Available online www.jocpr.com

Journal of Chemical and Pharmaceutical Research, 2022, 14(7):01-11

Research Article

ISSN: 0975-7384
CODEN(USA): JCPRC5

Application of Computational Strategies on Optimize Novel Benzimidazol Hybrids as Serin Proteases Inhibitors

Gajanan Sonwane[1*], Mayura kale [1], Vishnukant Mourya [2], Athira Baburaj[2]

[1]*Department of Pharmaceutical Chemistry, Rajarshi Shahu University, Maharashtra, India*

[2]*Department of Cellular and Molecular Biology, Marien Ngouabi University, Congo, Brazzaville*

Received: *24-Mar-2020, Manuscript No. JOCPR-20-8535;* ***Editor assigned:*** *27-Mar-2020, PreQC No. JOCPR-20-8535 (PQ);* ***Reviewed:*** *11-Apirl-2020, QC No. JOCPR-20-8535;* ***Revised:*** *03-Aug-2022, ManuscriptNo. JOCPR-20-8535 (R);* ***Published:*** *31-Aug-2022, DOI: 10.37532/0975-7384.2022.14(7).001.*

ABSTRACT

To understand pharmacophore properties of Pyrimidine-Benzimidazole hybrids derivatives and to design novel inhibitors of Serin protease, GQSAR approach was applied to analyze Group quantitative structure-activity relationship of 17 compounds. AutoDock 4.0 program was employed to locate the orientations and conformations of the inhibitors interacting with Serin protease. The interaction mode was demonstrated in the aspects of inhibitor conformation, hydrogen bonding, and electrostatic interaction. Similar binding conformations of these inhibitors and good correlations between the calculated binding free energies and experimental biological activities suggest that the binding conformations of these inhibitors derived from docking procedure were reasonable. Robust and predictive G-QSAR model was obtained by R_2 with Q_2 values of 0.9865 and 0.9405 for cross-validated respectively. The G-QSAR model built here will provide clear guidelines for novel inhibitors design based on the Pyrimidine-Benzimidazole hybrids derivatives against Serin protease.

Keywords: Pyrimidine-Benzimidazol hybrids*;* Serin protease*;* Docking*;* G-QSAR

INTRODUCTION

Uncontrolled expansion of cells in the body due to a numeral of abnormalities of the hereditary matter leads to the diseased state called “cancer”. According to the WHO, Cancer is a major offender in stealing lives of many

people around the globe with approximately 14 million new cases and 8.2 million cancer-related deaths. Cancer may be benign or malignant where malignant tumors possess the properties of dedifferentiation, metastasis and invasiveness [1]. The most common causes of cancer death are cancers.

- Lung (1.59 million deaths)
- Liver (745 000 deaths)
- Stomach (723 000 deaths)
- Colorectal (694 000 deaths)
- Breast (521 000 deaths)
- Oesophageal cancer (400 000 deaths)

The treatment for cancer is essentially chemotherapy, radiological therapy or surgery. Chemotherapy utilities by one or the other way as described beneath:

- Inhibit the synthesis of precursors of DNA
- Destruction of the DNA of the cell
- Breakdown the mitotic spindles or cause interruption in any of the steps involved in synthesis.

Benzimidazole is a most wanted target as a budding drug moiety since of its broad spectrum of activity. Here it is conjugated with pyrimidine for a better protuberance of its pharmacological activity. QSAR make available a new aspect of drug development since its appearance. Of course, it is a time-saving and cost-reducing enterprise, but the major expand being the improvement of new chemical entity with improved pharmacological activity and reduced toxicological activity [2-5]. QSAR, as the name suggests, is a correlation of the structure with its pharmacological activity. Many inventors are focusing on QSAR as a promising device for the drug discovery process [6].

Theory

The following cluster of descriptors has been tested to describe the toxicity of the compound of the first set and the values of these descriptors have been calculated on Dragon software.

Constitutional Descriptors

The following descriptors of this set have been used in the study: Molecular Weight (MW), Average Molecular.

Weight (AMW), Sum of atomic van der Waals volumes, scaled on Carbon Atom (Sv), Sum of atomic Sanderson electronegativities, scaled on Carbon Atom (Se), Sum of atomic polarizabilities, scaled on carbon molecule , Sum of Kier-vestibule Electrotopological States (Ss), Mean atomic van der Waals volume, scaled on Carbon Atom (Mv), Mean atomic Sanderson electronegativity, scaled on Carbon Atom (Me), Mean atomic polarizability, scaled on Carbon Atom (Mp), Mean Electrotopological State (Ms) and Rotatable Bond Fraction (RBF).28–32

Geometrical Descriptors

The following descriptors of this set have been used in the study: 3D-Wiener index (W3D), 3D-Balaban index (J3D), 3D-Harary index (H3D), Average Geometric Distance Degree (AGDD), D/D Index (DDI), Average, Distance/Distance Degree (ADDD), Gravitational Index G1 (G1), Gravitational Index G2 (G2), Radius of Gyration, Mass Weighted (Rgyr), Span R (SPAN), Average Span R (SPAM), 3D Petijean Shape Index (PJI3), Length-to Breadth relation by WHIM (L/Bw), absolute Eigen worth sum on geometry matrix (SEig) and d COMMA2 value/weighted by Atomic Masses (DISPm). 2.3 BCUT Descriptors [7].

The following descriptors of this class have been taken in the study: BEHm1, BEHm2, BEHm3, and BEHm4 have the highest eigenvalue no. 1-4 of Burden matrix/weighted by atomic masses, BELm1, BELm2, BELm3, BELm4 have the lowest eigenvalue no. 1-4 of Burden matrix/weighted by atomic masses [8]. BEHv1, BEHv2, BEHv3, BEHv4 have the highest eigenvalue no. 1-4 of Burden matrix/weighted by atomic van der Waals volumes and BELv1, BELv2, BELv3, BELv4 have the lowest eigenvalue of Burden matrix/weighted by atomic van der Waals volumes [9].

Empirical Descriptors

The following descriptors of this class have been studied: Hydrophilic factor (Hy), Aromatic Ratio (ARR), Ghose Crippen Molar Refractivity (MR), and Moriguchi Octanol-water partition coeff (MLOGP, log P), MW. 2.5 Topological Descriptors

The following descriptors of this class have been studied: Information index on molecular size (ISIZ), total information Index of Atomic Composition (IAC), first Zagreb index by valence vertex degrees (ZM1V), Narumi Simple Topological Index (SNar, log), Narumi Harmonic Topological Index (HNar), Total Structure Connectivity Index (Xt), Pogliani Index (Dz), Polarity Number (Pol), Gutman Molecular Topological Index (GMTI), Xu index (Xu), Super Pendentic Index (PI), Wiener W index (W), Harary H index (Har), detour index (w), Balaban J Index (J), Molecular Electrotopological Variation (DELS), Kier symmetry index (S0K), Kier Flexibility Index (PHI), Kier Benzene-Likeliness Index (BLI) and Lopping Centric Index (Lop). Galvez Topol Charge Indices.

The following descriptors of this class have been studied: GGI1, GGI2, GGI3 and GGI4 are the topological charge indices of order, and JGI1, JGI2, JGI3, JGI4 are the signify topological incriminate indices of order.The JGT is the global topological charge index [10].

MATERIALS AND METHODS

The molecular modeling was performed on the Molecular Design Suite on Windows 7 operating system.

Data Set

A set of 23 pyrimidine-benzimidazole derivatives in Figure 1 had been used to build a model for the study. These compounds have been reported to show anticancer activity [11]. The pIC50 values reported in the research article

had been used in the model building process. The structures were drawn on 2D draw application of the software. They were then converted to 3D format and optimized using Merk Molecular Force Field (MMFF) in Table 1. The samples were then subjected to calculation of the physicochemical and alignment descriptors [12].

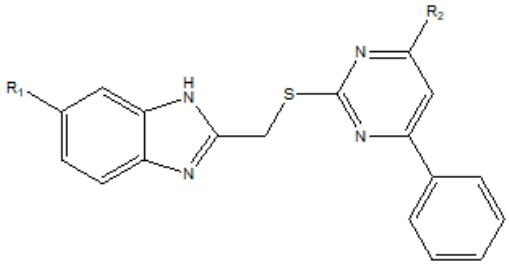
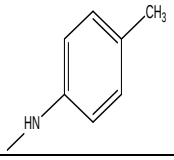
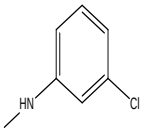
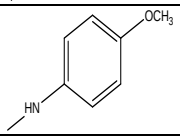
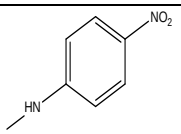
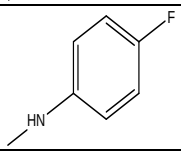
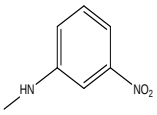
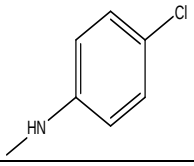
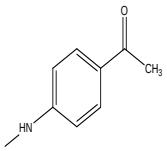
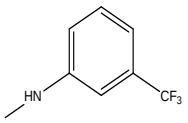
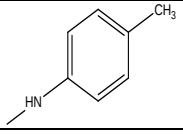
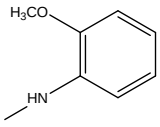
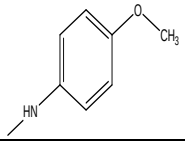
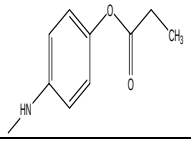
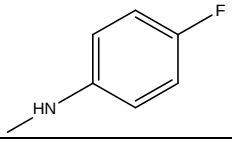
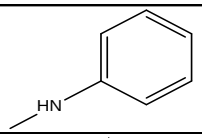
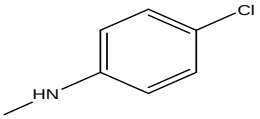
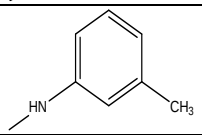
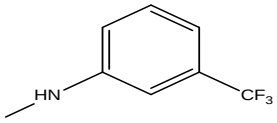
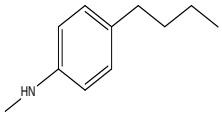
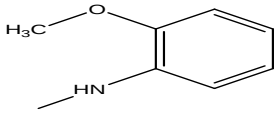
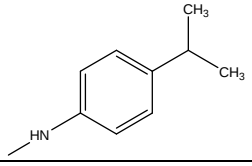
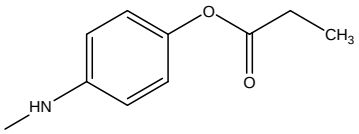
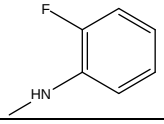


Figure 1: 2-((4-methyl-6-phenylpyrimidin-2-ylthio) methyl)-6-methyl-1H-benzo[d]imidazole derivatives.

Table 1: Structures and predicted activities of designed compounds.

Comp.	R_1	R_2	pIC_{50}	Comp.	R_1	R_2	pIC_{50}
5a	H		5.8761	5m	H		4.7765
5b	H		5.6925	5n	H		4
5c	H		5.6382	5o	H		4
5d	H		5.4894	5p	H		4
5e	H		5.4156	6a	Cl		5.9706
5f	H		4.3389	6b	Cl		5.9746
5g	H		4	6c	Cl		5.3665

5h	H		5.2132	6d	Cl		5.3555
5i	H		4.4782	6e	Cl		5.58
5j	H		4.819	6f	Cl		4.2747
5k	H		4.6976	6g	Cl		4
5l	H		5.139	-	-	-	-

Molecular Fragmentation

G-QSAR involves the correlation between the biological activity and the molecular fragment of interest. The fragmentation pattern is based on the chemical nature of series of compound i.e. Congeneric (template-based) and non-congeneric (chemically diverse). Here we are using congeneric series which makes use of a template with dummy atoms at the position of substitution [13]. This helps in planning the substitution of the compound with the atoms or group predicted to give better activity.

Data Selection

Here the data is divided into training and test set by sphere exclusion method. The division to training and test set is important for true external validation; otherwise the predictive ability of the samples will be calculated statistically based on the training set alone [14]. This would not provide a good prediction of activity, thereby making it significant to form test set. The dissimilarity value is inversely proportional to the size of the training set [15]. This step is followed by the model building.

Model Building

Partial Least Square (PLS), Multiple Linear Regression (MLR) and Principal Component Regression (PCR) are some of the approaches used for model building in QSAR [16]. These are used in concurrence with stepwise multiple regression which is helpful when the number of independent variables (descriptors) are greater than that of the molecules [17].

Validation of the Model

The model is validated using Leave-One-Out (LOO) or Leave-Some-Out (LSO) procedure which gives q_2 value. This value greater than 0.5 is considered as significant, robust and predictive. The compounds from the training set were

Removed individually, and the activity of each was predicted using the model fitted to the remaining molecules [18].

The process is repeated until all compounds in the training set are exhausted. The predictive power of the developed models was additional validated using the squared correlation coefficient (pred-r2) of the test set [19].

Docking Study

The crystal structure of Serine protease site model in complex with paromomycin (PDB entry 1KCT) was recovered from Brookhaven Protein Database. Molecular modeling of Pyrimidine-Benzimidazol hybrids derivatives was performed using software Autodock 4.0. Hydrogen atoms were added using standard geometrical parameters as implemented in Autodock [20]. The geometries of these compounds were subsequently optimized using Vlife MDS 4.5, Gasteiger Huckel charge with distance dependent dielectric, and conjugate gradient method with convergence criterion of 0.05 kcal/mol. Finally, Polar hydrogen and Kollman-allatom charges were added to Serin protease site accordingly [21]. The flexible docking program AutoDock 4.0 was used to tackle the interacting mode of a serial of inhibitors with Serin protease site model. Force field scoring, as used in AutoDock, tends to overestimate electrostatic energy. To temper an unreasonably strong attraction between the highly negative phosphate oxygen atoms in the RNA backbone and ammonium ions of the inhibitors [22].

The grid for energy evaluation was set in the center of gravity of the ligand with dimensions of 60 points and spacing of 0.375 A˚. Initial translation, quaternion, and torsion steps of 2.0 A˚, 50.0, and 50.08, respectively, were chosen with a reduction factor of 1 per cycle.No rotatable bonds in antibiotic were constrained in the course of docking. Standard Lamarckian Genetic Algorithm (LGA) parameters were used and the number of individuals in the population, energy evaluations, and generations was set to 100, 500,000, and 27,000 respectively [23]. Each docking simulation consisted of 50 independent docking runs. Finally, the 50 docked conformations for each derivative were extracted from the docking log files [24]. AutoDock found more than one possible binding modes for each Pyrimidine-Benzimidazol hybrids derivative. The docked complexes of Pyrimidine-Benzimidazol hybrids and Serin protease site model were selected according to the criteria of the most populated cluster combined with geometrical matching quality. We chose docking free energy, rather than total binding free energy, as the ranking criteria, which gave a better correlation of predicted and experimental interaction energy as shown in Tables 2 and 3. The conformations selected according to this criterion were then used as the starting point for further energetic minimization and geometrical optimization in Table 2 [25,26].

RESULTS AND DISCUSSION

In the present study, an automated procedure combing with docking and G-QSAR methods was successfully applied to a set of newly synthetic Pyrimidine-Benzimidazol derivatives against Serin protease. The conformations derived from docking procedure are reasonable and the hydrogen bond binding mode of neamine core is homologous when compared with the crystal structure of Pyrimidine-Benzimidazol in complex with Serin protease site model. Docking free energy evaluated for these derivatives by AutoDock program correlate well with the reported inhibitory activities as shown in Figures **2-**7. The binding mode between the Pyrimidine-Benzimidazol and Serin protease site model provide not only the insight into the nature of interaction, but also the theoretically evidence that neamine core is a

functionally conservative component and may involve the specificity of hydrogen bonding interaction compounds as shown in (Tables 3 and 4). Better understanding of the structure requirement of neamine core is helpful in the designing the new antibiotics.

Docking Result

pIC 50=R1-polarizabilityAHP(0.4154(±0.0259))+ R1-chiV4path cluster(-0.8167(±0.2588))+ R1 T_T_O_2(1.7179(±0.1399)+R1-chi4(-2.0311(±0.4558)+R1-PolarSurfaceAreaExcludingPandS(-0.2478(±0.0223)+ R1-T_T_C_4(-0.3747(±0.0189))+R1-T_C_N_3(2.3214(±0.3349))+ R1-T_T_Cl_1(-0.9443(±0.1035))+ R1-chi_2(1.9742(±0.1608))+R1-T_T_O_7(-4.2696(±0.3735))+0.0000).

Table 2: Docking results.

Binding energy	Ligand efficiency	Inhibit-constant	Inhibi-constant unit	Internal energy	Vdw-hb-desolv_energy	Electrostatic energy	Total-internal	Torsional energy	Unbound enegry	Refrms
-10.14	-0.4	36.93	nm	-11.07	-11	-0.1	-0.8	1.19	-0.55	21.1

Table 3: Model by Multiple linear regration correlation.

N	Degree-of-freedom	R_2	Q_2	F-test	R_2-se	Q_2-se	Pred-R_2	Pred-R_2se
19	8	0.9865	0.9435	58.3002	0.1283	0.2623	0.6148	0.4276

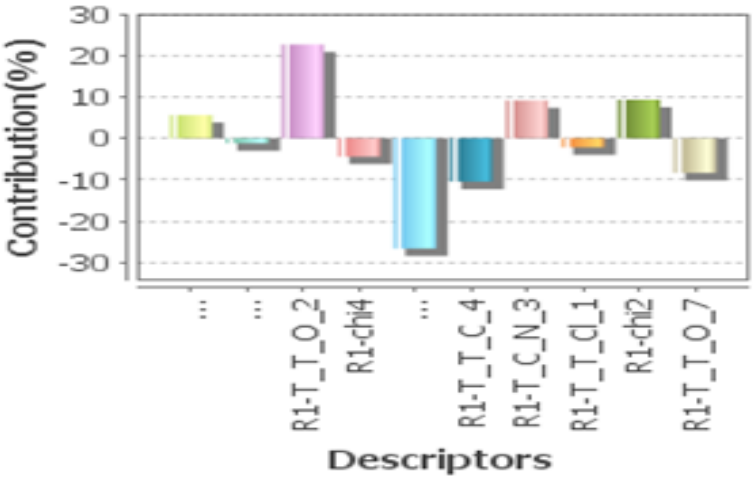


Figure 2: Activity contribution plot.

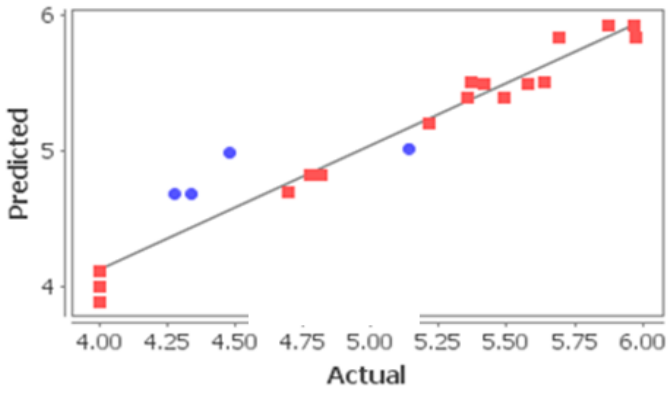


Figure 3: Activity Fitness plot. Note: Training Test.

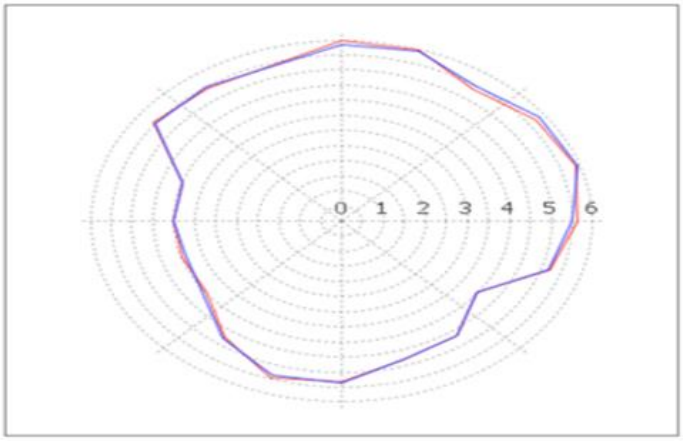


Figure 4: Activity plot of actual and predicted biological activity for training set.
Note: Actual Predicted.

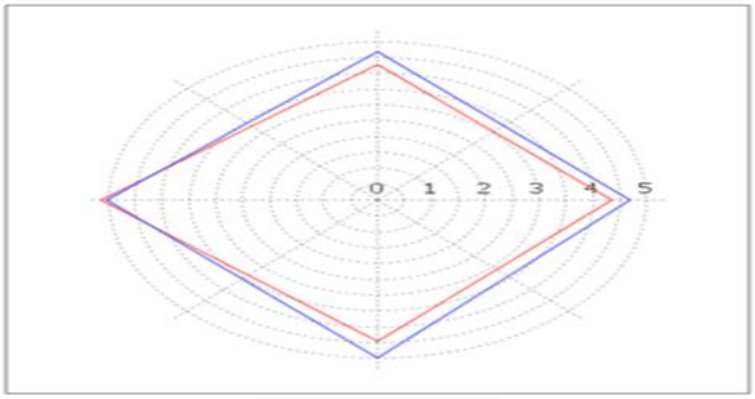


Figure 5: Activity plot of Actual and Predicted Biological activity for Test set.
Note: Actual Predicted.

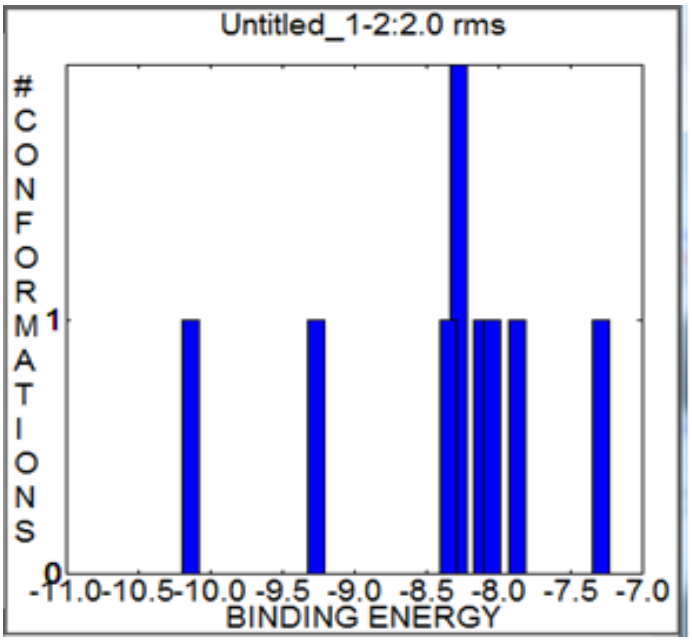


Figure 6: Binding energy of confirmation of novel Benzimidazol hybrids.

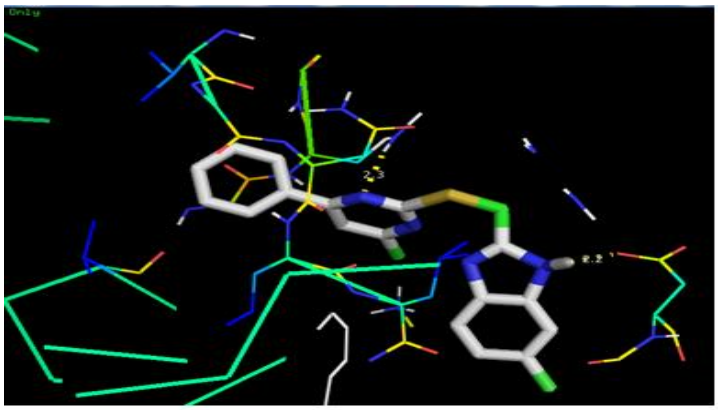


Figure 7: Binding interaction Serine protease site.

Table 3: ADME study of designed compounds.

ID	Value
BBB	1.01
Buffer_solubility_mg_L	139
$Caco_2$	45.2
CYP-2C19-inhibition	Inhibitor
CYP-2C9-inhibition	Inhibitor
CYP-2D6-inhibition	Non
CYP-2D6-substrate	Non

CYP-3A4-inhibition	Inhibitor
CYP-3A4-substrate	Non
HIA	94.9
MDCK	0.62
Pgp-inhibition	Non
Plasma-Protein_Binding	100
Pure_water_solubility_mg_L	1.79
Skin_Permeability	-3.3
SKlogD-value	4.31
SKlogP-value	4.31
SKlogS-buffer	-3.4
SKlogS-pure	-5.3

Table 4: Toxicity study of designed compounds.

ID	Value
Algae-at	0.01484
Ames-test	Mutagen
Carcino-Mouse	Negative
Carcino-Rat	Negative
Daphnia-at	0.0137
hERG-inhibition	Medium_risk
Medaka-at	0.00047
Minnow-at	0.00091
TA100-10RLI	Positive
TA100-NA	Negative
TA1535-10RLI	Negative
TA1535-NA	Positive

CONCLUSION

The binding mode between the Pyrimidine-Benzimidazol and Serin protease site model provide not only the insight into the nature of interaction, but also the theoretically evidence that neamine core is a functionally conservative component and may involve the specificity of hydrogen bonding interaction. Better understanding of the structure requirement of neamine core is helpful in the designing the new antibiotics. Based on active conformations extracted from docking procedure, high affinity and Pyrimidine-Benzimidazol -based inhibitors against Serin protease site.

GQSAR models were generated. Among them most significant model has squared correlation coefficient cross validated correlation coefficient (q_2) and predictive correlation coefficient 0.9865,0.9435 and 0.6148; respectively.R1-polar surface area excluding pands, descriptors which control the biological activity and are inversely related to it, R1-polarizability AH, pathCluster are descriptors which control the biological activity and are directly related to it. In the present study, an attempt has been made to identify the necessary structural and substituent sites which can be modified so as to improve the biological activity. From the present GQSAR, best models were generated among which any one can be used for predicting the activity of the newly designed compounds in finding some more potent molecules.

Finally, it is concluded that the work presented here will play an important role in understanding the relationship of physiochemical parameters with structure and biological activity. By studying the GQSAR model one can select the suitable substituent for further synthesizing bioactive compounds showing maximum potency.

ACKNOWLEDGEMENTS

Authors are thankful to principal government college of pharmacy, Aurangabad and Rajarshi Shahu College of Pharmacy buldana for their valuable CADD support

REFERENCES

1. Twombly RJ. *Natl Cancer Inst*. **2004**;96: 414-1415.
2. Macmillan. *Macmillan Cancer Inf*. **2014**;1: 15-20.
3. Burton W. *Cleeland CS Pain Pract.* **2001**; 1: 236-242.
4. W Wung. *BMJ*. **2013**;347:1-15.
5. Siegel R L, Miller KD, Jemal A, et al. *Cancer J Clin*. **2017**;67: 7-30.
6. Jakab Z, Galea G, Mauer-Stender K, et al. *Lancet Oncol*. **2014**;15: 47.
7. Gail MH, Brinton LA, Byar DP, et al. J *Natl Cancer Inst*. **1989**;81: 1879-1886.
8. Brawley OW. *J Natl Cancer Inst*. **2014**;106: 15-20.
9. Perera FP. *Sci*.**1997**;278: 1068-1073.
10. Chambers SK, Girgis A, Occhipinti S, et al. *J Turner B Morris and J Dunn Eur J Cancer Care (Engl)*. **2012**; 21: 213-223.
11. Rosen Sleep GM. *Med Clin*. **2007**; 2: 491-500.
12. Fung-Kee-Fung M, Oliver T, Elit L, et al. *Curr Oncol*. **2007**;14:195-208.
13. Bang YJ, Van Cutsem E, Feyereislova A, et al. *Lancet*. **2010**;376:687-697.
14. Surgery HG. *Lancet* **2006**; 24: 66-69.
15. Cramer RD. *J Comput Aided Mol Des*. 2012; 26: 35-38.
16. M Golbraikh, Shen Z, Xiao Y, et al. *J Comput Aided Mol Des*. **2003**;17: 241-253.
17. Polanski J, Bak A, Gieleciak R, et al. *J Chem Inf Model*. **2006**;46: 310-2318.
18. Salum LB and Andricopulo AD. *Expert Opin Drug Discov*. **2010**;5: 405-412.
19. Patel HM, Noolvi MN, Sharma MN, et al. *Med Chem Res*. **2014**; 23:4991-5007.
20. Polanski J, Bak A. *J Chem Inf Comput Sci*. **2003**; 43: 2081-2092.
21. Schubert W, Ugi I. *J Am Chem Soc*. **1978**;100: 37-41.
22. Khan KR, Sahu VK, Singh RK, et al. *Med Chem Res*. **2009**; 18:770-781.
23. González MP, Helguera AM, Medina R, et al. *Intern Elect J Mol Des*. **2004**; 3: 200-208.
24. Gosav S, Praisler M, Dorohoi D, et al. *Talanta*. **2006**; 70: 922-928.
25. Agatonovic-Kustrin S, Beresford R, Yusof APM *J Pharm Biomed Anal* **2001**; 25: 227-237.
26. Katritzky AR, Gordeeva E V. *J ChemInf Comput Sci*. **1993**; 33: 835–857.