



Review on QSAR using Anticancer Drug

Hemant AHIRWAR, Gabbar Kurmi, Rubeena Khan, Basant Khare, Anushree Jain, Prateek Kumar Jain, Bhupendra Singh Thakur*

Adina College of Pharmacy, ADINA Campus Rd, Lahdara, Sagar, MP, 470001

Article Info:

Article History:

Received 17 Sep 2022
Reviewed 09 Nov 2022
Accepted 24 Nov 2022
Published 15 Dec 2022

Cite this article as:

Ahirwar H, Kurmi G, Khan R, Khare B, Jain A, Jain PK, Thakur BS, Review on QSAR using Anticancer Drug, Asian Journal of Dental and Health Sciences. 2022; 2(4):59-63

DOI: <http://dx.doi.org/10.22270/ajdhs.v2i4.27>

*Address for Correspondence:

Bhupendra Singh Thakur, Adina College of Pharmacy, ADINA Campus Rd, Lahdara, Sagar, MP, 470001

Abstract

New drug discovery has been acknowledged as a complicated, expensive, time-consuming, and challenging project. It has been estimated that around 12 years and 2.7 billion USD, on average, are demanded for a new drug discovery via traditional drug development pipeline. How to reduce the research cost and speed up the development process of new drug discovery has become a challenging, urgent question for the pharmaceutical industry. Computer-aided drug discovery (CADD) has emerged as a powerful and promising technology for faster, cheaper and more effective drug design. Recently, the rapid growth of computational tools for drug discovery, including anticancer therapies, has exhibited a significant and outstanding impact on anticancer drug design, and has also provided fruitful insights into the area of cancer therapy. In this work, we discussed the Qualitative structure activity relationship, a computer-aided drug discovery process with a focus on anticancer drugs.

Keywords: New drug discovery, Computer-aided drug discovery, the Qualitative structure activity relationship, Anticancer

Introduction

Up to now, cancer remains a global and serious public health challenge. It is estimated that there are more than 200 different types of cancer, generally named according to the tissue where the cancer was recognized for the first time. Cancer is considered to be one of the significant causes for death in the 21st century and the most critical obstacle for the increase of global life expectancy. According to an analysis by the world health organization (WHO) in 2015, cancer is the second leading cause of death for patients younger than 70 years old in 91 countries and the third or fourth leading cause of death among 22 other countries¹. moreover, a global increase of 18.1 million new cancer cases and 9.6 million cancer-related deaths have been reported in a previous study², especially 70% of the death caused by cancer occur in low-income and middle-income countries. The fast growth of the cancer incidence and mortality has turned out to be global health challenges. How to reduce the cancer-related death rate has attracted significant attention from the government, society, medical industry, as well as scientific communities, expecting the rapid development of effective and safe drugs for cancer treatment. Despite of the impressive progress in biotechnologies and further understandings of the disease biology, the development of new, practical and innovative small molecule drugs remains an arduous, time consuming and expensive project, which requires collaborations from many expertise in multidisciplinary fields, including medicinal chemistry, computational chemistry, biology, drug metabolism, clinical research, etc. Furthermore, it has been

illustrated that the successful discovery and development of a new drug costs 12 years, and expensive investment³. Thus, novel drug development strategies with a reduced cost of time and money, as well as an enhanced efficiency are in high demand, which would contribute to a significant improvement in global health and life expectancy. Since the successful development of HIV protease inhibitor Viracept in the USA in 1997, which was the first drug design fully driven by its target structure⁴, computational methods have served as an essential tool in drug discovery projects and have been a cornerstone for new drug development approaches. This makes the drug developmental process faster and cheaper. Recently, the fast growth in computational power, including massively parallel computing on graphical processing units (GPUs), the continuous advances in artificial intelligence (AI) tools^{5,6}, have translated fundamental research into practical applications⁷ in the drug discovery field. This attracted considerable attention for their outstanding performance on providing new promising perspectives and solutions to overcome life threatening diseases. In this review, we aim at providing an overview of QSAR-method and anti-cancer therapy discovery in particular. We reviewed some of the most representative examples and clarified fundamental principles by exploring studies on anticancer drug designs with the help of QSAR methods.

Anti-cancer drug target prediction

Human contains approximately 30,000 genes, among which around 6,000 to 8,000 sites are estimated as potential pharmacological targets. However, less than 400 encoded

proteins have been proved to be effective for drug development until now^{8,9}. Cancer, compared to many other human diseases, now has a plethora of potential molecular targets for therapeutic development¹⁰. Traditional drug discovery mainly follows the paradigm of "one molecule - one target - one disease", without considering the interactions between drugs and proteins. However, an important fact that many complex diseases are relevant to a variety of target proteins¹¹⁻¹³ has been overlooked. Furthermore, unexpected drug functions derived from off-targets are accidental and uncontrollable activities because of the 'poly pharmacological' properties of certain drugs, which might result in undesirable side effects. Those are particularly pronounced for cancer drugs. On the other hand, there are some positive examples that benefit from the different pathways targeted by one given molecule. For example, sildenafil (viagra) was developed to treat angina, but now it is used for erectile dysfunction therapy¹⁴. There are several drugs, including anticancer drugs, whose corresponding target proteins (both primary and non-target) remain yet unidentified or unknown¹⁵. Furthermore, some attractive and potentially effective cancer targets remain outside of the scope of pharmacological regulation. Some of these targets such as phosphatases, transcription factors, and RAS family members have been described as undruggable, as they lack effective enzymatic active sites¹⁰. To make the full use of known drugs to treat new indications, the characterization of all potential new ligand binding sites has been illustrated as a key point in drug repositioning and repurposing. Therefore, new and highly qualitative bioinformatic target prediction methods are required for the accurate prediction of drug targets. Up to now, a wide range of drug target interactive web servers has been established, providing a series of drug-target databases and prediction tools. Moreover, various computational approaches have been used to study potential interactions between proteins and drugs. In particular, network-based models and ML-based models have emerged as important tools. A review by Chen et al. summarizes several available computational models for this application¹³. Interestingly, a method proposed by Campillos et al. that uses the similarity of drug side effects to determine whether multiple drugs could interact with the same target proteins attracted our attention¹⁶. Based on this research, Takarabe et al. took advantage of the US FDA's adverse event reporting system (AERS) to define the pharmacological similarity of all potential medicines and developed a novel system to predict large-scale interactions between unknown drug targets¹⁵. Notably, AERS was employed to predict interactions between drugs and targets for the first time. In 2010, Klipp et al. summarized several available computational models for network-based drug-target prediction¹⁷. Moreover, various biological data settings, including structures of bioactive compounds, sequences of target proteins, and information of ligand-target interactions, have been combined. A series of machine learning-based approaches have been demonstrated as efficient tools in detecting relationships among drug structures and corresponding target proteins from a large amount of data, such as supervised learning method¹⁸, bipartite graph learning method¹⁹, bipartite local model²⁰ and so on. A recent review by Mayr et al. compared the predictive performance of deep learning with other prediction approaches for multiple drug targets in the comparative studies of composite target prediction methods. As a result, feed-forward neural networks were identified with better performance in drug target prediction than other methods²¹. As above, since a large number of compounds and vigorous efforts are abandoned and wasted due to the off-target effects during the classical drug development procedure, a greatly enhanced development of target prediction in new drug exploration exhibited attractive advantages and further expansion in this area are still highly desirable.

Quantitative structure activity relationship

QSAR (Quantitative Structure Activity Relationship) is a ligand-based approach that relies on analyzing the biological activities of drugs using various molecular descriptors (MDs) or fingerprints (FPs). These models mathematically describe how the activities response to the targets according to the ligand's structural characteristics. QSAR was obtained by calculating the correlations between the properties of the ligand binding agent and the biological activity measured by experiments. Different ML and deep learning (DL) approaches have also been applied to develop QSAR models²²: including Support Vector Machine (SVM), Random Forest (RF), Polynomial Regression (PR), Multi Linear Regression (MLR), Artificial Neural Network (ANN). Unlike the pharmacophore models, QSAR models can measure biological activities quantitatively and can even find positive or negative effects according to certain characteristics of the molecule on its activity. QSAR has been applied to many other molecular design purposes, such as predicting the new molecule analog activity, optimizing lead, and predicting new structural leads in drug discovery. In the classical 2D-QSAR approaches, the biological activity is related to physical and chemical features consisting of steric, electronic, and hydrophobic characters of drugs, and the relationships are represented as mathematical equations²³. More advanced 3D-QSAR approaches, such as comparative molecular field analysis²⁴ and molecular similarity indexes in a comparative analysis²⁵, are based on the force field calculations. The structural information of molecules is needed, and developed models are represented in 3D contour maps facilitating the visualization and interpretation

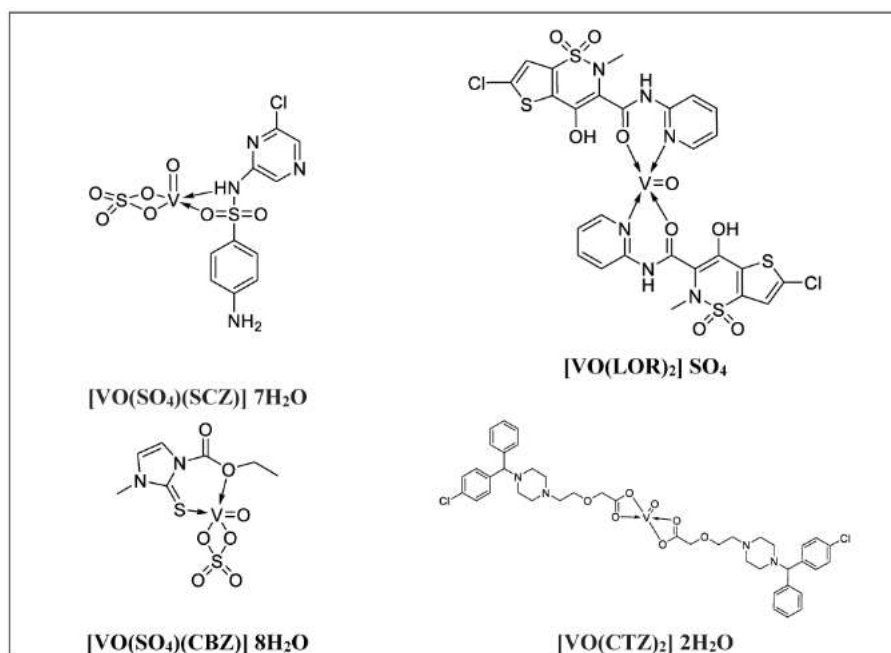
Successful stories of computational drug discovery

QSAR methods have proved to play an essential role in modern drug discovery. Since computational methods could cover almost all stages of the drug discovery pipeline, the applications of QSAR methods in anticancer drug discoveries have shown great advantages in terms of the required investment, resources, and time. More recently, QSAR methods have become a potent and powerful tool in several successful cases of anticancer drug development. Herein, we list several successful applications of QSAR methods for small molecule drugs, which have been applied to cancer treatment or are at later stages in the clinical trial. In a study of alomori et al., four oxidovanadium(IV) complexes have been synthesized and characterized by several spectroscopic methods. The IR spectra suggested that the ligands have bidentate coordination mode to the vanadium ion. In addition, the molar conductance, EPR, magnetic moment values, and electronic data support a square pyramidal structure for all complexes.

Computational studies were applied to demonstrate the optimization geometry and essential quantum parameters and confirm their biological efficiency. The metal complexes showed a square-pyramidal geometry arrangement around the metal ion which was agreeable with experimental results. The form coordination bonds length presented strength bonded between oxidovanadium(IV) and the investigated ligand to form stable complexes. The bioactivity study was varied and involved DNA-binding, molecular docking, QSAR, and cytotoxicity analyses. DNA binding results revealed two behaviors for the synthesized oxidovanadium(IV) complexes with an increase in CT-DNA concentration, hyperchromic with electrostatic or grooves binding modes and hypochromic signifying an intercalation binding mode. Molecular docking results showed that the [VO (CT)₂] 2H₂O complex exhibited significant interaction with colon cancer (3IG7) Protein with good selectivity. QSAR study has given significant information

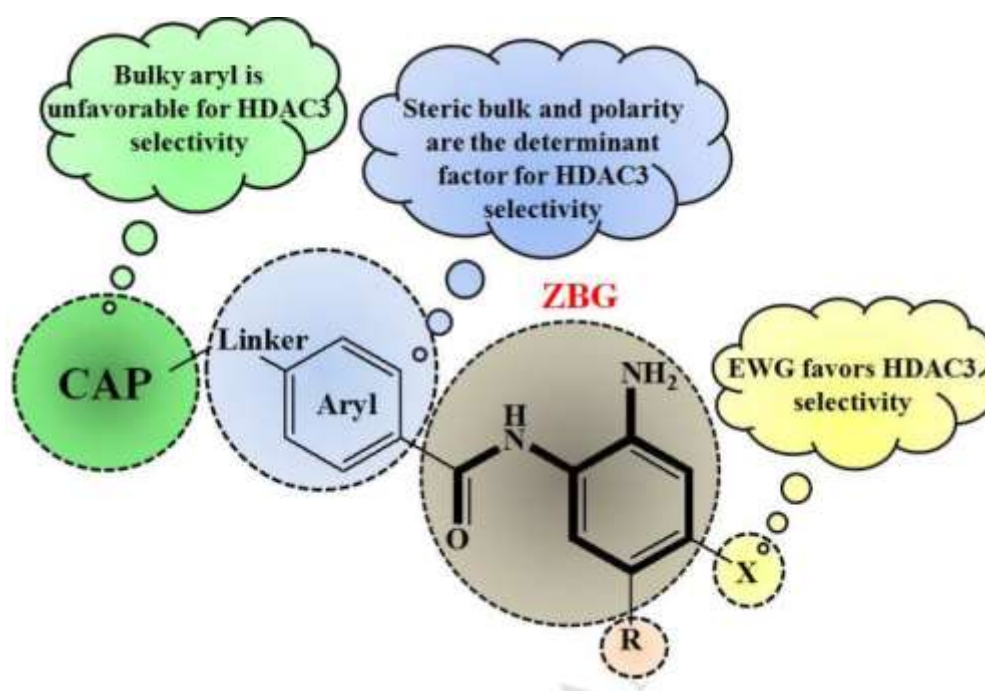
on biological activity by using the MLR method. QSAR model showed a good correlation between the predicted and the experimentally observed inhibitory activities. The results of validation indicate that the generated QSAR model possessed a high predictive power ($R^2 = 0.97$). Based on the molecular

docking and QSAR results, $[\text{VO}(\text{CTZ})_2] \cdot 2\text{H}_2\text{O}$ was selected and tested for its inhibitory activity against colon cancer cell line (HCT116). The selected complex showed higher anticancer activity than the standard cisplatin chemotherapy drug²⁶.



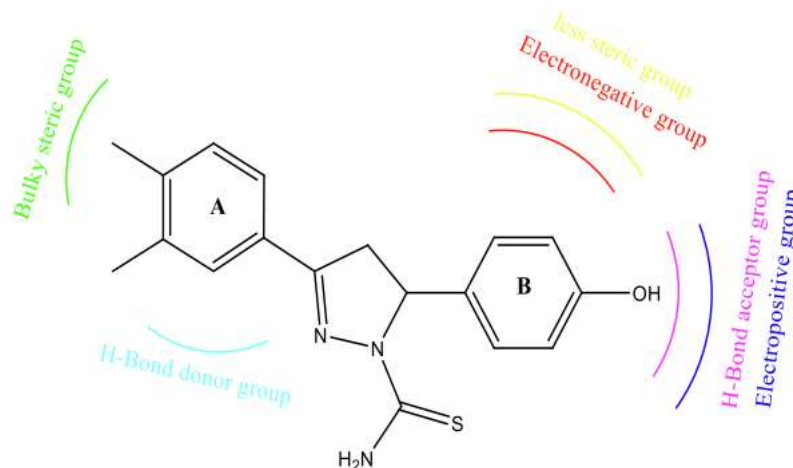
In another research QSAR and QAA studies have been conducted on diverse benzamide-derived HDAC3 inhibitors as the first initiative to explore the designing strategies of higher active and selective HDAC3 inhibitors over HDAC1 and HDAC2. QSAR models reveal that molecular size and shape along with the steric effect should have to be optimized to achieve higher HDAC3 inhibition. QAA models reflect that modification/substitution at the benzamide scaffold should be optimized in such a way so that these molecules possess lower steric bulk along with nonpolar features for achieving higher

HDAC3 selectivity over HDAC1 and HDAC2. However, the importance of spiro hydrophobic cap group, as well as electron withdrawing fluorine group at the benzamide scaffold, should be well-accounted for retaining higher HDAC3 selectivity over HDAC1. Moreover, less polar and less hydrophobic benzamides are preferred for HDAC3 selectivity over HDAC2. This detailed structural exploration will surely unveil a new vista of designing highly potent and selective benzamide-based HDAC3 inhibitors that may be a crucial weapon to battle against a variety of cancers²⁷.



An example of QSAR studies, to design new molecules with anticancer activity, two 3D QSAR models was developed: CoMFA ($Q_2 = 7.32$, $R_2 = 9.92$) and CoMSIA ($Q_2 = 7.32$, $R_2 = 9.92$) based on 42 Pyrazole derivatives. The results obtained for both models were satisfactory and validated by reliable methods. All the newly designed molecules respect the different drug-likeness rules (Lipinski, Veber, Ghose, Egan, Muegge, and ADMET prediction). Moreover, molecular

docking studies results on EGFR receptor indicated that the interaction results of the designed compounds T1-T5 in the cavity of the EGFR receptor (1 M17) show more type and the number of interactions (Hydrogen Bond interaction) compared to erlotinib as reference inhibitors of EGFR. Together, these results facilitate and guide the design and synthesis of novel and more potent Pyrazole derivatives with anticancer activity²⁸.



Conclusion

Cancer has become a tangible threat to human health. About 9.6 million people are estimated to die from the various forms of cancer each year, according to a statistic report. Cancer has become the second-largest disease that causes human death. However, developing a new drug molecule costs 12 years and 2.7 billion USD on average. The drug development for cancer even becomes more complicated, especially considering the molecular pharmacology is still not well understood. Hence, the discovery and development of new drugs is considered very expensive and time-consuming. In this respect, computational methods could be constructive for performing different tasks including protein-interaction network analysis, drug target prediction, binding site prediction, virtual screening, and many others. All these innovative methods could considerably facilitate the anti-cancer drug discovery. In recent years, with the advance of AI, more sophisticated methods, such as QSAR, retro-synthetic routine plan, drug scaffold generation, drug binding affinity predictions, were developed. The useful predictions generated by QSAR models combined with experimental validations could further speed up the anti-cancer drug development.

References

1. Yan B, Yang W, Han X, Han L. Crystal structures and antitumor activity evaluation against gastric carcinoma of two novel coordination polymers. *Main Group Chemistry*. 2019; 18(3):239-246. <https://doi.org/10.3233/MGC-180748>
2. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: A Cancer Journal for Clinicians*. 2018; 68(6):394-424. <https://doi.org/10.3322/caac.21492>
3. Kapetanovic I. Computer-aided drug discovery and development (CADD): in silico-chemico-biological approach. *Chemico-Biological Interactions*. 2008; 171(2):165-176. <https://doi.org/10.1016/j.cbi.2006.12.006>
4. Kaldor SW, Kalish VJ, Davies JF, Shetty BV, Fritz JE, Appelt K, Tatlock JH. Viracept (nelfinavir mesylate, AG1343): a potent, orally bioavailable inhibitor of HIV-1 protease. *Journal of Medicinal Chemistry*. 1997; 40(24):3979-3985. <https://doi.org/10.1021/jm9704098>
5. Chan HS, Shan H, Dahoun T, Vogel H, Yuan S. Advancing drug discovery via artificial intelligence. *Trends in Pharmacological Sciences*. 2019; 40(8):592-604. <https://doi.org/10.1016/j.tips.2019.06.004>
6. Yan B, Yang W, Han X, Han L. Crystal structures and antitumor activity evaluation against gastric carcinoma of two novel coordination polymers. *Main Group Chemistry*. 2019; 18(3):239-246. <https://doi.org/10.3233/MGC-180748>
7. Zhavoronkov A, Ivanenkov YA, Aliper A, Veselov MS, Aladinskiy VA, Aladinskaya AV, Aspuru-Guzik A. Deep learning enables rapid identification of potent DDR1 kinase inhibitors. *Nature Biotechnology*. 2019; 37(9):1038-1040. <https://doi.org/10.1038/s41587-019-0224-x>
8. Drews J. Drug discovery: a historical perspective. *Science*. 2000; 287(5460):1960-1964. <https://doi.org/10.1126/science.287.5460.1960>
9. Chen X, Yan CC, Zhang X, Zhang X, Dai F, Yin J, Zhang Y. Drug-target interaction prediction: databases, web servers and computational models. *Briefings in Bioinformatics*. 2016; 17(4):696-712. <https://doi.org/10.1093/bib/bbv066>
10. Lazo JS, Sharlow ER. Drugging undruggable molecular cancer targets. *Annual Review of Pharmacology and Toxicology*. 2016; 56:23-40. <https://doi.org/10.1146/annurev-pharmtox-010715-103440>
11. Hopkins AL. Network pharmacology: the next paradigm in drug discovery. *Nature Chemical Biology*. 2008; 4(11):682-690. <https://doi.org/10.1038/nchembio.118>
12. Yamanishi Y, Araki M, Gutteridge A, Honda W, Kanehisa M. Prediction of drug-target interaction networks from the integration of chemical and genomic spaces. *Bioinformatics*. 2008; 24(13):i232-i240. <https://doi.org/10.1093/bioinformatics/btn162>
13. Rajak H, Patel V, Jain DK, Dewangan PK, Sharma PC, Veerasamy R, Gupta AK, Dangi JS. Structure-activity relationships among novel 1, 3, 4-oxadiazole analogues for their anti-inflammatory activity. *Current Research in Pharmaceutical Sciences*. 2012; 03:142-148.
14. Ghofrani HA, Osterloh IH, Grimminger F. Sildenafil: from angina to erectile dysfunction to pulmonary hypertension and beyond. *Nature reviews Drug discovery*. 2006; 5(8):689-702. <https://doi.org/10.1038/nrd2030>

15. Takarabe M, Kotera M, Nishimura Y, Goto S, Yamanishi Y. Drug target prediction using adverse event report systems: a pharmacogenomic approach. *Bioinformatics*. 2012; 28(18):i611-i618. <https://doi.org/10.1093/bioinformatics/bts413>
16. Campillos M, Kuhn M, Gavin AC, Jensen LJ, Bork P. Drug target identification using side-effect similarity. *Science*. 2008; 321(5886):263-266. <https://doi.org/10.1126/science.1158140>
17. Klipp E, Wade RC, Kummer U. Biochemical network-based drug-target prediction. *Current Opinion in Biotechnology*. 2010; 21(4):511-516. <https://doi.org/10.1016/j.copbio.2010.05.004>
18. Srivastava N, Hinton G, Krizhevsky A, Sutskever I, Salakhutdinov R. Dropout: a simple way to prevent neural networks from overfitting. *The Journal of Machine Learning Research*. 2014; 15(1):1929-1958.
19. Li X, Chen H. Recommendation as link prediction in bipartite graphs: A graph kernel-based machine learning approach. *Decision Support Systems*. 2013; 54(2):880-890. <https://doi.org/10.1016/j.dss.2012.09.019>
20. Yıldırım MA, Goh KI, Cusick ME, Barabási AL, Vidal M. Drug-target network. *Nature biotechnology*. 2007; 25(10):1119-1126. <https://doi.org/10.1038/nbt1338>
21. Mayr A, Klambauer G, Unterthiner T, Steijaert M, Wegner JK, Ceulemans H, Hochreiter S. Large-scale comparison of machine learning methods for drug target prediction on ChEMBL. *Chemical Science*. 2018; 9(24):5441-5451. <https://doi.org/10.1039/C8SC00148K>
22. Mendenhall J, Meiler J. Improving quantitative structure-activity relationship models using Artificial Neural Networks trained with dropout. *Journal of Computer-Aided Molecular Design*. 2016; 30(2):177-189. <https://doi.org/10.1007/s10822-016-9895-2>
23. Hansch C, Fujita T. Additions and Corrections- ρ - σ - π Analysis. A method for the correlation of biological activity and chemical structure. *Journal of the American Chemical Society*. 1964; 86(24):5710-5710. <https://doi.org/10.1021/ja01078a623>
24. Cramer RD, Patterson DE, Bunce JD. Comparative molecular field analysis (CoMFA). 1. Effect of shape on binding of steroids to carrier proteins. *Journal of the American Chemical Society*. 1998; 110(18):5959-5967. <https://doi.org/10.1021/ja00226a005>
25. Klebe G, Abraham U, Mietzner T. Molecular similarity indices in a comparative analysis (CoMSIA) of drug molecules to correlate and predict their biological activity. *Journal of Medicinal Chemistry*. 1994; 37(24):4130-4146. <https://doi.org/10.1021/jm00050a010>
26. Alomari FY, Sharfalddin AA, Abdellattif MH, Domyati D, Basaleh AS, Hussien MA. QSAR Modeling, molecular docking and cytotoxic evaluation for novel oxidovanadium (iv) complexes as colon anticancer agents. *Molecules*. 2022; 27(3):649. <https://doi.org/10.3390/molecules27030649>
27. Adhikari N, Amin SA, Trivedi P, Jha T, Ghosh B. HDAC3 is a potential validated target for cancer: An overview on the benzamide-based selective HDAC3 inhibitors through comparative SAR/QSAR/QAAR approaches. *European Journal of Medicinal Chemistry*. 2018; 157:1127-1142. <https://doi.org/10.1016/j.ejmech.2018.08.081>
28. Singh A, Patel VK, Jain DK, Sharma PC, Gupta AK, Sharma AK, Rajak H. Histone Deacetylase Inhibitors: Current Status in Treatment of Leukemia. *Current Research in Pharmaceutical Sciences*. 2013; 03(04):102-107