a dataset of 57 acylureidoinolin derivatives to develop a 3D-QSAR model, which had Q2LOO = 0.641, and indicated that electrostatic and hydrophobic fields determine the activity of compounds. Thus, AKB has been the subject of QSAR research; however, the developed QSAR models find little usage due to a lack of generalizability, low predictive power, being based on small datasets comprising limited scaffolds, or a combination of these factors. Therefore, there is a need to develop a robust and balanced QSAR model based on a larger dataset, encompassing diverse structural scaffolds. Consequently, in the present work, a QSAR model has been developed that possesses high external predictive ability and extensive mechanistic interpretations supported by X-ray-resolved structures.

2. Results

As stated in Section 1, the focus was on developing a genetic algorithm–multilinear regression (GA–MLR) model with a combination of mechanistic interpretation and high

predictive power. We have discovered several structural features in the current investigation. The recently constructed seven-parameter model and its statistical validation parameters are as follows.

Model A: $pIC_{50} = 4.611 (\pm 0.224) + 0.559 (\pm 0.105) \times \text{fringNplaN4B} + 0.436 (\pm 0.11) \times \text{fsp3Csp2N5B} + 0.253 (\pm 0.038) \times \text{N_H_2B} + 0.164 (\pm 0.035) \times \text{fsp2Osp2C5B} + 0.1 (\pm 0.015) \times \text{da_lipo_5B} - 0.317 (\pm 0.056) \times \text{fringNC6B} - 0.262 (\pm 0.048) \times \text{fOringC6B}$.

Statistical parameters associated with model A: $R^2_{tr} = 0.815$, $RMSE_{tr} = 0.468$, $MAE_{tr} = 0.401$, $CCC_{tr} = 0.898$, $s = 0.473$, $F = 277.836$, $R^2_{cv} (Q2LOO) = 0.808$, $RMSE_{cv} = 0.477$, $MAE_{cv} = 0.408$, $CCC_{cv} = 0.895$, $Q2LMO = 0.807$, $R^2_{Yscr} = 0.016$, $Q2Yscr = -0.02$, $RMSE_{ex} = 0.446$, $MAE_{ex} = 0.373$, $R^2_{ex} = 0.814$, $Q2-F1 = 0.811$, $Q2-F2 = 0.811$, $Q2-F3 = 0.833$, $CCC_{ex} = 0.900$.

Model A is statistically robust, as shown by the high values of various statistical parameters, such as the coefficient of determination (R^2_{tr}) and cross-validated coefficient of determination for leave-one-out (R^2_{cv} or $Q2LOO$), the external coefficient of determination (R^2_{ex}), $Q2-F_n$ and the Concordance Correlation Coefficient (CCC_{ex}), etc., and the low values of lack-of-fit (LOF), root mean square error ($RMSE_{tr}$), and mean absolute error (MAE). As a result, model A has high external predictive ability [23–30], is devoid of random correlations [31,32], and meets suggested threshold values for key parameters. The Supplementary Materials contain the formulae to determine these parameters. A Williams plot was used to evaluate the model's applicability domain [33–36]. As a result, it complies with all the OECD-recommended standards and requirements for developing a valuable QSAR model. Different graphs associated with model A are depicted in Figure 2.

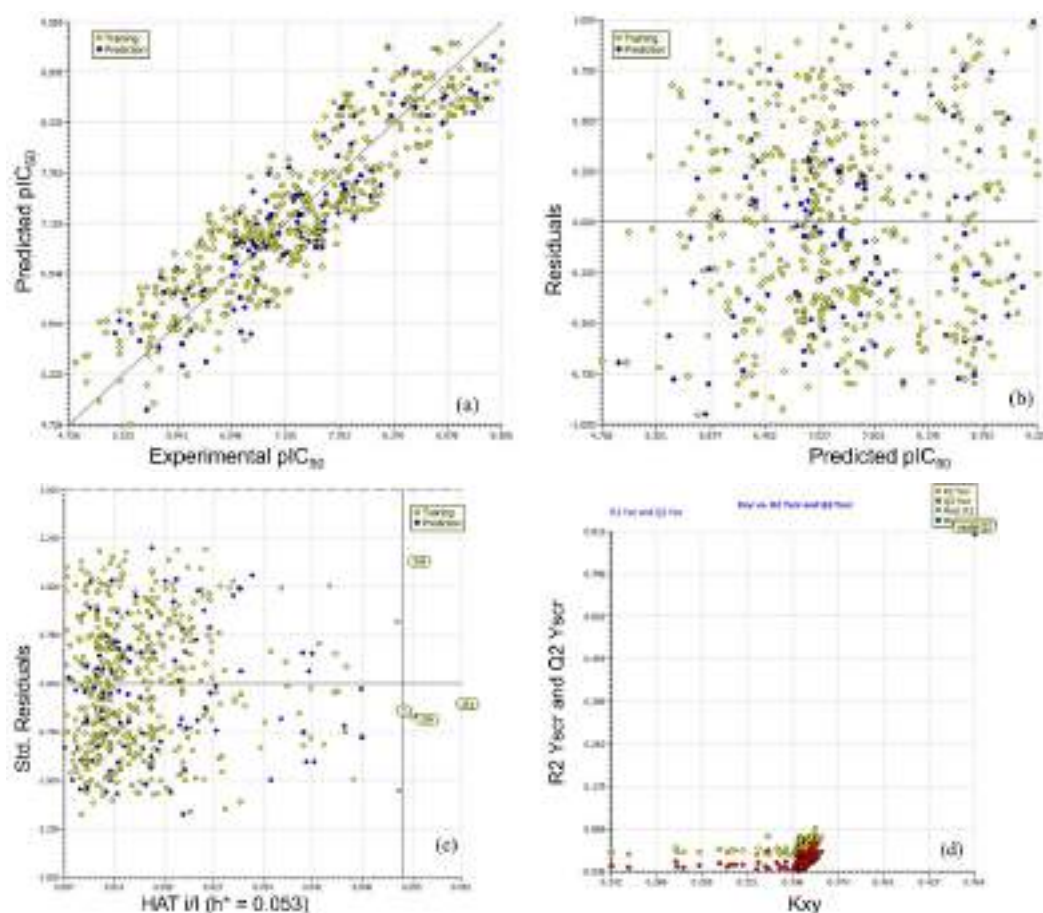


Figure 2. Different graphs related to model A: (a) experimental vs. predicted pIC_{50} (the solid line represents the regression line); (b) experimental vs. residuals; (c) Williams plot for applicability domain (the vertical solid line represents $h^* = 0.053$ and horizontal dashed lines represent the upper and lower boundaries for applicability domain); (d) Y-randomization plot.

There are seven descriptors in model A, which have been calculated by PyDescriptor [37] and tabulated in Table 1. Of the seven descriptors, five descriptors, viz. fringNplaN4B, fsp3Csp2N5B, N_H_2B, fsp2Osp2C5B, and da_lipo_5B, have positive coefficients in model A, implying that increasing their value could lead to a better activity profile, whereas the reverse is true for the remaining two descriptors, fOringC6B and fringNC6B, which have negative coefficients in model A. Each molecular descriptor, which is a numeric representation of structural features [37–39], has correlations with different types of pharmacophoric features, which govern the inhibitory profile. However, it is to be noted that a single structural feature can neither explain nor fully determine the final biological activity (IC₅₀) of a molecule. The biological activity IC₅₀, etc., is an outcome of a combination of different structural features and some unknown factors. Some features enhance the desired pharmacological activity, whereas others are responsible for reversing it. It is believed that two or more pharmacophoric groups concomitantly decide the biological activity (pharmacophore synergism).

Table 1. Different molecular descriptors present in model A and their descriptions.

Molecular Descriptor	Description
fringNplaN4B	Frequency of occurrence of planer nitrogen atoms exactly at 4 bonds from ring nitrogen atom
fsp3Csp2N5B	Frequency of occurrence of sp ² -hybridized nitrogen atoms exactly at 5 bonds from sp ³ -hybridized carbon atoms
N_H_2B	Total number of nitrogen atoms present within 2 bonds from hydrogen atoms
fsp2Osp2C5B	Frequency of occurrence of sp ² -hybridized carbon atoms exactly at 5 bonds from sp ² -hybridized oxygen atoms
da_lipo_5B	Total number of lipophilic atoms present within 5 bonds from H-bond donor cum acceptor atoms
fOringC6B	Frequency of occurrence of ring carbon atoms exactly at 6 bonds from oxygen atoms
fringNC6B	Frequency of occurrence of carbon atoms exactly at 6 bonds from ring nitrogen atoms

3. Discussion

Of the seven descriptors in model A, five descriptors, viz. fringNplaN4B, fsp3Csp2N5B, N_H_2B, da_lipo_5B, and fringNC6B, indicate the importance of different types of nitrogen atoms in determining the inhibitory activity for aurora kinase B. The same is true for carbon, which is present in four descriptors, viz. fsp3Csp2N5B, da_lipo_5B, fringNC6B, and fOringC6B. The relevance of oxygen is due to its presence in three descriptors, viz. fsp2Osp2C5B, da_lipo_5B, and fOringC6B. At the same time, it should be noted that the descriptors present in model A are highly interlinked; that is, increasing the value of one descriptor could significantly change the value of another descriptor. This leads to substantial changes in the biological profile of a molecule, pointing toward pharmacophore synergism, as molecular descriptors are mathematical representations of pharmacophores. For example, the values of descriptors fringNplaN4B and fringNC6B vary with the presence/absence of ring nitrogen atoms. Therefore, increasing the value of fringNplaN4B by escalating ring nitrogen atoms could also lead to a higher fringNC6B value. Therefore, in the present work, we have adopted an approach that involves the concomitant consideration of two or more molecular descriptors to explain the variance in the activity profile of matched molecular pairs (MMP). Accordingly, the molecular descriptors whose values have changed for MMP have been discussed concurrently with relevant examples in Section 3.

da_lipo_5B:

The descriptor da_lipo_5B is simultaneously associated with two important aspects of a molecule: lipophilic character and H-bonding-capable (donor and acceptor) atoms. It is to be noted that, in the present work, a carbon atom is non-lipophilic while calculating da_lipo_5B, if oxygen or nitrogen is attached to it. The average value of da_lipo_5B for the top one hundred active molecules (IC₅₀ = 0.26 to 4.3 nM) is 15.29, and the value for the

least active one hundred molecules ($IC_{50} = 611$ to $16,000$ nM) is 8.51. This reveals that the higher the number of lipophilic atoms within five bonds of a H-bond-capable atom, the higher the activity. This gives an initial impression that lipophilicity (mostly represented by $\log P$ [40]) is the only governing factor. However, the calculated $\log P$ ($clogP$), which represents molecular lipophilicity, has a weak correlation of 0.077 with pIC_{50} , whereas da_lipo_5B has a value of 0.533. Therefore, the conditional occurrence of lipophilic atoms in the vicinity of H-bonding-capable atoms is a better choice. A plausible reason could be the composition of the active site of AKB, which consists of the persistent presence of lipophilic residues such as Gly, Leu, Val, Phe, etc., between the acidic or basic residues such as Glu, Asp, Lys [22]. This is why an aurora kinase B inhibitor also requires the presence of H-bond-capable atoms, preferably with separation by five bonds and the concomitant occurrence of lipophilic atoms in their vicinity. This observation is confirmed by the reported X-ray-resolved structure of aurora kinase B (pdb: 4c2w [41]) (see Figure 3).

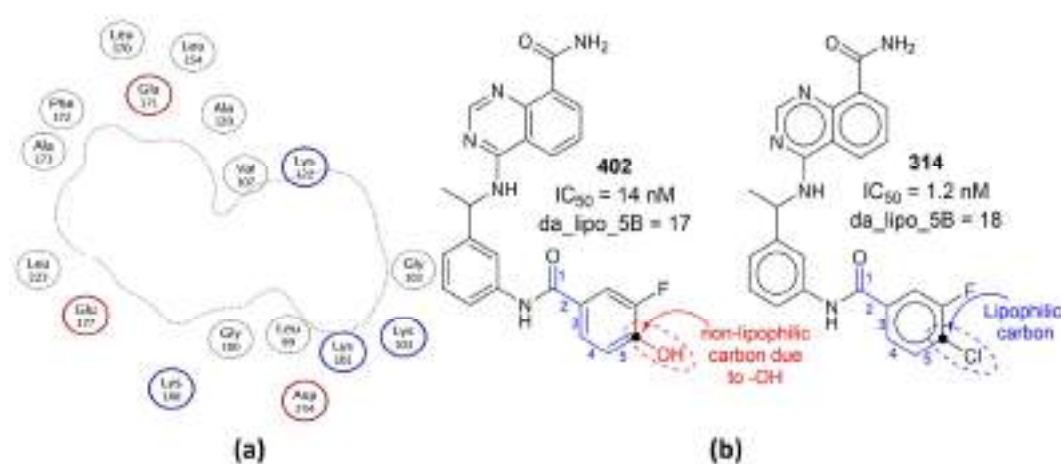


Figure 3. (a) A 2D depiction of active site of aurora kinase B (pdb: 4c2w). The dotted line represents the contour proximity of active site residues. Acidic and basic residues have been highlighted using red- and blue-colored circles. (b) Comparison of molecule 402 with 314 with respect to da_lipo_5B (blue-colored bonds and numbering).

The importance of da_lipo_5B highlights the significance of determining the numbers of donor cum acceptor atoms required to obtain better activity. The average value of donor cum acceptor atoms for the top one hundred active molecules ($IC_{50} = 0.26$ to 4.3 nM) is 3.21, and the value for the least active one hundred molecules ($IC_{50} = 611$ to $16,000$ nM) is 2.24. A comparison of the following pairs of molecules as representative examples further highlights the importance of da_lipo_5B : 314 with 402 (see Figure 3), 355 with 347, 206 with 207, 103 with 101, 103 with 99, 61 with 142, 57 with 58, etc.

fringNplaN4B:

fringNplaN4B stands for the frequency of occurrence of planer nitrogen atoms exactly at four bonds from a ring nitrogen atom. If the same planer nitrogen atom is also present at ≤ 4 bonds from the same or any other ring nitrogen atom through any path, then it is excluded while calculating fringNplaN4B. The importance of fringNplaN4B is reflected by the fact that the most active 110 molecules with IC_{50} values ranging from 0.26 to 5.9 nM have one or more combinations of planer and ring nitrogen atoms. The reverse is true for less active molecules ($IC_{50} = 16,000$ to 611 nM), with some exceptions, such as molecule numbers 213, 73, 71, 66, 20, etc. Moreover, it was observed that replacing fringNplaN4B with its corresponding equivalents, fringNplaN3B and fringNplaN5B, for three and five bonds led to a reduction in the performance of model A ($R^2 = 0.770$, for both). Moreover, fringNplaN3B and fringNplaN5B have a correlation of $R = 0.084$ and 0.028 with pIC_{50} , respectively, whereas fringNplaN4B is a better choice as a descriptor, with $R = 0.628$.

However, at first sight, it appears that, individually, ringN (number of ring nitrogen atoms) or nplanN (number of planer nitrogen atoms) could be an alternative to fringNplaN4B. However, both have a weak correlation of 0.207 and 0.374 with pIC50, respectively. Moreover, a loss in the statistical performance of model A on replacing fringNplaN4B with ringN ($R^2 = 0.772$) or nplanN ($R^2 = 0.770$) again confirmed the importance of fringNplaN4B. Therefore, a combination of ring and planer nitrogen atoms separated exactly by four bonds is an important structural feature to obtain a better pIC50 for AKB.

A literature survey reveals that for pyrrolopyrazole derivatives, a substituted 3-aminopyrazole moiety is important due to its ability to interact with the hinge region of the ATP binding site [2]. The three nitrogen atoms of the N-C-N-N pattern present in 3-aminopyrazole are responsible for binding with the receptor [2]. Unfortunately, it appears that the reported pattern is exclusive to pyrrolopyrazole derivatives bearing a substituted 3-aminopyrazole moiety. Interestingly, the terminal nitrogen atoms of the N-C-N-N pattern are actually ring and planer nitrogen atoms, thereby suggesting the possible presence of fringNplaN4B. However, in many active molecules of the present dataset bearing a substituted 3-aminopyrazole moiety, the value of fringNplaN4B is zero; this is because the planer nitrogen of the N-C-N-N pattern is also present within ≤ 4 bonds of the other ring nitrogen atom. However, in several active molecules for AKB, fringNplaN4B is present due to other scaffolds (see Figure 4). In other words, instead of the N-C-N-N pattern or a substituted 3-aminopyrazole moiety, an emphasis on the simultaneous presence of planer and ring nitrogen atoms separated by four bonds in the molecule is a better strategy to enhance the inhibitory profile against AKB. Hence, the present work successfully identified a novel aspect of a reported pattern (N-C-N-N) and extended it for other scaffolds.

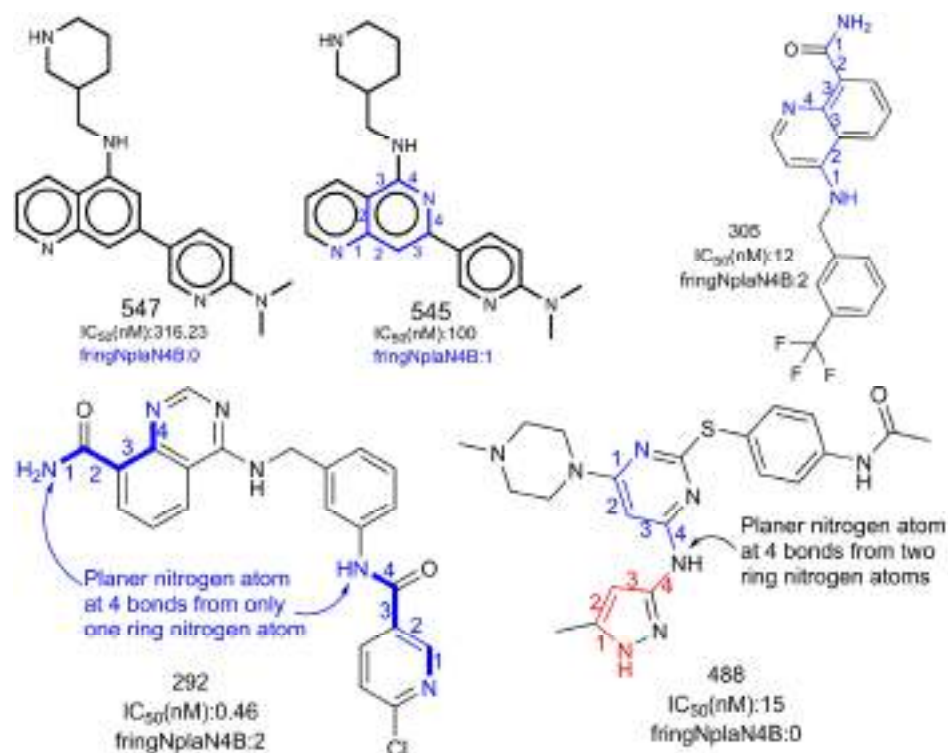


Figure 4. Representation of influence of fringNplaN4B on activity profile of AKB inhibitors. The numbers (blue/red) indicate the counting of number of bonds between ring and planer nitrogen.

N_H_2B:

The positive coefficient for N_H_2B indicates that the presence of hydrogen in the vicinity of nitrogen is beneficial to increase the inhibitory activity for aurora kinase B. In many molecules, N_H_2B exists due to the direct attachment of a hydrogen atom to a nitrogen atom (N-H) or due to hydrogen atoms bonded to carbon atoms adjacent to

nitrogen (N-CH_n fragment). N_H_2B favors two important structural features that could lead to a better inhibitory profile: (1) the presence of polar hydrogen atoms as N-H or N-CH_n fragments; (2) steric hindrance or bulkiness in the vicinity of nitrogen atoms, because hydrogen is the smallest among all the elements. The lesser the bulkiness around nitrogen atoms, the better the inhibitory profile. These two structural features in combination allow the polar interactions or H-bond formation between the ligand and the receptor. This observation, and the significance of N_H_2B as well as da_lipo_5B, is confirmed by the two forms of the ligand VX-680 (molecule number 14) in the pdb 4b8m [42].

The ligand VX-680 exists in two different forms, labeled as TA and TB in the present work, in the two chains of pdb 4b8m. From Figure 5 and Table 2, it is clear that the TA form consists of a higher number of hydrogen atoms than TB, especially in the vicinity of nitrogen atoms. This led to different values for N_H_2B for the two forms (see Figure 5). The form TA, having a higher N_H_2B value, has a higher number of interactions with the receptor, because the additional hydrogen atoms attached to the nitrogen atoms of the pyrazole (designated as N19 and N20) ring and aminopyrimidine (designated as N14) are responsible for H-bond interactions with Glu171, Phe172, and Ala173 (see Table 2). Meanwhile, these interactions are absent for TB, even though the respective atoms N19 and N14 of TB are more proximate to receptor atoms. The TB form has only one prominent interaction with the receptor due to the nitrogen (designated as N20) of the pyrazole ring in the form of a H-bond with Ala173.

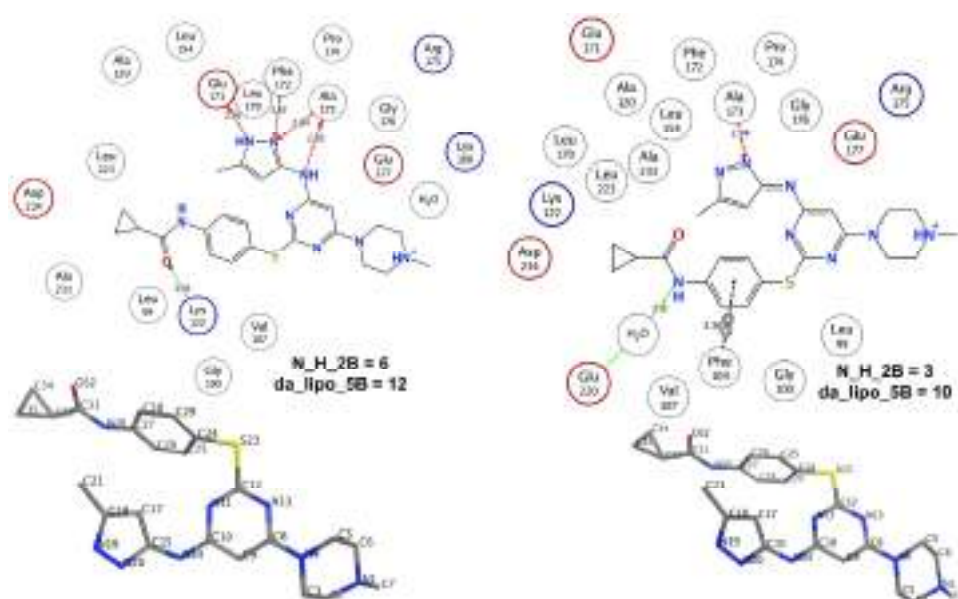


Figure 5. Pictorial representation of N_H_2B using VX-680 (pdb 4b8m) as an example.

Table 2. Distances of different atoms of TA and TB forms of VX-680 (molecule number 14) from the receptor atoms (pdb 4b8m).

TA Form				TB Form			
Residue	Residue Atom	Ligand Atom	Distance	Residue	Residue Atom	Ligand Atom	Distance
GLU171	O	N19	2.97	GLU171	O	N19	2.74
PHE172	CA	N20	3.47	PHE172	CA	N20	3.52
ALA173	N	N20	2.84	ALA173	N	N20	2.74
ALA173	O	N14	2.93	ALA173	O	N14	2.91
HOH2005	O	N13	3.32	HOH2005	O	N30	2.80

The following comparisons of molecules further highlight the importance of N_H_2B (see Figure 6): 108 with 75 and 101, 486 with 487 and 484, and 148 with 144, to list a few. A simple analysis of these examples indicates that the presence of a pyrazole ring leads to a better IC₅₀ for a molecule (see Figure 6). However, it has a negative correlation ($R = -0.177$) with pIC₅₀. A plausible reason appears from the present work suggesting that H-bond-capable polar groups are more suitable near the periphery of a molecule, rather than a pyrazole ring, to achieve good interactions with the receptor.

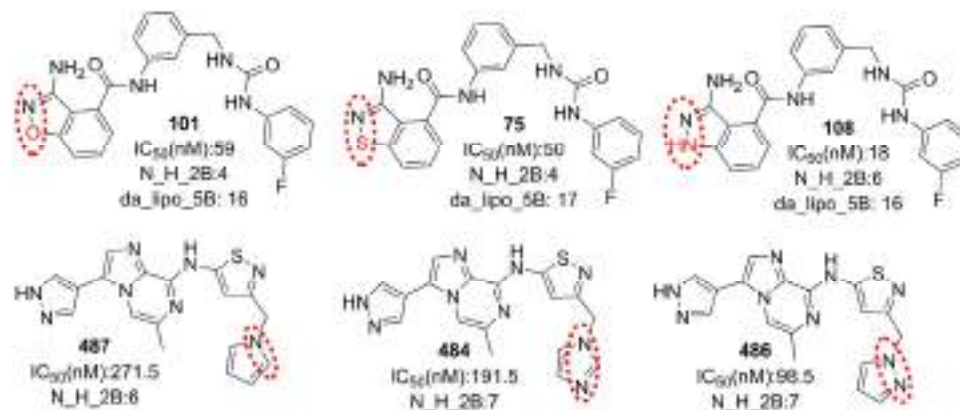


Figure 6. Representative examples to understand N_H_2B.

fsp3Csp2N5B:

The descriptor fsp3Csp2N5B is associated with two features, viz. sp²-hybridized nitrogen and sp³-hybridized carbon atoms. As it has a positive coefficient in model 1, increasing the numbers of such atoms favors the augmentation of pIC₅₀. At the same time, increasing fsp3Csp2N5B could influence the values of da_lipo_5B and N_H_2B, as these descriptors are associated with carbon and nitrogen too. Therefore, it indicates that pharmacophore synergism determines the final inhibitory ability of a molecule for AKB. This is clearly reflected when molecule 435 is compared with molecule 438.

The pdb 4c2v contains two different tautomeric forms of ligand YJA in two different chains, A and B. The influence of fsp3Csp2N5B along with N_H_2B is observed for the two tautomeric forms of co-crystallized ligand 'YJA' in the pdb 4c2v [41]. The two tautomeric forms show that YJA-T1 and YJA-T2 (see Figure 7) of ligand YJA have different values for fsp3Csp2N5B and N_H_2B (see Table 3). The online tautomer generator from Chemaxon (<https://disco.chemaxon.com/calculators/demo/plugins/tautomers/>, accessed on 28 October 2022) indicates that the ligand YJA can exist in seven different tautomeric forms. However, only two tautomeric forms, YJA-T1 and YJA-T2, predominate, with approximately 16 and 84 percent, respectively. The rest of the tautomeric forms have less than a 0.1% probability of existence.

A comparison of the interactions of YJA-T1 and YJA-T2 with the receptor and the solvent indicates that the two forms have established H-bonds with the similar amino acid residues of the receptor but with different distances (see Figure 8). The YJA-T2 has an additional H-bond with the solvent (HOH2108). Moreover, it has a higher number of interactions with the receptor and the solvent (H₂O) within 5 Å compared to YJA-T1. Thus, the increased value of fsp3Csp2N5B and N_H_2B for these two tautomeric forms correlates with a higher number of receptor atoms in the vicinity, which ultimately leads to an augmented number of interactions. Additional details related to the interactions of YJA-T1 and YJA-T2 with the receptor are available in Table S1 in the Supplementary Materials.

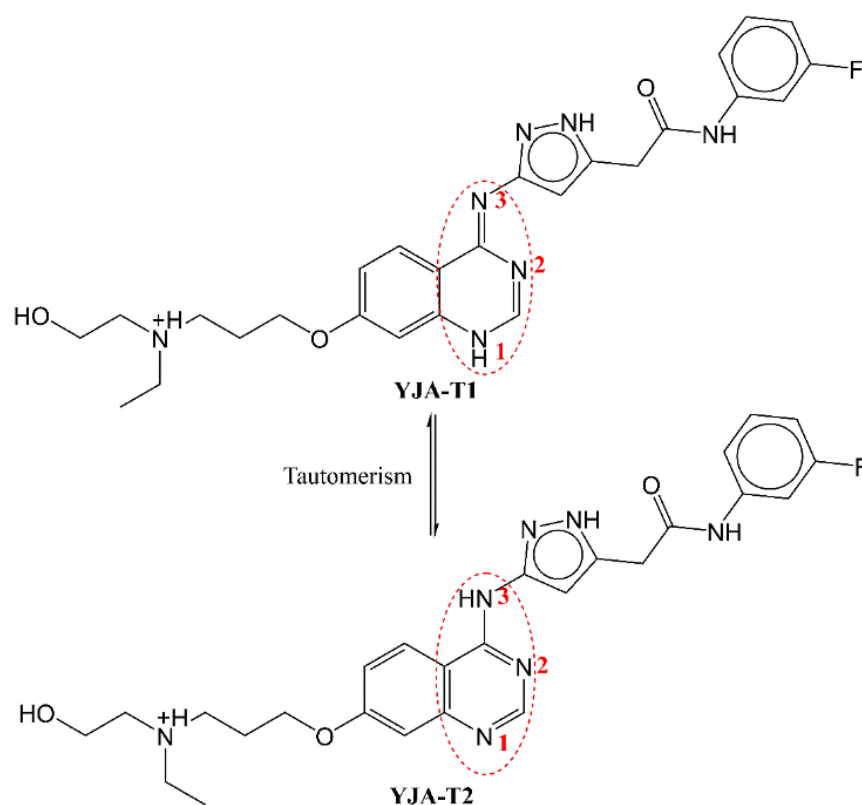


Figure 7. Tautomeric forms of ligand YJA (pdb 4c2v). The red colored numbers have been used for indication of nitrogen atoms involved in tautomerism.

Table 3. A comparison of two tautomeric forms, YJA-T1 and YJA-T2.

Tautomer with Descriptor Value	H-Bonds Formed with Distance (Å) with Angle (Donor-Hydrogen-Acceptor) (Cut-Off: 5 Å)	List of Receptor Heavy Atoms within 5 Å of N3 atom of Ligand (Residue-Atom-Distance in Å)	List of Receptor Heavy Atoms within 5 Å of N1 Atom of Ligand (Residue-Atom-Distance in Å)
YJA-T1 fsp3Csp2N5B = 0 N_H_2B = 6 fsp2Osp2C5B = 3	LYS122 at 1.668 with 159.8°, GLN145 at 2.251 with 142.4°, ALA173 at 1.952 with 163.9°	VAL107-CB-4.672, VAL107-CG1-4.351, VAL107-CG2-4.419, LU177-OE2-4.842, LEU223-CG-4.608, LEU223-CD1-3.627, LEU223-CD2-4.406	LEU99-CD1-4.259, ALA120-CB-4.501, GLU171-C-4.888, GLU171-O-4.058, PHE172-N-4.808, PHE172-CA-3.818, PHE172-C-3.832, PHE172-CB-4.641, PHE172-CG-4.403, PHE172-CD1-3.550, PHE172-CE1-4.156, ALA173-N-2.936, ALA173-CA-3.743, ALA173-C-4.208, ALA173-O-3.930, ALA173-CB-3.623, LEU223-CD1-4.121

Table 3. Cont.

Tautomer with Descriptor Value	H-Bonds Formed with Distance (Å) with Angle (Donor–Hydrogen–Acceptor) (Cut-Off: 5 Å)	List of Receptor Heavy Atoms within 5 Å of N3 atom of Ligand (Residue–Atom–Distance in Å)	List of Receptor Heavy Atoms within 5 Å of N1 Atom of Ligand (Residue–Atom–Distance in Å)
YJA-T2 fsp3Csp2N5B = 1 N_H_2B = 7 fsp2Osp2C5B = 3	LYS122 at 2.361 with 157.8°, GLN145 at 2.323 with 115.7°, ALA173 at 1.946 with 174.4°, HOH2108 2.222 with 106.7°	PHE104-CG-4.358, PHE104-CD2-3.203, PHE104-CE2-3.058, PHE104-CZ-4.124, VAL107-CB-4.591, VAL107-CG1-4.413, VAL107-CG2-4.142, LEU223-CD1-4.047, LEU223-CD2-4.948	LEU99-CD2-3.977, ALA120-CB-4.707, GLU171-C-4.734, GLU171-O-3.872, PHE172-N-4.690, PHE172-CA-3.669, PHE172-C-3.814, PHE172-CB-4.567, PHE172-CG-4.418, PHE172-CD1-3.618, PHE172-CE1-4.265, ALA173-N-2.953, ALA173-CA-3.799, ALA173-C-4.271, ALA173-O-3.915, ALA173-CB-3.635, LEU223-CD1-4.165

fsp2Osp2C5B:

The molecular descriptor fsp2Osp2C5B underlines the influence of a specific combination of sp²-hybridized carbon with sp²-hybridized oxygen in determining the inhibitory profile for AKB. The positive coefficient for fsp2Osp2C5B indicates that increasing such a combination of oxygen and carbon could lead to a better inhibitory profile. In the present dataset, there are 426 molecules with the presence of at least one such combination of oxygen and carbon. Likewise, the 200 most active molecules with IC₅₀ values in the range of 0.26 to 24 nM, except molecule numbers 36 and 469, also possess fsp2Osp2C5B >1. A comparison of molecule number 167 with 168 further strengthens this observation (see Figure 9).

A closer analysis revealed that the sp²-hybridized carbon with sp²-hybridized oxygen, required for the existence of fsp2Osp2C5B are, in general, aromatic carbon atoms and oxygen of the carbonyl group, especially the amide group, respectively. This further highlights the importance of aromatic rings—and in turn lipophilic atoms—as aromatic carbons are mostly lipophilic in nature. The need for an amide group in conjugation points out the necessity of a polar group to enhance the interactions with the receptor. The two tautomeric forms of YJA-T1 and T2 possess such a combination and it results in enhanced interactions with the receptor (see Figure 8). Obviously, a sp²-hybridized carbon atom will be at a respective distance of three and five bonds from the nitrogen and oxygen atoms of the same amide group; therefore, we also checked the importance of famdNsp2C3B (frequency of occurrence of sp²-hybridized carbon atoms exactly at three bonds from amide nitrogen atoms). It was observed that fsp2Osp2C5B and famdNsp2C3B have a correlation of 0.64 and 0.58, respectively, with pIC₅₀. Therefore, fsp2Osp2C5B is a better choice to be considered for future optimizations and activity predictions.

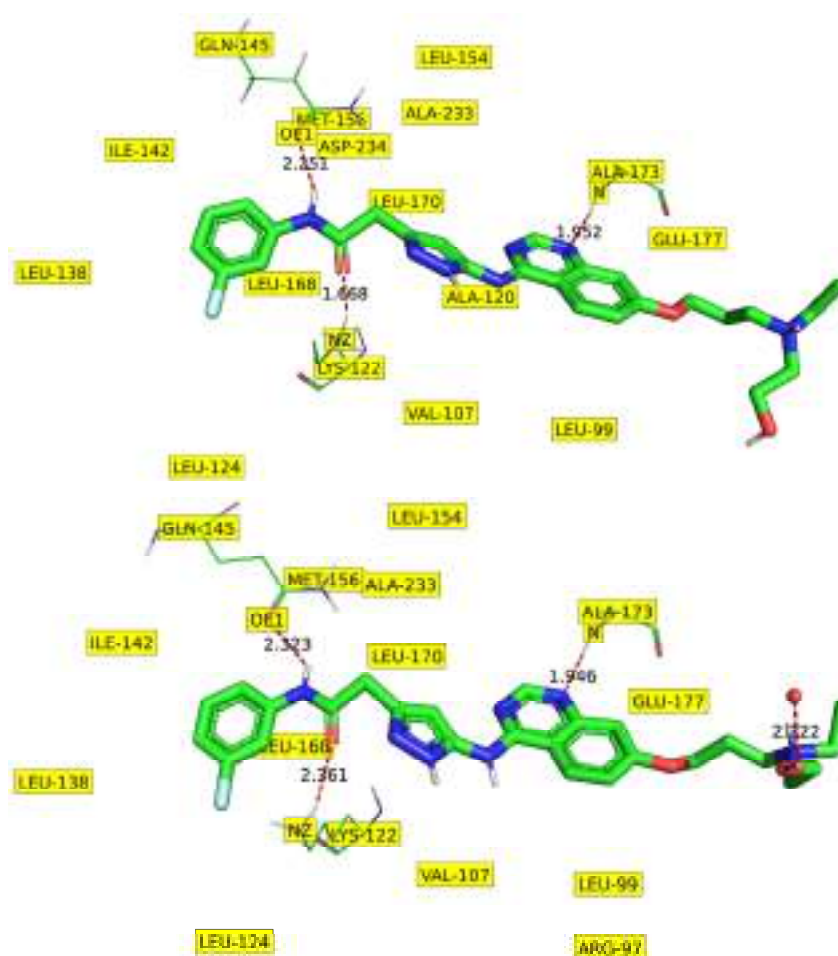


Figure 8. Depiction of prominent interactions of YJA-T1 and T2 with the receptor (pdb: 4c2v).

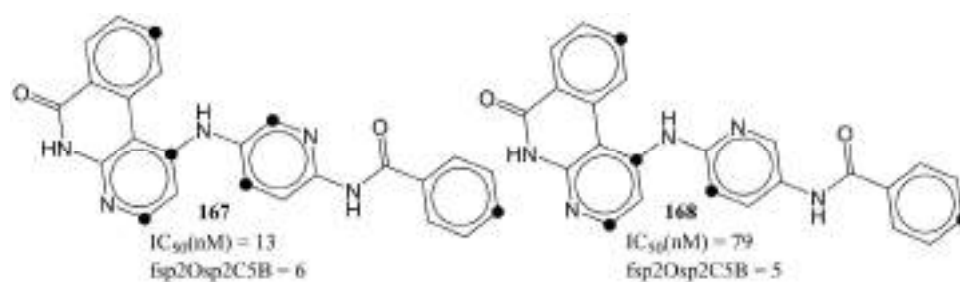


Figure 9. Representation of fsp2Osp2C5B using molecule numbers 167 and 168 as representative examples. The black circle represents the sp²-hybridized carbon at five bonds from sp²-hybridized oxygen.

fOringC6B:

The descriptor fOringC6B is associated with the simultaneous and conditional occurrence of polar (oxygen) and lipophilic characters (ring carbons) with an exact separation by six bonds. If a ring carbon is also present within five or less bonds of any other oxygen atom, then it is omitted while calculating fOringC6B. The molecular descriptor fOringC6B has a negative coefficient in model 1, which means that a higher number of such carbon atoms could reduce the inhibitory profile of a molecule for AKB. This is confirmed when the following pairs of molecules are compared: 526 with 511, 526 with 521, 204 with 205, 229 with 231, 477 with 485, and 256 with 257. The descriptor has been depicted in Figure 10. The red dots indicate the ring carbons, which contribute to fOringC6B at exactly six bonds

from the oxygen atom. The six bonds separating such carbon and oxygen atoms have been labeled with numbers.

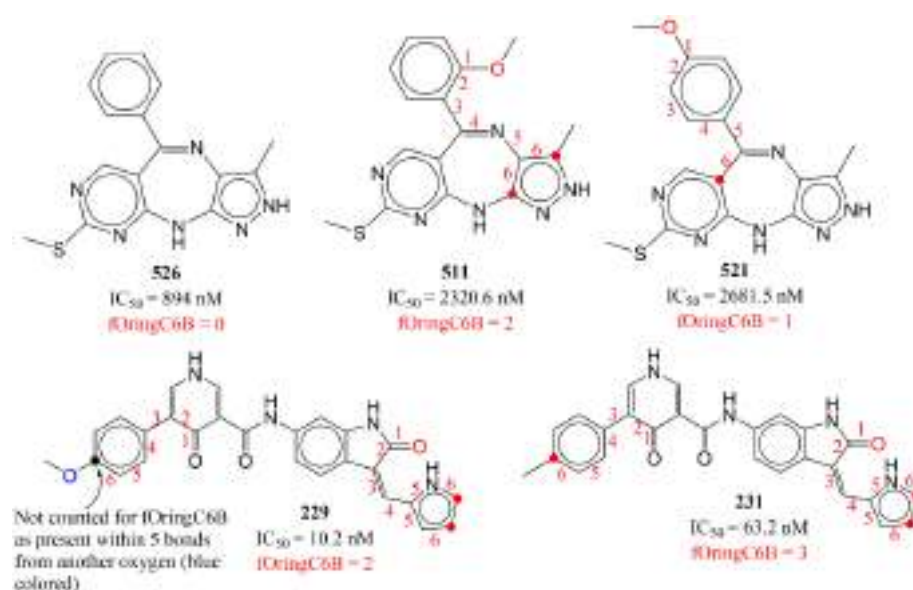


Figure 10. Representative examples for fOringC6B. The numbers (red) indicate the counting of number of bonds between ring carbon and oxygen atom.

It appears that reducing the number of ring carbon atoms is a feasible solution to achieve a lower value of fOringC6B, but this will affect negatively other descriptors, viz. da_lipo_5B, fsp2Osp2C5B. Instead, a solution is to reduce the number of oxygen atoms or alternatively increase their presence within five or less bonds of ring carbon atoms. The second solution is observed in the case of molecule number 229. The additional -OCH3 led to a decreased value of fOringC6B, because, while calculating fOringC6B, if a ring carbon atom was simultaneously present within six bonds of two or more oxygen atoms, it was excluded.

fringNC6B:

The molecular descriptor fringNC6B provides crucial information about the upper limit for separation required between the lipophilic (carbon atoms) and polar (nitrogen atoms) moieties to achieve a better activity profile. While calculating fringNC6B, if a carbon atom is also present within five bonds of any other ring nitrogen, then it is omitted. If a carbon atom is present exactly at a distance of six bonds from a ring nitrogen atom, then it contributes negatively; therefore, such a combination should be avoided. Reducing the bond gap between carbon and ring nitrogen is a feasible and justified solution, as other descriptors, viz. da_lipo_5B and fsp3Csp2N5B, also indicate the same. As stated earlier, a plausible reason for this could be the active site of AKB (see Figure 11). The influence of fringNC6B on activity is confirmed when following pairs of molecules are compared: 5 with 500, 5 with 506, 374 with 406, 507 with 514, to list a few.

As stated earlier, the descriptors present in model A are entangled. Therefore, changing one descriptor could result in changes in other descriptors. For example, the descriptors fringNplaN4B and fringNC6B indicate the importance of ring nitrogen atoms. The fringNplaN4B has a positive correlation with pIC50 but fringNC6B has the opposite relation. Therefore, increasing the value of fringNplaN4B by escalating the ring nitrogen atoms could also lead to a higher fringNC6B value. Hence, a balance of the appropriate number and types of nitrogen, carbon, and oxygen could lead to significant inhibitory activity for aurora kinase B.

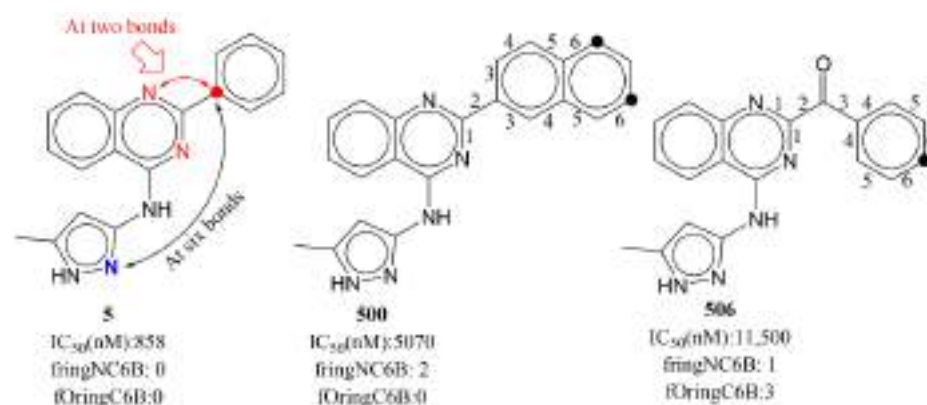


Figure 11. Depiction of fringNC6B using molecule numbers 5, 500, and 506 as representative examples. The carbon present at six bonds from ring nitrogen has been depicted using black dots. The numbers (black) indicate the counting of number of bonds between ring nitrogen and carbon.

4. Materials and Methods

In this work, we adhered to the OECD's and other researchers' suggested standards and recommendations [17–19,32,43,44] for a successful QSAR analysis. The various procedures for creating a model included meticulous dataset selection, data curation, 3D structure production for all molecules, computation and trimming of molecular descriptors, model creation and extensive validation, and mechanistic interpretation [45,46]. To eliminate bias and ensure proper model validation, these stages were carried out one at a time.

4.1. Selection of Dataset

The success and efficacy of a QSAR analysis in the drug discovery pipeline are significantly influenced by the size, composition, and structural diversity of the selected dataset used for the analysis [17–19,32,43,44]. As a result, a sizable dataset of 3398 reported AKB ligands was downloaded from BindingDB (<https://www.bindingdb.org/bind/index.jsp>, accessed on 14 January 2022). The dataset was then reduced to 561 molecules only after duplicates (average value for duplicates), salts, metal derivatives, rule-of-five violators, molecules with undefinable K_i values, etc., were eliminated during data curation [47]. The condensed dataset still included a variety of molecules, such as stereoisomers, positional and chain isomers, various heterocyclic and aromatic scaffolds, etc. Thus, it covered a broad chemical space. The experimental IC₅₀ ranged from 0.26 to 16,000 nM. The experimental IC₅₀ values were converted to pIC₅₀ for a better QSAR analysis ($-\log_{10} \text{IC}_{50}$). Figure 12 and Table 4 comprise some molecules that are very active and those that are least active, to help the readers to understand the structural variation present in the dataset.

Table 4. SMILES notation, IC₅₀ (nM), and pIC₅₀ (M) of five most and least active molecules of the selected dataset.

Mol ID	SMILES	IC ₅₀ (nM)	pIC ₅₀ (M)
339	<chem>O=C(Nc1cc(CNc2ncnc3c(C(=O)N)cccc23)ccc1)c1cnc(NC)cc1</chem>	0.26	9.585
326	<chem>O=C(Nc1cc(C(Nc2ncnc3c(C(=O)N)cccc23)C)ccc1)c1[nH]nc(C(C)C)c1</chem>	0.27	9.569
350	<chem>O=C(Nc1cc(CNc2ncnc3c(C(=O)N)cccc23)ccc1)c1[nH]nc2c1CCCC2</chem>	0.3	9.523
316	<chem>O=C(Nc1cc(C(Nc2ncnc3c(C(=O)N)cccc23)C)ccc1)c1cnc(C)cc1</chem>	0.32	9.495
383	<chem>O=C(Nc1cc(CNc2ncnc3c(C(=O)N)cccc23)ccc1)c1[nH]nc(C(C)C)c1</chem>	0.33	9.481
191	<chem>O=C1OCc2c(C)c(O)c(O)c(O)c12</chem>	8690	5.061
506	<chem>O=C(c1nc(Nc2n[nH]c(C)c2)c2c(n1)cccc2)c1cccc1</chem>	11,500	4.939

Table 4. Cont.

Mol ID	SMILES	IC ₅₀ (nM)	pIC ₅₀ (M)
202	<chem>O=C(C)c1scc(-c2cnc3[nH]c(-c4ccc(OC)cc4)nc3c2)c1</chem>	12,100	4.917
427	<chem>O=C(O)c1cnc(Nc2nccc(/C=C\3/C(=O)N(C)/C(=N/c4ccc(CC)cc4)/S/3)c2)cc1</chem>	12,505.05	4.903
194	<chem>O(C)c1c(Nc2ncc3c([nH]c(-c4c(C)onc4C)c3)c2)cccc1</chem>	16,000	4.796

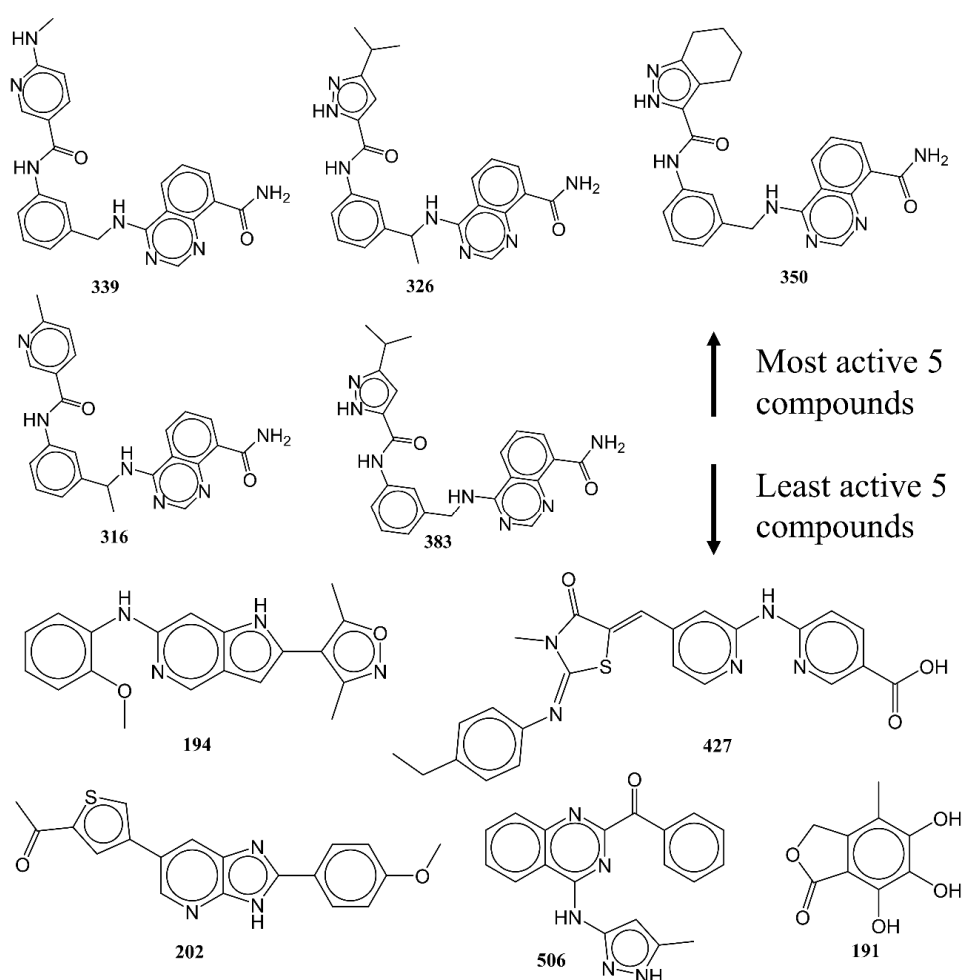


Figure 12. Representative examples from the selected dataset (five most active and five least active molecules).

4.2. Calculation of Molecular Descriptors and Objective Feature Selection (OFS)

The next step involved applying the proper methodology to convert SMILES notations into 3D-optimized structures. OpenBabel 3.1 [48] was used to translate SMILES to SDF for this. Then, utilizing PM3 as a force field for structure optimization and partial charge assignment, SDF was converted to MOL2 using MOPAC [49] 2016. After this, PyDescriptor [37] and PaDEL [50], which together offered more than 40,000 molecular descriptors for each molecule, were used for molecular descriptor calculation. Although using a large number of molecular descriptors increases the likelihood that a QSAR analysis will be effective, with a balance of predictive and mechanistic interpretation abilities, it also raises the risk of overfitting due to noisy redundancy in the descriptors or chance correlations. As a result, OFS was carried out using QSARINS 2.2.4 [51], which eliminated molecular descriptors that were nearly constant (for 90% of molecules) and highly inter-correlated ($|R| > 0.90$). After extensive OFS, only 1150 descriptors were finally included in the reduced set of molecular descriptors, but they nevertheless covered a wide descriptor space because they included

fingerprints, charged-based, 1D to 3D, and a good number of atom-pair descriptors. The likelihood of a mechanistic interpretation of the model increased because a significant portion of the descriptors could be readily interpreted in terms of structural traits.

4.3. Splitting the Dataset into Training and External Sets and Subjective Feature Selection (SFS)

SFS is one of the most important steps in the QSAR model-building process that involves choosing the right feature selection technique with an adequate number and set of molecular descriptors. Before developing the QSAR model, the dataset was randomly divided into a training set (80%, or 449 molecules) and a prediction set (20%, or 112 molecules), to allow for proper training and validation of the model. In order to eliminate bias, reduce information leakage [32], confirm the model's external predictive ability to predict for molecules other than the training set, and to improve the composition of the training and prediction sets, the dataset was randomly divided at a ratio of 80:20. The selection of molecular descriptors was done using the training set only. The prediction set, also known as the test set or external set, was used exclusively for judging the external predictive ability of the model.

To prevent over- and underfitting, the QSAR model must have an ideal number of molecular descriptors (variables). Consequently, the ideal number of descriptors for the model was identified using a straightforward graphical (or breaking point) method [45,46,52]. The value of Q^2_{LOO} typically increases considerably when a new variable (molecular descriptor) is added in stages to an MLR model until the desired elevation is reached. After this, the value of Q^2_{LOO} increases slightly or negligibly. As a result, the number of molecular descriptors that match the elevation point is ideal for creating a QSAR model. A graph of this is shown in Figure 13. The last elevation point in Figure 13 corresponds to seven molecular descriptors. Therefore, the genetic algorithm (GA) in combination with multi-regression (GA-MLR) method, using QSARINS 2.2.4, was used for the exhaustive search to identify seven molecular descriptors to develop the QSAR model. For GA-MLR, Q^2_{LOO} was used as the fitness parameter.

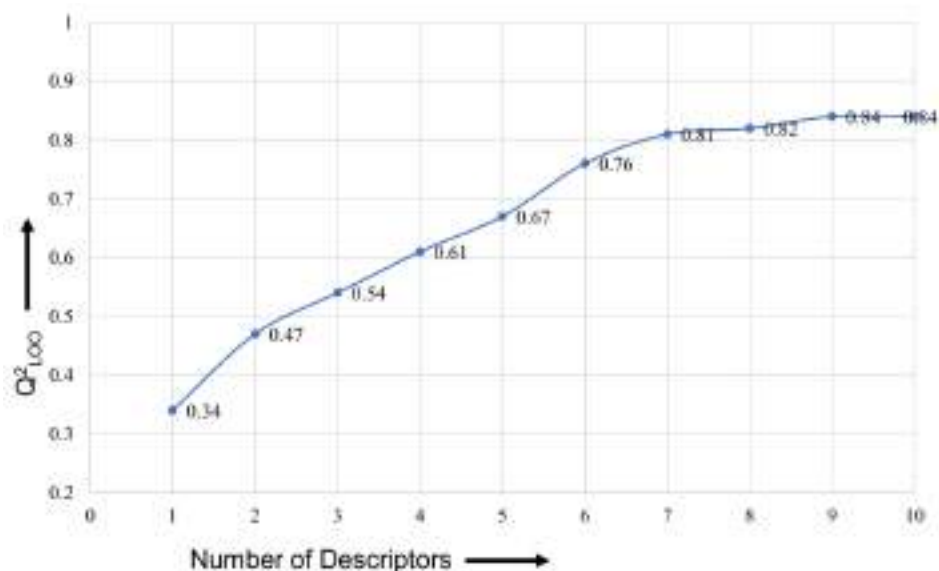


Figure 13. Plot of number of descriptors against leave-one-out coefficient of determination (Q^2_{LOO}) to identify the optimum number of descriptors.

4.4. Building Regression Model and Its Validation

Different combinations of various molecular descriptors were eventually found during the search for seven molecular descriptors for the QSAR model using GA-MLR. However, due to the statistical performance and the satisfaction of adhering to strict parameters and criteria, which have been recommended [17–19,23,27,32,33,44–46,52–57] by a significant

number of researchers, only one combination of molecular descriptors was chosen. The following threshold values and conditions were used to select the model:

$R^2_{tr} \geq 0.6$, $Q2LOO \geq 0.5$, $Q2LMO \geq 0.6$, $R^2 > Q2LOO$, $R^2_{ex} \geq 0.6$, $RMSE_{tr} < RMSE_{cv}$, $\Delta K \geq 0.05$, $CCC \geq 0.80$, $Q2-Fn \geq 0.60$, $r2m \geq 0.5$, $(1-r2/ro2) < 0.1$, $0.9 \leq k \leq 1.1$ or $(1-r2/r'o2) < 0.1$, $0.9 \leq k' \leq 1.1$, $|ro2 - r'o2| < 0.3$, $RMSE_{ex}$, MAE_{ex} , R^2_{ex} , $Q2F1$, $Q2F2$, $Q2F3$, and low $R2Y_{scr}$, $RMSE$ and MAE .

The model's application domain must be identified for additional validation. In order to assess the application domain of the QSAR model, we employed a Williams plot (standardized residuals vs. hat values).

5. Conclusions

In relation to different features influencing the inhibitory activity for AKB, the present analysis successfully highlighted the significance of different types of atoms, groups, patterns, and tautomerism. Additionally, it emphasized the significance of specific patterns of atoms of different hybridization and their inter-relations in determining the final activity. The conditional presence of lipophilic (carbon) atoms or groups with respect to nitrogen atoms was also successfully recognized by model A as being beneficial for obtaining higher inhibitory for AKB. The present work, for the first time, pointed out the role played by tautomerism for AKB inhibitors. Model A performed statistically well, which was indicative of its strong external prediction power. As the current work successfully recognized both previously described and novel pharmacophoric properties associated with AKB inhibition, the results are of immense use throughout the drug discovery pipeline for the development of lead/drug candidates against AKB.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/ijms232314527/s1>.

Author Contributions: V.H.M. and M.E.A.Z.: conceptualization, project design, and experimental studies; V.H.M. and S.A.A.-H.: drafting, resources, and funding management; M.M.R., S.A. and S.D.T.: data collection and curation, drafting, and data compilation; S.A., A.S. and A.A.A.-M.: draft revision and analysis. All authors have read and agreed to the published version of the manuscript.

Funding: The authors acknowledge the Deanship of Scientific Research at Imam Mohammad Ibn Saud Islamic University, Riyadh, Saudi Arabia, for its support of this research through research group number RG-21-09-76.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data are contained within the article and Supplementary Materials.

Acknowledgments: V.H.M. is grateful to Paola Gramatica (Italy) and her team for providing the free copy of QSARINS 2.2.4.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

SMILES	Simplified molecular-input line-entry system
GA	Genetic algorithm
MLR	Multiple linear regression
QSAR	Quantitative structure–activity relationship
WHO	World Health Organization
OLS	Ordinary least squares
QSARINS	QSAR Insubria
OECD	Organization for Economic Cooperation and Development

References

1. Du, R.; Huang, C.; Liu, K.; Li, X.; Dong, Z. Targeting AURKA in Cancer: Molecular mechanisms and opportunities for Cancer therapy. *Mol. Cancer* **2021**, *20*, 15. [[CrossRef](#)] [[PubMed](#)]
2. Garuti, L.; Roberti, M.; Bottegioni, G. Small Molecule Aurora Kinases Inhibitors. *Curr. Med. Chem.* **2009**, *16*, 1949–1963. [[CrossRef](#)] [[PubMed](#)]
3. Pollard, J.R.; Mortimore, M. Discovery and Development of Aurora Kinase Inhibitors as Anticancer Agents. *J. Med. Chem.* **2009**, *52*, 2629–2651. [[CrossRef](#)] [[PubMed](#)]
4. Jing, X.L.; Chen, S.W. Aurora kinase inhibitors: A patent review (2014–2020). *Expert Opin. Ther. Pat.* **2021**, *31*, 625–644. [[CrossRef](#)] [[PubMed](#)]
5. Willems, E.; Dedobbeleer, M.; Digregorio, M.; Lombard, A.; Lumapat, P.N.; Rogister, B. The functional diversity of Aurora kinases: A comprehensive review. *Cell Div.* **2018**, *13*, 7. [[CrossRef](#)]
6. Borisa, A.C.; Bhatt, H.G. A comprehensive review on Aurora kinase: Small molecule inhibitors and clinical trial studies. *Eur. J. Med. Chem.* **2017**, *140*, 1–19. [[CrossRef](#)]
7. Bavetsias, V.; Linardopoulos, S. Aurora Kinase Inhibitors: Current Status and Outlook. *Front. Oncol.* **2015**, *5*, 278. [[CrossRef](#)]
8. Kollareddy, M.; Zheleva, D.; Dzubak, P.; Brahmshatriya, P.S.; Lepsik, M.; Hajduch, M. Aurora kinase inhibitors: Progress towards the clinic. *Investig. New Drugs* **2012**, *30*, 2411–2432. [[CrossRef](#)]
9. Lok, W.; Klein, R.Q.; Saif, M.W. Aurora kinase inhibitors as anti-cancer therapy. *Anticancer Drugs* **2010**, *21*, 339–350. [[CrossRef](#)]
10. He, Y.; Fu, W.; Du, L.; Yao, H.; Hua, Z.; Li, J.; Lin, Z. Discovery of a novel Aurora B inhibitor GSK650394 with potent anticancer and anti-aspergillus fumigatus dual efficacies in vitro. *J. Enzym. Inhib. Med. Chem.* **2022**, *37*, 109–117. [[CrossRef](#)]
11. Keen, N.; Taylor, S. Mitotic drivers—Inhibitors of the Aurora B Kinase. *Cancer Metastasis Rev.* **2009**, *28*, 185–195. [[CrossRef](#)]
12. Kong, Y.; Bender, A.; Yan, A. Identification of Novel Aurora Kinase A (AURKA) Inhibitors via Hierarchical Ligand-Based Virtual Screening. *J. Chem. Inf. Model.* **2018**, *58*, 36–47. [[CrossRef](#)]
13. Durlacher, C.T.; Li, Z.L.; Chen, X.W.; He, Z.X.; Zhou, S.F. An update on the pharmacokinetics and pharmacodynamics of alisertib, a selective Aurora kinase A inhibitor. *Clin. Exp. Pharmacol. Physiol.* **2016**, *43*, 585–601. [[CrossRef](#)]
14. Imam, S.S.; Gilani, S.J. Computer Aided Drug Design: A Novel Loom to Drug Discovery. *Org. Med. Chem.* **2017**, *1*, 1–6. [[CrossRef](#)]
15. Baig, M.H.; Ahmad, K.; Roy, S.; Ashraf, J.M.; Adil, M.; Siddiqui, M.H.; Khan, S.; Kamal, M.A.; Provaznik, I.; Choi, I. Computer Aided Drug Design: Success and Limitations. *Curr. Pharm. Des.* **2016**, *22*, 572–581. [[CrossRef](#)]
16. Macalino, S.J.; Gosu, V.; Hong, S.; Choi, S. Role of computer-aided drug design in modern drug discovery. *Arch. Pharm. Res.* **2015**, *38*, 1686–1701. [[CrossRef](#)]
17. Gramatica, P. Principles of QSAR Modeling. *Int. J. Quant. Struct.-Prop. Relatsh.* **2020**, *5*, 61–97. [[CrossRef](#)]
18. Fujita, T.; Winkler, D.A. Understanding the Roles of the “Two QSARs”. *J. Chem. Inf. Model.* **2016**, *56*, 269–274. [[CrossRef](#)]
19. Cherkasov, A.; Muratov, E.N.; Fourches, D.; Varnek, A.; Baskin, I.I.; Cronin, M.; Dearden, J.; Gramatica, P.; Martin, Y.C.; Todeschini, R.; et al. QSAR modeling: Where have you been? Where are you going to? *J. Med. Chem.* **2014**, *57*, 4977–5010. [[CrossRef](#)]
20. Neaz, M.; Muddassar, M.; Pasha, F.; Cho, S.J. Structural studies of B-type Aurora kinase inhibitors using computational methods. *Acta Pharmacol. Sin.* **2010**, *31*, 244–258. [[CrossRef](#)]
21. Lan, P.; Chen, W.N.; Sun, P.H.; Chen, W.M. 3D-QSAR and molecular docking studies of azaindole derivatives as Aurora B kinase inhibitors. *J. Mol. Model.* **2011**, *17*, 1191–1205. [[CrossRef](#)] [[PubMed](#)]
22. Ashraf, S.; Ranaghan, K.E.; Woods, C.J.; Mulholland, A.J.; Ul-Haq, Z. Exploration of the structural requirements of Aurora Kinase B inhibitors by a combined QSAR, modelling and molecular simulation approach. *Sci. Rep.* **2021**, *11*, 18707. [[CrossRef](#)] [[PubMed](#)]
23. Gramatica, P. External Evaluation of QSAR Models, in Addition to Cross-Validation Verification of Predictive Capability on Totally New Chemicals. *Mol. Inform.* **2014**, *33*, 311–314. [[CrossRef](#)] [[PubMed](#)]
24. Chirico, N.; Gramatica, P. Real external predictivity of QSAR models. Part 2. New intercomparable thresholds for different validation criteria and the need for scatter plot inspection. *J. Chem. Inf. Model.* **2012**, *52*, 2044–2058. [[CrossRef](#)] [[PubMed](#)]
25. Chirico, N.; Gramatica, P. Real external predictivity of QSAR models: How to evaluate it? Comparison of different validation criteria and proposal of using the concordance correlation coefficient. *J. Chem. Inf. Model.* **2011**, *51*, 2320–2335. [[CrossRef](#)]
26. Gramatica, P. Principles of QSAR models validation internal and external. *QSAR Comb. Sci.* **2007**, *26*, 694–701. [[CrossRef](#)]
27. Gramatica, P. On the development and validation of QSAR models. *Methods Mol. Biol.* **2013**, *930*, 499–526. [[CrossRef](#)]
28. Rao, R.B.; Fung, G.; Rosales, R. On the Dangers of Cross-Validation. An Experimental Evaluation. In *Proceedings of the 2008 SIAM International Conference on Data Mining (SDM)*; Society for Industrial and Applied Mathematics: Philadelphia, PA, USA, 2008; pp. 588–596. [[CrossRef](#)]
29. Tropsha, A.; Gramatica, P.; Gombar, V.K. The Importance of Being Earnest Validation is the Absolute Essential for Successful Application and Interpretation of QSPR Models. *QSAR Comb. Sci.* **2003**, *22*, 69–77. [[CrossRef](#)]
30. Hawkins, D.M.; Basak, S.C.; Mills, D. Assessing model fit by cross-validation. *J. Chem. Inf. Comput. Sci.* **2003**, *43*, 579–586. [[CrossRef](#)]
31. Lučić, B.; Batista, J.; Bojović, V.; Lovrić, M.; Sović Kržić, A.; Bešlo, D.; Nadramija, D.; Vikić-Topić, D. Estimation of Random Accuracy and its Use in Validation of Predictive Quality of Classification Models within Predictive Challenges. *Croat. Chem. Acta* **2019**, *92*, 379–391. [[CrossRef](#)]

32. Masand, V.H.; Mahajan, D.T.; Nazeruddin, G.M.; Hadda, T.B.; Rastija, V.; Alfeefy, A.M. Effect of information leakage and method of splitting (rational and random) on external predictive ability and behavior of different statistical parameters of QSAR model. *Med. Chem. Res.* **2014**, *24*, 1241–1264. [[CrossRef](#)]
33. Kar, S.; Roy, K.; Leszczynski, J. Applicability Domain: A Step Toward Confident Predictions and Decidability for QSAR Modeling. In *Computational Toxicology*; Humana Press: New York, NY, USA, 2018; pp. 141–169. [[CrossRef](#)]
34. Roy, P.P.; Kovarich, S.; Gramatica, P. QSAR model reproducibility and applicability A case study of rate constants of hydroxyl radical reaction models applied to polybrominated diphenyl ethers and (benzo-)triazoles. *J. Comput. Chem.* **2011**, *32*, 2386–2396. [[CrossRef](#)]
35. Sushko, I.; Novotarskyi, S.; Korner, R.; Pandey, A.K.; Cherkasov, A.; Li, J.; Gramatica, P.; Hansen, K.; Schroeter, T.; Muller, K.R.; et al. Applicability domains for classification problems: Benchmarking of distance to models for Ames mutagenicity set. *J. Chem. Inf. Model.* **2010**, *50*, 2094–2111. [[CrossRef](#)]
36. Tropsha, A.; Golbraikh, A. Predictive QSAR modeling workflow, model applicability domains, and virtual screening. *Curr. Pharm. Des.* **2007**, *13*, 3494–3504. [[CrossRef](#)]
37. Masand, V.H.; Rastija, V. PyDescriptor: A new PyMOL plugin for calculating thousands of easily understandable molecular descriptors. *Chemom. Intell. Lab. Syst.* **2017**, *169*, 12–18. [[CrossRef](#)]
38. Todeschini, R.; Consonni, V. *Molecular Descriptors for Chemoinformatics*; Wiley-VCH: Weinheim, Germany, 2009.
39. Todeschini, R.; Consonni, V. *Handbook of Molecular Descriptors*; Wiley-VCH: Weinheim, Germany, 2000; Volume 11.
40. Di, L.; Kerns, E.H. *Drug-like Properties: Concepts, Structure Design and Methods: From ADME to Toxicity Optimization*, 2nd ed.; Elsevier/AP: Amsterdam, The Netherlands; Boston, MA, USA, 2016; 560p.
41. Sessa, F.; Villa, F. Structure of Aurora B-INCENP in complex with barasertib reveals a potential transinhibitory mechanism. *Acta Crystallogr. F Struct. Biol. Commun.* **2014**, *70 Pt 3*, 294–298. [[CrossRef](#)]
42. Elkins, J.M.; Santaguida, S.; Musacchio, A.; Knapp, S. Crystal structure of human aurora B in complex with INCENP and VX-680. *J. Med. Chem.* **2012**, *55*, 7841–7848. [[CrossRef](#)]
43. Masand, V.H.; Mahajan, D.T.; Gramatica, P.; Barlow, J. Tautomerism and multiple modelling enhance the efficacy of QSAR: Antimalarial activity of phosphoramidate and phosphorothioamidate analogues of amiprofos methyl. *Med. Chem. Res.* **2014**, *23*, 4825–4835. [[CrossRef](#)]
44. Masand, V.H.; Mahajan, D.T.; Ben Hadda, T.; Jawarkar, R.D.; Alafeefy, A.M.; Rastija, V.; Ali, M.A. Does tautomerism influence the outcome of QSAR modeling? *Med. Chem. Res.* **2014**, *23*, 1742–1757. [[CrossRef](#)]
45. Zaki, M.E.A.; Al-Hussain, S.A.; Bukhari, S.N.A.; Masand, V.H.; Rathore, M.M.; Thakur, S.D.; Patil, V.M. Exploring the Prominent and Concealed Inhibitory Features for Cytoplasmic Isoforms of Hsp90 Using QSAR Analysis. *Pharmaceuticals* **2022**, *15*, 303. [[CrossRef](#)]
46. Zaki, M.E.A.; Al-Hussain, S.A.; Al-Mutairi, A.A.; Masand, V.H.; Samad, A.; Jawarkar, R.D. Mechanistic Analysis of Chemically Diverse Bromodomain-4 Inhibitors Using Balanced QSAR Analysis and Supported by X-ray Resolved Crystal Structures. *Pharmaceuticals* **2022**, *15*, 745. [[CrossRef](#)]
47. Fourches, D.; Muratov, E.; Tropsha, A. Trust, but verify: On the importance of chemical structure curation in cheminformatics and QSAR modeling research. *J. Chem. Inf. Model.* **2010**, *50*, 1189–1204. [[CrossRef](#)]
48. O'Boyle, N.M.; Banck, M.; James, C.A.; Morley, C.; Vandermeersch, T.; Hutchison, G.R. Open Babel: An open chemical toolbox. *J. Cheminform.* **2011**, *3*, 33. [[CrossRef](#)]
49. Stewart, J.J.P. MOPAC: A semiempirical molecular orbital program. *J. Comput.-Aided Mol. Des.* **1990**, *4*, 1–103. [[CrossRef](#)]
50. Yap, C.W. PaDEL-descriptor: An open source software to calculate molecular descriptors and fingerprints. *J. Comput. Chem.* **2011**, *32*, 1466–1474. [[CrossRef](#)]
51. Gramatica, P.; Chirico, N.; Papa, E.; Cassani, S.; Kovarich, S. QSARINS: A new software for the development, analysis, and validation of QSAR MLR models. *J. Comput. Chem.* **2013**, *34*, 2121–2132. [[CrossRef](#)]
52. Bukhari, S.N.A.; Elsherif, M.A.; Junaid, K.; Ejaz, H.; Alam, P.; Samad, A.; Jawarkar, R.D.; Masand, V.H. Perceiving the Concealed and Unreported Pharmacophoric Features of the 5-Hydroxytryptamine Receptor Using Balanced QSAR Analysis. *Pharmaceuticals* **2022**, *15*, 834. [[CrossRef](#)]
53. Consonni, V.; Todeschini, R.; Ballabio, D.; Grisoni, F. On the Misleading Use of Q2F3 for QSAR Model Comparison. *Mol. Inf.* **2019**, *38*, e1800029. [[CrossRef](#)]
54. Golbraikh, A.; Muratov, E.; Fourches, D.; Tropsha, A. Data set modelability by QSAR. *J. Chem. Inf. Model.* **2014**, *54*, 1–4. [[CrossRef](#)]
55. Martin, T.M.; Harten, P.; Young, D.M.; Muratov, E.N.; Golbraikh, A.; Zhu, H.; Tropsha, A. Does rational selection of training and test sets improve the outcome of QSAR modeling? *J. Chem. Inf. Model.* **2012**, *52*, 2570–2578. [[CrossRef](#)]
56. Gramatica, P.; Cassani, S.; Roy, P.P.; Kovarich, S.; Yap, C.W.; Papa, E. QSAR Modeling is not Push a Button and Find a Correlation: A Case Study of Toxicity of (Benzo-)triazoles on Algae. In *Molecular Informatics*; Wiley Online: Hoboken, NJ, USA, 2012; Volume 31, pp. 817–835.
57. Huang, J.; Fan, X. Why QSAR fails: An empirical evaluation using conventional computational approach. *Mol. Pharm.* **2011**, *8*, 600–608. [[CrossRef](#)] [[PubMed](#)]



Artificial photosynthesis and current status of its application in generating fuels: A patent analytical study

Author information

- Dr.PradnyaNalawade
- VidyaBharatiMahavidyalaya, Department of Chemistry, Amravati, Maharashtra, India.
- pradnyapnalawade@gmail.com

Abstract:

Scientists around the world are working towards the goal of developing technologies to make use of energy from the sun to produce fuels for transport, industry and electricity generation. Fuels produced using solar energy and carbon dioxide from the environment would transform our energy options in the future by providing an alternative to fossil fuels.

This report analyses the artificial photosynthesis technology for producing fuels through the lens of intellectual property. It uses the scale and intensity of patent activity to provide an overview of innovation in this technology.

This study identified 53 innovations related to artificial photosynthesis. These results were categorized into four broad technology sub classes according to the type of system or area for fuel production and then analyzed to give a picture of the landscape in this area of research.

The applicants that made major contributions to this field have origins from the United States of America, Korea and Japan, whereas Israel, Spain, Russia has had nominal patent involvement in the contribution

to this field of study. Study on analysis of patent commercialization reveals that very few patents are exploited commercially.

Keywords: Artificial photosynthesis, renewable energy sources, Patents, sustainable energy

Declarations/ Conflict of Interest: There is no conflict of interest

Funding: 'Not applicable'

Introduction

Artificial photosynthesis is a term which describes the processes that, like natural photosynthesis, harvest sunlight and use this energy to chemically convert water and carbon dioxide into fuels [1]. The journey began in the 19th century with the discovery of the 'photovoltaic effect', whereby an electric current is produced in a material when exposed to light.

Scientists around the world are working towards the goal of developing technologies to exploit energy from the sun to produce fuels for transport, industry and electricity generation. Fuels produced using solar energy would transform our energy options in the future by providing an alternative to fossil fuels [2].

The purpose of this report is to explore the patent landscape of artificial photosynthesis technologies, specifically in the areas of solar energy harvesting and carbon dioxide reduction for the production of fuels.

The main focus of this report will be to investigate global patent trends within these technologies and more specifically to identify who is filing patent applications and where.

Another aim of this report is to see if the use of sunlight and atmospheric carbon dioxide for production of fuels as an emerging technology to solve the challenges of meeting the growing demand of fuel supply.

Patents can be used as indicators of research output. A patent is a right that is granted for any device, substance, method or process that is new, inventive, and useful Patent rights are legally enforceable and give the owner exclusive rights to commercially exploit the invention for a limited period of time.

It is a requirement of patent law that patent documents are published and that they fully disclose inventions. As a result of the disclosure requirement, patent literature reflects developments in science and technology. Patent documents include other useful information, such as international patent classifications and information about inventors and applicants.

Through the extraction and analysis of data associated with patent documents, it is possible to measure aspects of inventive activity such as scope, intensity, collaboration and impact. These metrics can be

developed across technology sectors and by various units of measurement, such as individuals (inventors), institutions (applicants), regions and countries. The above groups were manually classified sorted and provide more useful analysis than the IPC. The IPC was not suitable for our analysis, as the artificial photosynthesis technology is interdisciplinary.

Methodology

A patent landscape report is a comprehensive analysis based on patent documents in a specific technology area, covering global patenting activity or patenting in a certain geographical area. The first step in preparing such a landscape report is to understand the main objectives and the technology scope. The main objectives usually lead to a series of more specific questions to be addressed. The technology area is usually broken down into sub-areas representing various technical aspects. Next a state -of-the -art patent search is performed to collect relevant patent information. The patent search results are then analyzed to provide answers to those questions regarding the main objectives.

Search strategy and data formatting

We searched the Thomson Innovation database for patents relating to artificial photosynthesis for the period from 2002-2019. Following query was fired to obtain the data.

Search Query

1. CTB=(ARTIFICIAL PHOTOSYNTHESIS) OR CTB=(MIMICKING PLANTS FOR FUELS) AND CTB=(NANOMATERIALS);
2. CTB=(ARTIFICIAL ADJ PHOTOSYNTHESIS) OR CTB=(converting ADJ solar ADJ energy ADJ in ADJ to ADJ fuels) AND CTB=(SUNLIGHT);
3. CTB=(REDUCE ADJ CARBONDIOXIDE) AND CTB=(SUNLIGHT) AND CTB=(FUELS)

All the patent data was sorted according to relevance of topic and compacted into simple patent families [3]. A separate Excel sheet was prepared that includes all relevant simple patent families. Each simple family was identified and represented by the publication number of a single published family member, preferably the earliest published PCT publication number [4], if available. If the family has no PCT family member, the family member published in English was selected.

The patent data included the following bibliographic information for each family. Title

- Abstract
- Claims
- Applicant/assignee name(s)
- Inventor name(s)

- Priority information
- Earliest priority date (or application date in case no priority is claimed) , i.e. the date of first filing in the family
- Priority country of earliest priority
- Publication numbers of all members of the simple family
- IPC symbols
- Number of forward citations
- No. of backward citations
- Status of patent (Dead/alive)

The search results were categorized into the following groups

1. Artificial photosynthesis modules: patent families including systems which are used to convert water and carbon dioxide in to different types of fuels.
2. Photocatalysts: patent families including photocatalyst which are used to produce fuels using sunlight.
3. Photo-electrochemistry: patent families including electrochemical generation of fuels from water and carbon dioxide using sunlight.
4. Nanomaterials: patent families including nano-sized metal or metalloid particles used for generation of fuels from water, carbon dioxide or sunlight

The above groups were manually classified sorted and provide more useful analysis than the IPC. The IPC was not suitable for our analysis, as the field of artificial photosynthesis is more interdisciplinary.

Publication trend

Figure 1: Year wise Publication Trend of patents in the Artificial Photosynthesis technology

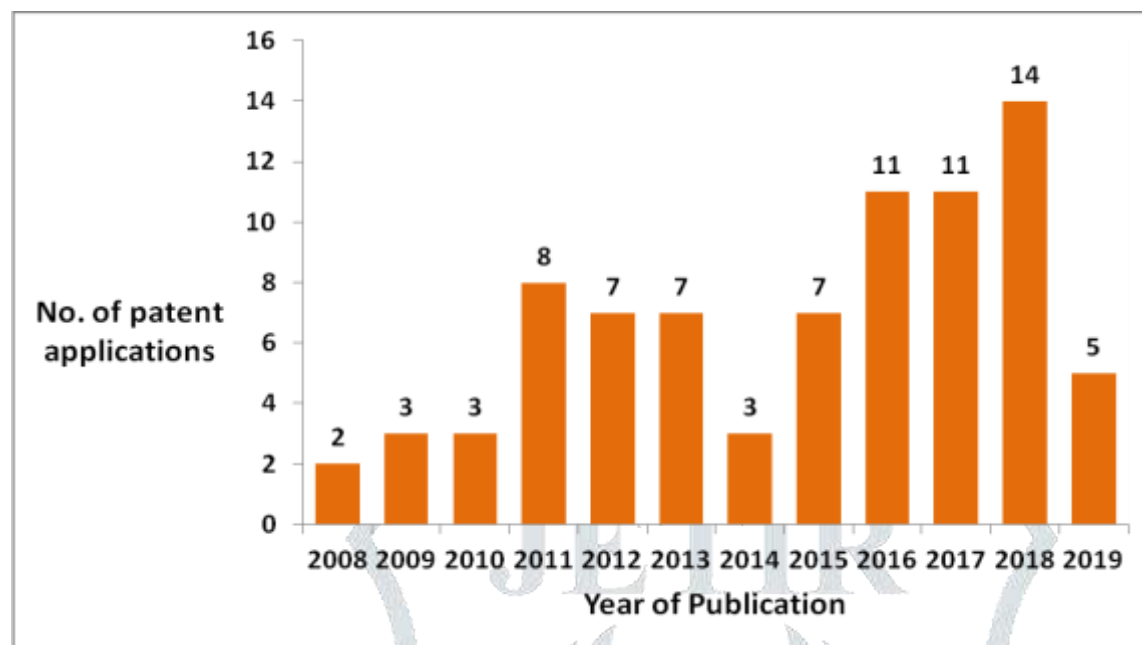


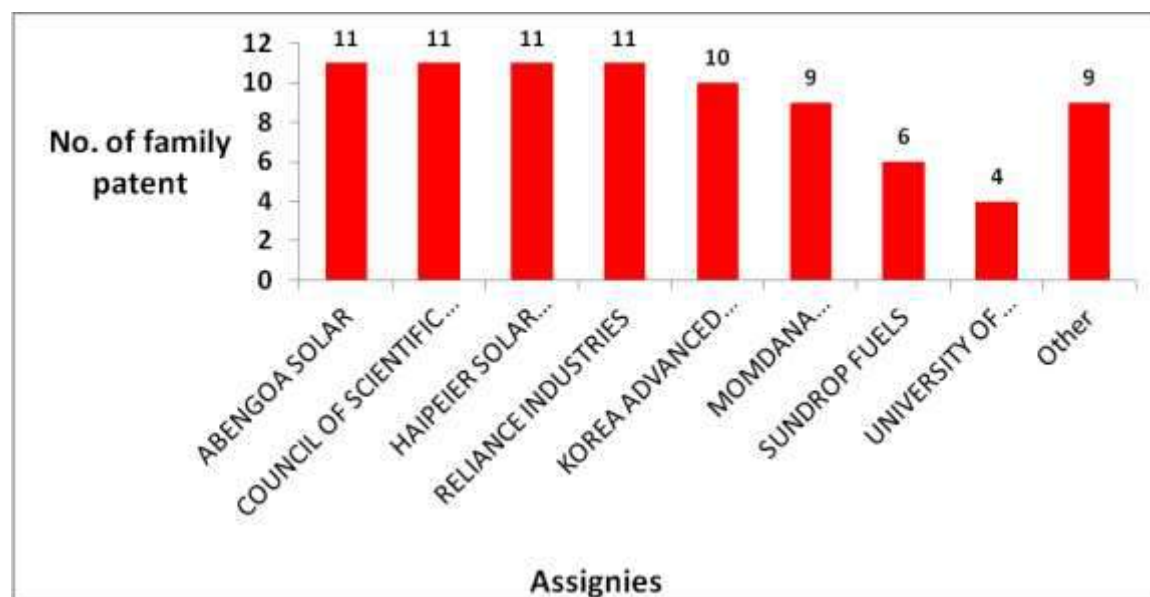
Figure 1 demonstrates the artificial photosynthesis patent applications over the period 2008-2019.

There has been steady work around this area throughout the last few decades but the real flow in the activity around this technology has happened in the year 2011-2018.

The total number of patent family applications per year in the field of agricultural nanomaterials has steadily increased from two patent applications in 2008 to 14 patent applications in 2018. In the year 2014 saw fewer publication of artificial photosynthesis patent applications, however from the 2015 onwards again growth in publication of patent applications was seen. The highest patent applications were published in 2018. The growing demand for fuel as well the global emphasis on reducing emissions and recycling of gaseous pollutant to save environment may have contributed to this more recent growth in patent filings.

Top assignees

From the 53 patent families related to artificial photosynthesis, 14 distinct applicants were identified. Figure 2 shows the top ranked applicants in this area. ABENGOA SOLAR, HAIPEIER SOLAR ENERGY HYPERSOLAR SAEED MUBIAN and COUNCIL OF SCIENTIFIC & INDUSTRIAL RESEARCH in collaboration with RELIANCE INDUSTRIES are the most active applicants, each with 11 patent family members.

Figure 2: Top Assignee

ABENGOA SOLAR is a subsidiary of Abengoa headquartered in Spain, is a multinational chemical company whose portfolio includes the development of photovoltaics, concentrated photovoltaics, or concentrated solar thermal technologies. ABENGOA SOLAR has patent families in the area of Photo-electrochemistry, and the most recent patent family has a priority date in 2008. In particular, WO2009095509 relates to liquid Low concentration solar plant and method for maximizing the electricity production of the photovoltaic modules thereof. This patent family has 11 applications.

COUNCIL OF SCIENTIFIC & INDUSTRIAL RESEARCH and RELIANCE INDUSTRIES are India based Government Research organization and an Indian private sector enterprise respectively. RELIANCE INDUSTRIES is involved in power generation, infrastructure, construction and defense. They have eleven patent family in the area of photo-electrochemistry and have priority date in 2011 with PCT publication WO2013046228 related to process for generation of hydrogen and syngas. KOREA ADVANCED INSTITUTE OF SCIENCE & TECHNOLOGY is a Korean Institute, has ten patent family members in the area of photocatalyst related to production of formic acid.

MONTANA STATE UNIVERSITY is a US based research Institute, has nine patent family members relating to Composite Nanomaterial for photocatalytic hydrogen production particularly PCT publication WO2013046228 relates to reaction for fixation of anthropogenic CO₂ into calcium carbonate using carbonic anhydrase (CA) as a biocatalyst. CA is being employed to accelerate the rate of hydration of CO₂ to form carbonate ions and proton. SUNDROP FUELS is US based bio-based fuel and chemical company, with headquarters in Longmont, Colorado has six patent family members with priority in 2010. More particularly PCT publication number WO2010036662 related to the products from a solar

assisted reverse-water-gas-shift reaction (RWGS) are used to create a liquid hydrocarbon fuel. A synthesis reactor uses any unconsumed hydrogen molecules and the resultant stabilized carbon monoxide molecules from the RWGS reaction in the hydrocarbon fuel synthesis process to create a liquid hydrocarbon fuel.

UNIVERSITY OF COLORADO again a public research university located in Boulder, Colorado, United States has four patent family members with priority in 2013.

Other major applicants are HIROYUKI KOBAYASHI, UCHICAGO ARGONNE, ZHEJIANG GONGSHANG UNIVERSITY, ZHEJIANG GONGSHANG UNIVERSITY, GOVERNING COUNCIL OF THE UNIVERSITY OF TORONTO collectively have nine patent family members.

Top filling country

Although ABENGOA SOLAR (Spain), and COUNCIL OF SCIENTIFIC & INDUSTRIAL RESEARCH in collaboration with RELIANCE INDUSTRIES (India) are the applicants with the largest number of patent family members, the United States is the country with the most applications. Figure 3 demonstrates top filling applicant origin. The top three countries hold almost 50 per cent of the total patent families.

The United States is a highest patent filling country in the field of artificial photosynthesis. Korea is the second largest filling country. Japan China, ranked third overall.

Figure 3: Top filling countries in the Artificial Photosynthesis technology

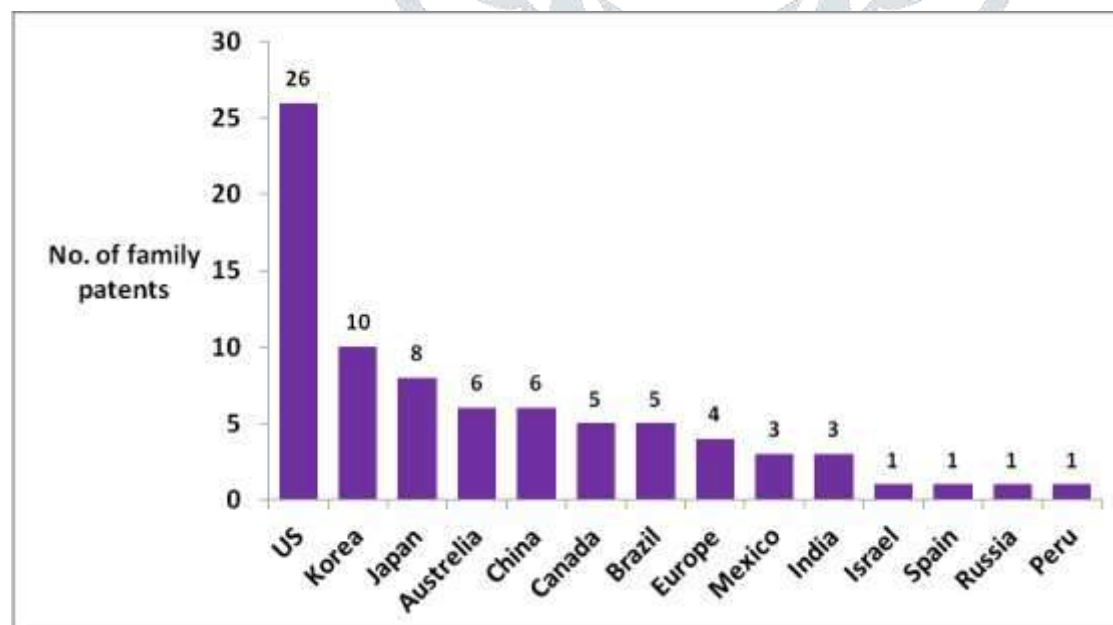
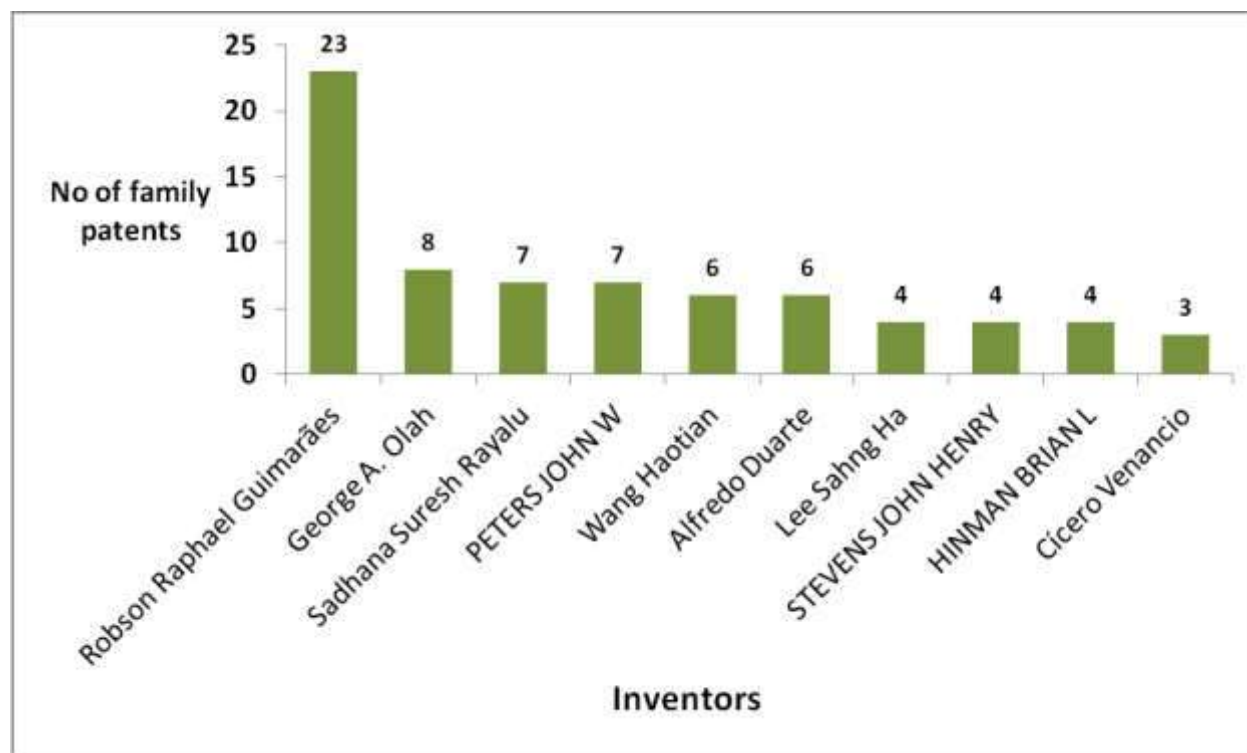


Figure 3 also shows that the remainder of the world are not patenting in large numbers in the area of artificial photosynthesis. Israel, Spain, Russia and Peru contribute list patents in the field of artificial photosynthesis.

Top inventors

An analysis of the inventors of the 53 patent family members identified about twelve productive Inventors with at least three patent family members.

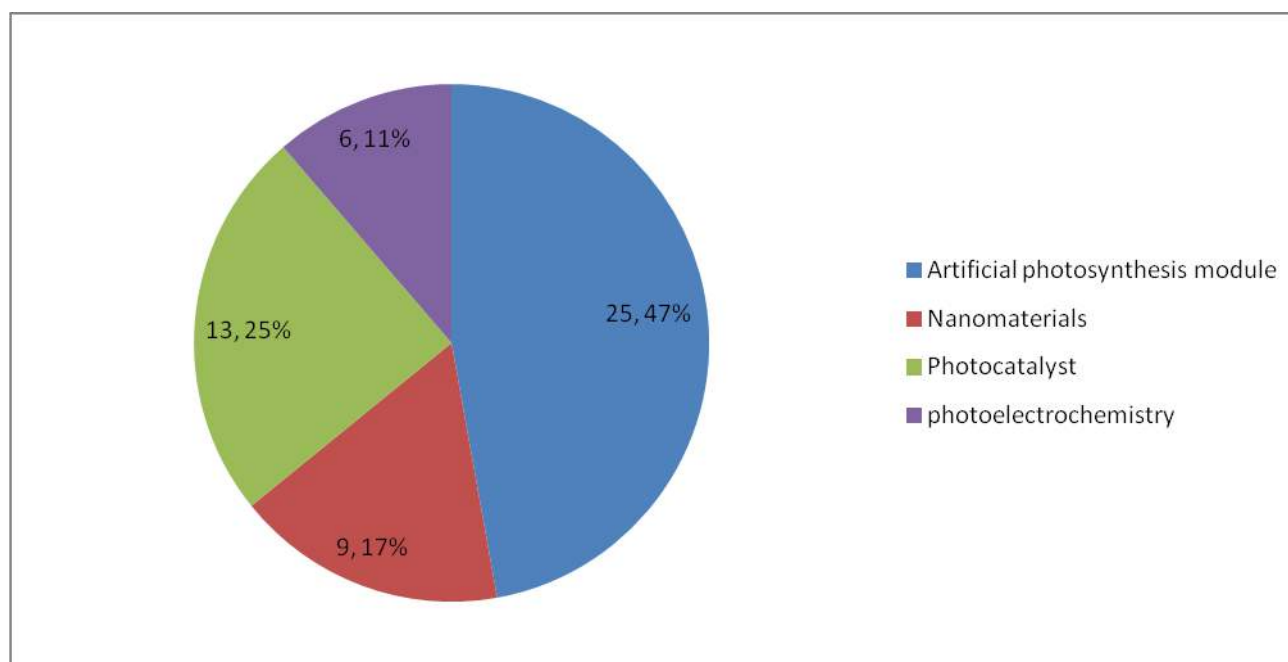
Figure 4: Top inventors in the Artificial Photosynthesis technology



ROBSON RAPHAEL GUIMARÃES was the leading inventor with 23 family patents filed across various countries. GEORGE A. OLAH was the second leading inventor, involved in eight family patents. SADHANA SURESH RAYALU and PETERS JOHN W were the third leading inventors, involved in seven patent inventions. Inventors WANG HAOTIAN, ALFREDO DUARTE were the next leading inventors with each of them having six patent family members. Lee Sahng Ha, STEVENS JOHN HENRY and HINMAN BRIAN L involved in four patent family members each and CÍCERO VENANCIO having three patent family members.

Technology sub-area analysis

The technology artificial photosynthesis was divided further in four sub-groups as shown in Figure 5. Each patent family was assigned a technology sub-area based on the type of fuel produced and the type of technology by reviewing the abstract and description associated with each of the patent families.

Figure 5: Share of various sub-areas in the artificial photosynthesis technology

An in-depth analysis of various areas related to artificial photosynthesis revealed that the “Artificial photosynthesis module” is the predominant area in artificial photosynthesis. The “Artificial photosynthesis module” accounts for 47% (or 25 patent family members) out of 53 patent family. The inventions classified in this category relates generally to modules for reduction of atmospheric carbon dioxide and water into fuel specifically, to carbonaceous fuel by utilizing sunlight.

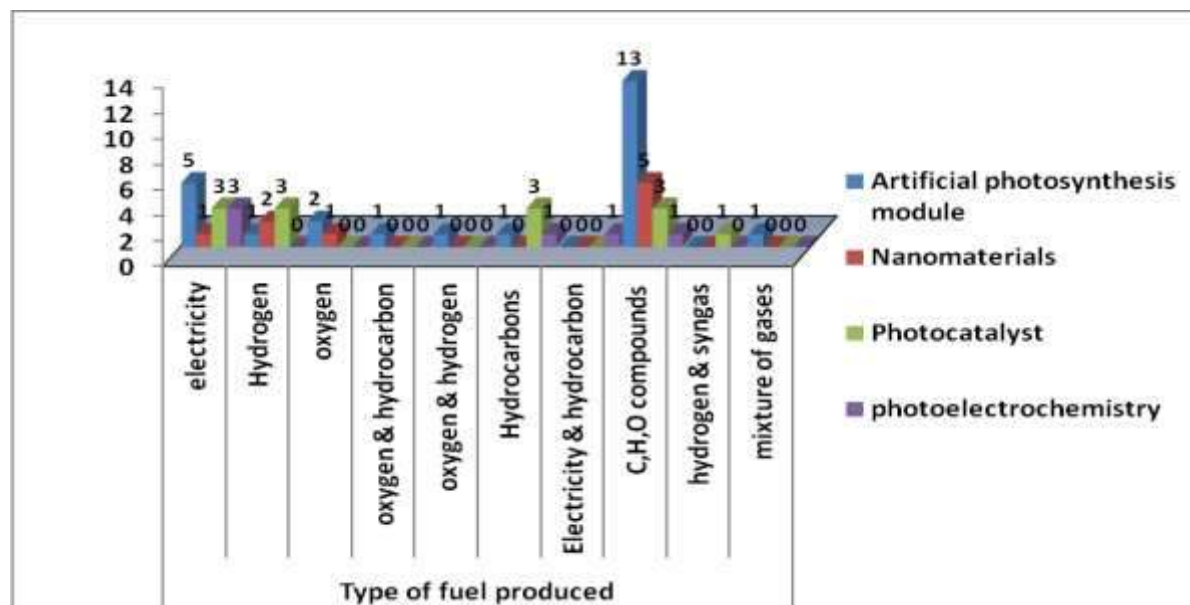
The second most important technology is that of “Photocatalyst”, which is disclosed in

13 patent families (accounting for 25% of the total 53 patent family members). All these patent family employed metal, metalloid and non-metals as catalyst for capturing the sunlight i.e. conversion of sunlight in to different types of fuels. Yet another 9 patent family disclose “nanomaterials” particularly Metal nanoparticles including bismuth, nickel cobalt serves as a charge carrier which enhances the photoefficiency of the solar cells. Semiconductor Nanomaterial including silicon nanowires, Titanium oxide, shee shaped carbon-nitride nanoparticles again acts as efficient catalyst for reduction of carbondioxide. Graphenenanomaterials have major contribution in the field of catalyst of solar energy conversion.

The “photo-electrochemistry” is another area which disclosed in six patent family members (11% of 53 patent family members); which relates to the photo-electrochemical conversion of solar energy into fuel energy with high efficiency by using CO₂ and water as starting materials.

Analysis based on type of fuel produced

Figure 6: Type of fuel produced in various sub-areas of the artificial photosynthesis technology



The obtained 53 patent family member were further sub-grouped on the basis of the type of Fuel produced by each of technology sub-area. Among 53 patents family carbon- hydrogen-oxygen compounds dominated the entire portfolio. As indicated on the graph,

13 out of 53 patent family members refer to the carbon-hydrogen-oxygen compounds under artificial photosynthesis module sub-area. It includes the alcoholic compounds such as methanol, ethanol which are used as additive for fuel or directly as fuel.

Other types of fuels including electricity, hydrogen, oxygen, hydrocarbons, mixture of gases are spread across all four technology subarea. There are 12 patents family which refer to the production of electricity collectively belongs to all four technology subarea. Remarkably, PCT publication WO2009095509 has total twelve family members, the invention related to low concentration solar plant and method for maximizing the electricity production of the photovoltaic modules. Hydrogen is the third leading fuel contributing six patent families. PCT publication WO2007086918 has nine family members across the world generally related to reaction of an electron donor with a composite material for photocatalytic hydrogen production by utilizing 1) a polymer gel;

2) a photocatalyst; and 3) a protein based H₂ catalyst. Hydrogen is one the major fuel product produced during photoelectrochemical decomposition of water.

Syngas and mixture of gas has less contribution to the technology, however these products are utilized for the production of hydrocarbons and alcoholic compounds [5].

Legal status of patents and technology commercialization analysis

Status of patent application is a proxy for assessment of commercialization of technology. Figure 7 indicates the legal status of the patents irrespective of countries they are filed. Out of 53 patent families across the world 39 patent families are in force. There are 14 patent families which are lapsed due to fees related issues or not being commercialized.

Figure 8 shows patent applications in major countries which are in force or dead at the date of conducting that patent search. By analyzing the numbers of patents still in force in these countries, it is possible to get a picture where the applicants consider the biggest markets for the inventions.

Figure 7: Legal status

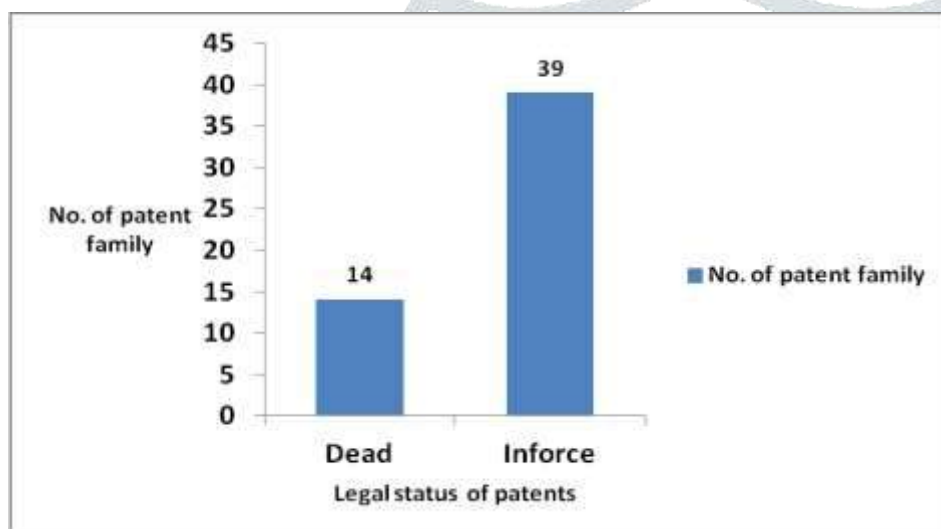


Figure 8: In Force and Lapsed patent applications since 2002

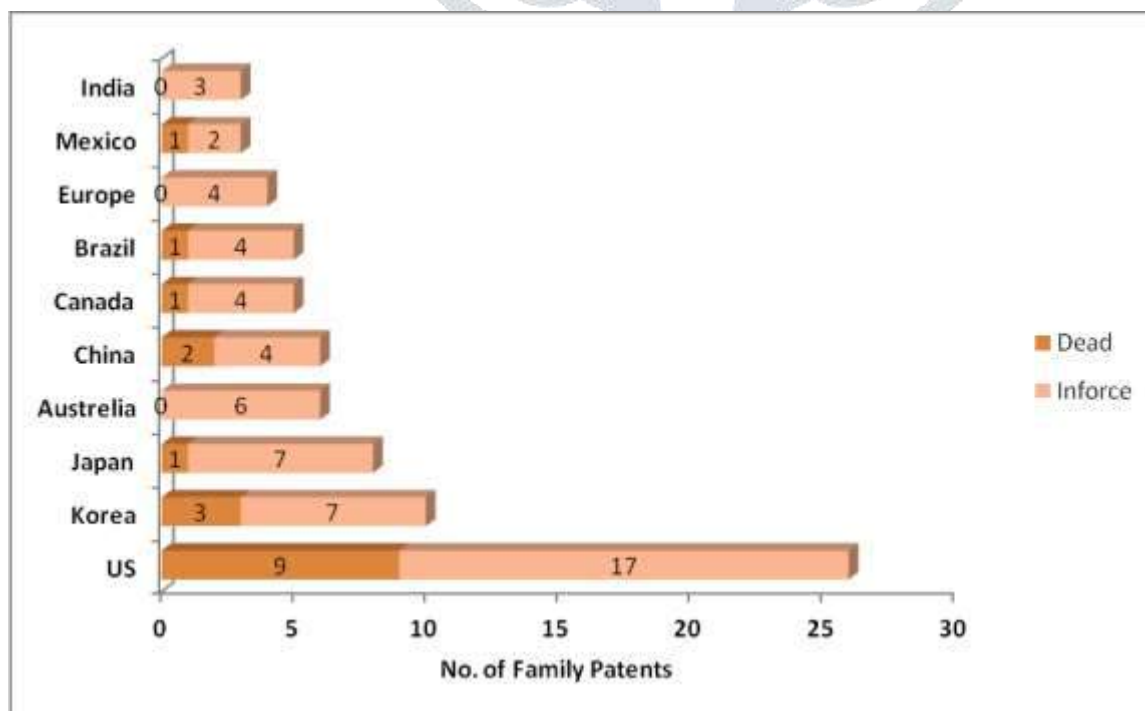


Figure 8 indicates that applicants are primarily protecting their inventions in the United States, Australia and Korea. Overall almost all major filing countries have kept their patents in force. As artificial Photosynthesis is an emerging technology therefore, it is challenging to move laboratory prototype systems to commercial technologies. However, if this new route of producing fuels is achieved on a large scale, it would transform our sustainable energy options by providing alternatives to oil, gas and coal as sources of fuel for transport, industry and electricity generation [6].

Scientists and engineers are working together on significant scientific and technological challenges to successfully scale up laboratory prototypes to a commercial scale [7]. PCT application number PCT/US2006/013742 having priority from US is a single patent family which is commercialized till date. It has nine patent family members across the world. The patent is related to the reduction process for converting CO₂ into methanol with the help of Nobel Laureate George A Olah (Inventor) [8]. The company CRI (Carbon Recycling International) George Olah Plant is located at Iceland. It has capacity of producing 5 million Liters of Methanol per year in the year 2015-2016. The recycling capacity of the plant is 5.5 thousand Cubic Meters of carbon dioxide per year. However, it took around 10 years to commercialize the technology [2006-2015(calculated from filling date)]. Further analysis revealed that there is no licensee for this technology [9].

Korea is second major filing country. The Korean Centre for Artificial Photosynthesis (KCAP) was launched at Sogang University in 2009, Funded by the National Research Foundation of Korea, it is working on materials development for photo-electrocatalysis and the development of systems for artificial photosynthesis, thus have both basic research and technical development with the goal of commercial exploitation[10]

There are programs dedicated to artificial photosynthesis research and innovation in the United States, the Netherlands and South Korea, as well as renewable energy research centers in China and Japan.

Recently, the US Department of Energy (DOE) set up three Energy Innovation Hubs, including a solar research facility. In July 2010, the DOE awarded \$122 million dollars over five years to researchers at California Institute of Technology in partnership with the DOE's Lawrence Berkeley National Laboratory for the formation of the Joint Center for Artificial Photosynthesis (JCAP). The preliminary objective of JCAP is to rapidly move artificial photosynthetic research out of the laboratory setting for commercial exploitation

[11].

There is also evidence of some industries like RELIANCE INDUSTRY, SUNDROP FUELS investing in artificial photosynthesis fuels research and innovation [12,13].

Major limitations on the practical application and commercial use of approaches to artificial Photosynthesis technology

1. Scalability of the invention is major obstacle to successful implementation of technology on commercial scale specifically with regard to photoelectrochemistry (electrolysers).
2. The area required for "light harvesting" is second major challenge.
3. On a larger scale, difficulties to store, distribute, and use hydrogen.
4. Higher cost of development and design of architectures which can be used for scaling- up from the small scale to large scale [14]

Conclusion

This report uses the scale and intensity of patent activity related to synthesis of fuel from sunlight and carbon dioxide using artificial photosynthesis technologies to provide an overview of innovation in the area. The report analysis on 53 patent families related to artificial photosynthesis, is concluded as below

1. Patent activity in the field has increased since the year 2011. A noticeable increase in patent families in the technology sub area of Artificial photosynthesis module and Photocatalyst has largely contributed to this.
2. Almost half of the patent families were in the category of Artificial photosynthesis module primarily on production carbon hydrogen and oxygen type of compounds.
3. In the area of Photocatalyst, the major technology focus was in improving production of hydrogen, hydrocarbon and carbon hydrogen and oxygen type of fuels with 13 patent families.
4. By identifying the origin of the applicants of patent families, the jurisdictions providing the majority of contributions to the field of study were determined. The majority of contributions arise from the United States of America, Korea, Japan and Australia. However, only one patent family originates from Israel, Spain, Russia and Peru.
5. This report identified few collaboration in the patent area includes collaboration between COUNCIL OF SCIENTIFIC & INDUSTRIAL RESEARCH and RELIANCE INDUSTRIES. The major patent filling assignees such as HAIPEIER SOLAR ENERGY HYPERSOLARSAEED MUBIAN UNIVERSITY OF CALIFORNIA and KOREA ADVANCED INSTITUTE OF SCIENCE & TECHNOLOGY found to exhibit a collaboration using the methods of this study.
6. The major markets are the United States of America and Korea, these countries had the highest proportion of patents still in force.

7. Carbon-Hydrogen_Oxygen containing fuel contributes more to the technology followed by electricity and hydrogen. However, very few patents were commercialized at industry level.

References

- [1] Photosynthetic energy conversion: natural and artificial, J. Barber, Chemical Society Review, 2009, 38, 185-196.
- [2] The latest state-of-the-art on artificial photosynthesis, Review article by Ibram Ganesh published in Chem Xpress 3, 131-148, (2014)
- [3] https://www.wipo.int/edocs/pubdocs/en/wipo_pub_946.pdf
- [4] About the Patent Cooperation Treaty (PCT) application:<http://www.wipo.int/pct/en/> [5] <https://en.wikipedia.org/wiki/Syngas>
- [6] Artificial Photosynthesis article by OhannesMessinge, Osamu Ishitani, and Dunwei Wang published in Sustainable Energy Fuels, 2018, 2, 1891–1892. [7] www.sciencedaily.com
- [8] <http://www.carbonrecycling.is/george-olah>
- [9] <https://portal.uspto.gov/pair/PublicPair>
- [10] G.Centi, S.Perathoner; Greenhouse Gas Sci. Technol., 1, 21 (2011).
- [11] Artificial Photosynthesis article published in The National Petroleum Council (NPC), 2012.
- [12] <http://www.sundropfuels.com/About%20Us/about>
- [13] Advances in renewable energy at Reliance Industries article Ajit Sapre published in 8th International Conference on Biofuels, Bioenergy & Bioeconomy, 2017 at Brazil
- [14] Powering the planet: Chemical challenges in solar energy utilization, N. S. Lewis and D. G. Nocera, PNAS Vol. 103, no 43, 15729, 2006.



SYNTHESIS AND ANTIMICROBIAL STUDIES OF NEWLY SYNTHESIZED 1-SUBSTITUTED-3-SUBSTITUTED PROPANE-1, 3-DIONES

Sushil Pagariya*¹, Rajendra Pathade¹, Ram Isankar¹, Pravin Bodkhe²

¹Department of Chemistry, Govt. Vidarbha Institute of Science and Humanities, Amravati, India

²Department of Chemistry, Vidyabharati Mahavidyalaya, Amravati, India

*Corresponding author: sushilpagariya@yahoo.co.in

ABSTRACT

In this current work, one new series of 1-Substituted-3-substitutedpropane-1, 3-diones or β -diketones 4(a-e) have been synthesized from 4-Hydroxy-3-methoxybenzaldehyde i.e. Vanillin. The structures of titled synthesized compounds of the series have been confirmed by usual chemical characteristics, elemental analysis, IR and NMR spectral studies. They have also been studied for their antimicrobial effects against the growth response of pathogenic microorganisms viz. *E. coli*, *S. aureus*, *A. flavus* through agar diffusion method.

Keywords: Propane-1, 3-diones, Vanillin, Spectral studies, Antimicrobial effects.

1. INTRODUCTION

Propane-1, 3-diones which are commonly referred as β -diketones are one of the important classes of organic compounds frequently encountered in synthetic chemistry [1-3]. They are significant intermediates not only as key building blocks for synthesis of core heterocycles such as pyrazole, isoxazole, triazole, flavone, benzodiazepine and pyrimidine [4-9] in medicinal chemistry, but also as an invaluable chelating ligand for various metals of transition and lanthanide series in material chemistry [10].

Aside from their synthetic importance, β -diketones have showed wide assortment of pharmacological activities such as antibacterial [11], antiviral [12], systematic insecticidal [13], antioxidant [14], prophylactic antitumor [15] as well as an anti-sunscreen agent that filters harmful ultra-violet (UV) radiation to protect skin [16]. Furthermore, β -diketones have examined as breast cancer chemo-preventive blocking agent [17], antiestrogenic [18] and anticarcinogenic [19] agent. In addition, β -diketones are well known to have a keto-enol tautomerism [20] and recently it has been reported that β -keto-enols are important pharmacophore for HIV-I integrase (IN) inhibitor [21].

Presence of such varying pharmacological activities in these molecules developed our interest to synthesize some new β -diketone molecules containing phenolic as

well as aldehydic group. With this view here, we have synthesized a new series of 1-Substituted-3-substituted derivatives of propane-1, 3-dione comprising moiety of vanillin, characterized them by usual chemical characteristics, elemental analysis and spectral techniques as well as investigated its antimicrobial activities through method of agar diffusion.

2. EXPERIMENTAL

All chemicals and solvents used during research were of highest purity purchased commercially from Merck, S.D. Fine and Alfa Aesar Company Ltd. The melting points of all the synthesized compounds were recorded by Thiele's melting point apparatus as uncorrected values. Elemental analysis was carried out on Thermo Scientific (Flash 2000) CHNS elemental analyzer. IR spectra were recorded over Shimadzu IRAffinity-1 instrument by means of KBr pallet. ¹H NMR spectra were scanned by Bruker Avance-II at 400 MHz using DMSO-d₆ as solvent and Trimethylsilane as an internal reference. ¹³C NMR spectrum of one sample (4a) was recorded on same instrument at 100 MHz.

2.1. Experimental process for synthesis of 1-Substituted-3-substitutedpropane-1,3-diones 4(a-e) [22]:

Synthesis of above titled compounds involves following preparatory steps:

2.1.1. Preparation of 4-Formyl-2-methoxyphenyl acetate (1)

Initially Vanillin (a) was refluxed with acetic anhydride along with sodium acetate for 1 hr. The reaction mixture was cooled and poured over crushed ice by which two distinct layers were formed. The lower organic layer was separated by separating funnel, washed number of times by distilled water and purified by distillation to get 4-Formyl-2-methoxyphenyl acetate (1). M.p. 78-80°C, Yield 86%.

2.1.2. Preparation of 5-Formyl-2-hydroxy-3-methoxyacetophenone (2)

4-Formyl-2-methoxyphenyl acetate (1) and anhydrous AlCl₃ (1:3) were heated in an oil bath at 120°C for 1 hr (Fries rearrangement). The cooled reaction mixture was decomposed by 10% ice cold HCl to form crude ketone which was then purified by dissolving it in glacial acetic acid followed by pouring the solution slowly in ice cold distilled water with continuous stirring to get 5-Formyl-2-hydroxy-3-methoxyacetophenone (2). M.p. 102-104°C, Yield 78%.

2.1.3. Preparation of 2-(Substitutedbenzoyloxy)-5-formyl-3-methoxyacetophenones 3(a-e)

0.04 mol 5-Formyl-2-hydroxy-3-methoxyacetophenone (2) and 0.05 mol corresponding substituted benzoic acid were dissolved in pyridine and POCl₃ was dropwise added with constant stirring below 10°C of temperature. The reaction mixture was kept at room temperature for overnight and then decomposed by ice cold 10% HCl solution. The solid product was separated, filtered, washed with distilled water followed by 10% NaHCO₃

washing and again with plenty of distilled water. Finally, they were recrystallized by hot ethanol to obtain 2-(Substitutedbenzoyloxy)-5-formyl-3-methoxyacetophenones 3(a-e) as:

2-(4'-Nitrobenzoyloxy)-5-formyl-3-methoxyacetophenone (3a), M.p.126-130°C, Yield 80%.

2-(4'-Methoxybenzoyloxy)-5-formyl-3-methoxyacetophenone (3b), M.p.116-120°C, Yield 71%.

2-(2'-Chlorobenzoyloxy)-5-formyl-3-methoxyacetophenone (3c), M.p.110-114°C, Yield 75%.

2-(4'-Chlorobenzoyloxy)-5-formyl-3-methoxyacetophenone (3d), M.p.124-128°C, Yield 76%.

2-(2',4'-Dichlorobenzoyloxy)-5-formyl-3-methoxyacetophenone (3e), M.p.134-136°C, Yield 84%.

2.1.4. Preparation of 1-(5'-Formyl-2'-hydroxy-3'-methoxyphenyl)-3-(substitutedphenyl)propane-1, 3-diones or β-diketones 4(a-e) via Baker-Venkataraman Rearrangement

0.05 mol 2-(Substitutedbenzoyloxy)-5-formyl-3-methoxyacetophenone 3(a-e) was dissolved in 40 ml of pyridine. The resulting solution was warmed up to 60°C and pulverized KOH was added gradually with continuous stirring. After 6-8 hrs, the reaction mixture was acidified by ice cold dilute HCl solution (1:1). The solid product was separated, filtered, washed with 10% solution of NaHCO₃ and then several times with distilled water. Finally, it was recrystallized from mixture of ethanol and acetic acid to get respective β-diketones namely 1-(5'-Formyl-2'-hydroxy-3'-methoxyphenyl)-3-(substitutedphenyl)propane-1,3-diones 4(a-e). The general experimental scheme for synthesis of above titled compounds is depicted in Fig. 1.

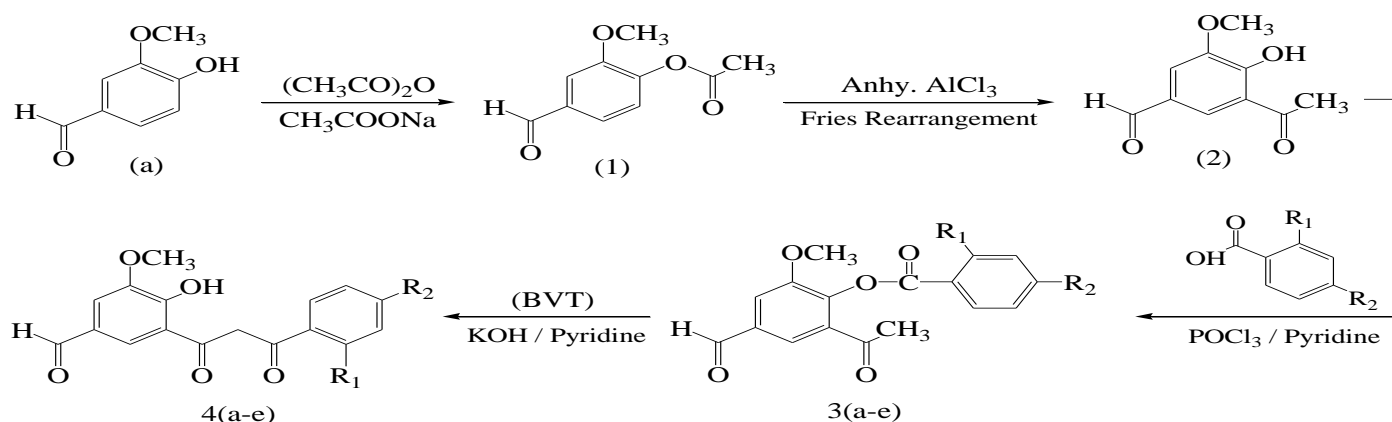


Fig. 1: Experimental scheme for synthesis of 1-(5'-Formyl-2'-hydroxy-3'-methoxyphenyl)-3-(substitutedphenyl)propane-1, 3-diones 4(a-e)

2.2. Antimicrobial study

In this section, all newly synthesized 1-(5'-Formyl-2'-hydroxy-3'-methoxyphenyl)-3-(substitutedphenyl)propane-1, 3-diones 4(a-e) were screened for their antimicrobial activities by method of Agar diffusion [23-24] in order to investigate their effects against growth response of pathogenic microorganisms *E. coli*, *S. aureus* and *A. flavus* at six concentrations ranging from 25 µg/ml to 1000 µg/ml. To prepare the solutions of above concentrations DMSO was used as solvent. Nutrient-agar and Czapek-Dox media were utilized respectively for antibacterial and antifungal analysis as well as standard Ciprofloxacin and Amphotericin drugs were utilized for the purpose of comparison.

2.2.1. Antibacterial analysis

First of all, the bacterial stock cultures were revived by inoculation in broth media and allowed to grown at 37°C for about 18 hrs. The plates of agar of above media were prepared and wells or holes were prepared in the plates. Each plate was inoculated with 18 hrs old cultures [100 µl, 10⁴ CFU] and spread evenly on the plate. Afterward, the wells or holes were filled with solutions of different concentration of compounds and standard drugs. All the plates were incubated at the temperature of 37°C for at least 24 hrs and zones of inhibition were measured as diameter (in mm).

2.2.2. Antifungal analysis

First of all, the fungal stock culture was revived by inoculation in broth media and allowed to grown at 27°C for about 48 hrs. The plates of agar of above media were prepared and wells or holes were prepared in the plates. Each plate was inoculated with 48 hrs old cultures [100 µl, 10⁴ CFU] and spread evenly on plate. Afterward, the wells or holes were filled with solutions of different concentrations of compounds and standard drugs. All the plates were incubated at 27°C temperature for near about 96 hrs and inhibition zones were measured as diameter (in mm).

3. RESULTS AND DISCUSSION

3.1. Spectroscopic data

The spectral data of IR, ¹³C NMR and ¹H NMR showed expected peaks or signals which correspond to various groups present in respective compound. Also, data on elemental analysis was found in full agreement with the proposed structure of compounds. The physical and

spectroscopic data of newly synthesized compounds 4(a-e) are summarized below.

3.1.1. 1-(5'-Formyl-2'-hydroxy-3'-methoxyphenyl)-3-(4'-nitrophenyl)propane-1, 3-dione (4a):

Brown solid; Yield 76%; M.p. 180-184°C; Elemental Anal. Calcd. for C₁₇H₁₃NO₇: C, 59.48; H, 3.82; O, 32.62. Found: C, 59.41; H, 3.75; O, 32.56. IR Spectra (KBr cm⁻¹): 3120 (phenolic OH), 2980 (aromatic C-H), 2850 (aliphatic C-H), 1695 (C=O), 1520 (aromatic C=C), 1285 (C-N). ¹H NMR Spectra (400MHz, DMSO-d₆, δ, TMS=0): 2.51 (s, 3H, -OCH₃), 3.43 (s, 2H, -CH₂), 3.80 (s, 1H, -OH), 8.15-8.32 (m, 6H, Ar-H), 13.54 (s, 1H, -CHO). ¹³C NMR Spectra (100MHz, DMSO-d₆, δ, TMS=0): 40 (-CH₂), 190 (-CHO), 195 (C=O), 123-136 (Ar-C), 149-165 (C=C).

3.1.2. 1-(5'-Formyl-2'-hydroxy-3'-methoxyphenyl)-3-(4'-methoxyphenyl)propane-1,3-dione(4b):

Dark brown solid; Yield 68%; M.p. 140-144°C; Elemental Anal. Calcd. for C₁₈H₁₆O₆: C, 65.85; H, 4.91; O, 29.24. Found: C, 65.72; H, 4.87; O, 29.03. IR Spectra (KBr cm⁻¹): 3340 (phenolic OH), 2980 (aromatic C-H), 2910 (aliphatic C-H), 1615 (C=O), 1515 (aromatic C=C), 1260 (C-O). ¹H NMR Spectra (400MHz, DMSO-d₆, δ, TMS=0): 2.51 (s, 3H, -OCH₃), 3.81 (s, 2H, -CH₂), 5.13 (s, 1H, -OH), 6.99-7.90 (m, 6H, Ar-H), 10.10 (s, 1H, -CHO).

3.1.3. 1-(5'-Formyl-2'-hydroxy-3'-methoxyphenyl)-3-(2'-chlorophenyl)propane-1,3-dione (4c):

Dark brown solid; Yield 72%; M.p. 133-136°C; Elemental Anal. Calcd. for C₁₇H₁₃ClO₅: C, 61.36; H, 3.94; O, 24.04. Found: C, 61.28; H, 3.87; O, 24.0. ¹H NMR Spectra (400MHz, DMSO-d₆, δ, TMS=0): 2.50 (s, 3H, -OCH₃), 3.38 (s, 2H, -CH₂), 4.10 (s, 1H, -OH), 6.77-7.65 (m, 6H, Ar-H), 9.88 (s, 1H, -CHO).

3.1.4. 1-(5'-Formyl-2'-hydroxy-3'-methoxyphenyl)-3-(4'-chlorophenyl)propane-1,3-dione (4d):

Brown solid; Yield 77%; M.p. 145-148°C; Elemental Anal. Calcd. for C₁₇H₁₃ClO₅: C, 61.36; H, 3.94; O, 24.04. Found: C, 61.30; H, 3.83; O, 24.0. IR Spectra (KBr cm⁻¹): 3340 (phenolic OH), 2970 (aromatic C-H), 2885 (aliphatic C-H), 1590 (C=O), 1385 (C-O), 875 (C-Cl). ¹H NMR Spectra (400MHz, DMSO-d₆, δ, TMS=0): 2.52 (s, 3H, -OCH₃), 3.36 (s, 2H, -CH₂), 4.30 (s, 1H, -OH), 7.23-7.74 (m, 6H, Ar-H), 10.9 (s, 1H, -CHO).

3.1.5. 1-(5'-Formyl-2'-hydroxy-3'-methoxyphenyl)-3-(2',4'-dichlorophenyl)propane-1,3-dione(4e):

Yellow-brown solid; Yield 80%; M.p. 168-170°C; Elemental Anal. Calcd. for C₁₇H₁₂Cl₂O₅: C, 55.61; H, 3.29; O, 21.79. Found: C, 55.58; H, 3.21; O, 21.67. ¹H NMR Spectra (400MHz, DMSO-d₆, δ, TMS=0): 2.51 (s, 3H, -OCH₃), 3.36 (s, 2H, -CH₂), 4.15 (s, 1H, -OH), 7.48-7.83 (m, 6H, Ar-H), 13.54 (s, 1H, -CHO).

3.2. Antimicrobial activity

In the present work, total five 1-Substituted-3-substituted derivatives of propane-1, 3-dione 4(a-e) were synthesized, purified by recrystallization and used individually to investigate their antimicrobial effects against pathogenic microorganisms viz. *E. coli*, *S. aureus* and *A. flavus*. The resulting data on antimicrobial activity of newly synthesized compounds 4(a-e) and standard drugs against *E. coli*, *S. aureus* and *A. flavus* with zone of inhibition in mm are tabulated in Table-(1-4) and their photographs are shown under Fig.-(2-4) respectively. From the results on antimicrobial activities, it was observed that, out of all these compounds 4(a-e), compound (4b) and (4c) were not showed any inhibition

zones against *E. coli* at all the tested concentrations, while compounds (4a), (4d) and (4e) has showed (3,5,8), (3,3,5) and (3,6,8) mm zones of inhibition at concentrations of 250, 500 and 1000µg/ml respectively. The minimum inhibitory concentration (MICs) at which these compounds (4a), (4d) and (4e) showed inhibition against *E. coli* was at 250 µg/ml but in case of compounds (4b) and (4c) for *E. coli*, MICs was not found. In case of *S. aureus*, compound (4a) showed (3, 5, 10, 11) mm zones of inhibition at 100, 250, 500 and 1000 µg/ml concentrations respectively with a MICs at 100 µg/ml. Compounds (4b) and (4d) showed 4 and 3 mm of inhibition zones at 1000 µg/ml concentration only while compound (4e) showed (3, 5, 7) mm of inhibition zones at 250, 500 and 1000 µg/ml respectively with MICs at 250 µg/ml. The compound (4c) not showed any inhibition zones for *S. aureus* at all the analyzed range of concentrations. The results on antifungal activity were shocked us because all these newly synthesized compounds 4(a-e) were not showed any zones of inhibitions at any concentrations against fungi *A. flavus*.

Table 1: Antibacterial activity of synthesized compounds 4(a-e) against *E. coli*

Compd. code	25 µg	50 µg	100 µg	250 µg	500 µg	1000 µg	MIC µg
4a	NI	NI	NI	3	5	8	250
4b	NI	NI	NI	NI	NI	NI	NF
4c	NI	NI	NI	NI	NI	NI	NF
4d	NI	NI	NI	3	3	5	250
4e	NI	NI	NI	3	6	8	250

Table 2: Antibacterial activity of synthesized compounds 4(a-e) against *S. aureus*

Compd. code	25 µg	50 µg	100 µg	250 µg	500 µg	1000 µg	MIC µg
4a	NI	NI	3	5	10	11	100
4b	NI	NI	NI	NI	NI	4	1000
4c	NI	NI	NI	NI	NI	NI	NF
4d	NI	NI	NI	NI	NI	3	1000
4e	NI	NI	NI	3	5	7	250

Table 3: Antibacterial activity of Ciprofloxacin drug against *E. coli* and *S. aureus*

Organism	25 µg	50 µg	100 µg	200 µg	400 µg	800 µg	MIC µg
<i>E. coli</i>	18	20	23	26	28	31	25
<i>S. aureus</i>	13	18	21	25	27	34	25

Table 4: Antifungal activity of Amphotericin drug against fungi *A. flavus*

Organism	25 µg	50 µg	100 µg	200 µg	400 µg	800 µg	MIC µg
<i>A. flavus</i>	NI	NI	NI	NI	7	10	400

NI: No Inhibition; NF: MIC not found; MIC: Minimum Inhibitory Concentration

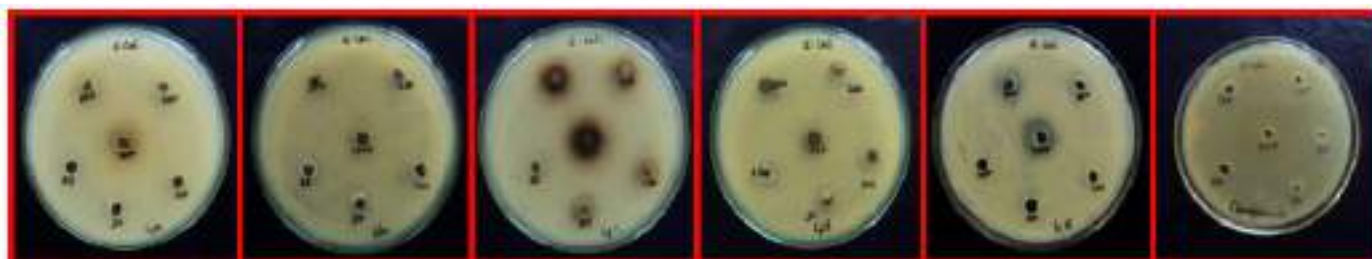


Fig. 2: Effects of synthesized compounds 4(a-e) and std. Ciprofloxacin on the growth response of *E. coli*

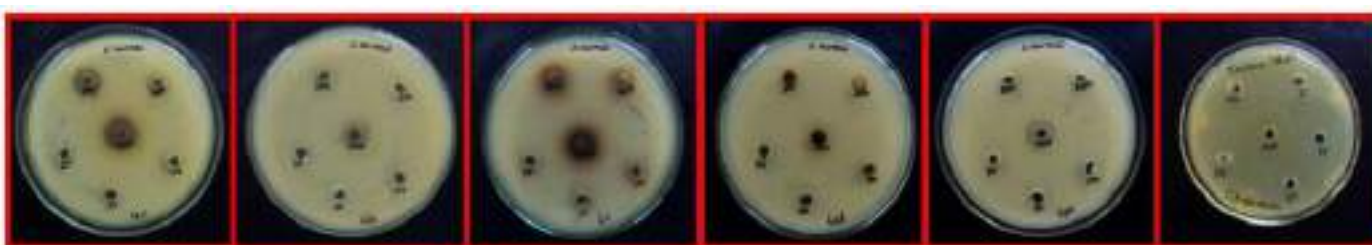


Fig. 3: Effects of synthesized compounds 4(a-e) and std. Ciprofloxacin on the growth response of *S. aureus*



Fig. 4: Effects of synthesized compounds 4(a-e) and std. Amphotericin on the growth response of *A. flavus*

4. CONCLUSION

In conclusion, a new series of 1-Substituted-3-substituted derivatives of propane-1, 3-dione 4(a-e) comprising 4-Hydroxy-3-methoxybenzaldehyde i.e. Vanillin moiety were synthesized successfully in satisfactory yield via Baker-Venkataraman Transformation (BVT) of corresponding substituted 2-benzoyloxyacetophenones 3(a-e) and their structures were elucidated or confirmed by chemical characteristics, elemental analysis and IR, ^{13}C NMR and ^1H NMR spectroscopic techniques. The results on antimicrobial studies reveals that, all the five compounds 4(a-e) were found to have low to moderate antibacterial effects against the growth response of pathogens *E. coli* and *S. aureus* as compared to std. Ciprofloxacin drug but in case of antifungal activity against a pathogen *A. flavus*, they were found to have negligible effects at all analyzed range of concentrations and said to be inactive.

5. ACKNOWLEDGEMENT

The authors thanks to The Director, GVISH, Amravati for providing laboratory facilities. We also thanks to SAIF, Panjab University, Chandigarh and BioGenics, Hubli, Karnataka to facilitate part of the research work.

Conflict of interest

There is no conflict of interest.

6. REFERENCES

1. Fargeas V, Baalouch M, Metay E, Baffreau J, Menard D, Gosselin P et al. *Tetrahedron*, 2004; **60(45)**:10359-10364.
2. Christoffers J, Oertling H, Fischer P, Frey W. *Tetrahedron*, 2003; **59(21)**:3769-3778.
3. Katritzky AR, Wang Z, Wang M, Wilkerson CR, Hall CD, Akhmedov NG. *J Org Chem*, 2004; **69(20)**:6617-6622.
4. Heller ST, Natarajan SR. *Org Lett*, 2006; **8(13)**:2675-2678.

5. Simoni D, Invidiata FP, Rondanin R, Grimaudo S, Cannizzo G, Barbusca E et al. *J Med Chem*, 1999; **42(24)**:4961-4969.
6. Valizadeh H, Amiri M, Khalili E. *Mol Divers*, 2012; **16(2)**:319-323.
7. Tang L, Zhang S, Yang J, Gao W, Cui J, Zhuang T. *Molecules*, 2004; **9(10)**:842-848.
8. Kumar R, Joshi YC. *J Chem Sci*, 2009; **121(4)**:497-502.
9. Kuzueva OG, Burgart YV, Saloutin VI, Chupakhin ON. *Chem Heterocycl Comp*, 2001; **37(9)**:1130-1135.
10. Garnovskii AD, Kharisov BI, Blanco LM, Garnovskii DA, Burlov AS, Vasilchenko IS, Bondarenko GI. *J Coord Chem*, 1999; **46(4)**:365-395.
11. Bennett I, Broom NJP, Cassels R, Elder JS, Masson ND, O'Hanlon PJ. *Bioorg Med Chem Lett*, 1999; **9(13)**:1847-1852.
12. Diana GD, Carabateas PM, Johnson RE, Williams GL, Pancic F, Collins JC. *J Med Chem*, 1978; **21(9)**:889-894.
13. Crouse GD, McGowan MJ, Boisvenue RJ. *J Med Chem*, 1989; **32(9)**:2148-2151.
14. Sugiyama Y, Kawakishi S, Osawa T. *Biochem Pharmacol*, 1996; **52(4)**:519-525.
15. Acton N, Brossi A, Newton DL, Sporn MB. *J Med Chem*, 1980; **23(7)**:805-809.
16. Andrae I, Bringham A, Bohm F, Gonzenbach H, Hill T, Mulroy L, Truscott TG. *J Photochem Photobiol B: Biol*, 1997; **37(1-2)**:147-150.
17. Singletary K, MacDonald C, Iovinelli M, Fisher C, Wallig M. *Carcinogenesis*, 1998; **19(6)**:1039-1043.
18. Lin CC, Tsai YL, Huang MT, Lu YP, Ho CT, Tseng SF, Teng SC. *Carcinogenesis*, 2006; **27(1)**:131-136.
19. Lin CC, Wei GJ, Huang MT, Ho CT. *J Food Drug Anal*, 2005; **13(3)**:284-288.
20. Dziemboska T, Rozwadowski Z. *Curr Org Chem*, 2001; **5(3)**:289-313.
21. Tchertanov L, Mouscadet JF. *J Med Chem*, 2007; **50(6)**:1133-1145.
22. Angaitkar JN, Bodkhe PS. *Int J Rec Scient Res*, 2016; **7(12)**:14549-14553.
23. Threlfall EJ, Fisher IS, Ward LR, Tschape H, Gerner-Smidt P. *Microb Drug Resist*, 1999; **5(3)**:195-200.
24. Prescott JF, Baggot JD, Walker RD. Antimicrobial susceptibility testing and interpretation of results. In: Antimicrobial therapy in veterinary medicine. 3rd ed. Ames, IA: Iowa State University Press; 2000. p.12-26.

A Short Review on Biomedical Signals

Gajbe R. J.

Department of Electronics, Vidya Bharati Mahavidyalaya, Amravati, M.S., India-444 602

Corresponding Author Mobile No.: +919423649592 Email: rjgajbe[at]gmail.com

Abstract: *This paper is briefly reviewed on types, examples, objective of Biomedical signals, which are observations of physiological activities of organisms, ranging from gene and protein sequences, to neural and cardiac rhythms, to tissue and organ images. Biomedical signal processing aims at extracting significant information from biomedical signals. With the aid of biomedical signal processing, biologists can discover new biology and physicians can monitor distinct illnesses.*

Keywords: Biomedical signals; examples, objectives; difficulties encountered

1. Introduction

The human body is a system which includes the nervous system, the cardiovascular system, and the musculoskeletal system, among others. The cardiac system performs the important task of rhythmic pumping of blood throughout the body to facilitate the delivery of nutrients, as well as pumping blood through the pulmonary system for oxygenation of the blood itself. Every system is consisting of several subsystems that carry on many physiological processes. Physiological processes are complex phenomena, including nervous or hormonal stimulation and control; inputs and outputs that could be in the form of physical material, neurotransmitters, or information; and action that could be mechanical, electrical, or biochemical. Most physiological processes are accompanied by or manifest themselves as signals that reflect their nature and activities. Such signals could be of many types, including biochemical in the form of hormones and neurotransmitters, electrical in the form of potential or current, and physical in the form of pressure or temperature [1-3].

Diseases or defects in a biological system cause alterations in its normal physiological processes, leading to pathological processes that affect the performance, health, and general well-being of the system. A pathological process is typically associated with signals that are different in some respects from the corresponding normal signals. If we possess a good understanding of a system of interest, it becomes possible to observe the corresponding signals and assess the state of the system.

In intensive-care monitoring, the ear drum temperature may sometimes be measured using an infra-red sensor. Occasionally, when catheters are being used for other purposes, a temperature sensor may also be introduced into an artery or the heart to measure the core temperature of the body. It then becomes possible to obtain a continuous measurement of temperature, although only a few samples taken at intervals of a few minutes may be stored for subsequent analysis [4, 5]. Let us now consider another basic measurement in health care and monitoring: that of blood pressure (BP). Each measurement consists of two values - the systolic pressure and the diastolic pressure. BP is measured in milli-meters of mercury (mm of Hg) in clinical practice, although the international standard unit for pressure is the Pascal [6-8].

2. Examples of Biomedical Signals

The preceding example of body temperature as a signal is a rather simple example of a biomedical signal. Regardless of its simplicity, we can appreciate its importance and value in the assessment of the well-being of a child with a fever or that of a critically ill patient in a hospital. The origins and nature of a few other biomedical signals of various types are described in the following subsections, with brief indications of their usefulness in diagnosis.

2.1 The action potential

This is the electrical signal that accompanies the mechanical contraction of a single cell when stimulated by an electrical current. Which cause by the flow of sodium, potassium, chloride and other ions across the cell membrane. It is the basic component of all bioelectrical signals. It provides information on the nature of physiological activity at the single-cell level. Recording an action potential requires the isolation of a single cell, and microelectrodes with tips of the order of a few micrometers to stimulate the cell and record the response [9].

2.2 The Electroneurogram (ENG)

The ENG is an electrical signal observed as a stimulus and the associated nerve action potential propagate over the length of a nerve. It may be used to measure the velocity of propagation (or conduction velocity) of a stimulus or action potential in a nerve [10]. ENG's may be recorded using concentric needle electrodes or silver - silver-chloride electrodes (Ag - AgCl) at the surface of the body [10].

2.3 The Electromyogram (EMG)

Skeletal muscle fibers are considered to be twitch fibers because they produce a mechanical twitch response for a single stimulus and generate a propagated action potential. Skeletal muscles are made up of collections of motor units, each of which consists of an anterior horn cell (or motoneuron or motor neuron), its axon, and all muscle fibers innervated by that axon. A motor unit is the smallest muscle unit that can be activated by volitional effort. The constituent fibers of a motor unit are activated synchronously. Component fibers of a motor unit extend

lengthwise in loose bundles along the muscle. In cross-section, the fibers of a given motor unit are interspersed with the fibers of other motor units [11-12].

2.4 The electrocardiogram (ECG)

The ECG is the electrical manifestation of the contractile activity of the heart, and can be recorded fairly easily with surface electrodes on the limbs or chest. The ECG is perhaps the most commonly known, recognized, and used biomedical signal. The rhythm of the heart in terms of beats per minute may be easily estimated by counting the readily identifiable waves. More important is the fact that the ECG wave shape is altered by cardiovascular diseases and abnormalities such as myocardial ischemia and infarction, ventricular hypertrophy, and conduction problems [13].

2.5 The electroencephalogram (EEG)

The EEG (popularly known as brain waves) represents the electrical activity of the brain. A few important aspects of the organization of the brain are as follows: The main parts of the brain are the cerebrum, the cerebellum, the brain stem (including the midbrain, pons medulla, and the reticular formation), and the thalamus (between the midbrain and the hemispheres). The cerebrum is divided into two hemispheres, separated by a longitudinal fissure across which there is a large connective band of fibers known as the corpus callosum. The outer surface of the cerebral hemispheres, known as the cerebral cortex, is composed of neurons (grey matter) in convoluted patterns, and separated into regions by fissures (sulci). Beneath the cortex lie nerve fibers that lead to other parts of the brain and the body [14-16].

2.6 Event-related potentials (ERPs)

The term **event-related potential** is more general than and preferred to the term **evoked potential**, and includes the ENG or the EEG in response to light, sound, electrical, or other external stimuli. Short-latency ERPs are predominantly dependent upon the physical characteristics of the stimulus, whereas longer-latency ERPs are predominantly influenced by the conditions of presentation of the stimuli.

2.7 The Electrogastrogram (EGG)

The electrical activity of the stomach consists of rhythmic waves of depolarization and repolarization of its constituent smooth muscle cells [17-19]. The activity originates in the mid-corpus of the stomach, with intervals of about 20 s in humans. The waves of activity are always present and are not directly associated with contractions; they are related to the spatial and temporal organization of gastric contractions. External (cutaneous) electrodes can record the signal known as the electrogastrogram (EGG). Chen et al. [20] used the following procedures to record cutaneous EGG signals.

2.8 The Phonocardiogram (PCG)

The heart sound signal is perhaps the most traditional biomedical signal, as indicated by the fact that the stethoscope is the primary instrument carried and used by physicians. The PCG is a vibration or sound signal related to the contractile activity of the cardio hemic system (the heart and blood together) and represents a recording of the heart sound signal. Recording of the PCG signal requires a transducer to convert the vibration or sound signal into an electronic signal: microphones, pressure transducers, or accelerometers may be placed on the chest surface for this purpose. The normal heart sounds provide an indication of the general state of the heart in terms of rhythm and contractility. Cardiovascular diseases and defects cause changes or additional sounds and murmurs that could be useful in their diagnosis [21-25].

2.9 The carotid pulse (CP)

The carotid pulse is a pressure signal recorded over the carotid artery as it passes near the surface of the body at the neck. It provides a pulse signal indicating the variations in arterial blood pressure and volume with each heart beat. Because of the proximity of the recording site to the heart, the carotid pulse signal closely resembles the morphology of the pressure signal at the root of the aorta; however, it cannot be used to measure absolute pressure [21].

2.10 Signals from catheter-tip sensors

For very specific and close monitoring of cardiac function, sensors placed on catheter tips may be inserted into the cardiac chambers. It then becomes possible to acquire several signals such as left ventricular pressure, right atrial pressure, aortic pressure, and intra cardiac sounds. While these signals provide valuable and accurate information, the procedures are invasive and are associated with certain risks [22, 23].

2.11 The speech signal

Human beings are social creatures by nature, and have an innate need to communicate. We are endowed with the most sophisticated vocal system in nature. The speech signal is an important signal, although it is more commonly considered as a communication signal than a biomedical signal. However, the speech signal can serve as a diagnostic signal when speech and vocal-tract disorders need to be investigated [25].

2.12 The Vibroarthrogram (VAG)

Considerable noise is often associated with degeneration of knee-joint surfaces. The VAG is the vibration signal recorded from a joint during movement (articulation) of the joint. Normal joint surfaces are smooth and produce little or no sound, whereas joints affected by osteoarthritis and other degenerative diseases may have suffered cartilage loss and produce grinding sounds. Detection of knee-joint problems via the analysis of VAG signals could help avoid unnecessary exploratory surgery, and also aid

better selection of patients who would benefit from surgery. The VAG signal, however, is not yet well understood, and is a difficult signal to analyse due to its complex non stationary characteristics [26-32].

2.13 Oto-acoustic emission signals

The oto-acoustic emission (OAE) signal represents the acoustic energy emitted by the cochlea either spontaneously or in response to an acoustic stimulus. The discovery of the existence of this signal indicates that the cochlea not only receives sound but also produces acoustic energy. The OAE signal could provide objective information on the micromechanical activity of the preneural or sensory components of the cochlea that are distal to the nerve-fiber endings. Analysis of the OAE signal could lead to improved noninvasive investigative techniques to study the auditory system. The signal may also assist in screening of hearing function and in the diagnosis of hearing impairment [33].

3.Objectives of Biomedical Signal Analysis

The representation of biomedical signals in electronic form facilitates computer processing and analysis of the data.

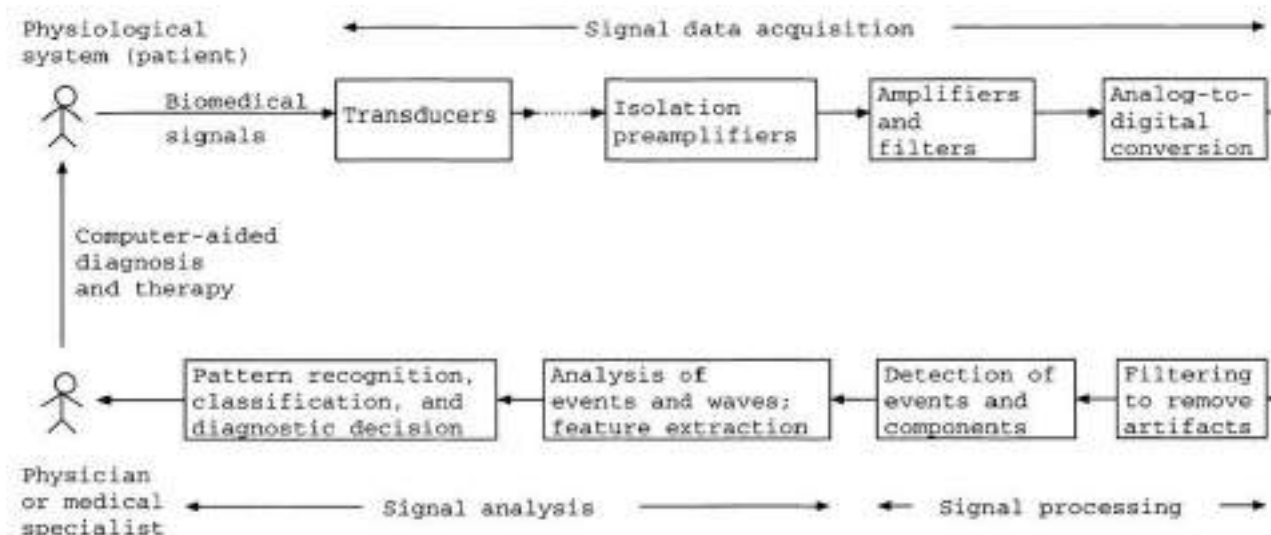


Figure 1.1: Computer-aided diagnosis and therapy based upon biomedical signal analysis

3.1 The human - instrument system:

The components of a **human – instrument system** [10-12] are:

- **The subject or patient:** It is important always to bear in mind that the main purpose of biomedical instrumentation and signal analysis is to provide a certain benefit to the subject or patient.
- **Stimulus or procedure of activity** Application of stimuli to the subject in active procedures requires instruments such as strobe light generators, sound generators, and electrical pulse generator.
- **Transducers:** electrodes, sensors.
- **Signal-conditioning equipment:** amplifiers, filters.
- **Display equipment:** oscilloscopes, strip-chart or paper recorders, computer monitors, printers.

Figure 1.1 illustrates the typical steps and processes involved in computer-aided diagnosis and therapy based upon biomedical signal analysis.

The major objectives of biomedical instrumentation and signal analysis [10-12] are:

- **Information gathering** - measurement of phenomena to interpret a system.
- **Diagnosis** - detection of malfunction, pathology, or abnormality.
- **Monitoring** - obtaining continuous or periodic information about a system.
- **Therapy and control** - modification of the behaviour of a system based upon the outcome of the activities listed above to ensure a specific result.
- **Evaluation** - objective analysis to determine the ability to meet functional requirements, obtains proof of performance, perform quality control, or quantify the effect of treatment.

Signal acquisition procedures may be categorized as being invasive or non invasive, and active or passive.

- **Recording, data processing, and transmission equipment:** analog instrumentation tape recorders, analog-to-digital converters (ADCs), digital-to-analog converters (DACs), digital tapes, compact disks (CDs), diskettes, computers, telemetry systems.
- **Control devices:** power supply stabilizers and isolation equipment, patient intervention systems.

The science of measurement of physiological variables and parameters is known as biometrics. Some of the aspects to be considered in the design, specification, or use of biomedical instruments [13-17] are:

- **Isolation of the subject or patient** - of paramount importance so that the subject is not placed at the risk of electrocution.

- **Range of operation** - the minimum to maximum values of the signal or parameter being measured.
- **Sensitivity** - the smallest signal variation measurable. This determines the resolution of the system.
- **Linearity** - desired over at least a portion of the range of operation. Any nonlinearity present may need to be corrected for at later stages of signal processing.
- **Hysteresis** - a lag in measurement due to the direction of variation of the entity being measured. Hysteresis may add a bias to the measurement, and should be corrected
- **Frequency response** - represents the variation of sensitivity with frequency.
- **Stability** - an unstable system could preclude repeatability and consistency of measurements.
- **Signal-to-noise ratio (SNR)** - power-line interference, grounding problems, thermal noise, and so on, could compromise the quality of the signal being acquired. A good understanding of the signal-degrading phenomena present in the system is necessary in order to design appropriate filtering and correction procedures.
- **Accuracy** - includes the effects of errors due to component tolerance, movement, or mechanical errors; drift due to changes in temperature, humidity, or pressure; reading errors due to, for example, parallax; and zeroing or calibration errors.

4. Difficulties Encountered in Biomedical Signal Acquisition and Analysis

4.1 Accessibility of the variables to measurement:

Most of the systems and organs of interest, such as the cardiovascular system and the brain, are located well within the body (for good reasons!). While the ECG may be recorded using limb electrodes, the signal so acquired is but a projection of the true 3D cardiac electrical vector of the heart onto the axis of the electrodes. Such a signal may be sufficient for rhythm monitoring, but could be inadequate for more specific analysis of the cardiac system.

4.2 Variability of the signal source:

It is evident from the preceding sections that the various systems that comprise the human body are dynamic systems with several variables. Biomedical signals represent the dynamic activity of physiological systems and the states of their constituent variables. The nature of the processes or the variables could be deterministic or random (stochastic); a special case is that of periodicity or quasi-periodicity.

4.3 Inter-relationships and interactions among physiological systems:

The various systems that compose the human body are not mutually independent; rather, they are inter-related and interact in various ways. Some of the interactive phenomena are compensation, feedback, cause-and-effect; collateral effects, loading, and take-over of function of a disabled system or part by another system or part.

4.4 Effect of the instrumentation or procedure on the system:

The placement of transducers on and connecting a system to instruments could affect the performance or alter the behavior of the system, and cause spurious variations in the parameters being investigated. The experimental procedure or activity required to elicit the signal may lead to certain effects that could alter signal characteristics. This aspect may not always be obvious unless careful attention is paid. may need some rest between procedures or their repetitions.

4.5 Physiological artefacts and interference:

One of the pre-requisites for obtaining a good ECG signal is for the subject to remain relaxed and still with no movement. Coughing, tensing of muscles, and movement of the limbs cause the corresponding EMG to appear as an undesired artefact. In the absence of any movement by the subject, the only muscular activity in the body would be that of the heart. When chest leads are used, even normal breathing could cause the associated EMG of the chest muscles to interfere with the desired ECG.

4.6 Energy limitations:

Most biomedical signals are generated at microvolt or millivolt levels at their sources. Recording such signals requires very sensitive transducers and instrumentation with low noise levels. The connectors and cables need to be shielded as well, in order to obviate pickup of ambient electromagnetic (EM) signals. Some applications may require transducers with integrated amplifiers and signal conditioners so that the signal leaving the subject at the transducer level is much stronger than ambient sources of potential interference.

4.7 Patient safety:

Protection of the subject or patient from electrical shock or radiation hazards is an unquestionable requirement of paramount importance. The relative levels of any other risks involved should be assessed when a choice is available between various procedures, and analyzed against their relative benefits. Patient safety concerns may preclude the use of a procedure that may yield better signals or results than others, or require modifications to a procedure that may lead to inferior signals. Further signal-processing steps would then become essential in order to improve signal quality or otherwise compensate for the initial loss.

References

- [1] Lathi BP. Signal Processing and Linear Systems. Berkeley-Cambridge, Carmichael, CA, 1998.
- [2] Oppenheim AV, Willsky AS, and Nawab SH. Signals and Systems. Prentice- Hall, Englewood Cliffs, NJ, 2nd edition, 1997.
- [3] Papoulis A. Signal Analysis. McGraw-Hill, New York, NY, 1977.

- [4] Papoulis A. Probability, Random Variables, and Stochastic Processes. McGraw-Hill, New York, NY, 1965.
- [5] Bendat JS and Piersol AG. Random Data: Analysis and Measurement Procedures. Wiley, New York, NY, 2nd edition, 1986.
- [6] Aulin JI and Chandrasekar V. Introduction to Probability and Random Processes. McGraw-Hill, New York, NY, 1997.
- [7] Ramsey FL and Schafer DW. The Statistical Sleuth - A Course in Methods of Data Analysis. Wadsworth Publishing Company, Belmont, CA, 1997.
- [8] Riffenburgh RH. Statistics in Medicine. Academic, San Diego, CA, 1993.
- [9] Bailar 111 JC and Mosteller F, editors. Medical Uses of Statistics. NEIM Books, Boston, MA, 2nd edition, 1992.
- [10] Webster JG, editor. Medical Instrumentation: Application and Design. Wiley, 489 New York, NY, 3rd edition, 1998.
- [11] Goodgold J and Eberstein A. Electrodiagnosis of Neuromuscular Diseases. Williams and Wilkins, Baltimore, MD, 3rd edition, 1983.
- [12] de Luca CJ. Physiology and mathematics of myoelectric signals. IEEE Transactions on Biomedical Engineering, 26:3 13-325, 1979
- [13] Rushmer RE Cardiovascular Dynamics. WB Saunders, Philadelphia, PA, 4th edition, 1976
- [14] Cooper R, Osselton JW, and Shaw JC. EEG Technology. Butterworths, London, UK, 3rd edition, 1980.
- [15] Kooi KA, Tucker RP, and Marshall RE. Fundamentals of Electroencephalography. Harper & Row, Hagerstown, MD, 2nd edition, 1978.
- [16] Hughes JR. EEG in Clinical Practice. Butterworth, Woburn, MA, 1982.
- [17] Verhagen MAMT, van Schelven LJ, Samsom M, and Smout AJPM. Pitfalls in the analysis of electrogastrographic recordings. Gastroenterology, 117:453-460, 1999.
- [18] Mintchev MP and Bowes KL. Capabilities and limitations of electrogastrograms. In Chen JDZ and McCallum RW, editors, Electrogastrography: Principles and Applications, pages 155-169. Raven, New York, NY, 1994.
- [19] Mintchev MP and Bowes KL. Extracting quantitative information from digital electrogastrograms. Medical and Biological Engineering and Computing, 34:244-248, 1996.
- [20] Chen JDZ, Stewart Jr. WR, and McCallum RW. Spectral analysis of episodic Rhythmic variations in the cutaneous electrogastrogram. IEEE Transactions on Biomedical Engineering, 40(2): 128-135, 1993
- [21] Rangayyan RM and Lehner RJ. Phonocardiogram signal processing: A review. CRC Critical Reviews in Biomedical Engineering, 15(3):211-236, 1988.
- [22] Tavel ME. Clinical Phonocardiography and External Pulse Recording. Year Book Medical, Chicago, IL, 3rd edition, 1978.
- [23] Luisada AA and Portaluppi F. The Heart Sounds - New Facts and Their Clinical Implications. Praeger, New York, NY, 1982.
- [24] Shaver JA, Salerni R, and Reddy PS. Normal and abnormal heart sounds in cardiac diagnosis, Part I: Systolic sounds. Current Problems in Cardiology, 10(3): 1-68, 1985.
- [25] Reddy PS, Salerni R, and Shaver JA. Normal and abnormal heart sounds in cardiac diagnosis, Part II: Diastolic sounds. Current Problems in Cardiology, 10(4):1-55, 1985.
- [26] Frank CB, Rangayyan RM, and Bell GD. Analysis of knee sound signals for non-invasive diagnosis of cartilage pathology. IEEE Engineering in Medicine and Biology Magazine, pages 65-68, March 1990.
- [27] Tavathia S, Rangayyan RM, Frank CB, Bell GD, Ladly KO, and Zhang YT. Analysis of knee vibration signals using linear prediction. IEEE Transactions on Biomedical Engineering, 39(9):959-970, 1992.
- [28] Moussavi ZMK, Rangayyan RM, Bell GD, Frank CB, Ladly KO, and Zhang YT. Screening of vibroarthrographic signals via adaptive segmentation and linear prediction modeling. IEEE Transactions on Biomedical Engineering, 43(1): 15-23, 1996.
- [29] Krishnan S, Rangayyan RM, Bell GD, Frank CB, and Ladly KO. Adaptive filtering, modelling, and classification of knee joint vibroarthrographic signals for non-invasive diagnosis of articular cartilage pathology. Medical and Biological Engineering and Computing, 35(6):677-684, 1997.
- [30] Rangayyan RM, Krishnan S, Bell GD, Frank CB, and Ladly KO. Parametric representation and screening of knee joint vibroarthrographic signals. IEEE Transactions on Biomedical Engineering, 44(1 1): 1068-1074, 1997.
- [31] Kernohan WG, Beverland DE, McCoy GF, Hamilton A, Watson P, and Mollan RAB. Vibration arthrometry. Acta Orthopædica Scandinavia, 61 (1):70-79, 1990.
- [32] Chu ML, Gradisar IA, and Mostardi R. A noninvasive electroacoustical evaluation technique of cartilage damage in pathological knee joints. Medical and Biological Engineering and Computing, 16:437-442, 1978.
- [33] Probst R, Lonsbury-Martin B, and Martin GK. A review of otoacoustic emissions. Journal of the Acoustical Society of America, 89(5):2027-2067, 1991

Importance of Artificial Intelligence in Healthcare: A Review

Gajbe R. J.

Department of Electronics

Vidya Bharati Mahavidyalaya, Amravati M.S. India-444 602

Abstract

Artificial intelligence (AI) has been developing very rapidly in the various area of biomedical software's algorithms, hardware implementation, and etc. In this review, we summarize the importance of AI, role of AI, in the various branches of biomedical research. The aim of this review is to keep track of new scientific accomplishments, to understand the availability of technologies, to appreciate the tremendous potential of AI. New progress and breakthroughs will continue to push the frontier and widen the scope of AI application, and fast developments are envisioned in the near future.

Keywords: AI, Necessity, Role, Importance

1. Introduction:

Artificial Intelligence (AI), where computers perform tasks that are usually assumed to require human intelligence, is currently being discussed in nearly every domain of science and engineering. Major scientific competitions like Image Net Large Scale Visual Recognition Challenges are providing evidence that computers can achieve human-like competence in image recognition. AI has also enabled significant progress in speech recognition and natural language processing. All of these advances open questions about how such capabilities can support, or even enhance, human decision making in health and health care. Two recent high-profile research papers have demonstrated that AI can perform clinical diagnostics on medical images at levels equal to experienced clinicians, at least in very specific examples [1,2].

Artificial intelligence (AI) is defined as the intelligence of machines, as opposed to the intelligence of humans or other living species. AI can also be defined as the study of "intelligent agents"—that is, any agent or device that can perceive and understand its surroundings and accordingly take appropriate action to maximize its chances of achieving its objectives. AI also refers to situations wherein machines can simulate human minds in learning and analysis, and thus can work in problem solving. This kind of intelligence is also referred to as machine learning (ML). Typically, AI involves a system that consists of both software and hardware. From a software perspective, AI is particularly concerned with algorithms. An artificial neural network (ANN) is a conceptual framework for executing AI algorithms. It is a mimic of the human brain—an interconnected network of neurons, in

which there are weighted communication channels between neurons. One neuron can react to multiple stimuli from neighboring neurons and the whole network can change its state according to different inputs from the environment [3-9].

The role of artificial intelligence in healthcare has been a huge talking point in recent months and there's no sign of the adoption of this technology slowing down, well, ever really. AI in healthcare has huge and wide reaching potential with everything from mobile coaching solutions to drug discovery falling under the umbrella of what can be achieved with machine learning. That being said, many healthcare executives are still too shy when it comes to experimenting with AI due to privacy concerns, data integrity concerns or the unfortunate presence of various organizational silos making data sharing next to impossible. However, the future of healthcare & the future of machine learning and artificial intelligence are deeply interconnected.

2. The Necessity of Artificial Intelligence:

"Artificial intelligence will be an important tool to help us all end the injustice in the uneven distribution of healthcare, and to make it more accessible and affordable for every person on Earth."

The use of Artificial Intelligence, or AI, is growing rapidly in the medical field, especially in diagnostics and management of treatment. To date there has been a wide range of research into how AI can aid clinical decisions and enhance physicians' judgement. In recent years, AI and machine learning have emerged as powerful tools for assisting diagnosis. This technology could revolutionise healthcare by providing more precise diagnoses.

The importance of AI as a component of the diagnostic process has been steadily increasing since building systems became more practical. There is ongoing enthusiasm for and hype about AI and both researchers and practitioners focus equally on this technology from multiple perspectives. There is no uniform definition for the term AI, but considered as “the ability of a machine to perform cognitive functions that we associate with human minds, such as perceiving, reasoning, learning, interacting with the environment, problem solving, decision-making, and even demonstrating creativity”. AI is generally associated with human-like behavior and covers a wide range of research areas, such as natural language processing or robotics. However, current practical applications, including healthcare and disease diagnostics, are narrowed down to a specific task, and are being developed using machine learning. Algorithms exploit medical data to generate predictions and continuously learn and develop over time by constantly processing new and updated data. Algorithms acquire information through different types of knowledge and input or over multiple years of experience. Therefore, AI empowered systems are able to process more knowledge compared with humans, possibly outperforming them for certain medical tasks [10-12].

Role of Artificial Intelligence

The field of healthcare is evolving at an increasing speed, and this is accompanied by a significant increase in the amount of data and challenges in terms of cost and patient outcomes, so AI applications have been used to reduce these challenges. Artificial Intelligence is very useful in solving problematic healthcare challenges and offers a number of advantages over traditional data analytics and clinical decision-making techniques.

The following are the most important examples of the role of artificial intelligence (AI) in healthcare and medicine:

AI in Medical Diagnosis:

Artificial Intelligence has the potential to revolutionize medical diagnostics. Unnecessary routine laboratory testing increases unnecessary financial costs. Therefore, artificial intelligence applications have been used to narrow the circle of laboratory analyzes that the patient may need. AI can detect the presence of early disease as soon as possible as it can automate a large portion of the manual work and speed up the diagnosis process.

Improving Clinical Workflow:

Artificial intelligence is currently being used to efficiently manage workflow and analyze imaging. AI can be used to improve clinical workflow, support better clinical insights, reduce clinical variability, aid in setting study priorities, and minimize physician burnout.

Artificial intelligence has the power to take over the time-consuming task of data input so that clinicians can focus on improving labor utilization, increasing daily productivity and providing the highest quality of care to patients.

Predicting ICU Transfers:

Unplanned transfer of patients to the ICU can have poor outcomes and sometimes even death in patients. Therefore, artificial intelligence has been used to reduce the percentage of these cases, by finding patients with severe cases. As artificial intelligence systems use patients' medical records, laboratory results, and their vital signs to manage patients' condition before it deteriorates, and forcing them to be transferred to the intensive care unit. Artificial intelligence systems can guide clinicians on where to start treatment.

Predicting Hospital Acquired Infections:

Artificial intelligence can standardize the diagnosis of infections with Infection Prevention and Control (IPC) implications, and facilitate the dissemination of IPC expertise. AI provides opportunities to improve diagnosis through objective pattern recognition. Using AI-driven models, clinicians can monitor high-risk patients, predict which patients are most likely to develop central-line infections and intervene to reduce risk.

Developing the Next Generation of Radiology Tools:

Artificial intelligence can help develop the next generation of imaging tools that will provide accurate information and detailed enough to replace the need for tissue samples in some cases. The next generation of artificial intelligence is expected to be more effective in the healthcare system and there will be further

improvements in performance. All of these developments promise to increase accuracy and reduce the number of routine tasks that exhaust time and effort [13-18].

3. Pros & Cons of Artificial Intelligence

- **Provides Real-Time Data**
Real-time analytics can help improve physician-patient relationships. Making vital patient data available through mobile devices can engage patients in their treatments. Mobile alerts can inform doctors and nurses of urgent changes in patient statuses and emergencies.
- **Streamlines Tasks**
For example, intelligent radiology technology is able to identify significant visual markers, saving hours of intense analysis. Other automated systems exist to automate appointment scheduling, patient tracking and care recommendations.
AI essentially allows hospitals to accept a wide array of plans, benefiting potential and existing patients.
- **Saves Time and Resources**
As more vital processes are automated, medical professionals have more time to assess patients and diagnose illness and ailment. AI is accelerating operations to save medical establishments precious productivity hours. In any sector, time equals money, so AI has the potential to save hefty costs.
- **Assists Research**
AI enables researchers to amass large swaths of data from various sources. The ability to draw upon a rich and growing information body allows for a more effective analysis of deadly diseases. Related to real-time data, research can benefit from the wide body of information available, as long as it's easily translated.
- **May Reduce Physician Stress**
Some latest research reports over half of the primary physicians feel stressed from deadline pressures and other workplace conditions. AI helps streamline procedures, automate functions, instantly share data and organize operations, all of which help relieve medical professionals of juggling too many tasks.

4. Limits of Ai In Medicine

- **Needs Human Surveillance**
Although AI has come a long way in the medical world, human surveillance is still essential. For example, surgery robots operate logically, as opposed to empathetically. Health practitioners may notice vital behavioral observations that can help diagnose or prevent medical complications.
- **May Overlook Social Variables**
Patient needs often extend beyond immediate physical conditions. Social, economic and historical factors can play into appropriate recommendations for particular patients. For instance, an AI system may be able to allocate a patient to a particular care center based on a specific diagnosis. However, this system may not account for patient economic restrictions or other personalized preferences.
- **May Lead To Unemployment**
Although AI may help cut costs and reduce clinician pressure, it may also render some jobs redundant. This variable may result in displaced professionals who invested time and money in healthcare education, presenting equity challenges.
- **Inaccuracies are Still Possible**
Medical AI depends heavily on diagnosis data available from millions of cataloged cases. In cases where little data exists on particular illnesses, demographics, or environmental factors, a misdiagnosis is entirely possible. This factor becomes especially important when prescribing particular medicine.
- **Susceptible To Security Risks**
As AI uses data to make systems smarter and more accurate, cyber-attacks will incorporate AI to become smarter with each success and failure, making them more difficult to predict and prevent. Once damaging threats outmaneuver security defences, the attacks will be much more challenging to address [19-22].

References :

1. Gulshan, V., Peng, L., Coram, M., Stumpe, M. C., Wu, D., Narayanaswamy, A., (2016). *Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs*. *Jama*, 316(22), 2402. <http://doi.org/10.1001/jama.2016.17216>

2. Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., & Thrun, S. (2017). *Dermatologist-level classification of skin cancer with deep neural networks*. *Nature*,542(7639),115–118. <http://doi.org/10.1038/nature21056>
3. Minsky M. (1965). *Steps toward artificial intelligence*. *Proc IRE*;49(1):8–30.
4. Weng J, McClelland J, Pentland A, Sporns O, Stockman I, Sur M, (2001).
5. *Autonomous mental development by robots and animals*. *Science*;291 (5504):599–600.
6. Wooldridge M, Jennings NR. Intelligent agents: theory and practice. (1995. *Knowl Eng Rev*;10(2):115–52.
7. Huang G, Huang GB, Song S, You K. (2015). *Trends in extreme learning machines: a review*. *Neural Netw*;61:32–48.
8. Hopfield JJ. (1982). *Neural networks and physical systems with emergent collective computational abilities*. *Proc Natl Acad Sci USA*;79(8):2554–8.
9. Watts DJ, Strogatz SH. (1998). *Collective dynamics of 'small-world' networks*. *Nature*;393(6684):440–2.
10. Zucker RS, Regehr WG. (2002). *Short-term synaptic plasticity*. *Annu Rev Physiol*;64:355–405.
11. Knijnenburg B, Willemsen M. (2016). *Inferring Capabilities of Intelligent Agents from Their External Traits*. *ACM Trans Interact Intell Syst [Internet]*.;6:1–25. <https://doi.org/10.1145/2963106>.
12. Luger E, Sellen A. (2016). *Like Having a Really Bad PA: The Gulf between User Expectation and Experience of Conversational Agents*. *Proc 2016 CHI Conf Hum Factors Comput Syst - CHI '16 [Internet]*. p. 5286–97. <https://doi.org/10.1145/2858036.2858288>.
13. Selz D. (2020). From electronic markets to data driven insights. *Electron Mark Electronic Markets*.;30:57–9. <https://doi.org/10.1007/s12525-019-00393-4>.
14. Miotto R, Wang F, Wang S, Jiang X, Dudley JT. (2018). *Deep learning for healthcare:review, opportunities and challenges*. *Brief Bioinform*;19(6):1236–46.
15. Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, Thrun S. (2017). *Dermatologist level classification of skin cancer with deep neural networks*. *Nature* ;542(7639):115–8.
16. Zhu W, Xie L, Han J, Guo X. (2020). *The application of deep learning in cancer prognosis prediction*. *Cancers*;12(3):603. (Basel).
17. Toma_sev N, Glorot X, Rae JW, (2019). *A clinically applicable approach to continuous prediction of future acute kidney injury*. *Nature* ;572:116–9.
18. Gulshan V, Peng L, Coram M. (2016). *Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs*. *JAMA*;316(22):2402–10.
20. Harutyunyan H, Khachatrian H, Kale DC, (2019). *Multitask learning and benchmarking with clinical time series data*. *Sci Data*;6:96.
21. Nurse E, Mashford BS, Yepes AJ, Kiral-Kornek I, Harrer S, Freestone DR. (2016).
22. *Decoding EEG and LFP signals using deep learning: heading TrueNorth*. In: *Proceedings of the ACM International Conference on Computing Frontiers*.; p. 259–66.
23. Kiral-Kornek I, Roy S, (2018). *Epileptic seizure prediction using big data and deep learning: toward a mobile system*. *EBio Medicine* ;27:103–11.
24. Guinney J, Saez-Rodriguez J. (2018). *Alternative models for sharing confidential biomedical data*. *Nat Biotechnol*; 36:391–2.
25. Schaffter T, Buist DSM, Lee CI, (2020). *Evaluation of combined artificial intelligence and radiologist assessment to interpret screening mammograms*. *JAMA Netw Open*;3(3):e200265.

AAYUSHI INTERNATIONAL INTERDISCIPLINARY RESEARCH JOURNAL (AIIRJ)
ISSN 2349-638x (Peer Review and Indexed Journal) IMPACT FACTOR 7.331
Devgiri Nagar, Ambejogai Road, Latur
Tq. Latur, Dist. Latur. Pincode 413512.
State Maharashtra, India.
Email ID's: editor@aiirjournal.com , aiirjpramod@gmail.com
Website: www.aiirjournal.com



Certificate of Publication

Awarded to

Gajbe R. J.

For Contributing Research Paper

“Importance of Artificial Intelligence in Healthcare: A Review”

In the

AAYUSHI INTERNATIONAL INTERDISCIPLINARY RESEARCH JOURNAL (AIIRJ)

Online Monthly Peer Review & Indexed Journal with ISSN 2349-638x (Impact factor 7.331)

for the month of **April 2022** with **Special Issue No. : 109**

Pramod Prakashrao Tandale
(Chief Editor)

Compressing and de-noising of Biomedical Signals

Gajbe R. J.

Department of Electronics

Vidya Bharati Mahavidyalaya, Amravati M.S. India-444 602

Abstract

Compressing and de-noising signals of ECG (Electrocardiograph), EEG (Electroencephalograph), GSR (Galvanic skin response), EOG (Electrooculography) and EMG (Electromyograph) is important in signal processing. Several steps are involved to process biomedical signals, among which the first step related to pre-processing, in which a noisy signal is processed for generating noise-free signal, which can be utilized for further operations. This work gives a detailed understanding of de-noising techniques those have been used. These techniques utilize the benefits of artificial neural networks (ANN), adaptive filtering, Remez Exchange Algorithm and finite impulse response (FIR) filtering. These techniques have been implemented for de-noising of biosignals, individually as well as combining with other techniques, for better results. The application of the method to simulated and biomedical signals shows its potential.

Keywords : Biomedical signals, analog, digital, FIR Filters, Remez Exchange Algorithm

1 Introduction:

Biomedical instrumentation is the branch of Medical Sciences where the medical instruments are to be studied. Biomedical instruments are used to record, analyze & process the biomedical signals. Biomedical signals are extracting the information from biological systems like brain, muscles or heart etc, under investigation. The extracting information may be so simple to note the pulse rate of a person from wrist or so complex like analyzing the information from heart by using ECG. There are various sources for extracting signals from the biological system [1-2].

They are:

- i. Bioelectric signal:- ECG (Electrocardiograph), EEG (Electroencephalograph), GSR (Galvanic skin response), EOG (Electrooculography) and EMG (Electromyograph) are the bioelectric signals. These signals are generated from the nerve cells or muscle cells. These signals are collecting from the cell membrane potential which may be excited under certain conditions to generate an action potential.
- ii. Bioacoustic Signals:- Lung sounds, heart sounds, bowel sounds, and joint sounds are the bioacoustic signals generated from the human body in easy and noninvasive way during its working.
- iii. Biochemical Signals:- Neurotransmitters are the chemical messengers from the body that control and regulate the body. These signals are extracted from the chemical measurement of living tissues & samples which are taken from living being.
- iv. Biomechanical Signals:- This signals are extracted from some mechanical function of the biological system. The flow of blood & pressure signals, all types of motion & displacement signals are the example of the biomechanical signals.
- v. Biomagnetic Signals:- MMG (Mechanomyogram) or MEG (Magneto-encephalogram) are the examples of biomechanical signals which are observed from the surface of muscles when it is contracted. Biomagnetic signals are obtained from the magnetic field produced by electrical currents occurred naturally in the brain of the living things.
- vi. Bio-optical Signals:- Bio optical signals can observe either naturally or the signals may be introduced to measure a biological parameter with an external light medium.
- vii. Bio-impedance Signals:- An electrical impedance signal can be obtained due to the changes in blood volume or blood resistivity from human being.

The biomedical signals are extracted from the biomedical instruments. Therefore it is useful to study of biomedical instruments.

2. Significance of Biomedical Signals

Biomedical signals are the collection of electrical signals acquired from any organ. This signal is normally a function of time. It may be in terms of its amplitude, frequency and phase. The analysis of these signals is very important for researchers as well as for careful medical diagnosis. It is also essential for proper treatment of

patients. If the signals are not properly diagnosed and analyzed, it will lead to wrong diagnosis and can be dangerous for the lives. Careful analysis of biomedical signals such as ECG, EMG, and EEG are very important for proper diagnosis of disease. These signals are noisy as well as artefacts which have to be removed for proper treatment of a patient. Presently, extensive efforts and research have been done in this area for developing better algorithms, upgrading existing methodologies, improving detection techniques to reduce noise, and to acquire accurate biosignals [3-4].

3. Sources of errors and significance of reduction of noise

Accurate measurement of biomedical signals depends on the properties of electrodes and its interaction with the skin, amplifier design and the conversion and subsequent storage of the biomedical signal from analog to digital form. Biomedical signals are influenced by electrical noise and some other factors.

3.1 Electrical noise and factors influencing biomedical signal:

The amplitude of biomedical signals is very small before being amplified. Biomedical signals collecting from any organ acquire noise while traveling through different tissues or blood vessels. Therefore it is important to understand the characteristics of the electrical noise. The electrical noise, which will be affected biomedical signals, can be categorized into the following types [5-6]:

- i. Power line interference
- ii. Base line drift
- iii. Electrode contact noise
- iv. Motion artifacts
- v. Muscle contraction
- vi. Instrumentation noise generated by electronic devices
- vii. Electrosurgical noise

Power line interference and base line drift: This noise could be occurred due to stray effect of the alternating current and loose contact of the electrodes to the patient as well as faulty electrodes. This error can cause due to power line interference or not properly earthen measuring instruments and the patients. This noise is reduced by using notch filter [7].

Inherent noise in electronics equipment: This noise is induced by the electronics equipments and cannot be eliminated. But it can be reduced by using high quality electronic components.

Ambient noise: This type of noise is procreated by the electromagnetic radiation from the surfaces of our body. It is almost impossible to avoid this. The ambient noise may have the amplitude that is one to three orders of magnitude greater than the biomedical signal.

Motion artifact: There are two major sources of motion artifact:

- 1) The electrode contact with the surface of our body and
- 2) The cable of electrode.

Motion artifact can be reduced by properly designed set up of the electronic circuitry.

Inherent instability of signal: The amplitude of biomedical signal is random in nature. It is influenced by the firing rate of the motor units. In most of the situation, firing frequency region is of the order of 0 to 20 Hz. This class of noise is considered to be unwanted and therefore there is a need of removal of the noise [4].

Electrosurgical noise: Electrosurgical noise is procreated by different medical equipment present within the patient care environment at frequencies between 100 kHz and 1 MHz, lasting for about 1 to 10 seconds [8]. The electrosurgical noise completely destroys the ECG signal and is presented in large amplitude [5].

The specific requirement of biomedical amplifiers [9];

- ECG amplifier
 - Frequency range is about 0.05- 100Hz.
 - Safety and protection: leakage current below safety standard limit of 10 μ A.
 - Electrical isolation from the power line and the (earth) ground.
 - Safety against high defibrillation voltages.

- EEG amplifier
 - Gain must deal with microvolt or lower levels of signals.
 - Components must have low thermal and electronic noise.
 - Other factors are similar to ECG.
- EMG amplifier
 - Slightly improved amplifier BW suffices.
 - Post-processing circuits are always needed (e.g. rectifier and integrator).
- EOG amplifier
 - High gain with very good low frequency or even DC response.
 - DC-drifting electrodes should be selected with great care.
 - Often active DC or drift cancellation or correction circuit may be necessary.

From the above background, it seems that the quality of biomedical signal can be improved by-

- a) Design the instrumentation amplifier with increase in Signal to Noise ratio (SNR), high common mode rejection ratio (CMRR), low input referred noise voltage and low dc offset as well as very low power consumption.
- b) The distortion of biomedical signal must be as small as possible by avoiding unnecessary filtering, notch filters and distortions of signal peaks.

4. Techniques used for reduction of noise

The amplitude of biomedical signals is very small. Therefore it is easily contaminated by noise. Due to noisy signal, there may be misleading of diagnosis and hence it is necessary to reduce noise from the original signals. Exclusive amplifiers designed with specific features and various filter circuits are used for reduction of noise from the biomedical signals.

Table 1.1: Electric potentials of biomedical signals [9]:

Biopotential	Frequency Range	Signal Amplitude	Electrode
Electrocardiogram (ECG)	0.05 - 100 Hz	1 - 5 mV	Surface
Electromyogram (EMG)	20 - 2000 Hz	0 - 10 mV	Surface, needle
Electroencephalogram (EEG)	0.5 - 40 Hz	0.001 - 0.01 mV	Surface
Electro-oculography (EOG)	DC -10 Hz	0.01 - 0.1 mV	Contact
Action potential of neurons	0 -10 KHz	50 - 90 mV	Glass pipette

4.1 Exclusive amplifiers:

The biomedical signals are extracted from organ of human beings and it is the input to the amplifier. This input consists of five components:

- i. The desired biopotential,
- ii. The undesired biopotential,
- iii. A power line interference signal of 50 Hz (60 Hz in some countries) and its harmonics,
- iv. Interference signals generated by the tissue/electrode interface, and
- v. Noise.

Essential factors to be considered for measuring equipments [8]:

- High amplification
- High CMRR (High differential gain and low common mode gain)
- High input impedance
- Low Noise
- Stability against temperature and voltage fluctuations
- Electrical safety, isolation and defibrillation protection.

Therefore proper design of the amplifiers is the extensive task. Table 1.2 shows the biopotential amplifiers design consideration for various biomedical signals.

Table 1.2 Biopotential amplifiers design consideration [10].

Biopotential	Exclusive amplifier design consideration	Additional features desired
EKG	Moderate gain, Bandwidth, noise, CMRR, Input Impedance	Electrical safety, isolation, defibrillation protection
EEG	High gain, very low noise filtering	Safety, isolation, low electrode-skin resistance
EMG	Gain and bandwidth of Op-Amps	Post acquisition data processing
EOG	DC and low drift	Electrode-skin junction potential. Artifact reduction

Now many companies are used CMOS technology for designing of high performance amplifiers, which provides increase in gain and CMRR along with low power consumption [4]. Along with high performance amplifiers, various filter circuits are also used for reduction of unwanted signals or noise. These filter circuits are classified in number of ways:

- i. Analog or digital Filters
- ii. Passive or active filters
- iii. Audio (AF) or radio frequency (RF) filters

4.2 Analog filters

Analog filters are designed to processed analog signals, while digital filters are processed analog signals using digital techniques. Op Amp based active filters are called analog filters. Passive filter uses passive elements such as resistors, capacitors and inductors, while active filter employs transistors or operational amplifiers in addition to the resistors and capacitors. RC filters are commonly used for low frequency signals or audio signals, whereas LC or crystal filters are used for high frequency signals or radio frequency signals. The crystal provides more stable operations at high frequency due to high value of figure of merit. In audio frequencies, inductors are not used because it is bulky, expensive and may dissipate more power. It also emits magnetic field [11].

Active filters offer the advantages over passive filters:

- a) Gain and frequency adjustment: The operational amplifier is capable of providing desired gain. Therefore input signal is not attenuated. Frequency response of active filter is excellent.
- b) No loading problem: The active filter uses Op Amp having high input resistance and low output resistance and therefore it does not cause any loading effect.
- c) Cost: Usually active filter is more economical than the passive filter. Because the circuit uses Op Amp which is not expensive and the absence of inductor.

Although active filters are most widely used in communication system and signal processing. Active filters are more extensively used in telephone, radio, television, radar, space satellites and **biomedical instruments**.

The most commonly used filters are;

- i. Low pass filter
- ii. High pass filter
- iii. Band pass filter
- iv. Band reject filter
- v. All pass filter

4.3 Digital Filters

Digital filters are important class of Linear Time Invariant (LTI) digital signal processing (DSP) systems. It is designed to modify the frequency characteristics of the input signal $x(n)$. Digital filters are extensively used because of certain advantages over Analog filters. Digital filters have the potential to achieve much better signal to noise ratio than Analog filters. Digital Filters have emerged as a strong option for removing noise, shaping spectrum and minimizing Inter-Symbol Interference (ISI) in communication architectures.

Digital Signal Processing (DSP) is used in video compression, digital set-top box, cable modems, digital versatile disk, portable video systems/computers, digital audio, multimedia and wireless communications, digital radio, digital still and network cameras, speech processing, transmission systems, radar imaging, acoustic beam formers, global positioning systems, and biomedical signal processing. The field of DSP is always been driven by the advances in DSP applications and in scaled Very-Large-Scale-Integrated (VLSI) technologies [12].

The design of digital filter involves five steps [13]:

- i. Filter Specification
- ii. Coefficient calculation
- iii. Realization
- iv. Analysis of Finite word length Effects
- v. Implementations

Digital Signal Processing Systems (DSPS) relate input signals $x(n)$ to the output $y(n)$. The relationship between the output sequence of this system and may be represented by the operator as $y(n) = F[x(n)]$. The Fig. 1.1 shows Discrete-time signal representation.



Figure 1.1: Discrete-time signal representation.

4.4 FIR Filter [14];

FIR filter is a digital filter without feedback. The block diagram of FIR filter is shown in the Fig. 1.2. If $x(n)$ is the input at n^{th} time, and $x(n-1)$ is the delayed input then the filter output $y(n) = x(n) + x(n-1)$,

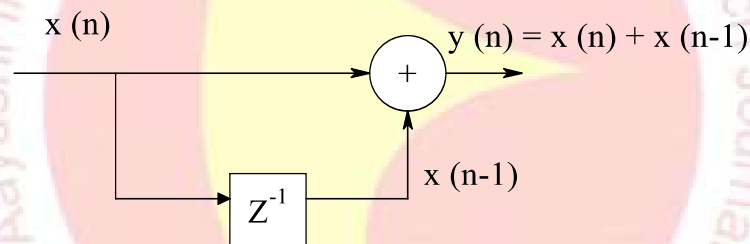


Figure 1.2: A simple FIR filter

The general transfer function of an FIR filter in the z transform is given by

$$H(z) = \sum_{l=0}^M b_l z^{-l} = H_0 z^{-M} \prod_{l=0}^M (z - z_l) \quad (1)$$

A filter with the above transfer function will be always stable, i.e. an input of finite amplitude will lead to an output that is also finite. This can be inferred from the location of the poles in the z -plane (the two-dimensional plane formed by plotting the real part of z along the abscissa, and the imaginary part along the ordinate). The poles correspond to the points in the z -plane where the denominator of $H(z)$ is zero; their counterparts are the zeros, where the numerator becomes zero. For the FIR filters, all poles are at the origin ($z=0$), and the zeros are at the z_l in the equation above. When the poles of a transfer function are within the unit circle (i.e. $z < 1$), then filter is stable; this clearly stated that the case of FIR filters.

The advantage of FIR filters is that they can be designed to have linear phase. FIR filters have a linear phase response if and only if its impulse response is symmetric or anti-symmetric, that is $h(n) = \pm h(M-n)$. Such filters may introduce delay in the output signal with respect to corresponding input signal, but all frequencies are delayed by the same amount. These filters thus distort the features of the signal that are not linear in their phase-response.

FIR filters can be designed by using optimization packages or by using approximations to the ideal infinite-impulse responses by cutting the finite length using the different types of tapered windows. The main drawback of FIR filters is that in order to satisfy demanding specifications, FIR filters require a relatively high

number of multiplications, additions and storage elements or large memory space. This makes FIR filters potentially more expensive than IIR filters in applications where the arithmetic operations or storage elements are costly, or need to be limited in number. FIR filters has the benefit of achieving linear phase, this motivates the widespread use of it.

4.5 FIR Filter Designing Methods:

Basically FIR Filter has three methods of designing [15] :

- i. Window Method
- ii. Frequency Sampling Method
- iii. Optimal Method (Remez Exchange Method)

4.5.1 Window Method:

The impulse response $h(k)$ is related to the inverse Fourier transform of $H(\omega)$. In this case, the pair of transformation is used which couples discrete time domain signals (sequence of samples) with a continuous spectral description.

Discrete Time Fourier Transformation is represented by the equation,

$$H(\omega) = \sum_{k=-\infty}^{\infty} h(k) e^{-j2\pi fTk} \quad (2)$$

There are many types of window functions:

- i. Kaiser Window
- ii. Hamming Window
- iii. Hanning Window
- iv. Rectangular window
- v. Blackmann window

By using Kaiser Window, it is quite possible to obtain separate control upon length or order of Filter. In Kaiser Window, there are two main parameters, the length of window and the shape parameter β . An important advantage of the window method is its Simplicity. It is simple to understand and simple to apply. It involves a minimum amount of computational efforts. The major disadvantage of it is its lack of flexibility.

The window method uses the Fourier series in conjunction with a class of functions known as window functions.

4.5.2 Frequency Sampling Method [15]:

When a desired frequency response has been specified, then the frequency-sampling method for FIR filter design is used which is the simplest and most direct technique. It consists of uniform sampling of the desired frequency response, and an inverse DFT technique is used to obtain the corresponding finite impulse response. But the results are not optimal because the response generally deviates from what is desired between the samples. When the desired frequency-response is under sampled then the resulting impulse response will be time aliased to some extent. It is important to evaluate the final impulse response by the use of a simulated DTFT (FFT with lots of zero padding) comparing to the originally desired frequency response.

4.5.3 Remez Exchange Algorithm [16]:

The Remez Exchange Algorithm is a standard method for filter designing. It minimizes the filter length and error between desired frequency response and actual frequency response. The Error function is defined as the difference between the ideal filter response and the practical filter response.

The magnitude of the error $|E(\omega)| = |W(\omega)[D(\omega) - P_c(\omega)]|$

Where

$D(\omega)$ is the ideal filter response or desired filter response,

$P_c(\omega)$ is the actual or practical filter response and

$W(\omega)$ is the weight function.

(3)

The Remez algorithm is not a general linear programming approach, but it is very robust, converges very rapidly to the optimal solution, and is widely used.

Optimal Filter Design Method is used to design a filter, in which filter coefficients are adjusted again and again until a particular error is minimized. The various methods of optimal filter design are as follows:

- i. Least square method.

- ii. Equiripple method.
- iii. Maximally flat.
- iv. Generalized equiripple.
- v. Constrained band equiripple.

5. Results and discussion

We started working with digital filter like FIR filter. We found that the source of noise doesn't have any fix band of interference, so we were planned for adaptive filtering algorithms. The algorithms differentiate in a way of updating the coefficient (h) of filter sample by sample inputs. After an extensive experimentation over the adaptive algorithms, it is observed that the performance of filters depend on the SNR values of input signal, μ value (Step size), number of taps and other factors. Finally, we choose Artificial Intelligent based filtering algorithms for dynamic solution. These techniques have only two processes- training and testing. Training process based on subset outcomes of adaptive filtering algorithm in initial stages, which may not require in later time even on change of source input as well, called trained filter / smart filter. Such intelligent filters give the freedom of selection of signal with different SNR values; also not bother about number of parameter settings which lead one more step towards the auto filter concept. The goal of our proposed work is to develop an algorithm which will use for removing noise excellently.

References:

- [1] **Raghibir Singh Khandpur (2001)**: "Hand Book of Biomedical Instrumentation, 2e". *Mc-Graw Hill Publication, New Delhi*,
- [2] **Raghibir Singh Khandpur (2019)** Compendium of Biomedical Instrumentation, John Wiley & Sons, Ltd. Volume 2, DOI:10.1002/97811192881
- [3] **Mizwan Z. and Kasturiwale H. (2014)**; "Study and Review of the Biomedical Signals With Respect To Different Methodologies" / *IJCSIT International Journal of Computer Science and Information Technologies, Vol. 5 (2), 1307-1309*.
- [4] **M. B. I. Reaz, M. S. Hussain, F. Mohd-Yasin (2006)**; "Techniques of EMG Signal Analysis: Detection, Processing, Classification and Application" / *Biol. Proced. on line, Vol.8: pp.11-35. doi.org/10.1251/bpo115*
- [5] **Sehamby R. and Singh B. (2016)**; "Noise Cancellation using Adaptive Filtering in ECG Signals: Application to Biotelemetry" / *International Journal of Bio-Science and Bio-Technology Vol.8, No.2, pp.237-244*.
- [6] **Kale S. N. and Dudul S. V. (2009)**; "Intelligent Noise Removal from EMG Signal Using Focused Time-Lagged Recurrent Neural Network" / *Applied Computational Intelligence and Soft Computing Volume 2009, Article ID 129761, 12 pages*.
- [7] **S. L. Joshi, R. A. Vatti, R. V. Tornekar (2013)**; "A Survey on ECG Signal Denoising Techniques" / *International Conference on Communication Systems and Network Technologies.*, doi.org/10.1109/csnt.2013.22
- [8] **Rajput A. (2017)**; "Result Analysis of Noise Removal in ECG Signal using Wavelet Decomposition Technique" / *International Journal of Computer Applications, Volume 170 – No.4*.
- [9] **Ramo T. (2006)**; "Chapter 74: Biopotentials and Electrophysiology Measurement" / *metrology.tkk.fi/courses/S-108.4010/2006/Biopotentials.ppt*.
- [10] **Nitish V. Thakor, (2017)** Chapter: Biopotentials and Electrophysiology Measurement, Book: Measurement, Instrumentation, and Sensors Handbook, CRC Press, Taylor and Francis Ltd., doi.org/10.1201/b15664-64
- [11] **Gayakwad R. A. (1999)**; "Op-Amps and Linear Integrated Circuits (3e)" / *Prentice Hall of India, pg. 290-293*.
- [12] **Gawande G. S. et al. (2012)**; "Performance analysis of FIR Digital Filter Design Techniques" / *International Journal of Computing and Corporate Research, Vol. 2, Issue 1*.
- [13] **Kumari J. and Nagaria D. (2015)**; "Performance Analysis of Low Pass FIR Filter using Different Methods" / *IJSRD - International Journal for Scientific Research & Development, Vol. 3, Issue 02*.
- [14] **Diniz P. S. R. et al.**; "ELECTRICAL ENGINEERING – Vol. I - Digital Signal Processing - Applications in Medicine" / *pg. 15-20*.
- [15] **Singh L. and Priya K. (2013)**; "FIR Filter Design and IIR Filter Design" / *International Journal of Computer Science And Technology Vol. 4, Issue 1*.
- [16] **A. Antoniou, (1991)**, Improved minimax optimisation algorithms and their application in the design of recursive digital filters, IEE Proceedings G Circuits, Devices and Systems, Vol. 138 issue 6, pp 724, doi.org/10.1049/ip-g-2.1991.0118

AAYUSHI INTERNATIONAL INTERDISCIPLINARY RESEARCH JOURNAL (AIIRJ)
ISSN 2349-638x (Peer Review and Indexed Journal) IMPACT FACTOR 7.331
Devgiri Nagar, Ambejogai Road, Latur
Tq. Latur, Dist. Latur. Pincode 413512.
State Maharashtra, India.
Email ID's: editor@aiirjournal.com , aiirjpramod@gmail.com
Website: www.aiirjournal.com



Certificate of Publication

Awarded to

Gajbe R. J.

For Contributing Research Paper

“Compressing and de-noising of Biomedical Signals”

In the

AAYUSHI INTERNATIONAL INTERDISCIPLINARY RESEARCH JOURNAL (AIIRJ)

Online Monthly Peer Review & Indexed Journal with ISSN 2349-638x (Impact factor 7.331)

for the month of **April 2022** with **Special Issue No. : 109**

Pramod Prakashrao Tandale
(Chief Editor)

Noise Reduction of Electromyography (EMG) signals by Artificial Neural Network (ANN) Technique

Gajbe R. J.

Department of Electronics, Vidya Bharati Mahavidyalaya, Amravati M.S. India-444 602

Abstract

Electromyography (EMG) signals can be used for clinical/ biomedical application and modern human computer interaction. EMG signals acquire noise while traveling through tissue, inherent noise in electronics equipment, ambient noise, and so forth. In this paper, it is shown that a Focused Artificial Neural Network ANN can elegantly solve to reduce the noise from EMG signal. It is seen that the performance of the proposed ANN clearly outperforms the best techniques. It is observed that ANN Filters work excellently as compared to adaptive filters and also observed that as the number of testing samples of biomedical signals.

Keywords: EMG, SNR, MSE, MATLAB, ANN

1 Introduction:

Biomedical signals are the collection of electrical signals and analysis of these signals is very important for researchers as well as for careful medical diagnosis. It is also essential for proper treatment of patients. If the signals are not properly diagnosed and analyzed, it will lead to wrong diagnosis and can be dangerous for the lives. Careful analysis of biomedical signals such as ECG, EMG, and EEG are very important for proper diagnosis of disease. These signals are noisy as well as artefacts which have to be removed for proper treatment of a patient. Presently, extensive efforts and research have been done in this area for developing better algorithms, upgrading existing methodologies, improving detection techniques to reduce noise, and to acquire accurate bio-signals [1-4].

The EMG signal is a biomedical signal that measures electrical currents generated in muscles during its contraction representing neuromuscular activities. The nervous system always controls the muscle contraction and relaxation. Hence, the EMG signal is a complicated signal, which is controlled by the nervous system and is dependent on the anatomical and physiological properties of muscles. EMG signal acquires noise while traveling through different tissues. The main reason for the interest in EMG signal analysis is in clinical diagnosis and biomedical applications. So far, research and extensive efforts have been made in the area, developing better algorithms, upgrading existing methodologies, and improving detection techniques to reduce noise and to acquire accurate EMG signals. Noise removal from noisy EMG signal is a filtering problem. Here the Neural Network model is trained to separate known noise from EMG signal. Neural Networks (NNs) have been efficiently used for nonlinear multivariable function approximation [5-7].

The main objective of our research work is to develop artificial intelligent model for reduction of noise from biomedical signal. This paper deals with intelligent removal of noise from the EMG signal using ANN-based Techniques.

2. Experimental : Performance measures by using following two factors;

i) Signal to Noise Ratio (SNR):

The output SNR (SNR_{out}), is calculated from power of the input signal $x(n)$ and an noise signal $e(n)$ and is given by,

$$SNR_{out} = 10 \text{ Log}_{10} \left(\frac{\text{Signal Power}}{\text{Noise Power}} \right)$$

Where, SNR_{out} is the ratio of the two powers expressed in decibels.

ii) Mean Square Error (MSE)

The formula for the mean square error is given by,

$$MSE = \frac{1}{N} \sum_{i=1}^N (\hat{y}_i - y_i)^2$$

Where, N = number of samples in the set of data,

y_i = Actual network output, and

\hat{y}_i = Desired network output.

2.1 Experimental Arrangement

i) Database Descriptions:

To design ANN Model, a sufficiently large amount of data is required for training and testing. We have collected standard data bases for biomedical signal from the <https://physionet.org>, <http://www.emglab.net> and etc.

ii) Software Specification Requirement: MATLAB:

MATLAB is very rich in command toolbox. Although MATLAB is intended primarily for numerical computing, an optional toolbox uses the MuPAD, symbolic engine, allowing access to symbolic computing capabilities. Matlab for designing artificial intelligent model. MATLAB has enormous Toolboxes. Out of which, we used Signal processing toolbox, Neural network toolbox, Fuzzy logic toolbox, Filter design toolbox, Curve fitting toolbox, and Data based toolbox for design the filters [8-9].

In this paper, we have simulated the MATLAB codes for the data conversion, adaptive filter algorithm, artificial intelligent ANN training and its testing.

iii) Software Implementation details:

The GUI has been constructed by using Matlab codes. Other set of codes have been used to run the various algorithms for noise removal in Matlab simulator. The main GUI contains four parts as shown in the following Fig. 1.1.

- 1] File input and its conversion
- 2] Adaptive filter algorithm and its Input parameter section
- 3] Output parameters section
- 4] Artificial intelligent noise removal section.

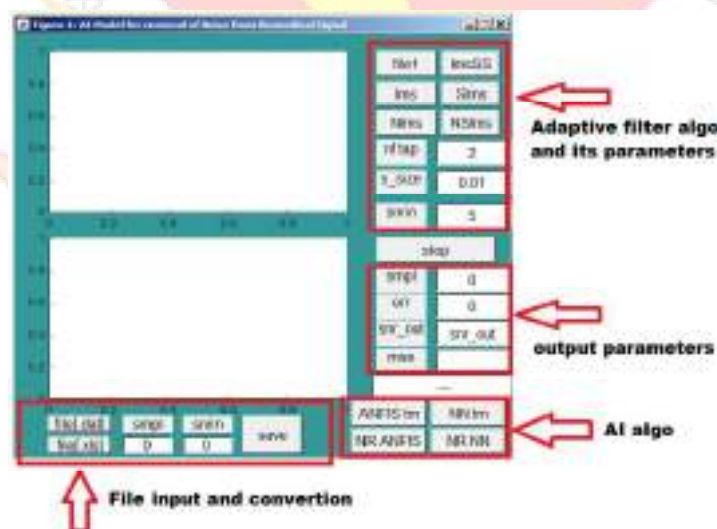


Figure 1.1: Artificial Intelligent Model for reduction of biomedical signal

2.2 Artificial intelligent noise removal algorithm: We have designed an artificial intelligent model for removal of noise from the biomedical signal by using Matlab coding on the basis of Sugeno architecture and run the

programme which creates graphical user interface. The Fig. 1.2 shows the block diagram of proposed artificial intelligent model.

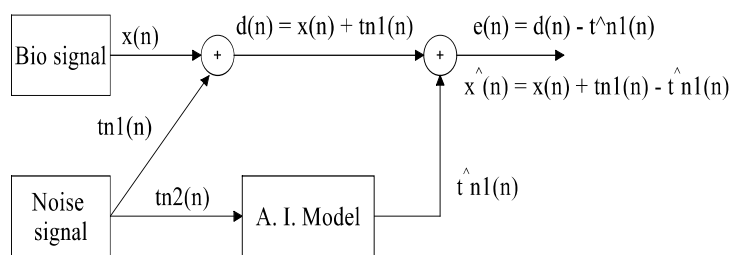


Figure 1.2: Block diagram of proposed A. I. Model

2.3 ANN based Noise Removal Algorithm:

Similar basic steps have been processed for training and weight storage of Neural Network Model as like an ANFIS. First we have trained the Neural Network by processing flowchart shown in Fig. 1.3 using ‘allwe’ file of relevant algorithm. Then processed the same steps used in ANFIS Model to save the weights, MF and MFT parameters. Once parameters have been saved then network was ready to test. Then we have tested neural network by performing steps as given in Flowchart of Fig. 1.4 on various biomedical signals.

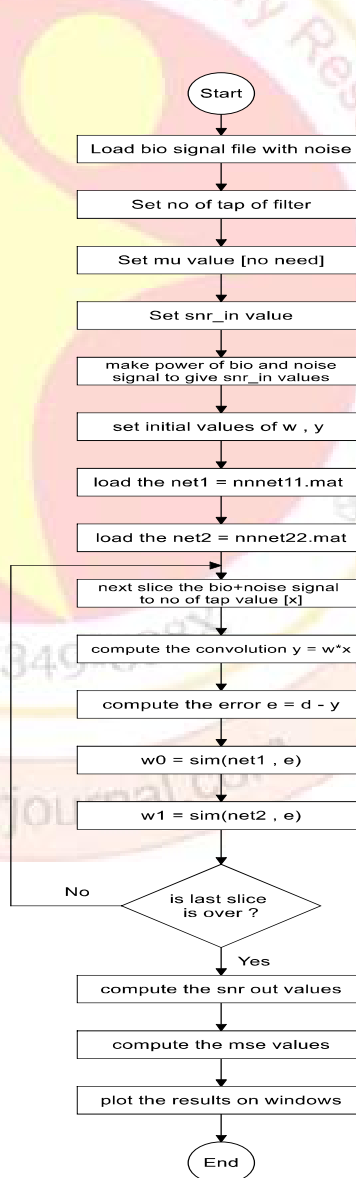
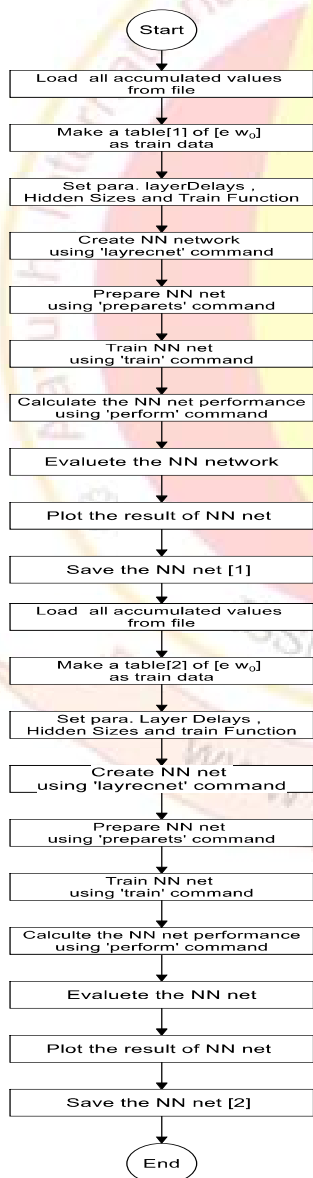


Fig.1.3: Flow chart of training of ANN Model

Fig.1.4: Flow chart of testing of ANN Model

2.4 Simulations

The major objective of this study was to investigate a noise removal filters. For this, simulations are carried out on EMG signals. The results are obtained on the output parameter section of MATLAB based GUI for removal of noise from biomedical signals [10-15].

We have taken various data-based files of EMG from cited resources given in database descriptions. First, we have converted these files into MATLAB format file that is .mat format for further processing and saved as .mat files.

Then we processed the adaptive filters on saved files. We have used LMS related algorithms namely LMS, Sign LMS, Sign LMS, Normalized LMS and Normalized LMS Sign for different step sizes, tap values and different SNR in values.

For the illustration of all these simulation process, in shortest way, we have taken the screen snap for LMS related algorithm mentioned above for ANN Model. After application of adaptive filtering process, we have gathered all needful data in the allwe_algo_name .mat file for the further training and testing of artificial intelligent algorithm such as noise removal Neural Network.

During the adaptive filtering, training and testing process, the all information regarding input parameters, output parameters and process status along with plots, etc. have available on the GUI panel. For ready outcome in notable way, we have summarized all the observations in tabular form.

The adaptive Filters based on LMS algorithms have been trained on ECG and EMG (Blue lines) signal (File name- 100.mat) of samples = 500. It is shown in Fig. 1.3

The various biomedical signals having sample size= 500: ECG and EMG (Blue lines) and biomedical signals with noise (Red lines) have been used for testing of Artificial intelligent and ANN model as shown in Fig. 1.4. a,b,c,d,. The performance parameters such as signal to noise ratio and MSE have been observed from GUI for various step sizes using different adaptive algorithms on ECG1 Signal as given in Table 1.1 and Table 1.2.

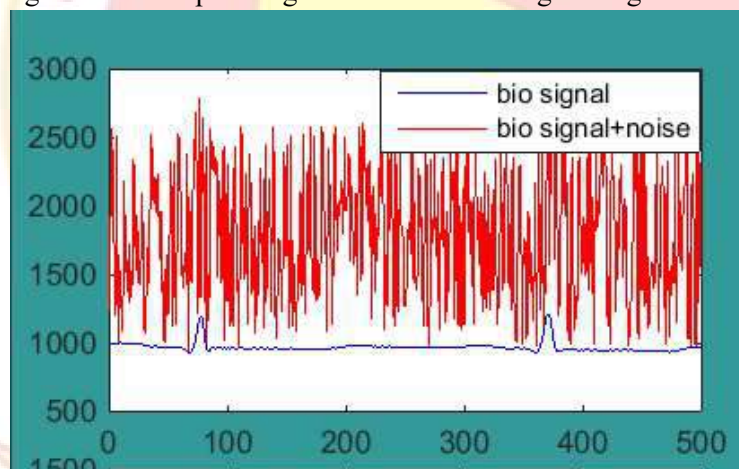
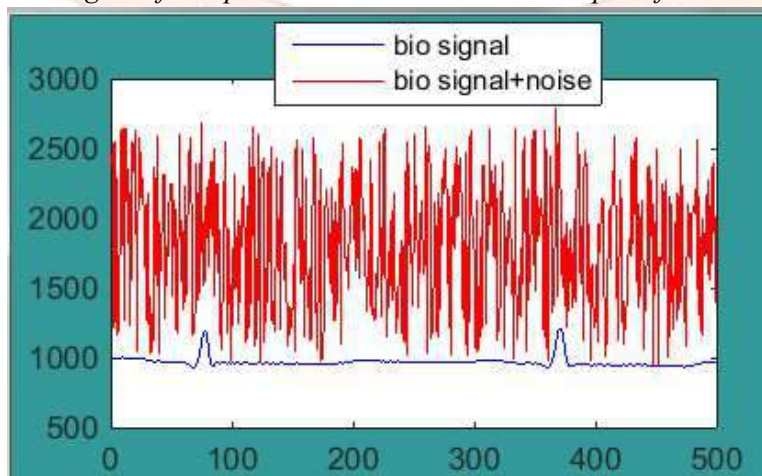


Figure 1.3: ECG signal of Samples= 500 used to train the adaptive filters LMS algorithms



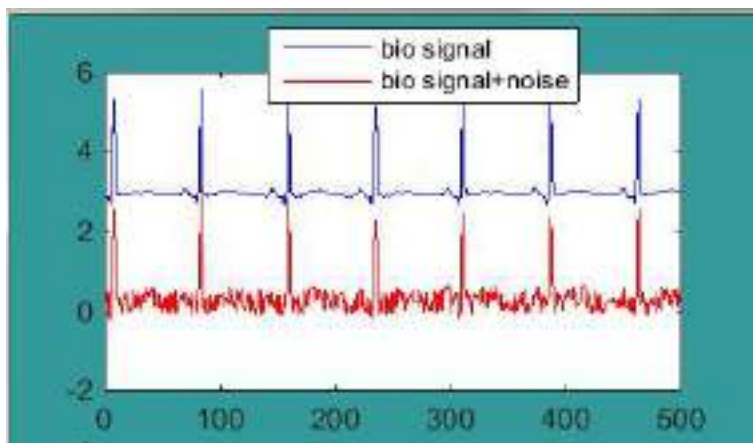


Figure 1.4 .a,b.: Blue line shows biomedical signals (ECGs and EMGs) and Red line shows biomedical signals with noise

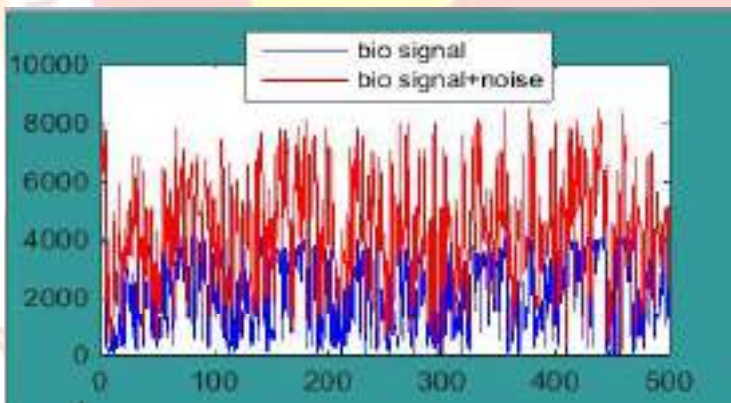
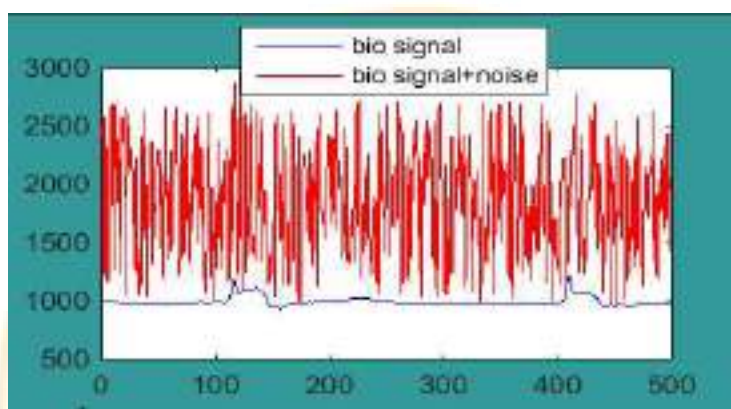


Figure 1.4.c,d: Blue line shows biomedical signals (ECGs and EMGs) and Red line shows biomedical signals with noise

Simulation Result for various Step under Adaptive Filters:

Table 1.1: Signal to noise ratio for various step sizes using different adaptive algorithms (on ECG1 Signal).

Step Size	SNR_out				
	LMS	NLMS	NSLMS	SLMS	SSLMS
1e-7	-0.19906	26.7339	26.7339	7.97304	26.7338
1e-8	0.21109	26.7339	26.7339	24.324	26.7339
1e-9	1.23573	26.7339	26.7339	26.7018	26.7339
1e-10	10.2783	26.7339	26.7339	26.7336	26.7339
1e-11	24.5255	26.7339	26.7339	26.7339	26.7339

1e-12	26.7039	26.7339	26.7339	26.7339	26.7339
1e-13	26.7336	26.7339	26.7339	26.7339	26.7339
1e-14	26.7339	26.7339	26.7339	26.7339	26.7339
1e-15	26.7339	26.7339	26.7339	26.7339	26.7339

Table 1.2: Mean square error for various step sizes using different adaptive algorithms

Step Size	MSE				
	LMS	NLMS	NSLMS	SLMS	SSLMS
1e-7	977169	1980.05	1980.05	148854	1980.09
1e-8	889108	1980.05	1980.05	3448.79	1980.05
1e-9	702247	1980.05	1980.05	1994.74	1980.05
1e-10	87545	1980.05	1980.05	1980.20	1980.05
1e-11	3292.4	1980.05	1980.05	1980.05	1980.05
1e-12	1993.78	1980.05	1980.05	1980.05	1980.05
1e-13	1980.19	1980.05	1980.05	1980.05	1980.05
1e-14	1980.05	1980.05	1980.05	1980.05	1980.05
1e-15	1980.05	1980.05	1980.05	1980.05	1980.05

Table 1.1 shows that SNR_{out} for various step sizes on LMS related algorithms. From this, we have observed that NLMS and NSLMS performed better at step sizes from 1e-07 to 1e-15. But overall, all algorithms have good performance on step size 1e-10, which governs the rate of convergence and speed of tracking ability. However, use of small step size is to ensure a small steady state error but it may decrease the convergence speed of the adaptive filter. However, increase in step size is to improve the convergence speed of the adaptive filter, but it might cause the adaptive filter to become unstable, so we have to select optimum value of step size [15,16]. The adaptive filter processed on LMS related algorithms using ECG 1 (100.mat file). The outcomes are depicted on GUI in Fig. 1.5.

Number of Samples used for Adaptive filters = 500.

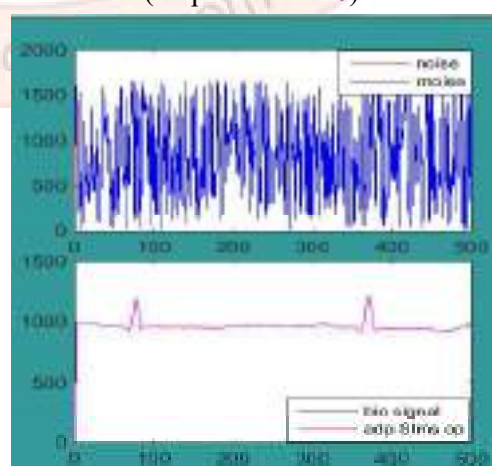
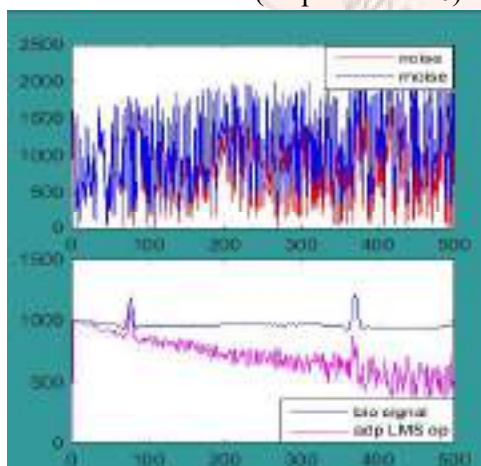
The Percent Root mean square Difference (PRD) indicates the reconstruction fidelity by point wise comparison with the original data. Another definition of error measure is called PRD and is given by:

$$PRD = \sqrt{\frac{\sum_{n=1}^N [x[n] - \bar{x}[n]]^2}{\sum_{n=1}^N [x[n] - \bar{x}]^2}} \times 100$$

Where $x[n]$ and $\bar{x}[n]$ are the original and reconstructed signals of length N, respectively and \bar{x} is the average value of the signal. PRD provides a numerical measure of the residual root mean square (RMS) error and the observed value of $prd_val = 0.0461$ [10-11].

LMS (Step size=1e-10)

SLMS (Step size= 1e-10)



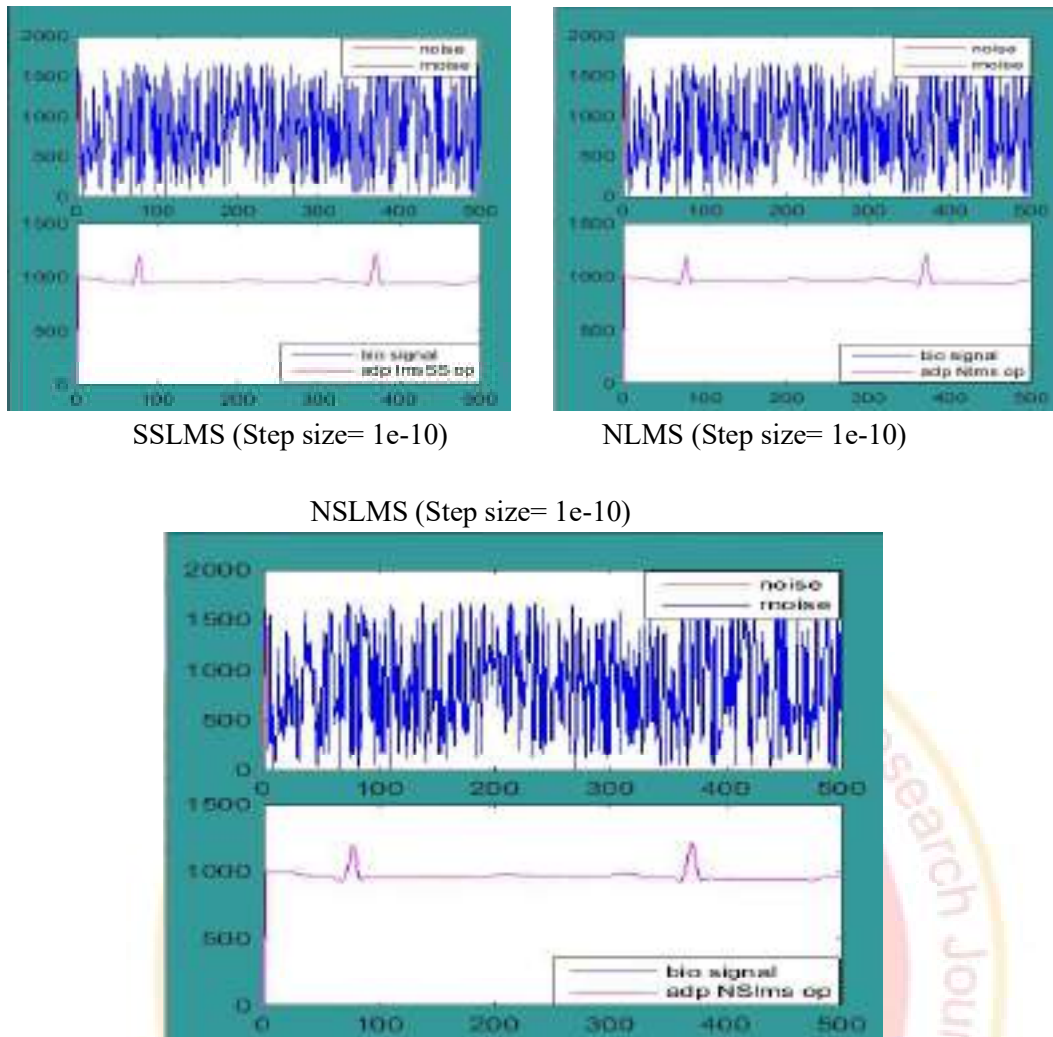


Figure 1.5: Simulation Result of Adaptive Filters for the ECG 1 signal (LMS, Sign LMS, Sign Sign LMS, Normalized LMS & Normalized LMS Sign algorithms)

Table 1.3: Report of Adaptive filters for various algorithms. (No. of Samples = 500)

Algorithm	Parameters [$\mu = 1e-10$ and No. of Taps= 2]		
	SNR in	SNR out	MSE
LMS	0	10.333	86450.1
SLMS	0	26.7336	1980.19
SSLMS	0	26.7339	1980.05
NLMS	0	26.7339	1980.05
NSLMS	0	26.7339	1980.05

From the Tables 1.1 and 1.2, we have to select proper value of step size which gives the good SNR out and minimum MSE. Thus we have preferred optimum value of step sizes which shows the best fit of straight line. Then we have developed adaptive filters for various LMS algorithms. The SNR out and MSE of various adaptive filters is shown in the Table 1.3.

3. Results and Discussion

The higher level of noise reduction is possible by performing adaptive noise cancellation techniques, otherwise this level of reduction of noise is difficult to achieve by using conventional filtering methods. As per the simulation results, it is observed that LMS algorithm has some limitations such as instability when the power

of input signal changes or the value of step size varies resulting in change of the rate of convergence. Thus SSLMS, NLMS and NSLMS algorithms are evolved from LMS algorithm to overcome the above-mentioned limitations. From the Table 1.1 and 1.2, it is inferred that the SSLMS, NLMS and NSLMS algorithm shows far greater stability with different step size. This combined with good convergence speed and relative computational simplicity. Thus, the SSLMS, NLMS and NSLMS algorithms are ideal for the real-time adaptive noise cancellation system. The amount of variation in non-stationary signals is difficult to control. LMS and SLMS have relatively poor performance for non-stationary as compared to SSLMS, NLMS and NSLMS and if we have to improve the performance of LMS and SLMS, the step size should have to reduce resulting in long convergence time. It is observed that ANN Filters works excellently as compared to adaptive filter and also observed that as the number of testing samples of biomedical signal (100.mat file) increases then the SNR value increases, while that of MSE decreases.

Thus, from the overall observations, we have concluded that the artificial intelligent networks ANN can be used to reduce noise from various biomedical signals intelligently. The SSLMS, NLMS and NSLMS are more precise and consistence as compared LMS and SLMS algorithms for wide range of step sizes [17-21].

References

- [1] N. Norali, M.H. Mat Som, "Surface Electromyography Signal Processing and Application: A Review", Proceedings of the International Conference on Man-Machine Systems (ICoMMS), 11 – 13 October 2009, Batu Ferringhi, Penang, MALAYSIA
- [2] M. B. I. Reaz, M. S. Hussain and F. Mohd-Yasin, "Techniques of EMG signal analysis: detection, processing, classification and Applications", Biol. Proced. Online 2006; 8(1): 11-35
- [3] Nurhazimah Nazmi, Mohd Azizi Abdul Rahman, Shin-Ichiroh Yamamoto, Siti Anom Ahmad, Hairi Zamzuri and Saiful Amri Mazlan, "A Review of Classification Techniques of EMG Signals during Isotonic and Isometric Contractions", Sensors 2016, 16, 1304
- [4] Peter Konrad, "A Practical Introduction to Kinesiological Electromyography", Version 1.4 March 2006, ISBN 0-9771622-1-4
- [5] Syed Zahurul Islam, Syed Zahidul Islam, Razali Jidin, Mohd. Alauddin Mohd. Ali, "Performance Study of Adaptive Filtering Algorithms for Noise Cancellation of ECG Signal", IEEE 2009.
- [6] Mohammad Zia Ur Rahman, Rafi Ahamed Shaik, D V Rama Koti Reddy, "Adaptive Noise Removal in the ECG using the Block LMS Algorithm" IEEE 2009.
- [7] Hong Wanl, Rongshen Ful, Li Shi, "The Elimination of 50 Hz Power Line Interference from ECG Using a Variable Step Size LMS Adaptive Filtering Algorithm" Life Science Journal, 3 (4), 2006.
- [8] J.D Bronzino and T.Ning, "Automatic classification of Respiratory signals", IEEE Proceedings/EMBS 11th International Conference, pp.669-670, Nov1989, Seattle, WA.
- [9] **Abraham A. M. (2004)**; "Fuzzy Logic for Control of Embedded System Applications"/ Elsevier Publications/ pg. 22, 33, 35-41, 40-78.
- [10] **Anon.**; "Fuzzy Logic Membership Function"/ <http://researchhubs.com/post/engineering/fuzzy-system/fuzzy-membershipfunction.html>.
- [11] **Anon**; "Reconstruction and Compression"/ http://shodhganga.inflibnet.ac.in/bitstream/10603/38609/10/10_chapter4.pdf
- [12] **Arar S. (2016)**; "FIR Filter Design by Windowing: Concepts and the Rectangular Window"/ <https://www.allaboutcircuits.com/author/steve-arar>.
- [13] **Gawali D. H. and Wadhai V. M. (2014)**; "Implementation of ECG Sensor for Real Time Signal Processing Applications"/ *International Conference on Advances in Electronics, Computers and Communications (ICAEECC)*, 978-1-4799-5496-4/14.
- [14] **Huang C. N. et al. (2004)**; "The review of applications and measurements in facial electromyography"/ *Journal of Medical and Biological Engineering*, 25(1):pg.15-20.
- [15] **Lee J. et al. (2009)**; "Performance Comparison of Variable Step-Size NLMS Algorithms"/ *Proceedings of the World Congress on Engineering and Computer Science 2009 Vol I., San Francisco, USA*.
- [16] MATLAB- wikipedia/ <https://en.wikipedia.org/wiki/MATLAB.on> dated 4th Aug. 2017.
MATLAB; <https://in.mathworks.com/help/nnet/ug/design-layer-recurrent-neural-networks.html>.

- [17] Poularikas A. D. and Ramadan Z. M. (2006); “Adaptive Filter Primer with MATLAB” / CRC press publications, Boca Raton, London, New York, pp. 101-166.
- [18] Syed Ateequr Rehman , R.Ranjith Kumar, “Performance Comparison of Adaptive Filter Algorithms for ECG Signal Enhancement” International Journal of Advanced Research in Computer and Communication Engineering vol. 1, Issue 2, 2012.
- [19] Suranai Pongpon Sri, Xiao-Hua Yu “An adaptive filtering approach for electrocardiogram (ECG) signal noise reduction using neural networks” Neurocomputing 117, page 206–213, 2013
- [20] Lubna Badri “Development of Neural Networks for Noise Reduction” The International Arab Journal of Information Technology, vol. 7, No. 3, 2010.
- [21] Demuth H. and Beale M., “Neural Network Toolbox-for Use with MATLAB User’s Guide, The Mathworks, Massachusetts.” 2002.



AAYUSHI INTERNATIONAL INTERDISCIPLINARY RESEARCH JOURNAL (AIIRJ)
ISSN 2349-638x (Peer Review and Indexed Journal) IMPACT FACTOR 7.331
Devgiri Nagar, Ambejogai Road, Latur
Tq. Latur, Dist. Latur. Pincode 413512.
State Maharashtra, India.
Email ID's: editor@aiirjournal.com , aiirjpramod@gmail.com
Website: www.aiirjournal.com



Certificate of Publication

Awarded to

Gajbe R. J.

For Contributing Research Paper

“Noise Reduction of Electromyography (EMG) signals by Artificial Neural Network (ANN) Technique”

In the

AAYUSHI INTERNATIONAL INTERDISCIPLINARY RESEARCH JOURNAL (AIIRJ)

Online Monthly Peer Review & Indexed Journal with ISSN 2349-638x (Impact factor 7.331)

for the month of **April 2022** with **Special Issue No. : 109**

Pramod Prakashrao Tandale
(Chief Editor)

Adaptive Neural-Based Fuzzy Inference System (ANFIS) : A Survey

Gajbe R. J.

Department of Electronics, Vidya Bharati Mahavidyalaya,
Amravati M.S. India-444 602

Abstract

The main aim of this study is to develop a flow prediction method, based on the adaptive neural-based fuzzy inference system (ANFIS) coupled with stochastic hydrological models. In comparison with the Mamdani model, the Sugeno model is based on automatic learning from the data and can accurately approximate a function using few rules. It has a stronger and more flexible illustration capability than the Mamdani model.

Keywords: ANFIS, Fuzzy Systems, Mamdani, Sugeno, Membership Function

Introduction: Fuzzy Systems

Fuzzy systems suggest a mathematical system that interprets the subjective human knowledge of the real time processes. This is a way to manipulate practical knowledge with some level of uncertainty. Fuzzy sets theory was introduced by Lofti Zadeh in 1965. The general architecture of a fuzzy system is depicted in Figure 1.1 The behavior of these systems is described by using a set of fuzzy rules such as *IF <premise> THEN <consequent>*. These system uses linguistics variables with symbolic terms. Each term represents a fuzzy set. The terms of the input space compose the fuzzy partition. The fuzzy system consists of three stages [1-4]:

1. The values of numerical inputs are mapped by a function according to a degree of compatibility of the respective fuzzy sets. This operation is called fuzzification.
2. The fuzzy system processes the rules in accordance with the firing strengths of the inputs.
3. The resultant fuzzy values are transformed again into numerical values. This operation is called defuzzification.

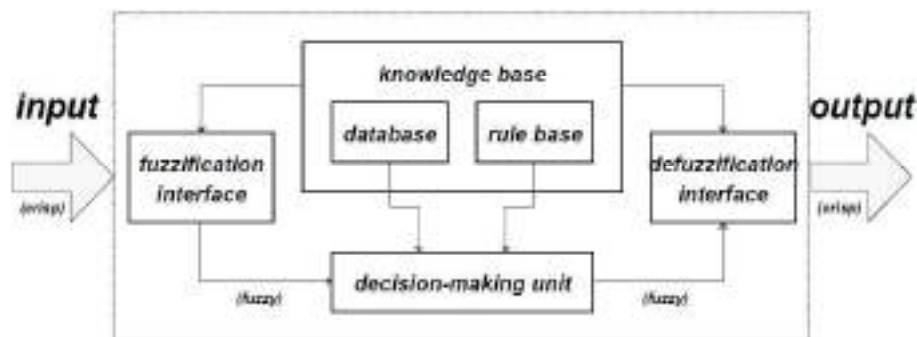


Figure 1.1: Basic Components of Fuzzy System

In an FIS, the knowledge base is comprised of the fuzzy rule base and the database. The database contains the linguistic term sets considered in the linguistic rules and the MFs defining the semantics of the linguistic variables, and information about domains. The rule base contains a collection of linguistic rules that are joined by AND/OR operator. Expert can provide his knowledge in the form of linguistic rules (Hao Yu et al.). The fuzzification process collects the inputs and then converts them into linguistic values or fuzzy sets. The decision logic which is called as fuzzy inference which generates output from the input, and finally the defuzzification process produces a crisp output for control action.

Interpretations of a certain rule or the rule base depend on the FIS model. The Mamdani and the Tagaki-Sugeno-Kang (TSK) or simply Sugeno model are two popular FISs. The Mamdani model is a non-additive fuzzy model that aggregates the output of fuzzy rules using the maximum operator, while the Sugeno model is an additive fuzzy model that aggregates the output of rules using the addition operator [1-4].

Mamdani inference model

Fuzzy variables are pasteurized by fuzzy logic rules with MIN and MAX operators. The fuzzy logic can be interpreted as the extended Boolean logic: for binary 0 and 1, the MIN and MAX operators, AND and OR

operators in Boolean logic, respectively. This has been depicted in the Table 1.1 and for fuzzy variables, the MIN and MAX operators [1-4].

Table 1.1: Binary Operations Using Boolean Logic and Fuzzy Logic

A	B	A AND B	MIN (A,B)	A OR B	MAX (A, B)
0	0	0	0	0	0
0	1	0	0	1	1
1	0	0	0	1	1
1	1	1	1	1	1

Table 1.2: Fuzzy Variables Operations Using Fuzzy Logic

A	B	MIN (A,B)	MAX (A, B)
0.3	0.5	0.3	0.5
0.3	0.7	0.3	0.7
0.6	0.4	0.4	0.6
0.6	0.8	0.6	0.8

In Mamdani model, let's assume that we have two fuzzy control rules:

R1: if x is A_1 and y is B_1 then z is C_1 and also

R2: if x is A_2 and y is B_2 then z is C_2

Fact: x is x_0 and y is y_0

Consequence: z is C

The inference procedure of the Mamdani model with the min (AND) and max (OR) operators and fuzzy inputs is shown in Figure 1.2.

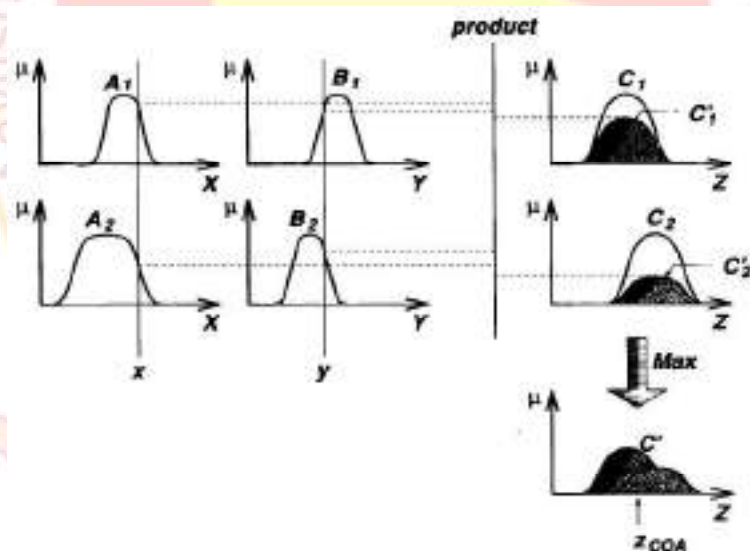


Figure 1.2: Mamdani Inference Model

The fuzzy implication is modeled by Mamdani's minimum operator and the sentence connective is also interpreted as propositions and defined by max operator[4].

The firing levels of the fuzzy rules denoted by α_i , $i = 1, 2$, are computed by

$$\alpha_1 = A_1(x_0) \wedge B_1(y_0), \alpha_2 = A_2(x_0) \wedge B_2(y_0)$$

The individual rule outputs are obtained by

$$C'_1(w) = (\alpha_1 \wedge C_1(w)), C'_2(w) = (\alpha_2 \wedge C_2(w))$$

Then the overall system output is computed by the individual rule outputs

$$C(w) = C'_1(w) \vee C'_2(w) = (\alpha_1 \wedge C_1(w)) \vee (\alpha_2 \wedge C_2(w))$$

Finally, to obtain a deterministic control action, employ any defuzzification strategy. For the instance, defuzzified value of fuzzy set is defined as its fuzzy centroid:

$$Z_o = \frac{\int z C(z) dz}{\int C(z) dz}$$

The Mamdani Model consists of five layers as depicted in Fig. 1.3.

Details are given below;

Layer 1: Input layer

Layer 2: Input membership or fuzzification layer

- Neurons represent fuzzy sets used in the antecedents of fuzzy rules determine the membership degree of the input.
- Activation function: The membership function.

Layer 3: Fuzzy Rule Layer

- Each neuron corresponds to a single fuzzy rule.
- Conjunction of the rule antecedents: *product*
- Output: the firing strength of fuzzy rule $R_i, \mu_{R_i} = \mu_{A_j} \times \mu_{B_k}$
- The weights in layer 3 and layer 4 : the normalized degree (certainty factors) of confidence of the corresponding fuzzy rules.

They are adjusted during training.

Layer 4: Output membership layer

- Disjunction of the outputs: $\mu_{C_i} = \mu_{R_j} \oplus \mu_{R_k} = \sum \mu_{R_j}$: *sum*.
- The integrated firing strength of fuzzy rule neurons R_j and R_k .
- Activation function: The output membership function.

Layer 5: Defuzzification layer

- Each neuron represents a single output. (Ex. - Centroid Method).

Learning:

- A various learning algorithm may be applied: Back propagation.
- Adjustment of weights and modification of input/output membership functions.
- Sum-Product composition and centroid defuzzification was adopted; a corresponding ANFIS was constructed easily.
- Extra complexity with max-min composition – no better learning capability or approximation power.

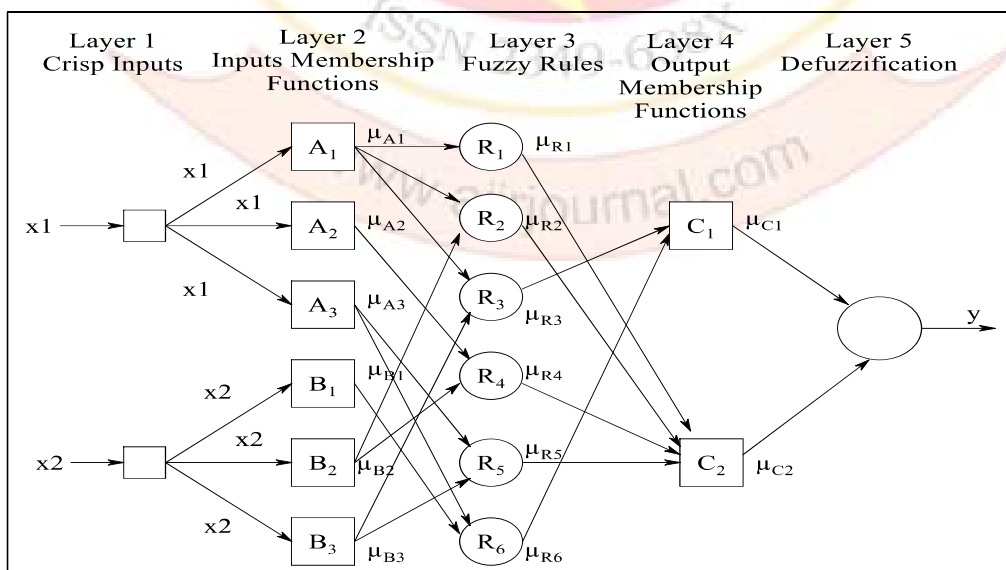


Figure 1.3: Neuro Fuzzy Equivalent System

The Mamdani model offers a high semantic level and a good generalization capability. It contains fuzzy rules built from expert knowledge. However, FISs based only on expert knowledge may result in insufficient accuracy. For accurate numerical approximation, the Sugeno model can usually generate a better performance.

Sugeno inference model

However, TSK or Sugeno architecture is more popular than Mamdani architecture where the defuzzification layer was replaced by normalization and weighted average. The TSK architecture does not require MAX operators, but a weighted average is applied directly to regions selected by MIN operators. TSK architecture is actually very straightforward method, in which the output weights are proportional to the average function values at the selected regions by MIN operators. The TSK fuzzy system works as a lookup table. Sugeno and Takagi use the following architecture which is graphically represented in Figure 1.4.

R₁: if x is A₁ and y is B₁ then z₁ = p₁x + q₁y + r₁

R₂: if x is A₂ and y is B₂ then z₂ = p₂x + q₂y + r₂

Fact: x is x₀ and y is y₀

Consequence: z

The firing levels of the fuzzy rules are computed by

$$w_1 = A_1(x_0) \wedge B_1(y_0), w_2 = A_2(x_0) \wedge B_2(y_0)$$

Then the crisp control action is demonstrated as

$$Z_0 = \frac{w_1 z_1 + w_2 z_2}{w_1 + w_2}$$

If we have n rules in our rule-base system then the crisp control action is calculated as

$$Z_0 = \frac{\sum_{i=1}^n w_i z_i}{\sum_{i=1}^n w_i}$$

Where w_i denotes the firing level of the ith rule, i = 1, 2 . . . , n

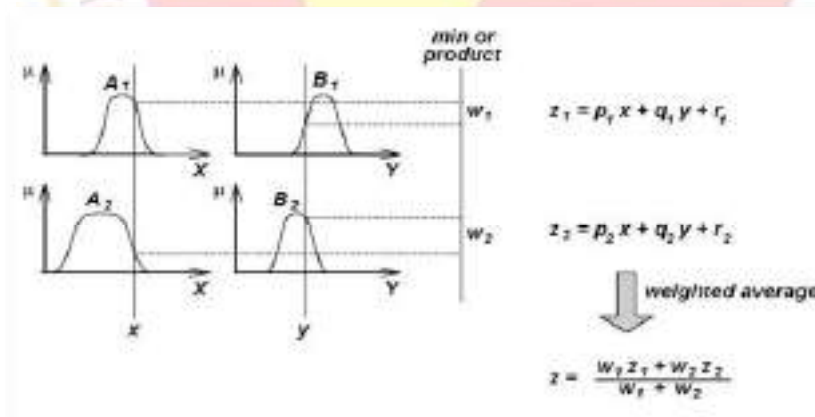


Figure 1.4: Sugeno Fuzzy Model

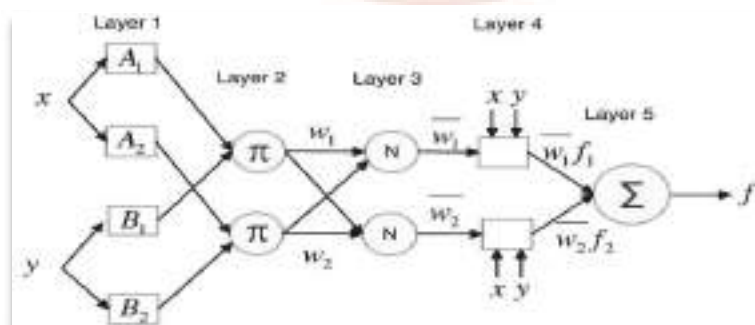


Figure 1.5: Equivalent ANFIS architecture

The TSK model normally selects z as first-order polynomial, hence the model is also called the first-order TSK model. When z is selected as constant, it is termed as the zero-order TSK model and can be regarded as a special case of the Mamdani model.

In comparison with the Mamdani model, the Sugeno model is based on automatic learning from the data and can accurately approximate a function using few rules. It has a stronger and more flexible illustration capability than the Mamdani model. In the Sugeno model, rules are extracted from the data, but the generated rules may have no meaning for experts. The Sugeno model has found more successful applications in building the fuzzy systems.

The Sugeno Model consists of five layers as depicted in Fig. 1.5, called ANFIS architecture. Details are given below;

Layer 1: Fuzzification layer

Every node I in the layer 1 is an adaptive node with a node function.

$O_{1,1} = \mu_{A_i}(x)$ for $i=1,2$ or : membership grade of a fuzzy set A_1, A_2 $O_{1,1} = \mu_{B_{i-2}}(y)$ for $i=3,4$.

Parameters in this layer: premise (or antecedent) parameters.

Layer 2: Rule layer

A fixed node labeled as Π where the output is the product of all the incoming signals: $O_{2,I} = w_i = \mu_{A_i}(x) \mu_{B_i}(y)$ for $i=1,2,\dots$, firing strength of a rule.

Layer 3: Normalization layer

A fixed node labeled N .

The i -th node calculates the ratio of the i -th rule's firing strength to the sum of all rules' firing strength: $O_{3,I} = \frac{w_i}{w_i + w_j}$ for $i=1,2$.

Outputs of this layer are stated as normalized firing strengths.

Layer 4: Defuzzification layer

An adaptive node with a node fn $O_{4,I} = \underline{w}_i f_i = \underline{w}_i (p_i x + q_i y + r_i)$ for $i=1,2,\dots$, where \underline{w}_i is the normalized firing strength from layer 3 and $\{p_i, q_i, r_i\}$ is the set of parameter pf this node are called the Consequent Parameters.

Layer 5: Summation neuron

A fixed node which computes the overall output as the summation of all incoming signals.

Overall output = $O_{5,1} = \sum \underline{w}_i f_i = \sum \underline{w}_i f_i / \sum \underline{w}_i$.

ANFIS learning:

A learning algorithm of the least-squares estimator including the gradient descent method.

Forward Pass: adjustment of consequent parameters, p_i, q_i, r_i .

Rule consequent parameters are known by the least-square estimator.

Find a least-square estimate of $k = [r_1 p_1 q_1 \dots r_n p_n q_n]$, k^* , that minimizes the squared error $e = |O_d - O|^2$.

$E = e^2 / 2 = (O_d - O)^2 / 2$.

The consequent parameters are adjusted while the antecedent parameters remain fixed.

Backward Pass: adjustment of antecedent parameters.

The antecedent parameters are tuned while the consequent parameters are kept fixed.

(Ex.) Bell activation function: $1 / [1 + ((x-c)/a)^{2b}]$.

Membership Function:

A fuzzy set is wholly characterized by its membership function (MF). Since most fuzzy sets in use have a universe of conversation X consisting of the real line R , it would not be practical to list all the pair defining a membership function. A more convenient and brief way to define an MF is to express it as a mathematical formula [5].

1. Triangular MF:

$$\text{triangle}(x; a, b, c) = \max \left[\min \left(\frac{x-a}{b-a}, \frac{c-a}{c-b} \right), 0 \right]$$

2. Trapezoid MF:

$$\text{trapezoid}(x; a, b, c, d) = \max \left[\min \left(\frac{x-a}{b-a}, 1, \frac{c-a}{c-b} \right), 0 \right]$$

3. Gaussian MF:

$$\text{gaussian}(x; \sigma, c) = \exp \left\{ - \left[\frac{x-c}{\sigma} \right]^2 \right\}$$

1. Generalized Bell MF:

$$\text{bell}(x; a, b, c) = \frac{1}{1 + \left| \frac{x-c}{a} \right|^{2b}}$$

2. Sigmoidal MF:

$$\text{sigmoid}(x; a, b, c) = \frac{1}{1 + \exp \left[a(x-c) \right]}$$

In the proposed model, we have used Generalized Bell Membership Function. So, brief discussion is given here. Figure 1. 7.6 depict the meaning of parameters in Generalized Bell Function. These parameters are referred as premise parameters.

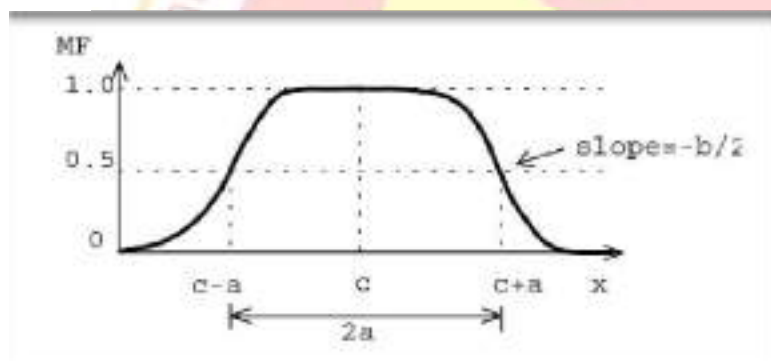


Figure 1.6: Meaning of parameters in Generalized Bell Function

The parameter b is usually positive. If it is negative, the shape of this MF becomes an upside-down bell. However, each of these parameters has a physical meaning: c determines the centre of the corresponding membership function; a is the half width; and b (together with a) controls the slopes at the crossover points.

Note that this MF is the overview of Cauchy distribution used in probability theory, so it is also referred to as the Cauchy MF. Because of their smoothness and concise notation, Gaussian and bell MFs are becoming increasingly popular for specifying fuzzy sets [2-3].

Advantages and disadvantages of Fuzzy Systems [8-10]:

The advantages of the fuzzy systems are:

- It has capacity to represent inherent uncertainties of the human knowledge with linguistic variables.
- It has ability to interact with the expert of the domain with the engineer designer.
- It represents the natural rules, so it can easily interpret the results.
- It can easily widen the base of knowledge through the addition of new rules.
- It is robust method in relation with the possible disturbances in the system.

Disadvantages of fuzzy system are:

- It is incapable to generalize or answer to what is written in its rule base.
- It is not robust in relation to the topological changes in the system. Topological changes will have to demand alterations in the rule base.
- It has to depend on the existence of a expert to determine the inference logical rules.

References:

1. **Walia N. et al. (2015)**; “ANFIS: Adaptive Neuro-Fuzzy Inference System- A Survey”/*International Journal of Computer Applications (0975 – 8887) Volume 123 – No.13.*
2. **Hao Yu.**; “ Chapter 20- Neuro-Fuzzy Systems”/ *pp.20.1- 20.9.*
3. **Jang J. S. R. (1997)**; “Neuro-Fuzzy and Soft Computing” /*Prentice-Hall, Upper Saddle River NJ, USA, pp. 335-368.*
4. **Tadeusiewicz R. et al.** / “Discovery of Neural Network Properties by means of C# programs”.
5. **Kale S. N. and Dudul S. V. (2009)**; “Intelligent Noise Removal from EMG Signal Using Focused Time-Lagged Recurrent Neural Network”/ *Applied Computational Intelligence and Soft Computing Volume 2009, Article ID 129761, 12 pages.*
6. **Mizwan Z. and Kasturiwale H. (2014)**; “Study and Review of the Biomedical Signals With Respect To Different Methodologies”/ (*IJCSIT International Journal of Computer Science and Information Technologies, Vol. 5 (2), 1307-1309.*
7. **Anon.**; “Fuzzy Logic Membership Function”/ <http://researchhubs.com/post/engineering/fuzzy-system/fuzzy-membership-function.html>.
8. **Anon.**; Adaptive Neuro-Fuzzy Inference System.
9. **Kumari J. and Nagaria D. (2015)**; “Performance Analysis of Low Pass FIR Filter using Different Methods”/ *IJSRD - International Journal for Scientific Research & Development, Vol. 3, Issue 02.*
10. **Gawali D. H. and Wadhai V. M. (2014)**; “Implementation of ECG Sensor for Real Time Signal Processing Applications”/ *International Conference on Advances in Electronics, Computers and Communications (ICA ECC), 978-1-4799-5496-4/14.*

AAYUSHI INTERNATIONAL INTERDISCIPLINARY RESEARCH JOURNAL (AIIRJ)
ISSN 2349-638x (Peer Review and Indexed Journal) IMPACT FACTOR 7.331
Devgiri Nagar, Ambejogai Road, Latur
Tq. Latur, Dist. Latur. Pincode 413512.
State Maharashtra, India.
Email ID's: editor@aiirjournal.com , aiirjpramod@gmail.com
Website: www.aiirjournal.com



Certificate of Publication

Awarded to

Gajbe R. J.

For Contributing Research Paper

“Adaptive Neural-Based Fuzzy Inference System (ANFIS) : A Survey”

In the

AAYUSHI INTERNATIONAL INTERDISCIPLINARY RESEARCH JOURNAL (AIIRJ)

Online Monthly Peer Review & Indexed Journal with ISSN 2349-638x (Impact factor 7.331)

for the month of **April 2022** with **Special Issue No. : 109**

Pramod Prakashrao Tandale
(Chief Editor)

Introduction to Biomedical Instruments

Gajbe R. J.

Department of Electronics, Vidya Bharati Mahavidyalaya, Amravati, Maharashtra, India

Article Info

Volume 9, Issue 3

Page Number : 106-110

Publication Issue

May-June-2022

Article History

Accepted : 01 May 2022

Published : 12 May 2022

ABSTRACT

This paper is mainly focused on Biomedical instruments , which are Basics of Medical Instruments, Block diagram of Medical Instrument. Biomedical signal processing aims at extracting significant information from biomedical signals. With the aid of biomedical signal processing, biologists can discover new biology and physicians can monitor distinct illnesses.

Keywords : Bioelectric signal, Block diagram, Microelectrodes

I. INTRODUCTION

Introduction to biomedical signals

Biomedical Instrumentation is the branch of medical science where the medical instruments are to be studied. Biomedical instruments are used to record, analyze & process the biomedical signals. Biomedical signals are extracting the information from biological systems like brain, muscles or heart etc, under investigation. The extracting information may be so simple to note the pulse rate of a person from wrist or so complex like analyzing the information from heart by using ECG. There are various sources for extracting signals from the biological system [1-2].

They are:

1. :-ECG (Electrocardiograph), EEG (Electroencephalograph), GSR (Galvanic Bioelectric signal skin response), EOG (Electrooculography) and EMG (Electromyograph) are the bioelectric signals. These signals are generated from the nerve cells or muscle cells. These signals are collecting from the cell

membrane potential which may be excited under certain conditions to generate an action potential.

2. BioacousticSignals:-Lung sounds, heart sounds, bowel sounds, and joint sounds are the bioacoustic signals generated from the human body in easy and noninvasive way during its working.

3. Biochemical Signals:- Neurotransmitters are the chemical messengers from the body that control and regulate the body. These signals are extracted from the chemical measurement of living tissues & samples which are taken from living being.

4. Biomechanical Signals:- This signals are extracted from some mechanical function of the biological system. The flow of blood & pressure signals, all types of motion & displacement signals are the example of the biomechanical signals.

5. BiomagneticSignals:-MMG (Mechanomyogram) or MEG (Magneto-encephalogram) are the examples of biomechanical signals which are observed from the surface of muscles when it is contracted. Biomagnetic signals are obtained from the magnetic field produced

by electrical currents occurred naturally in the brain of the living things.

6. Bio-optical Signals:- Bio optical signals can observe either naturally or the signals may be introduced to measure a biological parameter with an external light medium.

7. Bio-impedance Signals:- An electrical impedance signal can be obtained due to the changes in blood volume or blood resistivity from human being.

The biomedical signals are extracted from the biomedical instruments. Therefore it is useful to study of biomedical instruments.

Basics of Medical Instruments:

Fig. 1.1.1 shows the general block diagram of Biomedical Instrument. The detailed descriptions of each block are as follows [3-5]:

1. Primary Sensing Element - The measurand is given to the primary sensing element where the conversion takes place. This is done by a transducer which converts the measured quantity into a usable electrical output. The transduction may be of any from such that mechanical, electrical or optical.

2. Variable Conversion Element - The output of the primary sensing element is in the electrical form and is suitable for control, record and display. But it may be necessary to convert this output to some other suitable form for preserving the original information. This function is performed by the use of variable conversion element. A system may require one or more variable conversion elements.

3. Variable Manipulation Element - The biomedical signals have very small amplitude. So it is necessary to manipulate the signal preserving the original nature of it. Thus the voltage amplifier is the example of variable manipulation element. The element that follows the primary sensing element in a measurement system is called signal conditioning element. The variable conversion element and variable manipulation element are jointly called as Signal Conditioning Element.

4. Data Transmission Element - The transmission of data from one circuit to another is performed by the

data transmission element. In case of spacecraft, the control signals are sent from the control stations by using radio signals. The stage that follows the signal conditioning element and data transmission element collectively is called the intermediate stage.

5. Data Presentation Element - The readout or display devices which displays the required information about the measurand or measurement is called the data presentation element. The information of the measurement has to be monitor, control, and record for further analysis purposes. In case monitoring information, visual display devices are to be needed. For recording purpose, recorders like magnetic tapes, T.V equipment, and storage type CRT, printers, and so on are to be used.

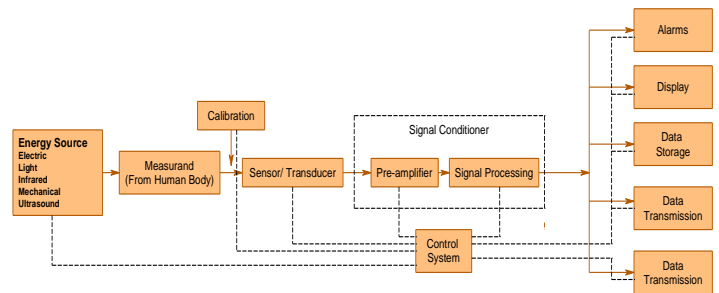


Figure 1 : General Block diagram of Medical Instrument

Biomedical signals are provided the information from biological systems under investigation. This is done by the use of excellent electrodes and powerful biomedical instruments having excellent signal to noise ratio. Therefore it is also necessary to study the recording electrodes used under the biomedical system.

Recording Electrodes used in the Biomedical Instrumentation:

Recording electrodes are the devices that convert ionic potentials into electronic potentials. The type of electrodes used for the measurements depends on the anatomical location of the organ. Bio-electrodes are the sensors that convert ionic conduction into electronic conduction. The purpose of bio-electrodes

is to acquire bioelectrical signals such as ECG, EMG, and EEG, etc.

Electrodes are mainly classified into two types [6]:

1. Perfectly polarized electrodes and
2. Perfectly non-polarized electrodes.

There are various electrodes that can be used to measure bioelectric signals. The three main classes of electrodes are Microelectrodes, Body Surface electrodes and Needle electrodes.

Microelectrodes

Microelectrodes are electrodes having fine tips to penetrate a single cell in order to obtain readings from the cell without damaging the nearby cells. The main function of microelectrodes is to obtain recording potential. Impedance of microelectrodes can play a very important role in the monitoring of low amplitude and high-resolution of the cells. Because of the smaller size of microelectrodes, it has high impedances in kilo ohms range.

It is classified into two types

- a. Metal type
- b. Micropipette type

a) Metal microelectrode:

Metal microelectrodes are prepared by electrolytic etching the tip of fine tungsten to the desired size and dimension. Then the wire is coated almost to the tip with any type of insulating material. The metal-ion interface takes place where the metal tip contacts the electrolyte. The main features of metal microelectrodes are

1. It has excellent signal to noise ratio
2. It is strong enough and easy to penetrate
3. It has high biocompatibility

b) Micropipette:

The micropipette type of microelectrode is a glass micropipette with its tip drawn out to the desired size. The micropipette is filled with an electrolyte which should be compatible with the cellular fluids. A micropipette is a small and extremely fine pointed pipette used in making microinjections. A commercial type of micropipette is shown in Fig. 1.1.2.



Figure 1.1. 2: Micropipette type Electrodes

Body Surface Electrodes:

Body surface electrodes are placed in contact with the skin of the subject to obtain bioelectric potentials. Body surface electrodes are of many types and sizes. The various types of body surface electrodes are discussed below.

Major body surface electrodes are

a) Immersion electrodes:

In early years, these types of bioelectric measuring electrodes were used. These electrodes were simply buckets of saline solution in which the subject placed such as hands or feet. So it was not a comfortable type of measurement and hence it was replaced with the plate electrodes.

b) Plate electrodes:

These electrodes were separated from skin surface by cotton pads soaked in a strong saline solution. These electrodes have generally smaller contact area. This type of electrodes has a tendency to lose their adhesive ability and therefore it causes errors in the result.

c) Floating electrodes:

These electrodes can eliminate the artifact errors which is a main problem with the plate electrodes. This is avoided by making direct contact of the electrodes with the skin. So the main advantage of floating electrodes is mechanical reliable. The conductive path between the metal and the skin is the electrolyte paste or jelly.

d) Disposable electrodes:

Normally plate electrodes, floating electrodes etc can be used more than one time. But it requires more care while cleaning after each use. Thus uses of disposable electrodes are now very popular which can be disposed after the use.

e) Suction electrodes:

These electrodes are well suited for the attachment of flat surfaces of body and the regions where the underlying tissues are soft. These electrodes have a small surface area. These types of electrodes are mainly used for the measurement of ECG. Suction electrodes are shown in Fig. 1.1.3. It is used a plastic syringe barrel to house suction tubing and input cables to an AC amplifier.



Figure 1.1.3: Suction electrodes

f) Ear clip & Scalp electrodes:

These electrodes are widely used in the measurement of EEG. Scalp electrodes can provide EEG signals easily by placing it over the naked head. For measurement of EEG, it requires 10 – 20 electrodes and scalp electrode usually used. They can avoid measurement errors and movement errors. During labour, internal monitoring may be needed and usually fetal scalp electrodes are placed under the baby’s scalp. It is also used to monitor baby’s heartbeat while still in uterus. A typical ear clip and scalp electrodes are shown in Fig. 1.1.4.



Figure 1.1.4: Ear clips and scalp electrodes

g) Needle Electrodes:

To reduce the interface and artifact error which are arise due to the electrode movement, a small sub-dermal needle electrodes can be used for the measurement of EEG, EMG etc. Actually the needle electrodes are used to penetrate the skin of the scalp for measurement of EEG. Generally they are simply inserted through a small section of the skin just beneath the skin parallel to it.

The needle electrodes for EMG measurement consist of fine insulated wires placed in such a way that their tips are in contact with the muscles, nerve or other tissues. The needle creates the hole necessary for insertion and the wires forming the electrodes are carried inside it. A typical EEG needle electrode is shown in Fig. 1.1.5.



Figure 1.1.5: Needle electrodes

The main advantage of needle electrode is that it is less sensitive to cause error due to movement than the surface electrodes. Also the needle electrodes have lower impedances when compared to surface electrodes as it makes direct contact with the sub-dermal tissues or intracellular fluid [7].

II. REFERENCES

- [1]. Anon. (2013); <http://www.ques10.com/p/6347/draw-block-diagram-for-generalized-measurement-sys/>
- [2]. Gayakwad R. A. (1999); "Op-Amps and Linear Integrated Circuits (3e)"/ Prentice Hall of India, pg. 290-293.
- [3]. Hao Yu et al.; " Chapter 20- Neuro-Fuzzy Systems"/ pp.20.1- 20.9.
- [4]. Haykin S. (1999); Neural Networks:A Comprehensive Foundation/ Prentice Hall, Upper Saddle River, NJ, pg. 3-4, 63-66.
- [5]. Khandpur R. S. (2001): "Hand Book of Biomedical Instrumentation, 2e". Mc-Graw Hill Publication, New Delhi, pg 35.
- [6]. Rajput A. (2017); "Result Analysis of Noise Removal in ECG Signal using Wavelet Decomposition Technique"/ International Journal of Computer Applications, Volume 170 – No.4.
- [7]. Ramo T. (2006); "Chapter 74: Biopotentials and Electrophysiology Measurement"/

metrology.tkk.fi/courses/S-108.4010/2006/Biopotentials.ppt.

Cite this article as :

Gajbe R J , "Introduction to Biomedical Instruments", International Journal of Scientific Research in Science and Technology (IJSRST), Online ISSN : 2395-602X, Print ISSN : 2395-6011, Volume 9 Issue 3, pp. 106-110, May-June 2022. Available at doi : <https://doi.org/10.32628/IJSRST229314>
Journal URL : <https://ijsrst.com/IJSRST229314>

Humidity Sensing Properties of PANI Doped Zinc Oxide Nanocomposites Thick Film Sensor**T R Ingle¹, R M Agrawal², G T Lamdhade³, R J Gajbe⁴**

1,3Department of Physics, Vidya Bharati Mahavidyalaya, Camp, C. K. Naidu Road, Amravati(M.S.)

2Department of Physics Shri R.L.T.College of Science, Civil Lines Road, Akola,(M.S)

4Department of Electronics, Vidya Bharati Mahavidyalaya, Amravati M.S. India

Corresponding Authors: Email: inglet@rediffmail.com (T.R.Ingle), gtilamdhade@rediffmail.com, oumgajanan@gmail.com (G.T.Lamdhade) Mob.+919403000509**Abstract**

Over the last decades a variety of chemical sensors have been developed based upon semiconductors, which monitor different characteristic sensor properties such as conductivities for electronic conductivity sensors. In the present work Polyaniline is prepared by polymerization of aniline under acidic condition. Zinc Oxide (ZnO) nanoparticle prepared by wet chemical method at room temperature using zinc nitrate and sodium hydroxide as starting material. ZnO nanoparticles were combined with conducting (PANI) polymer via polymerization in acidic aqueous solution to obtain a new type of inorganic – organic composites nanostructured. It is observed that PANI doped ZnO nanocomposites sensor shows a high response and sensitivity with good repeatability as compared to that of pure PANI and ZnO nanoparticle. The effect of hysteresis of the sensors, the effects of pure and composite oxide on sensitivity of the sensors were studied. The crystallinity and the crystallite size were examined by X-Ray Diffraction technique.

Keywords: Polyaniline, Zinc Oxide Nanocomposites, Humidity sensor.

1. Introduction

There is a growing demand for a sensing system that has high sensitivity, wide dynamic range, good stability, quick response, good reproducibility, simple structure and minimal cost. Metal oxide films sensitive to humidity have been reported earlier where sensing has been done using optical means. However, metal oxide humidity sensors depending upon measurements of electrical parameters require high temperature operation and consume significant amount of power. Humidity control and monitoring are of great interest to a wide area; these include moisture sensitive products, fresh and pack-age food, drug storage and environmental control for valuable Antiques or paintings etc. [1, 2]. Humidity sensors that are available in the market include dew point, infrared, catalytic and tin oxide-based sensors, which may be expensive, or require high temperature operation and consume significant amount of power and high cost of maintenance [3]. Much research has been focused on the development of humidity sensitive material [4–6]. Among these are the investigation of using conducting polymers such as polyaniline, polypyrrole, and polythiophene for humidity and gas sensing [7–9]. Advantages with polymers as sensing materials are light weight, flexible, low cost and simple fabrication process [10]. Pure polymer, polymer blends and polymer–inorganic composites have also been studied for the purposes, resulting in different degree of advancements in this area [11–16].

2. Synthesis of Material:

A) Synthesis of Polyaniline (PANI): In general is synthesized using two major polymerization approaches: electronic and chemical polymerization. In the present work polyaniline is synthesized by chemical polymerization method in which 0.2 M aniline hydrochloride is used as monomer unit. The synthesis is done by oxidative polymerization with 0.25 M ammonia peroxy sulphate in aqueous medium. Both solutions kept 1 hour at room temperature then mixed in beaker, briefly stirred. And left at rest to polymerized, next day, the PANI precipitate was collected on a filter, washed with three 100 ml portion of 0.2 M HCL and similarly with acetone. Polyaniline hydrochloride powder was dried in air and then in vacuum at 60°C. Polyaniline prepared under these reactions and processing condition are further referred to as standard sample.

B) Synthesis of Zinc oxide (ZnO): It is prepared by the aqueous solution of zinc nitrate. And is prepared by dissolving 0.2 M of zinc nitrate hexahydrate in 100 ml of distilled water. To this aqueous zinc nitrate solution 0.2 M sodium hydroxide is added and the reaction mixture was heated at 80°C along with stirring and the

process is carried out for four hour after which the white precipitate was obtained. The formed oxide wet precipitate is centrifuged Then the wet precipitate is washed with de-ionized water to remove impurity ions present in it and further heated in the oven at 150°C to dry the precipitate formed.

Characterization :The above synthesized PANI-ZnO composites are structurally and surface morphologically characterized by using different technique like X- ray diffraction (XRD),the x-ray diffraction patterns of the prepared samples are obtained by Siemens D 5000 X-ray diffractometer using CuK α radition ($\lambda = 1.717 \text{ \AA}$). The diffractograms are recorded in terms of 2θ in the range 40°-50° at ambient temperature with scanning rate of 2° per minute .The surface morphology of polyaniline and its composites are studied by using Leica's SEM (modal S 440) at 10kv.

Result And Discussions:

XRD Pattern

of ZnO

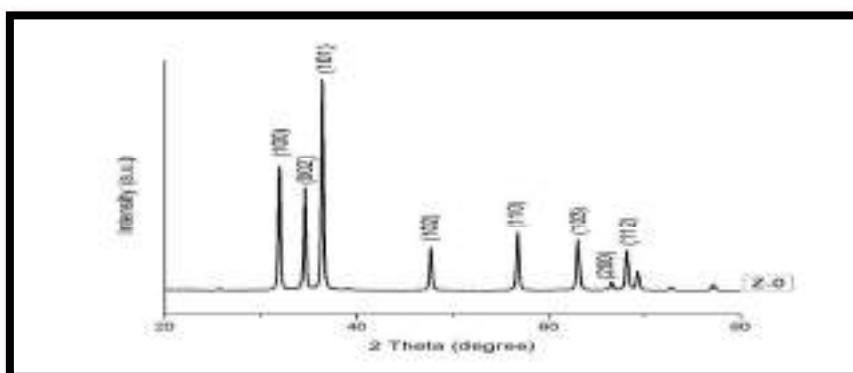


Fig. 1 XRD of Pure ZnO

XRD pattern of pristine zinc oxide (ZnO) nanostructure synthesized by liquid phase method via chemical wet reaction method were calculated at 800°C as shown in figure 1. The crystalline nature with 2θ peak lying at (100), (002), (101), (102), (110) and (103) planes. All the peaks match well the standard hexagonal wurtzite structure of zinc oxide (ZnO) with lattice constants $a = b = 0.3249 \text{ nm}$ and $c = 0.5206 \text{ nm}$ [JCPDS card no. 36-1451]. All the peaks are perfectly match with pure ZnO structure, which indicates the high purity of the obtained ZnO nanoparticle. The average crystalline size was found to be 37.38 nm calculated by Debye-Scherrer formula.

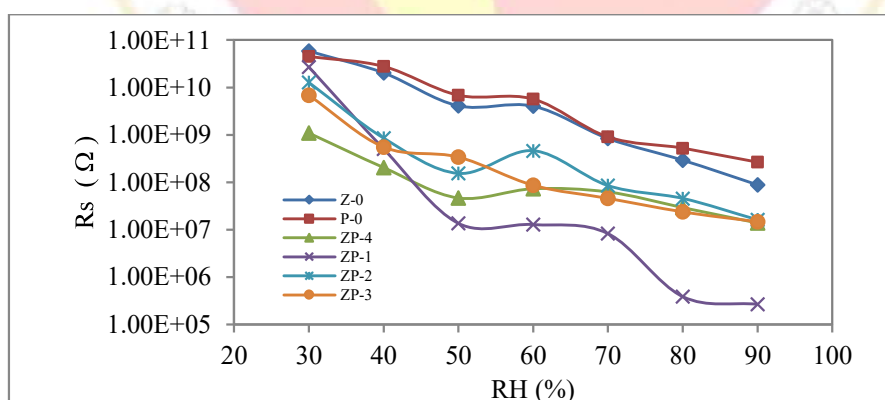
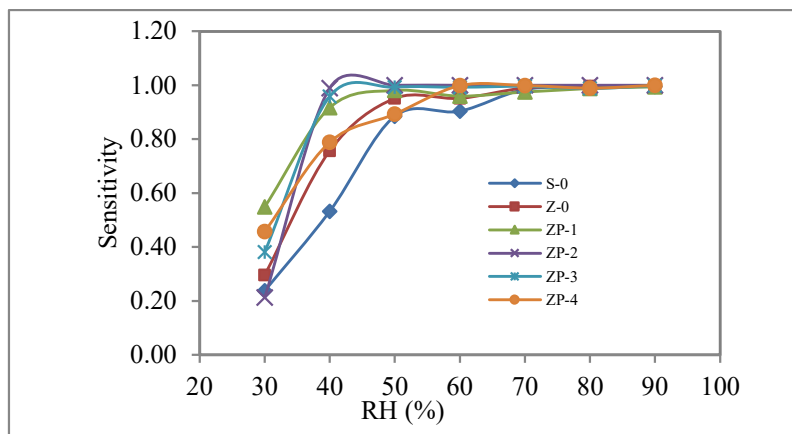


Figure: 2 Hysteresis plot

Hysteresis plot shows the variation between resistances of sample with respect to the relative humidity in increasing and decreasing order from 30 to 90 % RH as shown in the fig. 2. A very small hysteresis present during forward and reverse cycle of relative humidity, whereas a very significant average change observed in the value of resistance of sample, in the sample ZP-1 (10ZnO – 90PANI) the change in value of resistance is from $10^{11}\Omega$ to $10^5\Omega$, these is a remarkable change in the value of resistance.

Sensitivity**Figure: 3 Variation of sensitivity with Relative Humidity**

In the above samples the sensitivity is found to be increasing with the RH for all the samples of thick films and it is increasing up to some particular RH and then afterward it remains constant as shown in fig. 3 For higher RH the sensitivity is found to be higher in case of all samples of thick films. The sensitivity of ZP-1 (10ZnO-90PANI) is more than ZP-2, ZP-3, and ZP-4 samples and also from the pristine samples P-0 and Z-0. The (ZnO-PANI) composite sensors exhibits significantly higher sensitivity than sensor constructed specially from ZnO nanoparticles and PANI itself due to the formation of heterogeneous interface between them and more adsorption site was created to absorbed more water vapours.

5. Conclusions:

Nanostructured ZnO was successfully prepared via chemical precipitation method and PANI with IUPAC polymerization technique. Minimum crystallite size was found to be for ZnO is 37.38 nm. The Hysteresis plot shows very significant average change in the value of the resistance from $10^{11}\Omega$ to $10^5\Omega$ during forward and reversed cycles of sample ZP-1 (10ZnO-90PANI). The sensitivity is found to be increasing with the RH for all the samples of thick films and it is increasing up to some particular RH and then afterward it remains constant. Amongst all the prepared samples ZP-1 is more sensitivity than other prepared composite samples.

References:

- [1] Traversa E. Sens. Actuators B 23, 135-136., 1995
- [2] C.Y. Lee, G.B. Lee, Humidity sensors: a review, Sens. Lett. 3, 1-15. 2005
- [3] D. Hodgins, The electronic nose using conducting polymer sensors, Sens. Rev. 14 (4) (1994) 28-31.
- [4] V. K Khanna, V. K. and R.K., Nahar, Sens. Actuators 5, 187-198. 1984
- [5] D. Goram., And T., Marija., Sens. Actuators 18, 407-414. 1989
- [6] S.G Ansari., Z. A Ansari, M.R Kadam., Sens. Actuators B 21, 159-163. 1994
- [7] U Dellow. P. Keller. and Meyer J., Sens. Actuators A 61, 298-302. 1997
- [8] E Kan-Senchou. Kunaglee T. Jiin Liu F, Sens. Actuators B 56, 106-111. 1999
- [9] F Kock. and T Kockel., Chemical Engineering Journal, 76, 49-60. 2000.
- [10] R. Niranjana. and S Sathaye., Sens. Actuators B, 81, 64-67. 2001
- [11] W. Tai and J-H. Oh, Sens. Actuators B, 85, 154-157. 2002
- [12] K. Ogura, T. Saino, M. Nakayama, H. Shiigi, J. Mater. Chem. 7, 2363-2366., 1997
- [13] N. Parvatikar, S. Jain, C.M. Kanamadi., M.V.N.A., Prasad, J. Appl. Polym. Sci. 103, 653-658., 2007
- [14] S.T. McGovern, G.M. Spink, G.G. Wallace, Sens. Actuators B 107, 657-665. 2005
- [15] M.L. Singla., S. Awasthi, A. Srivastava., Sens. Actuators B 127, 580-585., 2007
- [16] M. Wan, Z.X. Wei, Z.M. Zhang, L.J. Zhang, K. Huang, Synth. Met. 135, 175-176. 2003

AAYUSHI INTERNATIONAL INTERDISCIPLINARY RESEARCH JOURNAL (AIIRJ)
ISSN 2349-638x (Peer Review and Indexed Journal) IMPACT FACTOR 7.331
Devgiri Nagar, Ambejogai Road, Latur
Tq. Latur, Dist. Latur. Pincode 413512.
State Maharashtra, India.
Email ID's: editor@aiirjournal.com , aiirjpramod@gmail.com
Website: www.aiirjournal.com



Certificate of Publication

Awarded to

T. R. Ingle, R. M. Agrawal, G. T. Lamdhade, R. J. Gajbe

For Contributing Research Paper

“Humidity Sensing Properties of PANI Doped Zinc Oxide Nanocomposites Thick Film Sensor”

In the

AAYUSHI INTERNATIONAL INTERDISCIPLINARY RESEARCH JOURNAL (AIIRJ)

Online Monthly Peer Review & Indexed Journal with ISSN 2349-638x (Impact factor 7.331)

for the month of **April 2022** with **Special Issue No. : 109**

Pramod Prakashrao Tandale
(Chief Editor)

Ba TiO₃ doped Polyaniline based Nanocomposites thick film sensor for humidity sensing Application

T R Ingle¹, R M Agrawal², G T Lamdhade³, R J Gajbe⁴

1,3Department of Physics, Vidya Bharati Mahavidyalaya, Camp, C. K. Naidu Road, Amravati(M.S.)

2Department of Physics Shri R.L.T.College of Science, Civil Lines Road, Akola,(M.S)

4Department of Electronics, Vidya Bharati Mahavidyalaya, Amravati M.S. India

Corresponding Authors:Email:ingletr@rediffmail.com,oumgajanan@gmail.com

Abstract

Polyaniline (PANI) and BaTiO₃-Pani composites were synthesized by chemical polymerization method using ammonium per sulphate (APS) as an oxidizing agent. This is a single step polymerization process to synthesize the conducting polymer. Thick films of PANI and BaTiO₃-Pani were fabricated by Screen – Printing followed by firing at 70° c for 30 min. BaTiO₃-Pani thick films resulted in humidity sensor. An exceptional sensitivity was found to the relative sensor at 80° c and no cross sensitivity was observed to other hazardous and polluting gases ever at higher concentration. . The effect of microstructure and dopant concentrations on the gas response, Hysteresis , sensitivity, of the sensor in the presence of humidity were studied and discussed.

Keywords: Polyaniline , Barium Titanate, BaTiO₃-Pani composites , Humidity sensor.

1. Introduction

The use of sensors by human being has been day by day increasing at an astounding rate in the last few years and modern society depends heavily on the use of the sensors for variety of purposes. Over the last decades a variety of chemical sensors have been developed based upon semiconductors, which monitor different characteristic sensor properties such as conductivities for electronic conductivity sensors, impedance for capacitance sensors, potentials for field effect sensors or temperatures for calorimetric sensors. For the determination of gas components, many of these devices make use of the same molecular detection principle. Depending on the operation temperature, their signals are caused by changes in the concentration of free electrons, dielectric constants, electrical fields and heats of adsorption or reaction. These changes result from physisorption, chemisorption, catalytic reactions, and surface or bulk defect reaction with particles from the gas phase.

There is a continuing need for accurate, reliable, inexpensive sensing systems for measuring relative humidity (RH), not only for human comfort but also for a broad spectrum of applications in chemical industry, process control, atmospheric sciences, agriculture etc. Humidity is one of the most common constituents present in the environment and its measurement is indispensable when it comes to monitoring of various environmental parameters. For instance, detecting organic pollutants in indoor atmosphere, organic vapour monitoring, maintenance of Green houses, performance of air/ smoke filters, hydrocarbon sensing are all affected by relative humidity conditions. Therefore, sensing and controlling relative humidity is of great importance [1].

In the recent years there has been significant progress in the field of polymer based humidity sensors. According to their sensing mechanisms these can be either resistive type or capacitive type. In addition to the traditional quaternary ammonium and sulfonate compounds, polymers containing phosphonium have also been studied for humidity sensing. Copolymers, mutually reactive copolymers and conjugated polymers have also been reported for humidity sensing. Conjugated polymers especially conducting polymers like polypyrrole, polyethylene, polypropylene etc. have shown humidity sensing properties [2]. Besides these metal-polymer nanocomposites for instance iron oxide-polypyrrole have also been reported for relative humidity sensing sensor. The present study deals with the humidity sensing application such as relative humidity, stability humidity, selectivity etc. of selected inorganic materials.

Humidity control and monitoring are of great interest to a wide area; these include moisture sensitive products, fresh and pack-age food, drug storage and environmental control for valuable Antiques or paintings etc. [3,4]. Humidity sensors that are available in the market include dew point, infrared, catalytic and tin oxide sensors, which may be expensive, or require high temperature operation and consume significant amount of power

and high cost of maintenance [5]. Much research has been focused on the development of humidity sensitive material [6–8]. Among these are the investigation of using conducting polymers such as polyaniline, polypyrrole, and polythiophene for humidity and gas sensing [9–11]. Advantages with polymers as sensing materials are light weight, flexible, low cost and simple fabrication process [12]. Pure polymer, polymer blends and polymer–inorganic composites have also been studied for the purposes, resulting in different degree of advancements in this area [13–19].

2. Synthesis of Material :

A) Synthesis of Polyaniline (PANI): In general is synthesized using two major polymerization approaches : electronic and chemical polymerization. In the present work polyaniline is synthesized by chemical polymerization method in which 0.2 M aniline hydrochloride is used as monomer unit . the synthesis is done by oxidative polymerization with 0.25 M ammonia peroxy sulphate in aqueous medium . both solution kept 1 hour at room temperature then mixed in beaker ,briefly stirred. And left at rest to polymerize, next day, the PANI precipitate was collected on a filter , washed with three 100 ml portion of 0.2 M HCL and similarly with acetone . polyaniline hydrochloride powder was dried in air and then in vacuum at 60°C. Polyaniline prepared under these reaction and processing condition are further referred to as standard sample[63].

B) Synthesis of Barium titanate (BaTiO₃): In preparation of barium titanate (BaTiO₃) 0.25 M Ba(NO₃)₂ solution and 0.25 M TiO(NO₃)₂ solution were dissolved in 2 N nitric acid solution in a beaker. About 0.6 M tartaric acid solution was then added to under constant magnetic string. The solution heated under continuous string to its boiling point until all the liquid evaporated. About 7 gm of ammonium nitrate was added towards the ends to avoid slurry formation. Brown fumes evolution takes places and fluffy mass were settled at the base of the beaker. The product is then dried in vacuum oven at 96°C for 2 hrs. So that moisture will removed from the final product and we will get dry product. Then this dry product was crushed into fine powder and finally this fine nanopowder of BaTiO₃ was calcinated at temperature 800°C for 5 hrs. in the auto controlled muffle furnace to remove the impurity form the product will be completely removed and get a final product of BaTiO₃ nanoparticle.

3. Experimental methodology for humidity: -

The humidity chamber/ stability chamber has stainless steel body purchased from Gayatri Scientific, Mumbai . The working size of humidity chamber is 50×43×430cm (H×D×W). The chamber having two open side door window namely inner side door and outside door, inner side is full length inner plexi glass door which is transparent and outer side door was metal door with magnetic gasket and lock . inside the chamber it has two trays, which are made of stainless steel. it has PID digital display temperature indicators which is outside at the top of the chamber , indicating the current temperature of the chamber .indicating the current temperature of chamber. The power sources of 230 V having frequency of 50Hz required. the temperature range is ambient to 80 C. The chamber has temperature accuracy of ± 1.0 C .

The range of humidity varying from 40% RH to 80% RH having the accuracy of ± 3% RH. To operate the chamber a program feed for the varying RH as well as different temperature. Which will sets the controlled the humidity of set temperature for given time. To measure the humidity characteristics , the sensor element was placed inside the humidity chamber by operating the program at different RH and temperature the humidity was varied . The sensor element was placed inside the humidity was measured by standard hygrometer which is already placed inside chamber . The electrical resistance of the film was measured by using Keithley source meter (2400). The humidity and temperature of the chamber can be controlled by using Bull's Eye control UK-100, it is a program controller for humidity and temperature verses time even controller .The UK-100 can be program by varying RH, temperature , time parameters. The total number of programs are 99 out of this are set program for set temperature 40 C by varying RH 1 steps NO. Of program 8, event 0001, are tabulated

4.Characterization :

The above synthesized PANI-BaTiO₃ composites are structurally and surface morphologically characterized by using different technique like X- ray diffraction (XRD), and scanning electron microscopy (SEM).the x-ray diffraction patterns of the prepared samples are obtained by Siemens D 5000 X-ray diffractometer using CuK α radiation ($\lambda = 1.717 \text{ \AA}$). The diffractograms are recorded in terms of 2θ in the range

40°-50° at ambient temperature with scanning rate of 2° per minute .The surface morphology of polyaniline and its composites are studied by using Leica's SEM (modal S 440) at 10kv.

X-Ray Diffraction

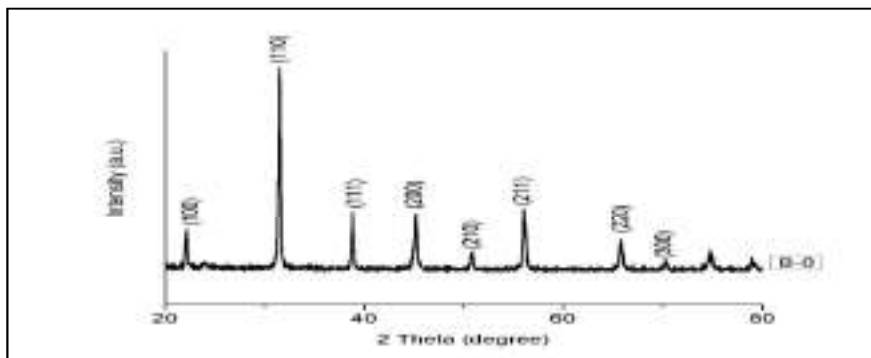


Figure: 1 XRD of (BaTiO₃) (B-0)

The XRD pattern of pristine Barium Titanate (BaTiO₃) nanostructure synthesized by liquid phase method via solid state method calcinated at 800°C as shown in figure 1. The crystalline nature with 2θ peak lying at (100), (110), (111), (200), (210) and (220) planes. All the peaks match well the standard perovskite type structure of Barium Titanate (BaTiO₃) with lattice constants $a = 3.992$ nm and $c = 4.036$ nm . All the peaks are perfectly match with pure BaTiO₃ structure, which indicates the high purity of the obtained BaTiO₃ nanoparticle. The average crystalline size was found to be 46.90 nm calculated by Deye-Scherrer formula.

Hysteresis plot :

Hysteresis plot shows the variation between resistances of sample with respect to the relative humidity in increasing and decreasing order from 30 to 90 % RH as shown in the figure. 2. A very small hysteresis present during forward and reverse cycle of relative humidity, where as a very significant average change observed in the value of resistance of sample, in the sample BP-1 (10BaTiO₃ – 90PANI) the change in value of resistance is from 10¹¹Ω to 10³Ω, these is a remarkable change in the value of resistance.

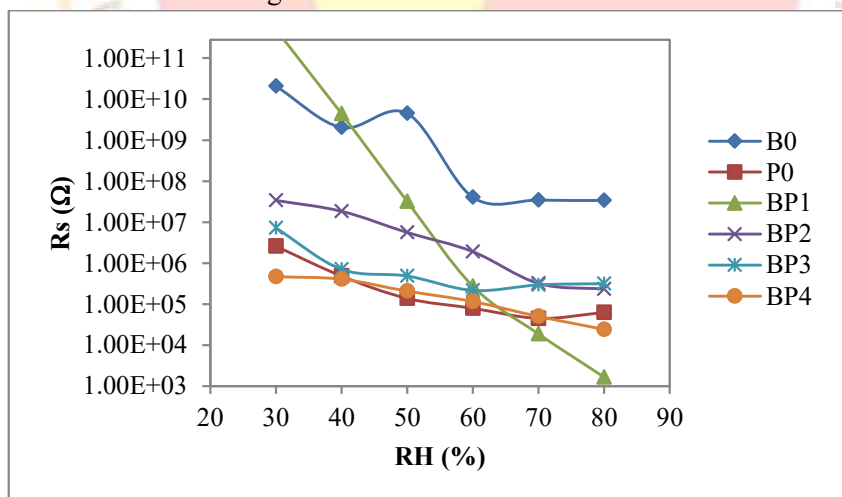


Figure: 2 Hysteresis plot

Sensitivity

In the above samples the sensitivity is found to be increasing with the RH for all the samples of thick films and it is increasing up to some particular RH and then afterward it remains constant as shown in figure 3. For higher RH the sensitivity is found to be higher in case of all samples of thick films. The sensitivity of BP-1 (10 BaTiO₃ -90PANI) is more than BP-2, BP-3, and BP-4 samples and also from the pristine samples P-0 and B-0. The (BaTiO₃-PANI) composite sensors exhibits significantly higher sensitivity than sensor constructed specially from Barium Titanate nanoparticles and PANI itself due to the formation of heterogeneous interface between them and more adsorption site was created to absorbed more water vapours.

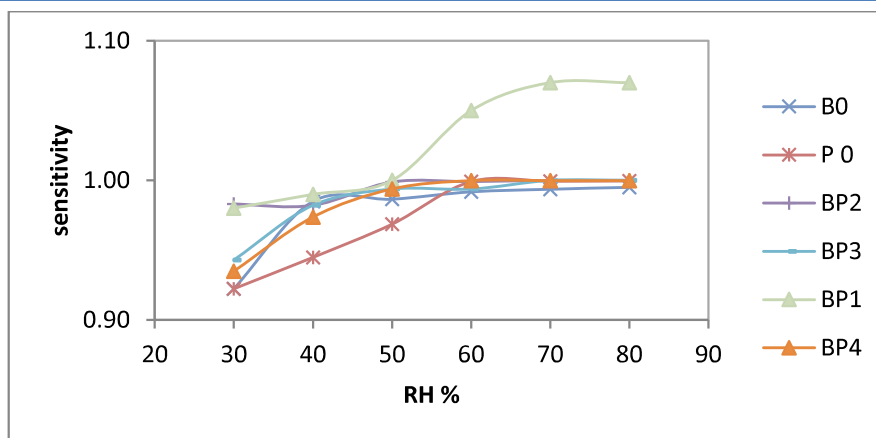


Figure: 3 Sensitivity curve

5. Conclusions

Nanostructured BaTiO₃ was successfully prepared via chemical precipitation method and PANI with IUPAC polymerization technique. Minimum crystallite size was found to be BaTiO₃ is 46..90 nm. The Hysteresis plot shows very significant average change in the value of the resistance from 10¹¹Ω to 10³Ω during forward and reversed cycles of sample BP-1 (10 BaTiO₃ -90PANI)). The sensitivity is found to be increasing with the RH for all the samples of thick films and it is increasing up to some particular RH and then afterward it remains constant. Amongst all the prepared samples BP-1 is more sensitivity than other prepared composite samples.

References:

- [1] Traversa E. *Sens. Actuators B* 23, (1995), 135-136.
- [2] Ansbacher F. and Jason A. C., *Nature* 24, (1993), 177-178.
- [3] E. Bracken, Combating humidity the hidden enemy in manufacturing, *Sens. Rev.* 17 (1997) 291–298.
- [4] C.Y. Lee, G.B. Lee, Humidity sensors: a review, *Sens. Lett.* 3 (2005) 1–15.
- [5] D. Hodgins, The electronic nose using conducting polymer sensors, *Sens. Rev.* 14 (4) (1994) 28–31.
- [6] S.K. Khijwania, K.L. Srinivasan, J.P. Singh, An evanescent-wave optical fiber relative humidity sensor with enhanced sensitivity, *Sens. Actuators B* 104 (2005) 217–222.
- [7] Z.M. Rittersma, Recent achievements in miniaturised humidity sensors are view of transduction techniques, *Sens. Actuators A* 96 (2002) 196–210.
- [8] Z. Chen, C. Lu, Humidity sensors: a review of materials and mechanisms, *Sens. Lett.* 3 (2005) 274–295.
- [9] N. Parvatikar, S. Jain, S.V. Bhoraskar, M.V.N.A. Prasad, Spectroscopic and electrical properties of polyaniline /CeO₂ composites and their application as humidity sensor, *J. Appl. Polym. Sci.* 102 (2006) 5533–5537.
- [10] T. Maddanimath, I.S. Mulla, S.R. Sainkar, K. Vijayamohan, K.I. Shaikh, A.S. Patil, S.P. Vernekar, Humidity sensing properties of surface functionalised polyethylene and polypropylene films, *Sens. Actuators B* 81 (2002) 141–151.
- [11] A.G. MacDiarmid, Polyaniline and polypyrrole: where are we headed? *Synth. Met.* 84 (1997) 27–34.
- [12] G. Harsanyi, Polymer films in sensor applications: a review of present uses and future possibilities, *Sens. Rev.* 20 (2000) 98–105.
- [13] R. Nohria, R.K. Khillan, Y. Su, R. Dikshit, Y. Lvov, K. Varahramyan, Humidity Sensor based on ultrathin polyaniline film deposited using layer-by-layer nano-assembly, *Sens. Actuators B* 114 (2006) 218–222.
- [14] K. Ogura, T. Saino, M. Nakayama, H. Shiigi, The humidity dependence of the electrical conductivity of a soluble polyaniline–poly(vinyl alcohol) composite film, *J. Mater. Chem.* 7 (1997) 2363–2366.
- [15] N. Parvatikar, S. Jain, C.M. Kanamadi, B.K. Chougule, S.V. Bhoraskar, M.V.N.A. Prasad, Humidity sensing and electrical properties of polyaniline /cobalt oxide composites *J. Appl. Polym. Sci.* 103 (2007) 653–658.
- [16] S.T. McGovern, G.M. Spink, G.G. Wallace, Micro-humidity sensors based on a processable polyaniline blend, *Sens. Actuators B* 107 (2005) 657–665.
- [17] M.L. Singla, S. Awasthi, A. Srivastava, Humidity sensing using polyaniline/Mn₃O₄ composite doped with organic/inorganic acids, *Sens. Actuators B* 127 (2007) 580–585.
- [18] M. Wan, Z.X. Wei, Z.M. Zhang, L.J. Zhang, K. Huang, Y.S. Yang, Studies on nanostructures conducting polymers via self-assembly method, *Synth. Met.* 135 (2003) 175–176.
- [19] S. Jain, S. Chakane, A.B. Samui, V.N. Krishnamurthy, S.V. Bhoraskar, Humidity sensing with weak acid-doped polyaniline and its composites, *Sens. Actuators B* 96 (2003) 124–129.

AAYUSHI INTERNATIONAL INTERDISCIPLINARY RESEARCH JOURNAL (AIIRJ)
ISSN 2349-638x (Peer Review and Indexed Journal) IMPACT FACTOR 7.331
Devgiri Nagar, Ambejogai Road, Latur
Tq. Latur, Dist. Latur. Pincode 413512.
State Maharashtra, India.
Email ID's: editor@aiirjournal.com , aiirjpramod@gmail.com
Website: www.aiirjournal.com



Certificate of Publication

Awarded to

T. R. Ingle, R. M. Agrawal, G. T. Lamdhade, R. J. Gajbe

For Contributing Research Paper

“Ba TiO₃ doped Polyaniline based Nanocomposites thick film sensor for humidity sensing Application”

In the

AAYUSHI INTERNATIONAL INTERDISCIPLINARY RESEARCH JOURNAL (AIIRJ)

Online Monthly Peer Review & Indexed Journal with ISSN 2349-638x (Impact factor 7.331)

for the month of **April 2022** with **Special Issue No. : 109**

Pramod Prakashrao Tandale
(Chief Editor)