HOSTED BY

ELSEVIER

Contents lists available at ScienceDirect

Saudi Pharmaceutical Journal

journal homepage: www.sciencedirect.com

Original article

3D QSAR study on substituted 1, 2, 4 triazole derivatives as anticancer agents by kNN MFA approach

Shailaja P. Desai [a,*], S.K. Mohite [b], Saad Alobid [c], M.G. Saralaya [a], Ashwini S Patil [a], Kuntal Das [d], Moneer E. Almadani [e], Syed Arif Hussain [f], Bader Hussain Alamer [g], Ebtesam Abdulrahman Jibreel [h], Ali Ibrahim Almoteer [i], Syed Mohammed Basheeruddin Asdaq [j,*]

[a] Annasaheb Dange College of Pharmacy, Ashta, Maharashtra, Walwa, Sangli 416301, India
[b] Department of Pharmaceutical Chemistry, Rajarambapu College of Pharmacy, Kasegaon, Sangli, Maharashtra 415409, India
[c] Department of Pharmacology and Toxicology, College of Pharmacy, King Saud University, Riyadh 11451, Saudi Arabia
[d] Department of Pharmacognosy, Mallige College of Pharmacy, #71 Silvepura Chikkabanavara Post, Bangalore 90, India
[e] Department of Clinical Medicine, College of Medicine, AlMaarefa University, Dariyah, Riyadh 13713, Saudi Arabia
[f] Respiratory Care Department, College of Applied Sciences, AlMaarefa University, Dariyah 13713, Riyadh, Saudi Arabia
[g] Department of Emergency Medical Services, College of Applied Sciences, AlMaarefa University, Riyadh, Saudi Arabia
[h] Department of Nursing, College of Applied Sciences, AlMaarefa University, Dariyah 13713, Riyadh, Saudi Arabia
[i] Department of Pharmacy, King Saud University Medical City, Riyadh, Saudi Arabia
[j] Department of Pharmacy Practice, College of Pharmacy, AlMaarefa University, Dariyah 13713, Riyadh, Saudi Arabia

ARTICLE INFO

Keywords:
3D-QSAR
Genetic algorithm
Anticancer agents
1,2,4-Triazole

ABSTRACT

Background and objectives: Researchers have recently focused on the biological and synthetic effects of 1, 2, and 4-triazole fused heterocyclic molecules because they have tremendous medicinal value. The objective of the present study was to carry out the 3D QSAR evaluation on the substituted 1,2, and 4 triazole derivatives for anticancer potential using k-Nearest Neighbor-Molecular Field Analysis (kNN-MFA) method.

Methods: Using the molecular design suite, a three-dimensional quantitative structure–activity relationship (3D-QSAR) analysis was undertaken on a series of 4-amino-5-(pyridin3yl)-4H-1, 2, and 4-triazole-3-thiol anticancer drugs (Vlife MDS). This study used a genetic algorithm and a manual selection approach on 20 substituted 1, 2, and 4-triazole derivatives. Based on the genetic algorithm (GA), the 3D-QSAR model was generated. Statistical significance and predictive capacity were evaluated using internal and external validation.

Results: The most significant model has a correlation coefficient of 0.9334 (squared correlation coefficient r2 = 0.8713), showing that biological activity and descriptors have a strong relationship. The model exhibited internal predictivity of 74.45 percent (q2 = 0.2129), external predictivity of 81.09 percent (pred r2 = 0.8417), and the smallest error term for the predictive correlation coefficient (pred r2se = 0.1255). The model revealed steric (S 1047—0.0780—0.0451S 927) and electrostatic (E 1002) data points that contribute remarkably to anticancer activity. A molecular field study demonstrates a link between the structural features of substituted triazole derivatives and their activities.

Conclusion: The good-to-moderate anticancer potential of compounds confirms the significant pharmacological role of 1,2,4-triazole derivatives. These results could lead to the identification of potential chemical compounds with optimal anticancer activity and minimal side effects.

Peer review under responsibility of King Saud University.

* Corresponding authors at: Department of Pharmacy Practice, College of Pharmacy, AlMaarefa University, Dariyah 13713, Riyadh, Saudi Arabia.
E-mail addresses: shailaja.desaiadcbp@gmail.com (S.P. Desai), mohitk@gmail.com (S.K. Mohite), saaalobaid@ksu.edu.sa (S. Alobid), drkkdsd@gmail.com (K. Das), mmadani@um.edu.sa (M.E. Almadani), pulmoarif@gmail.com (S. Arif Hussain), bamer@um.edu.sa (B. Hussain Alamer), ijibreel@um.edu.sa (E. Abdulrahman Jibreel), Aalmotair@ksu.edu.sa (A. Ibrahim Almoteer), sasdaq@gmail.com, sasdag@um.edu.sa (S. Mohammed Basheeruddin Asdaq).

https://doi.org/10.1016/j.jsps.2023.101836
Received 17 September 2023; Accepted 18 October 2023
Available online 24 October 2023

1. Introduction

In recent decades, researchers have focused on the synthetic and biological implications of fused heterocyclic derivatives of 1, 2, and 4-triazoles. Many therapeutically significant pharmacological compounds have been incorporated in the 1, 2–4-Triazole Ring System, which provides a number of antimicrobial and CNS stimulants (Singh and Singh, 2009). Among these heterocycles are the mercapto and the ion-substituted 1, 2, 4 Triazole. The mercapto- and thione-substituted 1, 2, and 4-triazole ring systems have been extensively considered. The range of biological activities with various derivatives have been reported to have antibacterial, antifungal, anticancer, diuretic, and hypoglycemic properties (Dilmaghani et al., 2015; Wakale et al., 2013). Triazole inhibits the 14 alpha-Demethylase that prevents ergosterol production. The nitrogen in Triazole is bound with heme iron of the CYP450 prosthetic group. The remaining antifungal azoles develop bonding interactions with the apoprotein, determining the drug's relative selectivity for specific apoprotein (Vanden et al., 1989). Computational chemistry has made enormous advances recently, posing new hurdles to drug development through a statistical method (Sliwoski et al., 2014; Genheden et al.,2017). QSAR is a dependable method for establishing a quantitative relationship between biological activity and descriptors and representing the physicochemical properties of compounds in a sequence (Mandlik et al., 2016; Noolvi et al., 2013; Thereza A et al., 2022; Kwon et al.,2019). It helps to determine the biological activity of newly developed series that contributes to the process of drug discovery (Bhole et al., 2009; Golbraikh et al., 2017). A series of 3D-QSAR studies have been conducted on several compounds to predict their biological functions at specific targets. For instance, the CoMSIA approach has been used by Daoui and his colleagues to develop a 3D-QSAR model of several compounds derived from Magnolia officinalis targeting EGFR tyrosine kinase in non-small cell lung cancer cell line (Daoui et al.,2022). In addition, another study has identified a compound derived from imidazole as a lead compound in targeting breast cancer cells (Rashid M et al., 2023). The mode of the current work is to search for novel 1, 2, and 4- Triazole that has the tendency to be used as an anticancer drug. A series of anticancer drugs known as 4- amino-5-(pyridin3yl)-4H-1, 2, 4 triazole-3-thiol were chosen for QSAR research. KNN-MFA, GA- kNN-MFA, like many 3D QSAR approaches, requires a sufficient alignment of a collection of molecules (Jain et al., 2008; Sharma et al., 2013; Samad et al., 2014). Steric and electrostatic energy is calculated using a charge + 1 methyl probe at the grid's lattice sites which were generated. The relationships using the kNN approach are generated with the help of interaction energy values at the grid locations. Training sets constructed in this approach cover all descriptor space areas of representative points. The test set is used to evaluate the QSAR model's prediction ability based on the training set, which is not used in model formation. Once the training and test sets have been constructed, the kNN method is applied to descriptors generated over the grid.

2. Materials and methods

2.1. Ligand preparation

The 2D sketch application was used to create the constructions, which were then translated into 3D structures. To optimize the structures, the Dreiding Force Field method (Sasaki and Yamashita, 2021) and the Modified Qeq Charge in dielectric characteristics were utilized, with a convergence requirement being 0.01 (root mean square gradient), 10,000 cycles, and 1.0 serving as the dielectric constant. The limits of electrostatic energy were determined to be 30.0 Kcal/mol, whereas the limits of steric energy were determined to be 10.0 Kcal/mol.

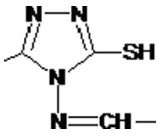
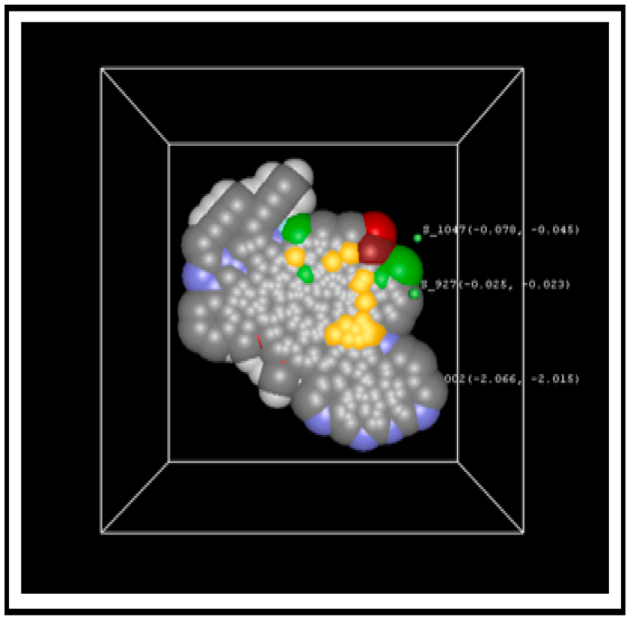


Fig. 1. Template used for alignment of molecules.

(a) Space fill model



(b) Ball and stick model

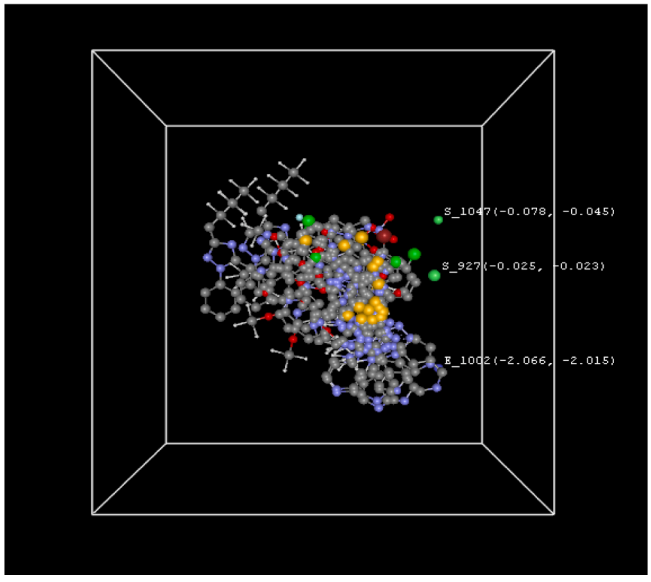


Fig. 2. Template-based alignment of molecules.

2.2. Molecular alignment

Proper molecule alignment is required for investigating 3D QSAR that depicts the structural variety of a given set of compounds (Raj et al., 2011). Template (Fig. 1) was used and the most active molecule $5a_1$ was

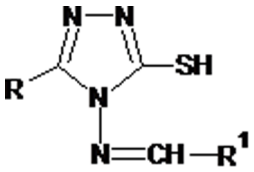


Fig. 3. 4- Amino-5-(pyridin3yl)-4H-1, 2, 4 triazol-3-thiol derivatives with IC 50 values.

used as a reference molecule. Template-based derivatives were aligned (Fig. 2).

2.3. Data set

The basic nucleus of 20 molecules containing 4- amino-5-(pyridin3yl)-4H-1, 2, 4-triazol-3-thiol was produced and used for 3D-QSAR analysis utilizing the regression analysis approach (Fig. 3). Data set containing training set (16 compounds) and a test set (four compounds) was constructed to evaluate the QSAR model externally.

2.4. 3D QSAR model generation by GA- kNN-MFA

Partial Least Squares (PLS) investigations and molecular modeling were carried out using an HCL computer equipped with a genuine Intel Pentium Dual Core processor and Windows XP operating system. The program known as Molecular Design Suite was utilized for these tasks (MDS). The descriptor molecular field is computed using grid points located all over the space around the molecule. This section provides a general description of the manner in which each molecule will bind to the active site. Using a methyl probe with a charge of + 1, steric, electrostatic, and hydrophobic interaction energies were calculated at the grid's lattice points. For synthesized compounds, the k-Nearest Neighbor-Molecular Field Analysis (kNN-MFA) methodology with Genetic Algorithm (GA) (Raj et al., 2011; Jain and Nag, 2012) was used, which was found to be the most relevant and acceptable method for performing 3D QSAR. A simple approach to distance learning is used in the kNN methodology. An unidentified number is assigned a classification using this method based on in the training set, the vast majority of its k-Nearest Neighbors. Appropriate distance measures are used to determine proximity. The GA variable selection approach simulates a physical process known as annealing, which involves heating a system to a high temperature and then gradually cooling it to a predefined temperature.

2.5. Validation of the models

2.5.1. Validation of the models

Both internal and external sources were used to validate the models (Crisan et al., 2019). This internal validation (cross validation), depends upon the extraction of compound from training set to predict its pharmacological action. Repetition in removal of compounds was continuing until all compounds get removed and predicted for their pharmacological actions. The value of the predicted correlation coefficient, also known as pred r2, is used in the process of performing external validation [(pred r2)]. The following equation is used to do this calculation. yi and y* represent the actual and anticipated activities of the ith compound in the test set, respectively, whereas y mean represents the average activity of all of the compounds in the training set. Both summations are applicable to each and every one of the test chemicals. The obtained value for the pred r2 can be considered suggestive.

$$pred\_r^2 = 1 - \sum (yi - y^*)^2 / \sum (yi - ymean)^2$$

Table 1
Results of 3D QSAR analysis using GA- kNN – MFA method.

Statistical parameters	Training set Model A	Training set Model B	Training set Model C
	Test Set a	Test set b	Test set c
K nearest neighbour	3	3	3
N	13	15	18
Degree of freedom	10	10	14
q^2	0.4843	0.7356	0.2129
q^2_se	0.966	7.4777	0.1152
pred_r^2	0.2102	0.9345	0.8417
pred_r^2se	0.1452	28.1821	0.1255
	(+)vely contribution		(-)vely contribution
Descriptors (Test set c)			E_1002 −2.0659—2.0147 S_1047 −0.0780—0.0451 S_927 −0.0255—0.0229

Where, N = no. of molecules, k = no. of descriptors in a model, r^2 = coefficient of determination, q^2 = cross-validated r^2, pred_r^2se = r^2 of external test set.

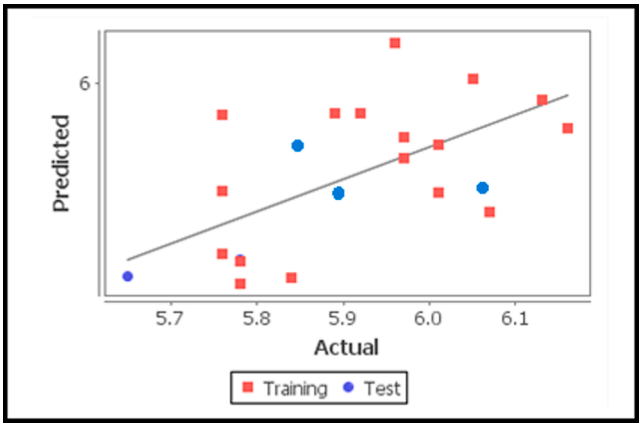


Fig. 4. Fitness Plot for Training and Test set.

3. Results

3.1. 3D QSAR studies using GA- kNN – MFA method

Training set (16 compounds) and test set (4 compounds) containing 4- amino-5-(pyridin3yl)-4H-1, 2, 4 triazole-3-thiol rings were constructed using genetic algorithm and sphere exclusion methods in Vlife MDS suite. Training and test sets were based upon Unicolumn statistics. The outcome of our findings indicates that the test could be used for interpolation. A genetic algorithm was applied for building QSAR models. In 3D descriptor models, the minimal structural characteristics of 4- amino-5-(pyridin3yl)-4H-1, 2, 4 triazol-3-thiol derivatives were obtained. When all these models were compared, model "C'' was judged to be the best, with a pred r^2 score of 0.8417. The contribution of electronic and steric properties of compounds was shown to be optimal for biological activity at lattice positions E 1002, S 1047, and S 927, respectively (Table 1). Statistically model is good with respect to r^2, q^2, and pred-r^2. The straight-line graph (Fig. 4) exhibited actual and predicted values of the model generated. The radar graph confers the idea of the best model selection through comparison. The observed and predicted activities of the training set and test set were shown in red and blue in the radar graph (Fig. 5). Overlay of observed and predicted activities of the training set (A) predicts a better r^2 value which can be confirmed by Table 2.

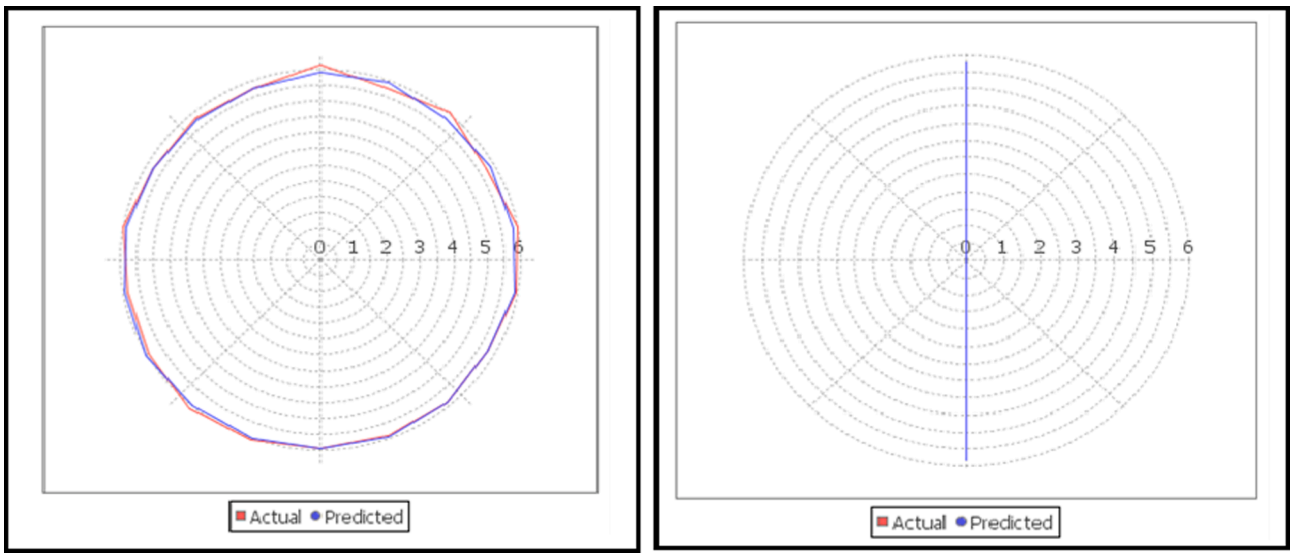


Fig. 5. Radar Graph between actual and predicted biological activity of training set (A) and test set (B).

Table 2
Actual value, Predicted value, and Residual value of synthesized compounds.

Compounds	Actual value	Predicted value	Residual value
$5a_1$	92.87	80.23	12.64
$5a_2$	59.1	64.36	−5.26
$5a_3$	68.23	59.12	9.11
$5a_4$	82.71	72.80	9.81
$5a_5$	67.23	70.17	−2.94
$5a_6$	92.79	79.59	13.2
$5a_7$	75.83	75.30	0.53
$5a_8$	70.23	72.45	−2.22
$5a_9$	73.22	71.51	2.21
$5a_{10}$	62.3	63.61	−1.31
$5a_{11}$	94.24	80.19	14.05
$5a_{12}$	62.45	72.42	−9.97
$5a_{13}$	75.83	75.57	0.26
$5a_{14}$	81.22	90.30	−9.08
$5a_{15}$	48.88	45.34	3.54
$5a_{16}$	46.09	52.71	−6.62
$5a_{17*}$	46.09	52.71	−6.62
$5a_{18*}$	62.45	72.42	−9.97
$5a_{19*}$	46.12	52.13	−6.3
$5a_{20*}$	59.23	58.41	0.82

4. Discussion

Quantitative structure–activity relationships, also known as QSAR, have been utilized for many years to establish correlations between the physicochemical properties of chemical substances and the biological activities of those substances in order to develop a statistical model that is reliable for predicting the activities of new chemical entities. Formalism can be summed up in its central tenet, which states that discrepancies in the biological activity of substances can be accounted for by changing in the structural properties of the substances (Verma et al., 2010). In traditional QSAR studies (Garro et al., 2015), the affinities of ligands to the sites at which they bind, inhibition constants, rate constants, and other biological end points have been shown to be correlated with atomic, group, or molecular properties such as lipophilicity, polarizability, electronic and steric properties (Hansch analysis), or certain structural features (Free-Wilson analysis). In earlier research, 1,2,4-antifungal, antibacterial, anticancer, anticonvulsant, antituberculosis, antiviral, an-tiparasitic, analgesic, and anti-inflammatory actions were discovered, along with structure–activity relationships for 1,2,4-triazole derivatives (Aggarwal and Sumran, 2020). Previous SAR investigation (Kumari et al., 2021) of triazole derivatives show that substitution on positions 3, 4, and 5 of the triazole ring can be changed, but the groups linked to the nitrogen atom at the 4th position inflict the biggest change in physicochemical attributes and biological profile. It promotes hydrogen bonding and is also resistant to metabolic degradation, which may be advantageous in improving solubility and binding bimolecular targets. In the present study, anti-cancer activity was established for 20 substituted 1, 2, 4-triazole derivatives. We used steric, electrostatic field descriptors to evaluate the activity of new molecules. Negative moderate electrostatic data points revealed low to optimum electronegative substituents for anticancer activity, while negative low to optimum steric data points revealed low to optimum bulkier moieties for anticancer activity (Slavov et al., 2007).

5. Conclusion

The results of 3D QSAR studies developed the models that showed a relationship between biological activity and statistical significance. These models contain descriptors calculated from the training and test sets provided. Negative optimum electrostatic data indicated the requirement of low to optimum electronegative substituents enhancing biological activity, and negative low average values of steric data points indicated that low to optimum groups are necessary for anticancer activity. Prediction of biological activity serves as a screening tool for a vast compound library, allowing potent, optimal compounds with high predicted activity to be selected. As a result of calculating residuals for all models using the GA-kNN-MFA QSAR method, model “C” was found to be the best model. The good to moderate anticancer potential of compounds, as well as established facts and hypotheses about the anticancer action of these compounds, balanced our attempt at 3D-QSAR studies of 1,2,4-Triazole derivatives. These results could lead to the identification of potential chemical compounds with optimum anticancer activity and minimal side effects. Further laboratory and clinical studies need to be conducted to confirm our findings and to elucidate their clinical relevance.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors would like to acknowledge the financial support offered by the Researchers Supporting Project number (RSPD2023R853) at King Saud University, Riyadh, Saudi Arabia.

Funding

The authors are thankful for the financial support offered by the Researchers Supporting Project number (RSPD2023R853) at King Saud University, Riyadh, Saudi Arabia. The authors also would like to express their gratitude to AlMaarefa University in Riyadh, Saudi Arabia, for providing funding for this study.

References

Aggarwal, R., Sumran, G., 2020. An insight on medicinal attributes of 1,2,4-triazoles. Eur. J. Med. Chem. 205, 112652.

Bhole, R.P., Bhusari, K.P., 2009. 3-D QSAR study of p-hydroxybenzohydrazide derivatives as Anti Escherichia coli agents. Dig. J. Nanomater. Biostructures. 4 (2), 329–338.

Crisan, L., Iliescu, S., Ilia, G., Funar-Timofei, S., 2019. Quantitative structure-property relationship modeling of phosphoric polyester char formation. Fire Mater. 43 (1), 101–109.

Daoui, O., El Mouhi, R., Barghady, F., et al., 2022. 3D-QSAR modeling, Molecular Docking and drug-like properties investigations of novel heterocyclic compounds derived from Magnolia Officinalis as Hit Compounds against NSCLC. Moroccan. J. Chem. 10 (4).

Dilmaghani, K.A., Nasuhi Pur, F., Nezhad, M.H., 2015. Synthesis and antibacterial evaluation of new thione substituted 1,2,4-triazole schiff bases as novel antimicrobial agents. Iran. J. Pharm. Res. 14 (3), 693–699.

Garro Martinez, J.C., Vega-Hissi, E.G., Andrada, M.F., et al., 2015. QSAR and 3D-QSAR studies applied to compounds with anticonvulsant activity. Expert Opin. Drug Discov. 10 (1), 37–51.

Genheden S, Reymer A, Saenz-Méndez P, et al. Chapter 1. Computational Chemistry and Molecular Modelling Basics. In Computational Tools for Chemical Biology 2017; pp. 1-38.

Golbraikh A, Wang XS, Zhu H, et al. Predictive QSAR modeling: Methods and applications in drug discovery and chemical risk assessment. In: Handbook of Computational Chemistry 2017; pp. 2303-2340.

Jain, A.K., Manocha, N., Ravichandran, V., et al., 2008. Three-dimensional qsar study of 2, 4 - disubstituted-phenoxy acetic acid derivatives as a Crth 2 receptor antagonist : using the k-nearest neighbor method. J. Nanomater. 3 (3), 147–158.

Jain, S.K., Nag, S., 2012. 3D QSAR analysis on arylbenzofuran derivatives as histamine H3 receptor inhibitors using K nearest neighbor molecular field analysis (KNNMFA). Int. J. PharmTech Res. 4 (4), 1691–1702.

Kumari, M., Tahlan, S., Narasimhan, B., Ramasamy, K., Lim, S.M., Shah, S.A., Mani, V., Kakkar, S., 2021 Dec. Synthesis and biological evaluation of heterocyclic 1, 2, 4-triazole scaffolds as promising pharmacological agents. BMC Chem. 15, 1–6.

Kwon, S., Bae, H., Jo, J., et al., 2019. Comprehensive ensemble in QSAR prediction for drug discovery. BMC Bioinf. 20, 521. https://doi.org/10.1186/s12859-019-3135-4.

Mandlik V, Bejugam PR, Singh S. Application of Artificial Neural Networks in Modern Drug Discovery. In: Artificial Neural Network for Drug Design, Delivery and Disposition. Edition: 1. Chapter: Section 2, Chapter 6. 2016; pp. 123-139.

Noolvi, M.N., Patel, H.M., 2013. A comparative QSAR analysis and molecular docking studies of quinazoline derivatives as tyrosine kinase (EGFR) inhibitors: A rational approach to anticancer drug design. J. Saudi Chem. Soc. 17 (4), 361–379.

Raj, N., Jain, S.K., 2011. 3d QSAR studies in conjunction with k-nearest neighbor molecular field analysis (k-NN-MFA) on a series of substituted 2-phenyl-benzimid-azole derivatives as an anti allergic agents. Dig. J. Nanomater. Biostruct. 6 (4), 1811–1821.

Maryam Rashid, 1 Ayesha Maqbool, 1 Nusrat Shafiq. et al. The combination of multi-approach studies to explore the potential therapeutic mechanisms of imidazole derivatives as an MCF-7 inhibitor in therapeutic strategies. Frontier in Chemistry 2023; 11: 1197665.

Samad, A., Naffaa, M.M., Bakht, M.A., et al., 2014. Target based designing of anthracenone derivatives as tubulin polymerization inhibiting agents: 3D QSAR and docking approach. Int. J. Med. Chem. 2014, 658016.

Sasaki, K., Yamashita, T., 2021 Feb 12. Modification and validation of the DREIDING force field for molecular liquid simulations (DREIDING-UT). J. Chem. Inf. Model. 61 (3), 1172–1179.

Sharma, M.C., Sharma, S., Sahu, N.K., et al., 2013. 3D QSAR kNN-MFA studies on 6-substituted benzimidazoles derivatives as Nonpeptide Angiotensin II Receptor Antagonists: A rational approach to antihypertensive agents. J. Saudi Chem. Soc. 17, 167–176.

Singh, R.J., Singh, D.K., 2009. Syntheses, characterization and biological activity of some 1, 2, 4-triazole derivatives. E-J. Chem. 6 (3), 796–800.

Slavov, S., Atanassova, M., Galabov, B., 2007. QSAR analysis of the anticancer activity of 2,5-disubstituted 9-aza-anthrapyrazoles. QSAR Comb. Sci. 26 (2), 173–181.

Sliwoski, G., Kothiwale, S., Meiler, J., et al., 2014. Computational methods in drug discovery. Pharmacol. Rev. 66 (1), 334–395.

Soares, T.A., Nunes-Alves, A., Mazzolari, A., et al., 2022. The (Re)-Evolution of Quantitative Structure-Activity Relationship (QSAR) studies propelled by the surge of machine learning methods. J. Chem. Inf. Model. 62 (22), 5317–5320. https://doi.org/10.1021/acs.jcim.2c01422.

Vanden, B.H., Marichal, P., Gorrens, J., et al., 1989. Biochemical approaches to selective antifungal activity. Focus on azole antifungals. Mycoses 32 (1), 35–52.

Verma, J., Khedkar, V.M., Coutinho, E.C., 2010. 3D-QSAR in drug design–a review. Curr. Top. Med. Chem. 10 (1), 95–115.

Wakale, V.S., Pattan, S.R., Vishal, T., 2013. Therapeutic importance of 1, 2, 4-triazole : A review. Int. J. Res. Pharm. Biomed. Sci. 4 (3), 985–1001.