IJPSR (2022), Volume 13, Issue 6 (Research Article)

INTERNATIONAL JOURNAL OF PHARMACEUTICAL SCIENCES AND RESEARCH

Received on 31 August 2021; received in revised form, 03 November 2021; accepted, 17 November 2021; published 01 June 2022

QUANTITATIVE STRUCTURE-ACTIVITY RELATIONSHIP (QSAR) STUDIES OF SUBSTITUTED BENZIMIDAZOLE ANALOGUES WITH POTENT ANTIBACTERIAL ACTIVITY

Shambhavi Parab, Chandani Muleva, Jinal Shah, Bhairavi Murkute and Madhura Vaidya *

Department of Pharmaceutical Chemistry, MET Institute of Pharmacy, Bhujbal Knowledge City, Bandra Reclamation, Bandra West, Mumbai - 400050, Maharashtra, India.

Keywords:

Benzimidazole, Antimicrobial, Quantitative structure-activity relationship, Descriptor

Correspondence to Author:
Dr. P. Madhura Vaidya

Associate Professor, Department of Pharmaceutical Chemistry, MET Institute of Pharmacy, Bhujbal Knowledge City, Bandra Reclamation, Bandra West, Mumbai - 400050, Maharashtra, India.

E-mail: madhurav_iop@met.edu

ABSTRACT: Computational chemistry is a robust and economical tool for designing and developing potent therapeutic molecules. Benzimidazole and its analogues are compounds with a unique pharmacophore that derives its application as broad-spectrum antimicrobial agents. A quantitative structure-activity relationship (QSAR) study was performed using a data set of 28 benzimidazole analogues with antimicrobial activity. Descriptors were generated using various free software's such as PaDEL-Descriptor, SwissADME and OCHEM (Online chemical modelling environment). The statistical analysis was performed using the Stats. Blue software by multiple linear regression (MLR) method for determining the relationship between dependant variable and various independent variables. The best QSAR model had an r^2 value of 0.6773, and the predictive r^2 value for external validation was 0.7150. The results show that a positive correlation is established between the descriptors, TPSA (Topological polar surface area), H-bond acceptors, iLOGP (Implicit LOGP), GGI4 (Galvez topological charge indices of order 4), thereby achieving the most accurate outputs.

INTRODUCTION: Benzimidazole is a versatile pharmacophore in medicinal chemistry to design and develop novel molecules of therapeutic importance. It is a heterocyclic organic compound formed by fusing an imidazole nucleus with a benzene ring containing 2 nitrogen atoms at positions 1 and 3, exhibiting an amphoteric nature [1]. The chief bioactive candidate, benzimidazole, exhibits a range of biological activities including anti-bacterial, anti-oxidant, anti-inflammatory, anti-cancer, anthelmintic, anti-ulcer, antipsychotic, antiprotozoal and anti-bacterial antifungal.

QUICK RESPONSE CODE

DOI: 10.13040/IJPSR.0975-8232.13(6).2358-66

This article can be accessed online on www.ijpsr.com

DOI link: http://dx.doi.org/10.13040/IJPSR.0975-8232.13(6).2358-66

The action of benzimidazole derivatives on different targets like DNA, transpeptidase, microtubules, and fumarate reductase enzymes leads to the inhibition of various organisms. Benzimidazole molecules with antibacterial activity are known to form covalent adducts with the membrane-bound bacterial transpeptidase enzymes penicillin-binding proteins (PBPs). Thereby preventing the formation of cell walls, eventually leading to cell wall decomposition and death [2].

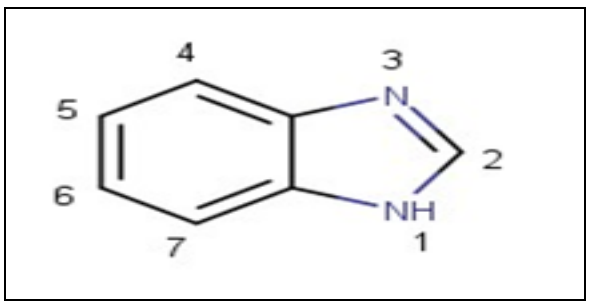


FIG. 1: 1H-1, 3-BENZODIAZOLE

In terms of the SAR, the prime factors that contribute to the antibacterial activity are electron-withdrawing substituents, chlorine, or bromine at position 5 of the benzimidazole ring and -$CH(CH_3)NH_2$ or -CH_2Cl at position 2. Derivatives with a branched methyl group (-$CH(CH_3)NH_2$) at position 2 result in enhanced antibacterial activity as compared to the CH_2NH_2 group at the same position. Generation of new potent microbial inhibitors is possible by fine-tuning and appropriate modification of substituents placed at positions 1, 2 and 5, toxicological evaluation would be necessary to confirm their utility for medicinal use [3]. Quantitative structure-activity relationship (QSAR) is a statistical modelling method for predicting the biological activity of compounds based on the statistical and mathematical relationships obtained [4]. It correlates the topology of a molecule with the various physicochemical descriptors, such as TPSA, HBA, etc and the biological potency of unknown ligands with known ligands. Multiple linear regression and partial least square regression are the conventional QSAR methods capable of decoding linear relationships. There are different types of descriptors: 1D descriptors, 2D descriptors, and 3D descriptors depending on the following parameters: molecular weight, LogP, number of functional groups, topological indices, geometrical parameters, molecular surfaces, quantum chemistry descriptors. During the past few years, several *in-silico* studies and quantitative structure-activity relationships (QSAR) models have been developed for benzimidazole analogues [5, 6]. Identification of the chemical structure by a diversity of molecular descriptors is an important step in QSAR studies. Only a small subset of these calculated descriptors can generate the QSAR model of interest. This study is aimed to develop a robust and accurate model for the antimicrobial activity of benzimidazole analogues.

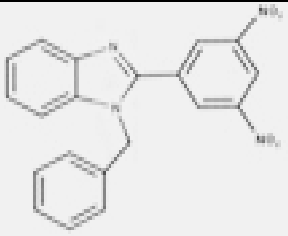
The QSAR models generated multiple linear regressions (MLR) as linear methods.

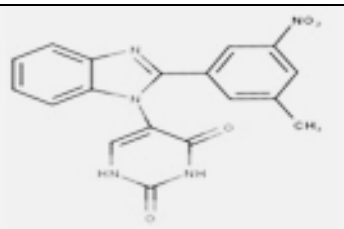
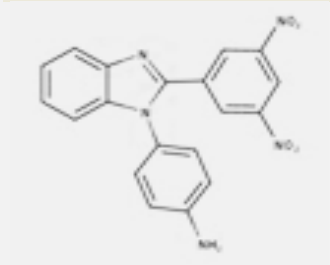
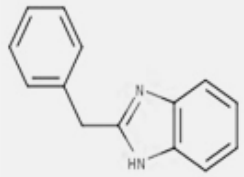
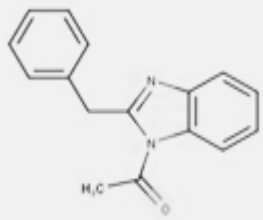
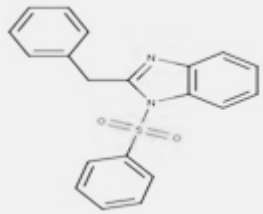
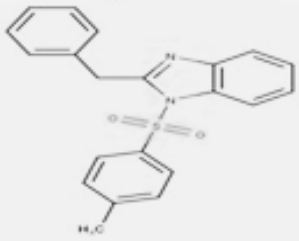
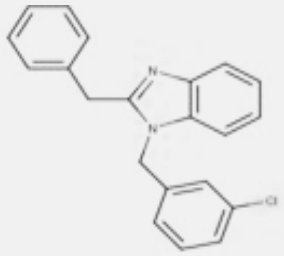
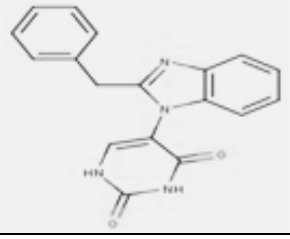
MATERIALS AND METHODS: The complete computerized Quantitative structure-activity relationship (QSAR) studies for benzimidazole as the basic pharmacophore was carried out using HP Laptop, with Intel(R) core processor and Windows 10 operating system. The chemical structures of various benzimidazole derivatives were sketched and optimized using Marvin Sketch version 21. The obtained chemical structures were then converted into SMILES using an online tool import/export smiles, Marvin for JS version 6.2, ChemAxon.

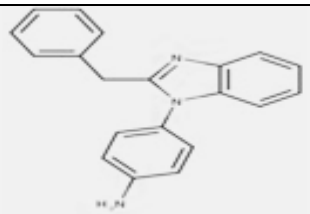
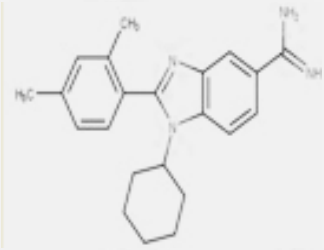
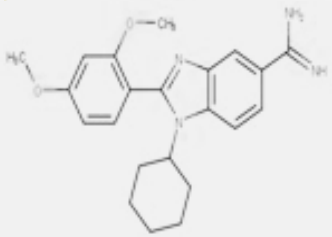
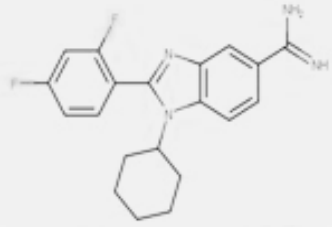
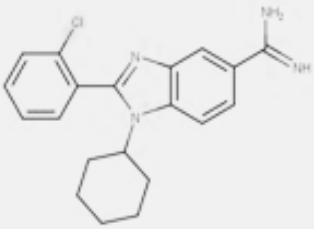
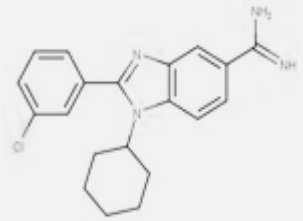
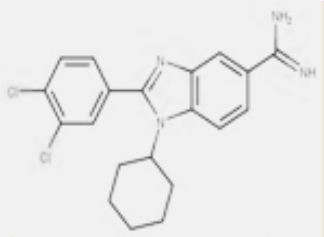
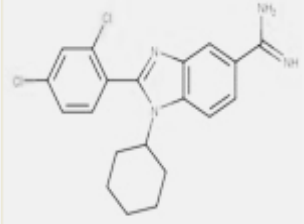
Descriptor Calculation: The molecular descriptors were calculated for the optimized data set molecules. Various online available software's such as online chemical modelling environment (OCHEM) [7], Padel-descriptor [8] and SwissADME [9] were employed to calculate and optimize the 2-dimensional molecular descriptors. Of all the descriptors obtained HBA (Hydrogen bond acceptor), TPSA (Topological polar surface area), iLOGP and GGI4, were considered extensively for the generation of the QSAR equation. A variety of other descriptors were also obtained, including apol, VBAC, nAtomP, AMR, Kier1, Kier2, Kier3, XLOGP3, WLOGP, ALOGP, SCH-6, SCH-7, SC-3, SC-5, GGI3, GGI4, SpAD_D, VCH-7, SPC-5, GGI2, SpDiam_D, SPC-4, SpMax_D.

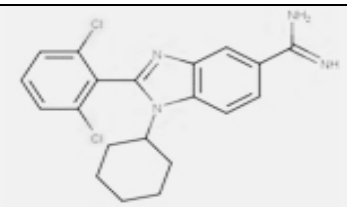
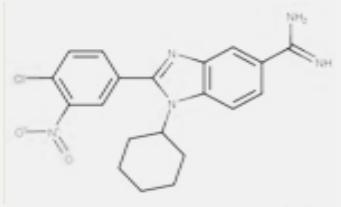
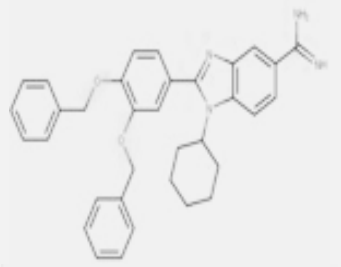
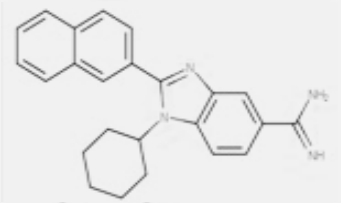
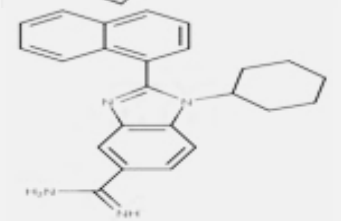
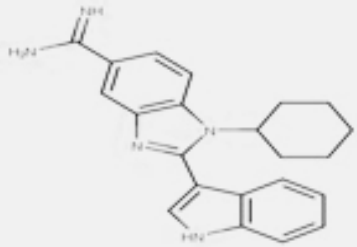
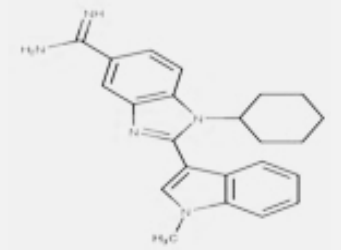
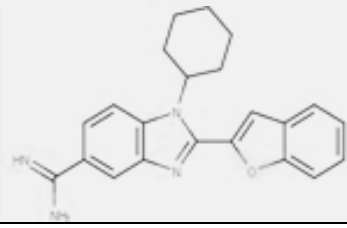
Experimental Data: For the generation of QSAR models, firstly, 32 benzimidazole analogues were selected from various reported literature for studying their antibacterial activity [10, 11, 12]. Further for the 2D-QSAR model development, the pMIC values were calculated as Log (1/MIC), where the literature reported MIC values were first converted into μM/mL.

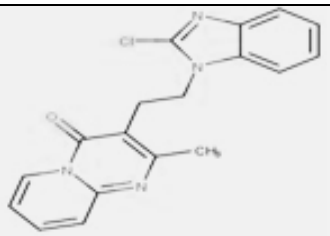
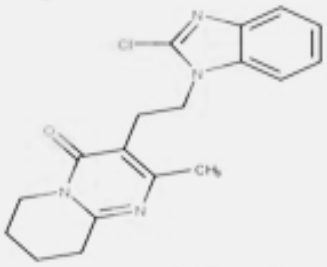
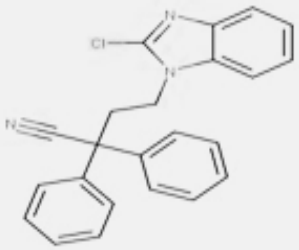
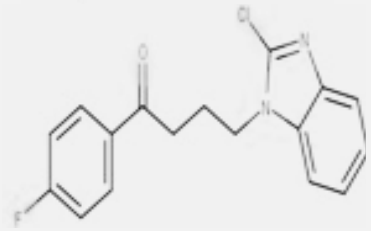
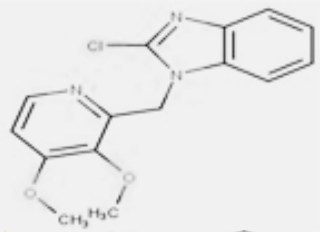
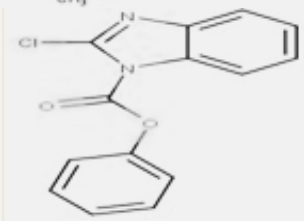
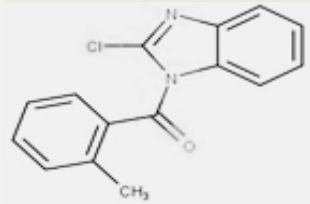
TABLE 1: COMPOUNDS USED IN QSAR MODEL GENERATION [10, 11, 12]

Compound code	S. no.	Structure	pMIC (*S. aureus*)	pMIC (*P. aeruginosa*)	Smiles
MR.	1[a]		1.0788	-	[O-] [N+] (=O) C1=CC (=CC(=C1)C1=NC2=CC=CC= C2N1CC1=CC=CC=C1)[N+]([O-])=O

4e	2[a]		1.3665	-	CC1=CC(=CC(=C1)[N+]([O-])=O)C1=NC2=CC=CC=C2N1C1=CNC(=O)NC1=O
4f	3[b]		1.0798	-	NC1=CC=C(C=C1)N1C2=CC=CC=C2N=C1C1=CC(=CC(=C1)[N+]([O-])=O)[N+]([O-])=O
5	4[a]		1.1249	1.125	C(C1=NC2=C(N1)C=CC=C2)C1=CC=CC=C1
6a	5[a]		5.2146	-	CC(=O)N1C(CC2=CC=CC=C2)=NC2=C1C=CC=C2
6b	6[b]		1.3487	0.7462	O=S(=O)(N1C(CC2=CC=CC=C2)=NC2=C1C=CC=C2)C1=CC=CC=C1
6c	7[a]		-	0.4622	CC1=CC=C(C=C1)S(=O)(=O)N1C(CC2=CC=CC=C2)=NC2=C1C=CC=C2
6d	8[a]		1.0277	1.0273	ClC1=CC=CC(CN2C(CC3=CC=CC=C3)=NC3=C2C=CC=C3)=C1
6e	9[a]		0.4059	0.7071	O=C1NC=C(N2C(CC3=CC=CC=C3)=NC3=C2C=CC=C3)C(=O)N1

6f	10[a]		0.3793	0.6802	NC1=CC=C(C=C1)N1C(CC2=CC=CC=C2)=NC2=C1C=CC=C2
N1	11[a]		1.3224	0.7208	CC1=CC(C)=C(C=C1)C1=NC2=C(C=CC(=C2)C(N)=N)N1C1CCCCC1
N2	12[a]		0.7719	0.7719	COC1=CC(OC)=C(C=C1)C1=NC2=C(C=CC(=C2)C(N)=N)N1C1CCCCC1
N3	13[a]		-0.1598	0.4423	NC(=N)C1=CC2=C(C=C1)N(C1CCCCC1)C(=N2)C1=C(F)C=C(F)C=C1
N4	14[a]		1.0424	0.4405	NC(=N)C1=CC2=C(C=C1)N(C1CCCCC1)C(=N2)C1=C(Cl)C=CC=C1
N5	15[a]		1.644	0.7414	NC(=N)C1=CC2=C(C=C1)N(C1CCCCC1)C(=N2)C1=CC(Cl)=CC=C1
N6	16[b]		1.6861	0.4808	NC(=N)C1=CC2=C(C=C1)N(C1CCCCC1)C(=N2)C1=CC(Cl)=C(Cl)C=C1
N7	17[b]		1.6861	0.4808	NC(=N)C1=CC2=C(C=C1)N(C1CCCCC1)C(=N2)C1=C(Cl)C=C(Cl)C=C1

N8	18[a]		1.6861	0.4808	NC(=N)C1=CC2=C(C=C1)N(C1CCCCC1)C(=N2)C1=C(Cl)C=CC=C1Cl
N9	19[a]		1.3958	0.7937	NC(=N)C1=CC2=C(C=C1)N(C1CCCCC1)C(=N2)C1=CC(=C(Cl)C=C1)[N+]([O-])=O
N10	20[a]		1.8210	0.9187	NC(=N)C1=CC2=C(C=C1)N(C1CCCCC1)C(=N2)C1=CC(OCC2=CC=CC=C2)=C(OCC2=CC=CC=C2)C=C1
N11	21[a]		1.6635	0.4592	NC(=N)C1=CC2=C(C=C1)N(C1CCCCC1)C(=N2)C1=CC=C2C=CC=CC2=C1
N12	22[a]		-0.1429	0.7602	NC(=N)C1=CC2=C(C=C1)N(C1CCCCC1)C(=N2)C1=C2C=CC=CC2=CC=C1
N13	23[b]		1.6497	0.7471	NC(=N)C1=CC2=C(C=C1)N(C1CCCCC1)C(=N2)C1=CNC2=C1C=CC=C2
N14	24[a]		1.6676	0.4627	CN1C=C(C2=NC3=C(C=CC(=C3)C(N)=N)N2C2CCCCC2)C2=C1C=CC=C2
N15	25[a]		1.6517	0.7484	NC(=N)C1=CC2=C(C=C1)N(C1CCCCC1)C(=N2)C1=CC2=CC=CC=C2O1

5a	26[a]		0.433	0.734	CC1=C(CCN2C(Cl)=NC3=CC=CC=C23)C(=O)N2C=CC=CC2=N1
5b	27[b]		0.359	0.5351	CC1=C(CCN2C(Cl)=NC3=CC=CC=C23)C(=O)N2CCCCC2=N1
5c	28[a]		0.4735	0.5704	ClC1=NC2=CC=CC=C2N1CCC(C#N)(C1=CC=CC=C1)C1=CC=CC=C1
5d	29[a]		0.5007	0.4038	FC1=CC=C(C=C1)C(=O)CCCN1C(Cl)=NC2=CC=CC=C12
5e	30[a]		0.6865	0.3856	COC1=C(OC)C(CN2C(Cl)=NC3=C2C=CC=C3)=NC=C1
5f	31[a]		0.1347	0.0377	ClC1=NC2=CC=CC=C2N1C(=O)OC1=CC=CC=C1
5g	32[a]		0.7388	0.4325	CC1=C(C=CC=C1)C(=O)N1C(Cl)=NC2=CC=CC=C12

Training set [a], test set [b]

Evaluation and Validation of QSAR Models: The chosen benzimidazole analogues were divided into a training set (22) and test set (6) based on the proximity of the representative test set to the representative training set concerning activity and descriptor space. Four molecules (5, 7, 13, and 22) with significantly different biological activity data were removed from the data set. The statistical

analysis was performed using the Stats. Blue online software by Multiple linear regression (MLR) method for determining the relationship between dependant variable and various independent variables. The training set was employed for the development of the QSAR model whereas the test set was used to challenge the robustness of the model by determining the predictive ability of the model. Good predictive power for the QSAR model is the one that has a predictive r^2 value greater than 0.5.

RESULTS AND DISCUSSIONS: The equations were generated using a training set consisting of 22 molecules and a test set (3, 23, 17, 16, 27, and 6) that consisted of 6 molecules **Table 3**. From the QSAR equation, the r^2 value is 0.6773, and the predictive r^2 value for external validation is 0.7150 **Table 2**. The equation shows a positive correlation between Topological Polar Surface Area (TPSA), Implicit logP (iLOGP), and GGI4 descriptors.

The models were used to predict pMIC values of the training set and test set **Table 3**. **Fig. 2A** and **Fig. 2B** indicate the plots of actual versus predicted pMIC of the training set and test set compounds. The difference between the actual and predicted pMIC is indicated as residual value **Table 3**. The lower residual values indicate good predictivity of the developed QSAR model.

TABLE 2: BEST QSAR MODEL

Equation	**pMIC=-1.2008-0.1959*#H-bond acceptors + 0.0132* TPSA +0.4211 *Ilogp +0.4692* GGI4**
r^2	0.6773
r^2_{adj}	0.6013
Residual standard error	0.3317 on 17 degrees of freedom
Overall F-statistic	8.9183 on 4 and 17 degrees of freedom
Overall p-value	0.0005
Predictive r^2	0.9961 (internal validation) 0.7150 (external validation)

TABLE 3: ACTUAL AND PREDICTED PMIC VALUES OF THE TRAINING SET AND TEST SET

Compound Number	**Actual pMIC**	**Predicted pMIC**	**Residuals**
1[a]	1.079	0.9360	0.143
2[a]	1.367	1.2822	0.084
3[b]	1.0798	1.0169	0.0629
4[a]	1.125	0.9951	0.130
6[b]	1.349	0.7588	0.590
8[a]	1.028	1.0915	-0.064
9[a]	0.406	0.9867	-0.581
10[a]	0.379	1.1397	-0.760
11[a]	1.322	1.4817	-0.159
12[a]	0.772	1.2452	-0.473
14[a]	1.042	1.3177	-0.275
15[a]	1.644	1.3795	0.265
16[b]	1.6861	1.4338	0.2523
17[b]	1.6861	1.4298	0.2563
18[a]	1.042	1.3681	0.318
19[a]	1.644	1.4170	-0.021
20[a]	1.042	1.8288	-0.008
21[a]	1.644	1.4927	0.171
23[b]	1.6497	1.3871	0.2626
24[a]	1.668	1.4757	0.192
25[a]	1.652	1.4626	0.189
26[a]	0.433	0.8296	-0.397
27[b]	0.359	0.9258	-0.5668
28[a]	0.474	0.8875	-0.414
29[a]	0.501	0.3894	0.111
30[a]	0.687	0.3260	0.360
31[a]	0.135	0.4199	-0.285
32[a]	0.739	0.6489	0.090

Training set a, test set b.

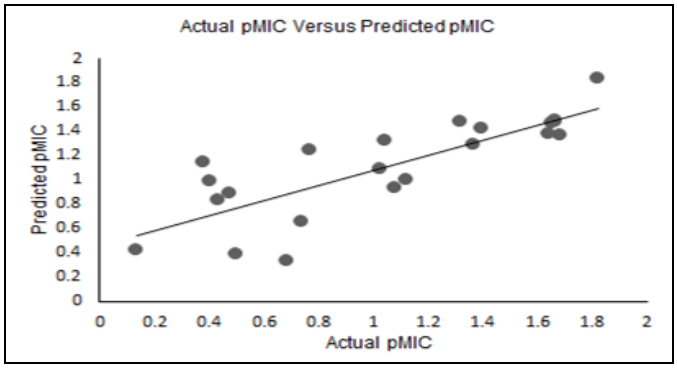


FIG. 2A: GRAPH OF ACTUAL VERSUS PREDICTED PMIC OF THE TEST SET

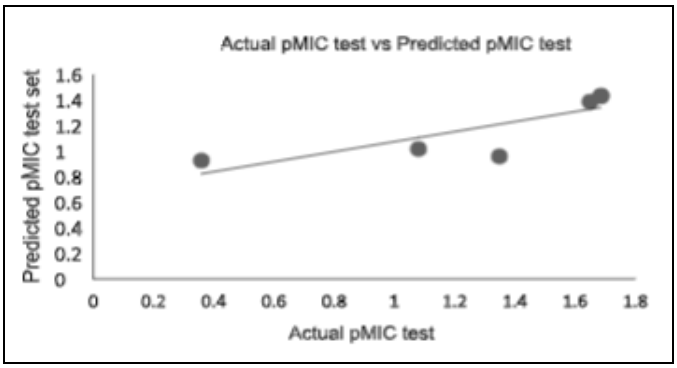


FIG. 2B: GRAPH OF ACTUAL VERSUS PREDICTED PMIC OF THE TEST SET

H-bond acceptor Count is negatively correlated with pMIC values. Hydrogen Bond Acceptor (HBA) atoms are stated to have a lone pair of electrons [13]. The study shows that derivatives involving a sulfo group, a hydrogen bond acceptor group in the side chain, show lower activity [14]. Topological polar surface area (TPSA) of a molecule is defined as the sum of the surface area of all polar atoms or molecules, mainly oxygen and nitrogen, along with their attached hydrogen atoms [15] Antibacterial activity is found to be linearly correlated with TPSA, in this series, higher polarity apparently tends to increase the antibacterial activity. iLOGP signifies a ratio of solute concentration in octanol and water [16]. GGI4 is a descriptor with a topological charge index of order four. GGI4 gave maximum contribution in the model, as its contribution positively affects the model, the steady increase in this descriptor value can further improve the pMIC value. GGI4 shows the ability to represent the molecular charge distribution by comparing them with the dipole moment of heterogeneous set of hydrocarbons. Thus an increase in the number of heterogeneous hydrocarbons will lead to elevation of the bioactivity of the compound [17].

CONCLUSION: The research focuses on an in-depth study on benzimidazole and its analogues as potent antimicrobial agents. The conducted literature survey highlighted the therapeutic efficacy of benzimidazole derivatives employed for the treatment of various diseases. Appropriate modification and fine-tuning of substituents at positions 1, 2 and 5 results in generation of newer derivatives that can be attractive research candidates. Multiple linear regression and partial least square regression, the conventional QSAR methods capable of decoding linear relationships, are employed in the study. Descriptors required to generate QSAR equations were generated using various free software like OCHEM, PaDEL descriptor, and SwissADME. The final QSAR model with the GGI4 descriptor providing the highest contribution in the model gives an excellent correlation between the dependent and independent variables. The positive correlation of TPSA, iLOGP, GGI4 exhibits that an increase in the value of these descriptors will increase the biological activity of the molecules. On the contrary, the negative correlation of the H-bond acceptor indicates that a decrease in its value will increase the overall biological activity of the molecules. This, therefore, suggests the ability of the benzimidazole derivatives as potential candidates for further research.

ACKNOWLEDGEMENT: The authors would like to thank MET Institute of Pharmacy, Mumbai for the encouragement and valuable assistance.

CONFLICTS OF INTEREST: The authors declare no conflicts of interest.

REFERENCES:

1. Singh VK and Parle A: The intriguing benzimidazole: a review. International Journal of Pharmaceutical Science & Research 2019; 10(4): 1540-52.
2. Kumar U, Narang R, Nayak SK, Singh SK and Gupta V: Benzimidazole: Structure-Activity Relationship and Mechanism of Action as Antimicrobial Agent. Research Journal of Pharmacy and Technology 2017; 10(7): 2400.
3. Alasmary F, Snelling A, Zain M, Alafeefy A, Awaad A and Karodia N: Synthesis and Evaluation of Selected Benzimidazole Derivatives as Potential Antimicrobial Agents. Molecules 2015; 20(8): 15206-15223.
4. Muratov EN, Bajorath J, Sheridan RP, Tetko IV, Filimonov D, Poroikov V and Tropsha A: QSAR without borders. Chem Society Reviews 2020; 49(11): 3525-3564.

5. Zivkovic J, Kalauzović S, Milosavljević M & Kalauzović K: *In-silico* evaluation of selected benzimidazole derivatives in the discovery of new potent antimicrobial agents. Acta Medica Medianae 2019; 58(1): 106-115.
6. Chintakunta R and Meka G: Synthesis, *In-silico* studies and antibacterial activity of some novel 2-substituted benzimidazole derivatives. Future Journal of Pharmaceutical Sciences 2020; 6(1): 128.
7. Sushko I, Novotarskyi S, Körner R, Pandey AK, Rupp M, Teetz W and Tetko IV: Online chemical modeling environment (OCHEM): Web platform for data storage, model development and publishing of Chemical Information. Journal of Computer-Aided Molecular Design 2011; 25(6): 533-554.
8. Yap CW: PaDEL-descriptor: An open source software to calculate molecular descriptors and fingerprints. Journal of Computational Chemistry 2011; 32(7): 1466-1474.
9. Daina A, Michielin O and Zoete V: SwissADME: a free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. Scientific Report 2017; 7(1): 42717.
10. Gund DR, Tripathi AP and Vaidya SD: Synthesis and antimicrobial activity of some novel N-substituted benzimidazoles. European J of Chem 2017; 8(2): 149-154.
11. Erol M, Celik I, Temiz-Arpaci O, Goker H, Kaynak-Onurdag F and Okten S: Synthesis, molecular docking and ADME prediction of some new benzimidazole carboxamidines derivatives as antimicrobial agents. Medicinal Chemistry Research 2020; 29(11): 2028-2038.
12. Ajani OO, Tolu-Bolaji OO, Olorunshola SJ, Zhao Y and Aderohunmu DV: Structure-based design of functionalized 2-substituted and 1,2-disubstituted benzimidazole derivatives and their *in-vitro* antibacterial efficacy. Journal of Advanced Research 2017; 8(6): 703-712.
13. Karas LJ, Wu CH, Das R, & Wu JI: Hydrogen bond design principles. Wiley interdisciplinary reviews. Computational Molecular Science 2020; 10(6): 1477.
14. Ai Y, Wang ST, Sun PH and Song FJ: Combined 3D-QSAR Modeling and Molecular Docking Studies on Pyrrole-Indolin-2-ones as Aurora A Kinase Inhibitors. International J of Molecular Sci 2011; 12(3): 1605-1624.
15. Caron G & Ermondi G: Molecular descriptors for polarity: The need for going beyond polar surface area. Future Medicinal Chemistry 2016; 8(17): 2016-0165.
16. Daina A, Michielin O & Zoete V: iLOGP: A Simple, Robust and Efficient Description of n-Octanol/Water Partition Coefficient for Drug Design Using the GB/SA Approach. J of Chem In and Model 2014; 54: 3284–3301.
17. Olasupo SB, Uzairu A, Shallangwa G and Uba S: QSAR analysis and molecular docking simulation of norepinephrine transporter (NET) inhibitors as anti-psychotic therapeutic agents. Heliyon 2019; 5(10): 02640.

How to cite this article:
Parab SB, Muleva CT, Shah JD, Murkute BC and Vaidya MP: Quantitative structure activity relationship (QSAR) studies of substituted benzimidazole analogues with potent antibacterial activity. Int J Pharm Sci & Res 2022; 13(6): 2358-66. doi: 10.13040/IJPSR.0975-8232.13(6).2358-66.