Contents lists available at UGC-CARE

International Journal of Pharmaceutical Sciences and Drug Research

[ISSN: 0975-248X; CODEN (USA): IJPSPP]

Available online at www.ijpsdronline.com

Research Article

In-silico Studies of Heterocyclic Benzoxazole Derivatives as an Anticancer Agent: Molecular Docking, 2D and 3D QSAR

Smita J. Pawar*, Dhanashri Zope, Amol P. Kale

PDEA's Department of Pharmaceutical Chemistry, Seth Govind Raghunath Sable College of Pharmacy, Pune, Maharashtra, India.

ARTICLE INFO

Article history:

Received: 05 September, 2023

Revised: 20 October, 2023

Accepted: 27 October, 2023

Published: 30 November, 2023

Keywords:

Benzoxazole, Docking, Thymidylate synthase, 2D QSAR, 3D QSAR, Anticancer, MLR, kNN MFA.

DOI:

10.25004/IJPSDR.2023.150612

ABSTRACT

In-silico molecular docking studies and QSAR study of benzoxazole derivatives synthesized by Kakkar *et al*. was done. Comparative studies of docking of 5-flurouracil and 20 compounds revealed considerable interactions, indicating the affinity of newly synthesized compounds for thymidylate synthase. The statistically significant 2D-QSAR models were developed using a molecular design suite (VLifeMDS 4.6). The study was performed with 20 compounds (data set) using sphere exclusion (SE) algorithm, random selection and manual selection methods used for the division of the data set into training and test set. Multiple linear regression [MLR] methodology with stepwise (SW) forward-backward variable selection method was used for building the QSAR models. The results of the 2D-QSAR models were further compared with 3D-QSAR models generated by k-Nearest Neighbor Molecular Field Analysis (kNN-MFA), investigating the substitutional requirements for the favorable anticancer activity against HCT 116 cell line and providing useful information in the characterization and differentiation of their binding sites. The results may be useful for further designing benzoxazole derivatives as anticancer agents prior to synthesis.

INTRODUCTION

Cancer is the leading cause of death in developed countries and the second leading cause of death in developing countries.[1] The risk of recurrence is still very significant even though surgical resection may be curative. In addition to surgical resection, adjuvant or neoadjuvant use of chemotherapeutic drugs alone or in conjunction with radiotherapy continues to be the mainstay of treatment regimens for high-risk patients. Unfortunately, only a modest drop in mortality is seen when the aforementioned standard therapy procedures are used, and the probability of developing a disease recurrence is still very significant.[2] The main issues in treating cancer are cytotoxicity and genotoxicity of anticancer medications against normal cells, which increase the risk of subsequent malignancy. One of the standard drugs for treatment of colorectal cancer is 5-fuorouracil (5-FU). However, it is associated with many side effects as it affects the cancer cells and the normal cells. In order to overcome the undesirable side effects of available anticancer agents, novel chemotherapeutic agents are needed for more effective cancer treatment. Therefore, finding and developing medications that can effectively trigger apoptosis while having the least negative effects on cancer cells is of significant interest.[3]

Chemotherapy is still a crucial component of cancer treatment since it kills cancer cells without harming healthy cells, and it has had great success thanks to the development of numerous new medications. As a result, various therapeutic attack kinds have been researched.[4-8] Drugs that perturb microtubule/tubulin dynamics are used widely in cancer chemotherapy. Despite of this progress, the discovery of most potent anticancer agents is a challenging issue in cancer chemotherapy for future generations. Therefore, there is a critical need to research and create innovative anticancer drugs with various

***Corresponding Author:** Dr. Smita Pawar

Address: PDEA's Department of Pharmaceutical Chemistry, Seth Govind Raghunath Sable College of Pharmacy, Pune, Maharashtra, India.

Email ✉**:** sjpawar4477@gmail.com

Tel.: +91-9970705513

Relevant conflicts of interest/financial disclosures: The 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.

modes of action. Creating drug molecules with specified features has long been a prized objective for chemists, especially when developing novel medicines for improved medical care.[9]

Benzoxazole moiety is a medicinally relevant scaffold and is identified as a pharmacophore structure. The benzoxazole nucleus can be found in molecules used in research to assess new products with promising biochemical activity, including anticancer activities,[10-12] and they serve as a topoisomerase-I poison[13] and show antibacterial,[14] antifungal,[15] antimicrobial,[16] and anti-measles virus activities.[17] Benzoxazole ring occurs in a number of natural products such as salvianen and pseudosalvianen.[19,20] Derivatives of benzoxazole have gained much importance because of its wide applications in the medicinal sector. In nature, benzoxazole and their derivatives are widely distributed and are known to have significant physiological effects. Benzoxazole derivatives can bind to a variety of targets with strong affinity.[21,22] It is particularly crucial as a building block for synthesizing numerous physiologically active natural compounds. Kakkar *et al.* designed and synthesized benzoxazole-based derivatives and evaluated their anticancer activity. All the synthesized compounds showed remarkable anticancer activity against HCT 116 human cancer cell line.[23]

The two conventional core pillars of the drug discovery process are empirical research technique and fortuitous chance. Trial-and-error synthesis techniques and random screening for biological activity take a lot of time and are not cost-effective. Additionally, health risks and therapeutic effects are evaluated using various *in-vivo* techniques and tests, with animal models for these purposes involving many ethical problems.[24-25] The long-term goal of employing computer tools to speed up the drug discovery process was made possible by the development of computer technology to calculate the characteristics of molecules. The QSAR method saves us in this situation. Without the inconveniences of trial-and-error synthesis and random screening for activity and toxicity issues, quantitative structure-activity relationship modeling enables us to develop predictive biological activity models as a function of molecular and structural data of the target therapeutic molecules. The QSAR method correlates the biological properties of a compound to molecular descriptors obtained from the molecule's chemical properties. It is a mathematical correlation between the activity profile and the chemical structures of the molecule that has been statistically validated.[26-29]

Drug exerts its biological activity by binding to the pocket of receptor molecule, usually protein. In their binding conformations, the molecules exhibit geometric and chemical complementarily, both of which are essential for successful drug activity. The computational process of searching for a ligand that is able to fit both geometrically and energetically into the binding site of a protein is called molecular docking. Molecular docking helps study drug/ligand or receptor/protein interactions by identifying the suitable active sites in the protein, obtaining the best geometry of ligand-receptor complex and calculating the energy of interaction for different ligands to design more effective ones. The target or receptor is either experimentally known or theoretically generated through knowledge-based protein modeling or homology modeling. The molecular docking tool has been developed to obtain a preferred geometry of interaction of ligand-receptor complexes having minimum interaction energy based on different scoring functions.[30]

Traditional computer-assisted quantitative structure-activity relationship (QSAR) studies, which were invented by C. Hansch *et al.* in 1962 which helps to correlate the bioactivity of compounds with structural descriptors and also have been proven to be one of the useful techniques for accelerating the drug design process. We have used multiple linear regression methodology (MLR) to conduct 2D QSAR on benzoxazole derivatives in order to gain additional insights into the structure-activity relationships of these derivatives and comprehend the mechanism of their substitutional specificity. Cross-validation tests, randomization tests, and external test set prediction were used to determine the significance of the QSAR models. Before synthesizing new anticancer drugs, the robust 2D models may be helpful in further designing new candidates.[31]

Three-dimensional quantitative structure-activity relationship (3D QSAR) facilitates the evaluation of three-dimensional molecular fields around molecules and generates a relationship of these fields' values with the activity. One technique for establishing a connection between activity and molecular field is the k-nearest neighbor (kNN) method, which interprets the findings and offers suggestions for developing new compounds. For Molecular Field Analysis (kNN-MFA), a set of molecules must be properly aligned. A typical rectangular grid is then created, enclosing the molecules. Using the methyl probe of charge+1, the steric and electrostatic energies are calculated at the grid's lattice points. For relationship generation utilizing the kNN approach, these interaction energy levels at the grid points are taken into account. We developed a 3D QSAR model using the kNN approach, and this model generated additional leads for the synthesis of effective anticancer drugs.[32,33]

In the present research work, a data set of 20 molecules showing good inhibitory activity against HCT 116 human cancer cell line was subjected to docking, 2D and 3D QSAR analyses, in search of newer and potent anticancer agents. Significant models were generated statistically, and the most robust models for 2D and 3D QSAR were obtained using MLR and stepwise variable selection kNN-MFA approach, respectively using V-Life Molecular Design Suite software version 4.6.

Materials and Methods

Selection of Data Set

Table 1 lists a set of 20 selected compounds from a series of novel benzoxazole analogs reported by Kakkar *et al.*, 2019.[23] To make it more suited for QSAR studies, the biological activities reported in IC_{50} (µg/mL) for anticancer activity were transformed to algorithmic IC_{50} values. These derivatives were substituted with different groups (like hydrogen, fluorine, chlorine, methyl, and alkyl chain, methoxy many more) on different positions and every group has its own contribution in the physicochemical characteristics and biological activities of the designed compounds.[34]

Sketching of Molecules

The construction of the molecular structures is the next step in constructing a model. V- Life MDS software was used to sketch the structures, which were also energy reduced. Through using same software, the energy-minimized structures were used to generate molecular descriptors such as electronic, geometric, hydrophobic, and topological features.[35] Table 1 shows the substitution, IC_{50} and -log (IC_{50})values of the compound chosen for the MLR model.

Docking Study

Molecular docking techniques consist in finding the low-energy binding modes of a ligand within the active site of a macromolecule and evaluating the binding energy with a score.[36] Explanation of the selectivity of small sets of ligands has been attempted with accurate but time-consuming techniques.[37-39] Otherwise, automated docking methods may be used to estimate the thymidylate synthase affinity for large molecular databases. The objective of this work had been to use structural information of the target to confirm the binding affinity of synthesized molecules to thymidylate synthase as well as study binding nature. Docking studies were carried out in Vlife molecular docking suite 4.6 by using Biopredicta.

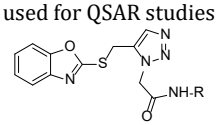
Optimization of Protein

Docking studies were carried out using thymidylate synthase. The 3H9K isoform of human thymidylate synthase was downloaded from Protein Data Bank website. The tetramer is converted into the monomer. Water molecules, cofactors, and heme were deleted from protein. The reference molecule was extracted. Hydrogens were added in molecule and energy was minimized using Merck molecular force field (MMFF).

Optimization of Ligand

3D structure is optimized, conformers were generated (Monte Carlo method), and the least energy conformers were selected. For the preparation of ligand, the structures of all the compounds used against HCT 116 human cancer cell line in *in-vitro* assay were energy minimized using MMFF.

Table 1: The chemical structures, IC_{50} and -log (IC_{50}) of compounds used for QSAR studies



S. No	*Substitutions (R)*	*IC_{50}*	*-log IC_{50}*
1	3-nitrophenyl	97.5	-1.989
2	4-nitrophenyl	73.1	-1.864
3	4-methoxyphenyl	108.7	-2.036
4	2-chloro-4-nitrophenyl	22.5	-1.352
5	2,4,5-trichlorophenyl	85.3	-1.931
6	2-methyl-3-chlorophenyl	84.6	-1.927
7	2-methyl-3-chlorophenyl	72.0	-1.857
8	4-ethylphenyl	254.2	-2.405
9	2,4-dimethylphenyl	66.1	-1.820
10	2-chlorophenyl	175.1	-2.243
11	4-fluorophenyl	130.4	-2.115
12	4-bromophenyl	225.1	-2.352
13	3,4-dichlorophenyl	230.3	-2.362
14	2-bromophenyl	90.0	-1.954
15	3,5-dimethylphenyl	40.7	-1.610
16	3-bromophenyl	38.3	-1.583
17	2,3-dimethylphenyl	177.9	-2.251
18	2-bromophenyl	148.7	-2.172
19	3-chlorophenyl	50.0	-1.699
20	4-chlorophenyl	200.1	-2.301

Docking

Batch docking of optimized molecule was performed on the optimized receptor. Cavity number 1 for the 3H9K receptor was selected where the docking would be performed. The flexible GA batch docking did docking studies.

2D QSAR Analysis

Calculation of descriptors

Numbers of descriptors were calculated after optimization or minimization of the energy of the data set molecules. Various types of physicochemical descriptors were calculated: Individual (Molecular weight, H-Acceptor count, H-Donor count, X log P, slog P, SMR, Polarizability, etc.), Element count (N, C, S count etc.), Distance based topological (DistTopo, Connectivity Index, Wiener Index, Balaban Index), Estate numbers ($SsCH_3$count, $SdCH_2$count, $SssCH_2$count, StCH count, etc.), Estate contribution ($SsCH_3$-index., $SdCH_2$- index, $SssCH_2$-index , StCH index), and Polar surface area. The invariable descriptors (the descriptors that are constant for all the molecules) were removed, as they do not contribute to QSAR.

Generation of training and test set

In order to evaluate the QSAR model, data set was divided into training and test set using random selection and manual selection method. The training set is used to develop the QSAR model for which biological activity data are known. Test set is used to challenge the QSAR model developed based on the training set to assess the model's predictive power, which is not included in model generation.

Random Selection Method: To construct and validate the QSAR models, internal and external, the data sets were divided into training and test sets randomly. The ratio of 70:30 (training: test) was applied to data set.

Manual data selection method: Data set is divided manually into training and test sets on the basis of the result obtained in random selection method.

Generation of 2D QSAR model by multiple linear regression

Two-dimensional quantitative structure activity relationship (2D QSAR) studies using the multiple linear regression (MLR) method were performed on a series of benzoxazole derivatives as anticancer agents using software VLife MDS. MLR is a method used for modeling linear relationship between a dependent variable Y (Activity) and independent variable X (2D/3D descriptors). MLR is based on least squares. MLR estimates values of regression coefficients (r^2) by applying least squares curve fitting method. The model creates a relationship in the form of a straight line (linear) that best approximates all the individual data points. In regression analysis, the conditional mean of dependant variable (Activity) Y depends on (descriptors) X.

3D QSAR Analysis

We attempted to obtain further insights into the structural requirement of benzoxazole derivatives as anticancer agents by applying 3D-QSAR using kNN-MFA method. Molecular alignment was performed which is useful for studying shape variation with respect to the base structure i.e., benzoxazole moiety selected for alignment. Further, the aligned molecules were used for 3D QSAR studies. The method used for alignment was template based. A template structure was defined and used as a basis for alignment of a set of molecules. The geometries of the aligned molecules were stored in the align molecules subfolder in the align folder. All the aligned molecules were opened from Align Molecules subfolder in the Align folder. The 20 aligned molecules are shown in Fig. 1.

Calculation of descriptors

The 3D QSAR worksheet was launched at the beginning. The subfolder containing all aligned molecules was opened. The data of biological activity (IC_{50}) was inserted in the column. This was followed by the calculation of the molecular local shape field descriptors for finding a relationship with the activity and various parameters for field calculation selected are shown below:

Field Type: Electrostatic, steric, and hydrophobicity

Charge Type: Gasteiger-Marsili

For performing a robust QSAR analysis, descriptors that show variation for all the molecules are important. A descriptor that is constant for all the molecules will not contribute to QSAR and, hence should be removed from the worksheet. Thus, invariable columns were removed.

Data selection of training and test set

The training and test set was selected by manual selection method, ensuring that the molecules have uniform spread (training and test) in terms of both activity and chemical space. The dependent variable selected was biological activity as a negative log of IC_{50}; the remaining variables were considered independent variables.

Generation of 3D QSAR model by kNN MFA model

This method employs a stepwise variable selection procedure combined with kNN to optimize the number of nearest neighbors (k) and the selection of variables from the original pool as described in simulated annealing. The kNN-MFA model was generated using stepwise forward-backward variable selection method. The relative positions of the local fields around aligned molecules that are important for activity variation in the model were observed by clicking the Show Points button in the 'Model Summary' dialog box. The best model was selected on the basis of various statistical parameters. The quality and predictability of the model was estimated from the cross-validated squared correlation coefficient and predicted r^2.

Result and Discussion

Docking Studies

In this comparative docking experiment of benzoxazole compounds with known anticancer drug 5-fluorouracil the dock score was calculated -4.742203 kcal/mole. Van der Waals interactions, H-bond interactions, and hydrophobic interactions are some of the elements considered while predicting the ligand's potency or affinity to the receptor. The higher the affinity of molecule towards the receptor, the more negative the value of the energy of binding. The increased Van der Waals interaction indicates that the ligand structure contains more bulky groups due to

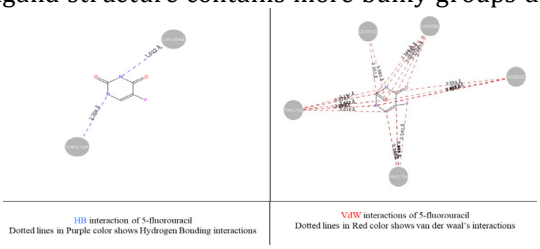


Fig. 1: 2D-view of the reference molecule show the hydrogen bonding and van der waal's interaction with thymidylate synthase.

which Van der Waals interactions were formed. If the charge interactions are present, it assists in determining more appropriate binding and thus exhibits more affinity to the receptor, adding more efficacies. Obtained results were evaluated in terms of docking score into the active site of (3H9K). The software provides facility of the batch docking of the optimized ligand molecules with the simulated receptor. All ligands are docked an active site i.e., cavity for docking of receptor 3H9K. We compared the dock score of each selected molecule and reference molecule, which shows the variation in the docking score, so we predicted that molecules with minimum dock score show more affinity for thymidylate synthase inhibition and Dock score shown in Table 2. Table 2 shows that ligands 4 and 16 have the betterdock scores, or the lowest binding energy in Kcal/mol among all the compounds. These molecules have a higher affinity against the active site of the receptor. It is understandable that molecules with a lower dock score and binding energy confirm more affinity towards the receptor.

Interactions of Ligands with Receptor

Molecule 5-Fluorouracil

This is a reference molecule. It is anticancer agent. The low dock score of this molecule is -4.742203 kcal/mol with thymidylate synthase (3H9K). The ligand showed hydrogen bond interaction with LYS169A and THR170A as well as van der waal's interaction with LYS169A, PRO172A, THR170A, HIS261D and LEU221D. These interactions are important for increased affinity for thymidylate synthase.

Molecule (4)

The low dock score of this ligand is -4.154462 kcal/mol against the binding site of the thymidylate synthase receptor. In Fig. 2 the residual interactions of compound 4 revealed one (1) conventional H-bonds, four (4) hydrophobic, eleven (11) electrostatic (Van Der Waal's) interaction types with different amino acid residues in the active site of the targeted receptor. The NH group attached to the carbonyl group in molecule 04 can interact with the residues of LYS169A with a distance of 2.392 Å. It was also noteworthy that the hydrophobic interaction also existed between the hydrophobic residues of compound 04 and the active site of the ILE168A, LYS169A, THR241A, LEU243A at a distance of 4.884, 4.734, 4.894 and 4.786 Å, respectively. The eleven Van der Waal'sinteraction were formed with THR51D (3.701), ARG50D (3.822), GLY242A (3.364), THR241A (3.716), PRO172A (3.175 Å), ASN171A (3.566 Å), THR170A (3.378 Å), LEU234A (2.965 Å), LYS169A (3.195 Å), ILE240A (3.26 Å), and ILE168A (3.029 Å) at different interaction distances accordingly.

Molecule (16)

The low dock score of this ligand is -4.219219 kcal/mol against binding site of thymidylate synthase receptor. In

Table 2: Docking score of the benzoxazole derivatives with the structure of thymidylate synthase (3H9K)

Compound	*Dock score*	*Binding energy*
1	-3.822572	-640.6586
2	-4.096091	-649.4137
3	-3.322664	-639.9574
4	-4.154462	-651.7083
5	-4.546775	-636.2337
6	-3.849419	-639.9083
7	-3.76660	-636.0455
8	-3.576736	-632.3346
9	-3.939365	-636.845
10	-4.138222	-636.8219
11	-3.772533	-629.5271
12	-3.998726	-631.0459
13	-4.021897	-632.1873
14	-4.144334	-636.521
15	-4.077910	-634.5393
16	-4.219219	-649.9165
17	-4.338950	-640.9052
18	-4.123791	-634.0953
19	-4.035604	-629.9877
20	-4.081383	-630.1479
5-Fluorouracil	-4.742203	-1117.7934

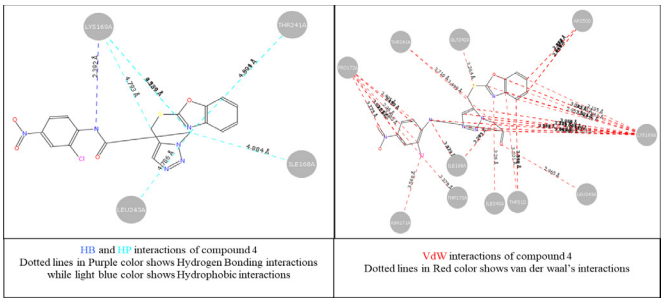


Fig. 2: 2D-view of molecule 4 show the hydrogen bonding, hydrophobic and van der waal's interaction with thymidylate synthase.

Fig. 3 the residual interactions of compound 16 revealed one (1) conventional H-bonds, four (4) hydrophobic, and 12 electrostatic (Van der Waal's) interaction types with different amino acid residues in the active site of the targeted receptor. The NH group attached to the carbonyl group in molecule 16 can interact with the residues of ILE168A with a distance of 2.442 Å. Also, the hydrophobic interaction also existed between the hydrophobic residues of compound 16 and the active site of the LYS169A, THR170A, PRO172A at a distance of 4.8553, 4.582, and 4.88 Å, respectively. The twelve Van der Waal's interaction were formed with THR170A, LEU243A, VAL203A, HIS256D, ASP218D, ARG175E, ILE168A, PRO172A, SER206A, THR241A, LYS169A, and TYR258D at interaction distances of 2.844, 3.181, 2.992, 3.835, 3.496, 2.703, 3.983, 3.385, 3.977, 2.894, 2.835, and 3.782 Å, respectively.

2D QSAR Model

Different sets of 2D-QSAR models were generated using the MLR analysis in conjunction with stepwise forward-backward variable selection method. Different training

and test set were constructed using random and manual selection method. The best QSAR model was selected based on values of r^2 (squared correlation coefficient), q^2 (cross-validated correlation coefficient), pred_r^2 (predicted correlation coefficient for the external test set), F (Fisher ratio) reflects the ratio of the variance explained by the model and the variance due to the error in the regression. The F–test values indicate that the model is statistically significant. The statistical values of the corresponding best model are reported in Table 3.

Test set: 2, 11, 13, 14, 16, 17, 19

The best equation obtained by multiple linear regression is as follows:

Biological activity (IC_{50}) = 0.1071(± 0.0001) T_T_T_4 + 0.1523(0.009) T_C_N_7 -0.2974(0.0166) SssCH_2E-index -0.7617(0.0117) T_N_N_7 -5.4863(1.1947) SAAverage.

Interpretation of the 2D-QSAR model

From equation, 2D-QSAR model explains 99.32% (r^2 = 0.9932) of the total variance in the training set, which means that the model's failure probability is 1 in 10000. This model was considered as the model showed an internal predictive power (q^2: 0.8166) of 81% and a predictivity for the external test set (pred_r^2:0.5828) of about 58%.The descriptors T_T_T_4 and T_C_N_7 indicates the Count of number of any atom separated from any other heteroatom by four bonds distance in a molecule and count of number of carbon atoms (single, double or triple bonded) separated from any nitrogen atom(single, double or triple bonded) by 7 bonds distance in a molecule, respectively. The positive contribution of 40.07 and 18.26% of these descriptors revealed the increase of anticancer activity of benzoxazole with the presence of NO_2 group such as compound 1, 2, and 4. Higher values of these descriptors lead to good anticancer activity while lower values lead to reduced anticancer activity. Moreover, negative coefficient value of descriptors SssCH_2E-index [number of CH_2 group connected with two single bonds], T_N_N_7 [count of number of nitrogen atoms separated from any other nitrogen atom by 7 bonds in a molecule],

Table 3: Statistical results for MLR method

Parameter	*MLR method*
n	13
Degree of freedom	7
r^2	0.9932
q^2	0.8166
F test	203.2395
r^2_se	0.0329
q^2_se	0.1701
pred_r^2	0.5828
pred_r^2se	0.4512
Descriptors (range)	T_T_T_4 (0.1071) T_C_N_7 (+ 0.1523) SssCH_2E-index (-0.2974) T_N_N_7 (-0.7617) SAAverage (-5.4863)

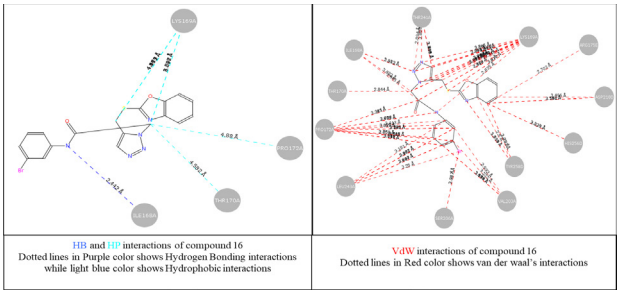


Fig. 3: 2D-view of the molecule 16 show the hydrogen bonding, hydrophobic and van der waal's interaction with thymidylate synthase

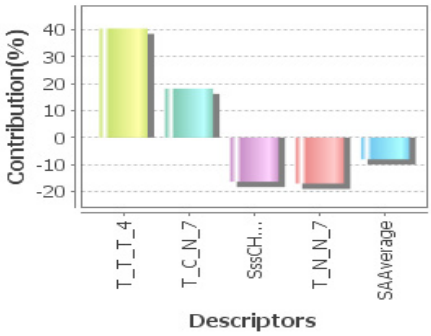


Fig. 4: Contribution chart for 2D-QSAR model showing contribution of different descriptors

SAAverage [average hydrophobicity function value] on the anticancer activity indicated that lower value leads to better anticancer activity whereas higher value leads to decreased anticancer activity. The contribution chart for model is represented in Fig. 4 reveals that the two descriptors T_T_T_4, T_C_N_7 contributed 40.07, 18.26%, respectively. Three more descriptors SssCH2E-index, T_N_N_7, SAAverage, are contributing inversely 16.65, 18.49, and 7.56%, respectively to biological activity.

Predictivity of MLR model

Predictivity of the model was evaluated by predicting the activity of the molecules belonging to the training set (internal predictivity) as well as molecules in the test set (external predictivity) as shown in Table 4. (Red color = test set). Data fitness plot and Graphical representation of the model's predictive power are shown in Figs 5 and 6, respectively.

3D-QSAR Model

3D QSAR models were developed for set of 20 compounds using stepwise kNN MFA method for anticancer activity. The steric and electrostatic descriptors specify the regions, where variation in the structural features of different compounds in training set leads to increase or decrease in activities.

In the present study the benzoxazole nucleus considered as a template and each molecule was superimposed on the template. 3D QSAR model was derived using the kNN descriptors as independent variables and IC_{50}

Table 4: Actual and predicted activities for 20 compounds based on the best 2D-QSAR models

compound	*Observed value*	*Predicted value*	*Residual value*
1	-1.989	-2.104	-0.115
2	-1.864	-2.102	-0.238
3	-2.036	-2.208	-0.172
4	-1.352	-1.654	-0.302
5	-1.931	-1.839	0.092
6	-1.927	-1.758	0.169
7	-1.857	-1.760	0.097
8	-2.405	-2.327	0.078
9	-1.820	-1.778	0.042
10	-2.243	-2.353	-0.11
11	-2.115	-2.272	-0.157
12	-2.352	-2.353	-0.001
13	-2.362	-1.967	0.395
14	-1.954	-1.839	0.115
15	-1.610	-1.694	-0.084
16	-1.583	-1.978	-0.395
17	-2.251	-1.568	0.683
18	-2.172	-2.14	0.032
19	-1.699	-1.839	-1.14
20	-2.301	-2.379	-0.078

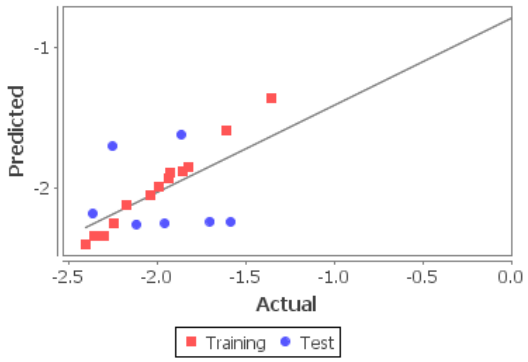


Fig. 5: Data fitness plot for 2D-QSAR model

concentration as a dependent variable. 3D QSAR models were built in V-Life molecular design software using SWkNN method. Various models have been developed by changing training and test set data. Selection of best model was performed based on statistical parameters of the models. The corresponding best model is reported in Table 5.

Test set: 2, 11, 13, 14, 16, 17, 19

Interpretation of 3D QSAR model

The kNN MFA model generated from template-based alignment showed Q2 of 0.8264 with two descriptors, namely S_1120 and E_808. Steric and electrostatic field energy of interaction between template and compounds at their corresponding spatial grid points of 1120 and 808. Which showed relative position and ranges of the corresponding important steric and electrostatic files in the model. As far as S_1120 steric field is concerned, a negative range indicates that negative steric potential was favorable for increased activity; hence, less bulky group

Table 5: Statistical results of kNN molecular field analysis

Parameter	*kNN MFA method*
n	13
Degree of freedom	10
q^2	0.8264
q^2_se	0.1310
pred_r^2	-1.2354
pred_r^2se	0.3913
Descriptors (range)	S_1120 (-0.1268 -0.1166) E_808 (0.0734 0.1147)

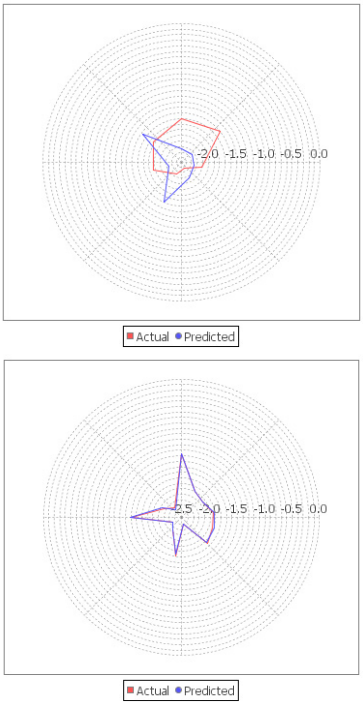


Fig. 6: Graph between actual and predicted biological activity of test and training set for 2D-QSAR model

was preferred in that region. Thus, a green sphere (S_1120) near the benzoxazole ring's substituent suggests a less bulky group is required in this region. One electrostatic descriptor indicated by blue sphere (E_808) is observed in the model and contributes positively, suggesting that more electropositive (electron withdrawing) groups are favorable in this position. Fig. 7 shows the 3D alignment of molecules with the important steric and electrostatic points contributing to the 3D QSAR model.

Predictivity of kNN MFA Model

Predictivity of the model was evaluated by predicting the activity of the molecules belonging to the training set (internal predictivity) as well as molecules in the test set (external predictivity) as shown in Table 6. (Red color = test set). Data fitness plot and Graphical representation of the model's predictive power are shown in Figs. 8 and 9, respectively.

Table 6: Actual and predicted activities for 20 compounds based on the best 3D-QSAR models

Compound	*Observed value*	*Predicted value*	*Residual value*
1	-1.989	-1.839	0.15
2	-1.864	-1.721	0.143
3	-2.036	-2.144	-0.108
4	-1.352	-1.476	-0.124
5	-1.931	-1.919	0.012
6	-1.927	-1.862	0.065
7	-1.857	-1.862	-0.005
8	-2.405	-2.397	0.008
9	-1.820	-1.857	-0.037
10	-2.243	-2.246	-0.003
11	-2.115	-2.288	-0.173
12	-2.352	-2.328	0.024
13	-2.362	-2.164	0.198
14	-1.954	-2.247	-0.293
15	-1.610	-1.559	0.051
16	-1.583	-2.245	-0.662
17	-2.251	-1.643	0.608
18	-2.172	-2.18	-0.008
19	-1.699	-2.244	-0.548
20	-2.301	-2.327	-0.026

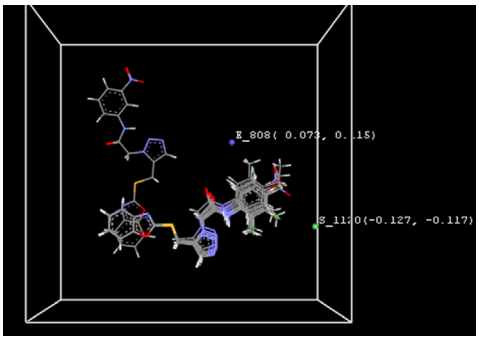


Fig. 7: 3D-alignment of molecules (Ball and stick model) with the important steric and electrostatic points contributing 3D-QSAR model

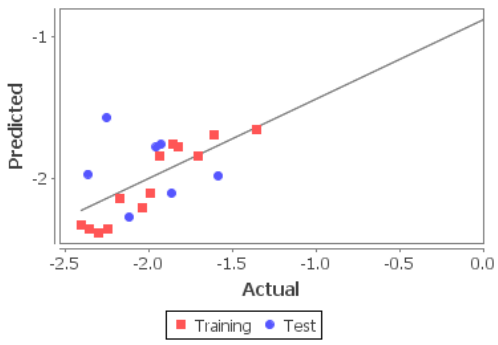


Fig. 8: Data fitness plot for 3D-QSAR model

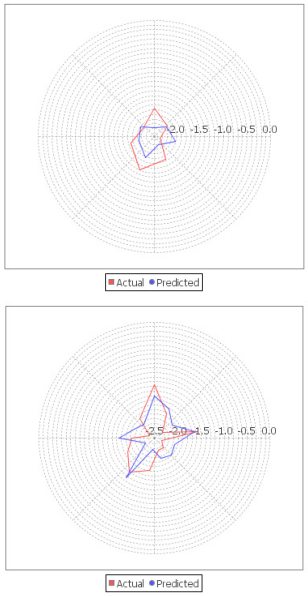


Fig. 9: Graph between actual and predicted biological activity of test and training set for 3D-QSAR model

Conclusion

According to the docking simulation, the hydrophobic Van der Waals, H-bond interactions create stable compounds of ligands with receptors. According to Table 2, the ligands 4 and 16 were shown to have the lowest dock scores of -4.154462 and -4.219219, respectively and the lowest binding energy of -651.7083 and -649.9165 kcal/mol, respectively, indicating that they had a higher affinity for the receptor's active site. It is evident that molecules with a low dock score and binding energy have a greater affinity for the receptor. Statistically significant 2D/3D-QSAR models were generated with the purpose of deriving structural requirements for the anticancer activities of some novel benzoxazole derivatives against HCT 116. The best 2D QSAR models indicate that the descriptors T_T_T_4, T_C_N_7positively participate for the anticancer activity whereas, $SssCH_2E$-index, T_N_N_7 and SAAveragenegatively participate for the anticancer activity. kNN-MFA investigated the substitutional requirements for the receptor-drug interaction and constructed the best 3D-QSAR models by kNN MFA method, providing that negative steric potential i.e., less bulky groups was favorable for increased activity whereas positive electrostatic potential is needed for potential anticancer activity. In conclusion, the information provided by the robust 2D/3D-QSAR models use for the design of new molecules and hence, it helped to identify the lead that could be optimized further for better anticancer potential.

Acknowledgement

The authors acknowledge the support from Pune District Education Association's Seth Govind Raghunath Sable College of Pharmacy, Saswad (Maharashtra) 412301.

References

1. The Global Burden of Disease: 2004 Update; World Health Organization: Geneva, Switzerland, 2008.
2. Wang Z, Shi XH, Wang J, Zhou T, Xu YZ, Huang TT, Li YF, Zhao YL, Yang L, Yang S. Synthesis, Structure-Activity Relationships and Preliminary Antitumor Evaluation of Benzothiazole-2-thiol Derivatives as Potential Anticancer Agents. Bioorg. Med. Chem. Lett. 2011; 21:1907–1101.
3. Garcia M, Jemal A, Ward E. M, Center M. M, Hao Y, Siegel R. L, Thun. M. J. Global cancer facts & Fig.s. American Cancer Society. 2007.
4. Hanahan D, Weinberg RA. The Hallmarks of cancer. Cell. 2000; 100: 57.
5. Stratton MR, Campbell PJ, Futreal PA. The Cancer Genome. Nature. 2009; 458: 719.
6. Hanahan D, Weinberg RA. Hallmarks of Cancer: The Next Generation. Cell. 2011; 144: 646.
7. Boyle FT, Costello GF, Cancer therapy: A Move to the Molecular Level. Chem Soc Rev. 1998; 27: 251.
8. Gibbs JB, Mechanism-based Target Identification and Drug Discovery in Cancer Research. Science. 2000; 287: 1969.
9. Lu Y, Chen YJ, Xiao M, Li W, Miller DD. An Overview of Tubulin Inhibitors that Interact with the Colchicine Binding Site. Pharm Res. 2012; 29: 2943.
10. Don MJ, Shen CC, Lin YL, Syu WJ, Ding YH, Sun CM. Nitrogen Containing Compounds from Salvia Miltiorrhiza. Nat Prod. 2005; 68: 1066.
11. Viirre RD, Evindar G, Batey RA. Copper-catalyzed Domino Annulation Approaches to the Synthesis of Benzoxazoles under Microwave-accelerated and Conventional Thermal Conditions. J Org Chem. 2008; 73: 345.
12. Osmaniye D, Çelikateş BK, Saglık BN,Levent S,Cevik UA,Cavusoglu BK, Ilgın S, Ozkay Y,Kaplancıkli ZA. Synthesis of some New Benzoxazole Derivatives and Investigation of their Anticancer Activities. Eur J Med Chem. 2021; 210: 112979.
13. Cheng CC, Liu DE, Chou TC. Design of Antineoplastic Agents on the Basis of the "2-Phenylnaphthalene-Type" Structural Pattern. I. Synthesis of Substituted 3- Phenylquinazolones, Benzoxazolo[2,3-b] quinazolones and Benzothiazolo[2,3- b]quinazolones. Heterocycles. 1993; 35: 775.
14. Elnima EI, Zubair MU, Al-Badr AA. Antibacterial and antifungal activities of benzimidazole and benzoxazole derivatives. Antimicrobial Agents and Chemotherapy, 1981; 19: 29-32.
15. Kaplancikli Z. A, Turan-Zitouni G, Revial G, Guven K. Synthesis and Study of Antibacterial and Antifungal Activities of Novel 2-[[(benzoxazole/benzimidazole-2- yl)sulfanyl] acetylamino] thiazoles. Archives of Pharmacal. Research. 2004; 27: 1081–1085.
16. Sun LQ, Chen J, Bruce M, Deskus JA, Epperson JR, Takaki K, Johnson G, Iben L, Mahle C. D, Ryan E, Xu C. Synthesis and Structure-activity Relationship of Novel Benzoxazole Derivatives as Melatonin Receptor Agonists. Bioorg. Med. Chem. Lett. 2004; 14: 3799.
17. Ertan T, Yildiz I, Tekiner-Gulbas B, Bolelli K, Temiz-Arpaci O, Yalcin I, Aki E, Ozkan S, Kaynak F. Synthesis, biological evaluation and 2D-QSAR analysis of benzoxazoles as antimicrobial agents. Eur. J Med. Chem. 2000; 44: 501.
18. Alper-Hayta S, Arisoy M, Temiz-Arpaci O, Yildiz I, Aki E, Oezkan S, Kaynak F. Synthesis and Anticancer Activity of New ((Furan-2-yl)-1,3,4-thiadiazolyl)-1,3,4- oxadiazole Acyclic Sugar Derivatives. Eur. J Med. Chem. 2008; 43: 2568.
19. Sun A, Prussia A, Zhan W, Murray EE, Doyle J, Cheng LT, Yoon JJ, Radchenko EV, Palyulin VA, Compans RW, Liotta DC, Plemper RK, Snyder JA. Nonnucleoside Inhibitor of Measles Virus RNA-Dependent RNA Polymerase Complex Activity. J Med. Chem. 2006; 49: 5080.
20. Borrel C, Thoret S, Cachet X, Guenard D, Tillequin F, Koch M, Michel S. New antitubulin derivatives in the combretastatin A4 series: synthesis and biological evaluation. Bioorg. Med. Chem. 2005; 13: 3853.
21. Katritzky A. QSAR modeling, synthesis, and bioassay of diverse leukemia RPMI-8226 cell line active agents. Bioorg. Med. Chem. Lett. 2010; 45: 5183-99.
22. Rapatri V, Chitre T, Bothara K. Novel 4-(morpholin-4-yl)-N-(arylidene) Benzohydrazides: Synthesis, Antimycobacterial activity, and QSAR Investigations. Eur. J Med. Chem. 2009; 44: 3954-60.
23. Kakkar S, Kumar S, Narasimhan B, Lim S. M, Ramasamy K, Mani V, Ali Shah S. A. Design, Synthesis, and Biological Potential of Heterocyclic Benzoxazole Scafolds as Promising Antimicrobial and Anticancer Agents. Chemistry Central Journal. 2018; 12: 96.
24. Sastry MG, Adzhigirey M, Day T, Annabhimoju R, Sherman W. Protein and Ligand Preparation: Parameters, Protocols, and Influence on Virtual Screening Enrichments. J. Comput. Aided Mol. Des. 2013; 27(3): 221-234.
25. Kumar S, Singh J, Narasimhan B, Shah S.A.A, Lim S.M, Ramasamy K, Mani V. Reverse Pharmacophore Mapping and Molecular Docking Studies for Discovery of GTPase HRas as Promising Drug Target for Bis-pyrimidine Derivatives. Chem. Cent. J. 2018; 12: 106.
26. Katritzky A. QSAR modeling, synthesis, and bioassay of diverse leukemia RPMI-8226 cell line active agents. Bioorg. Med. Chem. Lett. 2010; 45: 5183-99.
27. Rapatri V, Chitre T,Bothara K. Novel 4-(morpholin-4-yl)-N-(arylidene) benzohydrazides: Synthesis, Antimycobacterial Activity, and QSAR Investigations. Eur. J Med. Chem. 2009; 44: 3954-60.
28. Karthiga S, Velmurugan D. Molecular modeling, QSAR and pharmacophore studies on anti-viral, anti-malarial and anti-inflammatory bioactive compounds from marine sources. Asian J Pharma Clin Res. 2015; 8: 36-43.
29. VLife Molecular Design Suite version 3.5; VLife Sciences Technologies Pvt. Ltd., Pune, India; 2010.
30. Sharma V, Sharma P.C, Kumar V. In silico Molecular Docking Analysis of Natural Pyridoacridines as Anticancer Agents. Adv. Chem. 2016; 1-9.
31. Hansch C, Fujita T. p-σ-π Analysis. A method for the correlation of biological activity and chemical structure. J. Am. Chem. Soc. 1964; 86(8): 1616-1626.
32. Life Molecular Design Suite 3.0, VLife Sciences Technologies Pvt. Ltd; Baner Road: Pune, Maharashtra, India. www.Vlifescien ces. com. Accessed Jan 2019
33. Nandi S, Bagchi M. 3D-QSAR and Molecular Docking Studies of 4-anilinoquinazoline Derivatives: A Rational Approach to Anticancer Drug Design. Mol Divers. 2010, 14, 27.
34. Bora P, Kumar A, Kumar Singh A, Singh H, Narasimhan B, Kumar P. Molecular Docking and QSAR Studies of Indole Derivatives as Antifungal Agents. Cur. Chinese Chemistry. 2023; 3:1-12.
35. Ou-Yang S, Lu J, KongX, Liang Z, Luo C, Jiang H. Computational drug discovery. Acta Pharmacol. Sin.2012; 33(9): 1131-1140.
36. Habeeb AG, Parveen P, Knaus EE. Design and synthesis of 4, 5-diphenyl-4-isoxazolines: Novel Inhibitors of Cylooxygenase-2 with Analgesic and Anti-Inflammatory Activity. J Med. Chem. 2001; 44: 2921-7.
37. Palomer A, Perez J. J, Navea S, Llorens O, Pascual J, Garcia L. ModelingCyclooxygenase Inhibition: Implication of active site Hydration on the Selectivity of Ketoprofen Analogues. J Med. Chem. 2000; 43: 2280-4.
38. Price ML, Jorgensen WL. Analysis of Binding Affinities for Celecoxib Analogues with COX-1 and COX-2 from Combined Docking and monte carloSimulations and Insight into the COX-2/COX-1 Selectivity. J Am. Chem. Soc. 2000; 94: 55-66.
39. Price ML, Jorgensen WL. The Rationale for the Observed COX-2/COX-1 Selectivity of Celecoxib from Monte Carlo Simulation. Bioorg. Med. Chem. Lett. 2001; 11:1541-4.

HOW TO CITE THIS ARTICLE: Pawar SJ, Zope D, Kale AP. *In-silico* Studies of Heterocyclic Benzoxazole Derivatives as an Anticancer Agent: Molecular Docking, 2D and 3D QSAR. Int. J. Pharm. Sci. Drug Res. 2023;15(6):780-788. **DOI:** 10.25004/IJPSDR.2023.150612