Bulletin of Environment, Pharmacology and Life Sciences
Bull. Env. Pharmacol. Life Sci., Vol 12 [12] November 2023 : 350-357

Online ISSN 2277-1808
Journal's URL: http://www.bepls.com
CODEN: BEPLAD

REVIEW ARTICLE **OPEN ACCESS**

A Mini Review on Descriptors Used in QSAR Study

Phool Singh Yaduwanshi[1]*, Sheelu Singh[1], Indra Kumar Sanodiya[2] Shivkant Patel [3], Manglesh Kumar[2], Samisha Sharma[1]

[1]* IES Institute of Pharmacy, IES University, Bhopal, (M.P.), Pin 462044, India
[2]IITM, IES University, Bhopal, (M.P.), Pin 462044, India
[3] Department of Pharmacy, Sumandeep Vidyapeeth Deemed to University,Piparia Vadodara Gujrat, India
***Corresponding Author:** Phool Singh Yaduwanshi
Email- phoolsinghyaduwanshi@gmail.com

ABSTRACT

The scientific fields of biology, engineering, and chemistry use regression or classification models referred to as QSAR. The biological activity of the compounds may be the QSAR response variable. The physico-chemical characteristics of the substances or their theoretical molecular descriptors serve as predictors in QSAR modelling. First off, QSAR models summarise a chemical data set's purported link between chemical structures and biological activity. QSAR models also foretell the behaviours of new substances. The given quantity of a drug needed to produce a particular biological response can be used to quantify biological activity. Additionally, a mathematical relationship, or quantitative structure-activity relationship, can be found when physicochemical attributes or structures are stated in numerical form. If rigorously validated, the mathematical formula can then be utilised to forecast the modelled reaction of different chemical structures.
Keywords: *QSAR, Drug Design and descriptors.*

Received 01.10.2023 Revised 19.10.2023 Accepted 23.11.2023

INTRODUCTION

Various chemical and physical molecular properties are used in quantitative structure activity relationships (QSARs), a subset of structure properties correlations, to characterize the relationship between structure and property. QSARs are used in drug receptor interactions [1-4]. Since the development of low-cost, high-speed computing technology in recent years, these QSARs are widely used in medicinal chemistry [5].They rely on the capacity to analyse connections between physical attributes and biological function. Biological activity is defined as a linear free energy relationship in this method is unquestionably helpful for anticipating active molecules before simple cumulative synthesis, as well as for examining, albeit indirectly, the nature of the forces that interact between a drug and a receptor. The fact that the ligand occupies three dimensions allows the QSAR technique to be expanded [6]. In essence, QSAR is the application of techniques to determine a correlation between the structural descriptors of chosen molecules and their activities. They can be determined empirically or with the aid of computation techniques. The activities that are employed in QSAR to compare with its descriptors typically involve chemical dimensions and biological testing. Many domains related to drug design are currently making use of QSAR techniques. However, there were several issues that came up when scientists tried to use Hammett-type connections for biological systems, which revealed that the other structural descriptors were equally crucial for QSAR analysis [7, 8]. According to Corwin Hansch's research on the significance of lipophilicity as a metric for compounds, the octanol(oil)-water partition coefficient (log P) is a measure of how well certain compounds perform biologically [8]. It has been shown that using this metric to estimate the compound's bioavailability is useful. In order to link a structural characteristic like lipophilicity to the actions of compounds, various correlations were subsequently constructed.
log P is showing below:

$$=\text{Log}\left(\frac{1}{C}\right)=a \log P + Y \qquad \ldots\ldots\ldots (1.1)$$

Where,
log P = Partition Coefficient
y = Constant value

C = concentration (molar) of compound producing standard response (e.g., ED_{50}, LD_{50}, IC_{50}, etc.).

Technically, the QSAR method makes use of characteristics that have been ascribed to a variety of chemicals groups and can be used to change the structure of drugs. Any description of a compound can be interpreted as the degree to which its group actually contributes to each individual property of the parent compound. Consequently, it is acknowledged that choosing the parameters is a crucial step in QSAR analysis.

The several descriptors that are typically used in QSAR research are as follows.

Thermodynamic Variables

Melting Point: The capillary method and the use of a thistle tube can be used in experiments to estimate it for compounds with higher melting points.

Partition Coefficient (LogP): (octanol (oil)/ water partition coefficient), which measures a molecule's hydrophobicity, is a technical measure. The following equation provides a scientific explanation for it.

$$\text{Log } P_{oct/water} = \text{Log } \frac{[\text{solute}]_{oct}}{[\text{solute}]_{unionized/water}} \quad \ldots\ldots(1.2)$$

The ratio of the concentration of the unionised chemical dispersed between two solutions of different natures (organic and aqueous) can be used to define the value of Log P. By altering the aqueous phase's pH level so that the dominant form of the molecule doesn't become ionised, the log P value of ionizable solutes can be found [9, 10].

Heat of Formation: The degree of a molecule's relative thermal stability is determined by the amount of enthalpy needed to construct the molecule from each of its component atoms. It is estimated using quantum-chemical methods and is useful for modelling chemical reactions, intermolecular interactions, and conformational analysis, among other things.

Molar Refractivity (MR): The refractivity of molar is useful since it may be used to determine a compound's volume and polarisation susceptibility.

The volume in part of this equation is given by the compound's molecular weight divided by its density, and the term known as $(n2 - 1)/(n2 + 1)$ provides a corrective factor that can be used to gauge how easily polarizable a substituent is. The presence of pi electrons or a lone pair of electrons in a compound's substituent has major implications. It can be inferred that the substituent will attach to a polar surface if the molar refractivity value in the QSAR equation is positive, whereas a negative sign suggests steric hindrance at the site of the substituent's binding.

Energy Stretching:

Stretching energy is the term used to describe the energy required to stretch the bonds. By considering that the bond connecting the two atoms operates as a mechanical spring that obeys Hooke's rule, it is possible to calculate the stretching bond energy for pairs of atoms that are connected by a single link.

Torsion Energy:

Stretching energy, denoted by the symbol ETorsion, is mathematically expressed as the bond energy resulting from conformational changes in the bond.

Energy VDW:

The Van der Waals energy of a molecule refers to the energy expended during a molecule's interaction with its receptor. Van der Waal energy, denoted as EvdW, can be derived directly from the Leonard-Jones potential equation.

Electronic Parameters

Energy Bend:

E bend is calculated as: E bend is defined as the change in bond energy caused by modification of the molecule's bond angles.

HOMO Energy:

In a molecule with electrons present, HOMO is regarded as having the greatest energy level. It assists in determining numerous characteristics as well as a molecule's molecular reactivity. When compared to molecules with inferior HOMO energy levels, those with superior HOMO energy levels are better at giving away their electrons and are hence a little more reactive. Thus, it is clear that the HOMO energy descriptor can be used to calculate a molecule's nucleophilicity.

LUMO Energy

The lowest as energy level of a molecule, or LUMO, is one in which no electrons are present. As a result, it might be acknowledged as a crucial limitation in understanding the characteristics and molecular reactivity of a molecule under study. A molecule is said to receive the involved electrons at its LUMO energy level while acting as a Lewis acid theory in the creation of a bond. When compared to molecules with greater LUMO energy levels, molecules with inferior LUMO energy levels are better at receiving their electrons. So it makes sense that the LUMO energy descriptor can be used to determine a molecule's electrophilicity [13].

Steric Parameters

The Ovality:

This technical language can be understood as the range of the molecular core's cross sectional deviation from perfect ovality.

A shell's or molecule's internal ovality can be defined as 2 ((a-b))/((a+b)), where a and b stand for the lengths of the major and minor axes, respectively. In terms of mathematics.

Moment of Dipole: The dipole moment aids in determining a molecule's orientation and strength in an electrostatic field. The Debye units (D) are used to calculate this 3D electronic description.

Connolly Solvent Accessible Area: It is a steric descriptor which helps in determining the molecule's surface area that is interacting with the solvent. This descriptor allows for a negative coefficient in the model, which suggests a growth in the size of the implicated substituent and, as a result, clarifies that the accessible surface area of the molecular solvent does not operate to promote the activity of the molecule under investigation.[14,15]

Connolly Molecular Surface Area: The molecular surface (MS), which is a single sheet that is composed of the merger surface and the candidate surface, is continuous. The merger surface is part of the van der Waals surface that a probe sphere can reach. The candidate surface, however, is the probe's interior surface when it comes into touch with two or more atoms. It goes by the name of Connolly surface.

Connolly Solvent Excluded Volume: The volume that is limited inside the molecular surface when it is in use or in contact is known as the Connolly solvent excluded volume. The limit of the grouping of all potential probes that do not pierce the molecule itself is known as the solvent-excluded surface (SES), often referred to as the molecular exterior.

Principle of Movement of Inertia (X,Y,Z): The total body's moment of inertia with respect to one of the major axes (X, Y, or Z) is the definition of principle moment of inertia. A series of straight lines through the centre of mass are used in the calculation.[16,17]

Discovery and layout

From a blind screening approach that seeks to find chemical hits largely through a process usually referred to as "rational" drug discovery and layout, the search for hit molecules with the ability to work as tablets has substantially improved. The ACE inhibitor Capoten (captopril), which was created in the 1980s, was the first medication to maximise the use of structural information. The HIV protease inhibitor (Viracept), which was introduced in 1997, was the first drug with a design that was totally pushed through the form of the target accepted for the American market. These results signalled the beginning of an intense search for computational methodologies, techniques for increasing and developing new drugs, and better, faster, and less expensive treatments. [18,19]

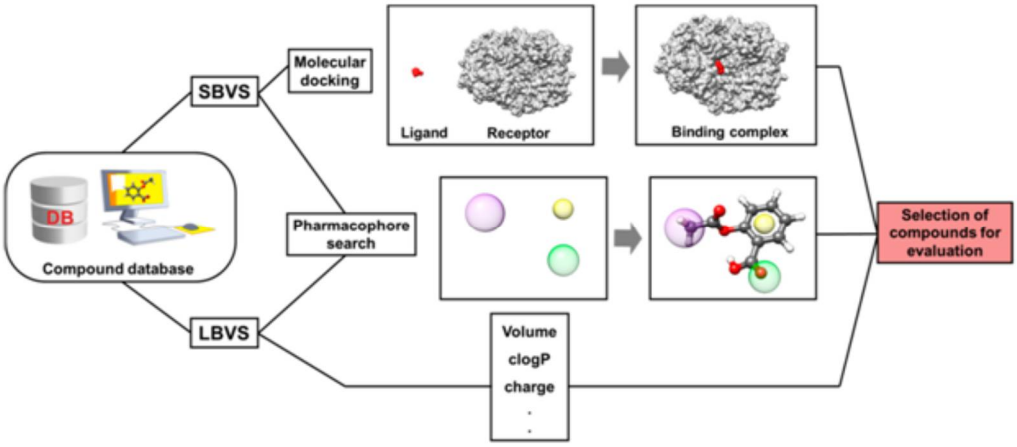

Figure 1.2 Computer-aided drug discoveries

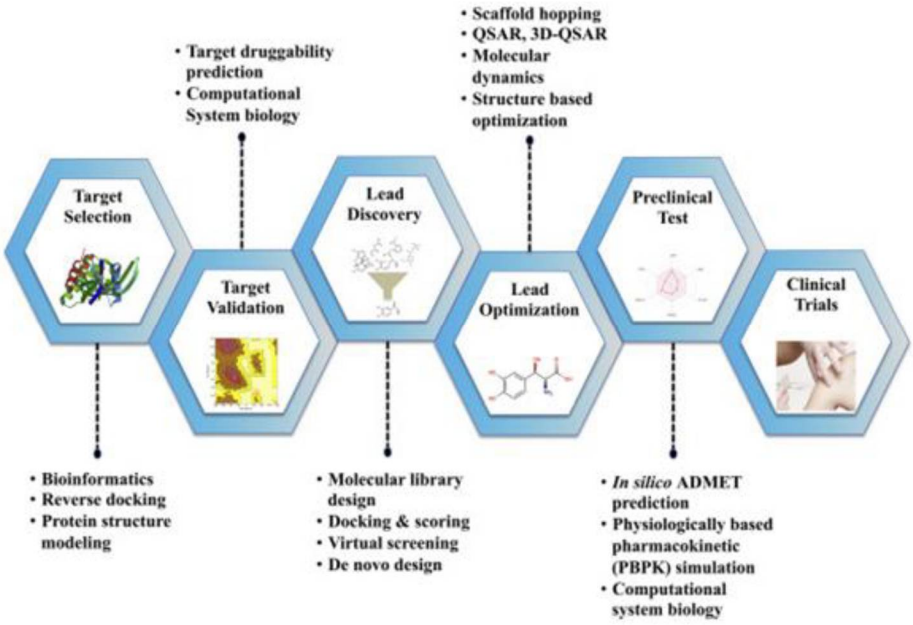

Figure 1.3 Steps involved in Computer based drug discovery process

Ligand and shape based methods

Evidence of the effectiveness of computational drug design in the field of drug development includes a wide spectrum of novel therapeutic entities that are presently undergoing clinical investigation.

Ligand-based totally strategies

Utilise present knowledge of active compounds to make predictions about novel chemical entities that will behave similarly. To identify the minimal essential structural properties a molecule must have in order to bind to the target of attention, a library of compounds can be used to build a pharmacophore model from a single considered energetic molecule. Utilising fingerprint-based similarity searches, which represents molecules as bit strings that signal the presence or absence of predetermined structural characteristics, the energetic molecule is frequently compared to the library.

Structure primarily based strategies

The fact that active ligands are not necessary beforehand is one advantage of the structure-based drug creation techniques. It is possible to build novel ligands with therapeutic properties from a drug's three-dimensional structure. Structure-based techniques help the development of novel medicines since they help create and refine the original lead molecule. It has been a popular practise in virtual screening to combine ligand- and structure-based techniques because it has been hypothesised that doing so can increase the advantages and minimise the drawbacks of each strategy.[20.21]

Ligand-based methods

The idea of molecular similarity states that the effectiveness of a novel drug design is often demonstrated by pointing to the characteristics of existing ligands. Molecules with high structure similarity are more likely to have same hobby data. This technique, which is thought of as a backdoor strategy for drug discovery, can also be used when the three-dimensional shape of an objective is ambiguous or uncertain [22,23].

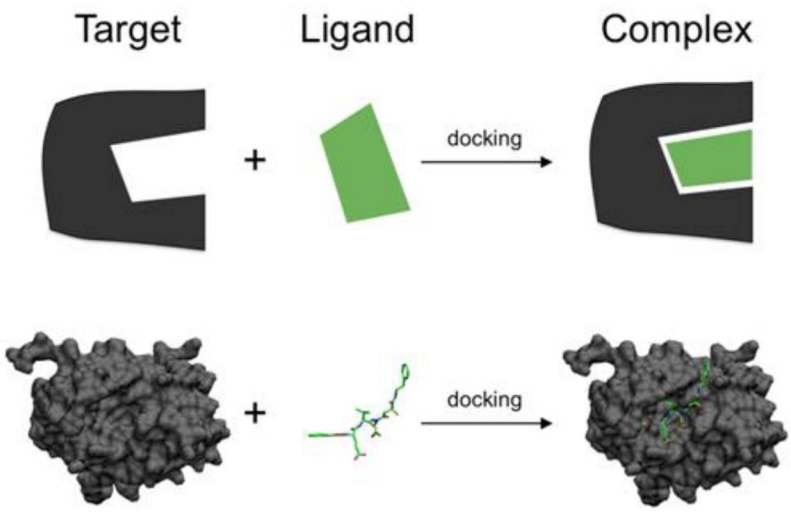

Figure 1.4 Ligand based method of molecular docking

Structure-primarily based techniques

In this technique, utilising the target as a mould, the interaction with any minuscule molecules present in the chemicals library is computationally simulated. Then, only people whose physiques looked better on the binding web page are chosen. The ligand's steric complementarity can be adjusted to boost the affinity for the receptor if the binding site is known. In fact, by targeting specific ligand portions that fit poorly in the complex's active site online, chemical modifications to decrease the active capacity by decreasing Vander Waals contacts can be hypothesised, increasing complementarity with the receptor. Similar to this, the chelate structure can be purposefully changed to enhance the receptor's electrostatic complementarity. [24,25].

Molecular Docking

Molecular docking is a well-liked method of molecular simulation that looks at how the ligand and target interact. The target and ligand's real interaction is virtually simulated by the docking technique. Additionally, it forecasts the optimum ligand orientation and shape within the binding site. Docking determines the preferred alignment of one small molecule to a target in order to produce a stable complex. There are various processes involved. Small molecules are first docked onto the target's active web page using docking methods. Three tasks are primarily carried out by docking programmes. Potential ligands are selected from a library of chemical compounds using docking applications. They can also predict the binding characteristics of known or potential ligands. Finally, these systems determine compounds that are significantly more likely to engage the therapeutic target by computing probable binding affinities based on the anticipated binding pose. Large chemical libraries have been effectively screened by docking algorithms, which have resulted in a more manageable subset that is rich in binders. In circumstances where there are real interactions, the anticipated ligand posture often matches well with empirically determined protein chelate complexes. Shape and based general approaches have been successful in discovering new drugs, despite the fact that binding posture prediction is one of its strengths [26].

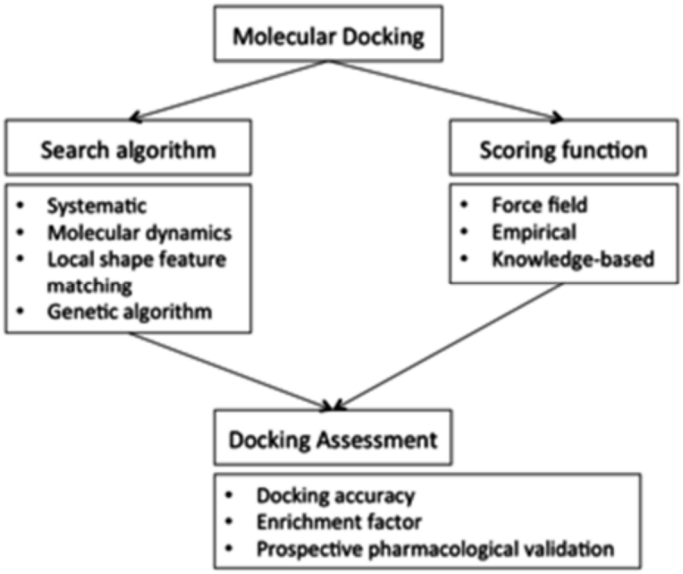

Figure 1.5 Molecular Docking methods

Structure -based (pharmacophore) modeling

According to the methodology of shape-based (pharmacophore) modelling, which comprises an assessment of the complementary chemical activities of the active website and their spatial interactions, a subsequent pharmacophore model is constructed with preset features. In this example, it is shown that a pharmacophore can be created fully (without ligand) or partially (with ligand) based on research into the target binding website (macromolecule). The fact has made it feasible to digitally screen huge libraries of chemicals using a pharmacophore at minimal cost and quickly [27].

Virtual screening

Virtual screening entails picking compounds using computer techniques from vast databases as opposed to direct physical screening them. Using this method, active part of compounds that might alter a certain molecular route can be promptly discovered. Virtual screening is far superior to physical screening in many ways. It takes much less time and assistance overall. Additionally, virtual screening can be used to evaluate drugs that aren't yet on the market; if any intriguing compounds are discovered, they can be obtained or created [28].

Successful aspect of computational drug discovery:

Rapid improvements in CADDD technologies and methods have accelerated drug development process by make it possible to quickly and economically screen and synthesise vast libraries of compounds. Many jobs related to the hunt for improved or innovative tablets have successfully used these techniques during the past few years. Examples of selected inhibitors that were developed using computational chemistry and logical drug design techniques [29].

APPLICATION

The statistical significance and prognostication of QSAR models determine their use. The molecule being predicted must fall inside the model's application domain in order for the QSAR prediction to be legitimate. The training set's molecular composition, along with the model descriptors and modelled response, make up the applicability domain, a theoretical region of the chemical universe. The leverage approach can be used to determine whether a novel chemical falls within the range of application. Interpolation from training sets using Jawors is another strategy. Stan Forth and colleagues' cluster-based approach

CONCLUSION

The QSAR models can be employed for predicting the actions of untested compounds, among other uses. Through molecular modelling, simulation, and virtual screening of various possibilities before synthesis, computer-aided technologies contribute in the rational design of drugs. The concept, a brief history, and all the components involved in given modeling were covered in this review article.

ACKNOWLEDGMENT

All the authors are thankful to IES University, Bhopal for their valuable support.

REFERENCES

1. Huang Y, Li X, Xu S, Zheng H, Zhang L, Chen J, Hong H, Kusko R and Li R (2020) Quantitative Structure–Activity Relationship Models for Predicting Inflammatory Potential of Metal Oxide Nanoparticles, Environmental Health Perspectives, 128:6
2. Rojas C, Tripaldi P, Pérez-González A, Duchowicz P and Pis Diez R (2018) A retention index-based QSPR model for the quality control of rice, Journal of Cereal Science, 10.1016/j.jcs.2017.11.004, (79):303-310
3. Mu Y, Wang Z, Wu F, Zhong B, Yang M, Sun F, Feng C, Jin X, Leung K and Giesy J (2018) Model for Predicting Toxicities of Metals and Metalloids in Coastal Marine Environments Worldwide, Environmental Science & Technology, 10.1021/acs.est.7b06654, 52(7): 4199-4206
4. Lo Y, Rensi S, Torng W and Altman R (2018) Machine learning in cheminformatics and drug discovery, Drug Discovery Today, 10.1016/j.drudis.2018.05.010
5. Saavedra L, Romanelli G and Duchowicz P (2018) Quantitative structure-activity relationship (QSAR) analysis of plant-derived compounds with larvicidal activity against Zika Aedes aegypti (Diptera: Culicidae) vector using freely available descriptors , Pest Management Science, 10.1002/ps.4850, 74(7):1608-1615
6. Gupta S and Mallick S (2018) Modelling the water–plant cuticular polymer matrix membrane partitioning of diverse chemicals in multiple plant species using the support vector machine-based QSAR approach, SAR and QSAR in Environmental Research, 10.1080/1062936X.2017.1419985, 29,(3):171-186
7. Bitam S, Hamadache M and Hanini S (2018) Prediction of therapeutic potency of tacrine derivatives as BuChE inhibitors from quantitative structure–activity relationship modelling, SAR and QSAR in Environmental Research, 10.1080/1062936X.2018.1423640, 29(3):213-230
8. Wei M, Yang X, Watson P, Yang F and Liu H (2018) Development of QSAR model for predicting the inclusion constants of organic chemicals with α-cyclodextrin, Environmental Science and Pollution Research, 10.1007/s11356-018-1917-2, 25 (18):17565-17574
9. Cheng Z, Yang B, Chen Q, Ji W and Shen Z (2018) Characteristics and difference of oxidation and coagulation mechanisms for the removal of organic compounds by quantum parameter analysis, Chemical Engineering 10.1016/j.cej.2017. 09.065, (332):351-360
10. Cohen J, Rice J and Lewandowski T (2017) Expanding the Toolbox: Hazard-Screening Methods and Tools for Identifying Safer Chemicals in Green Product Design, ACS Sustainable Chemistry & Engineering , 10.1021/acssuschemeng.7b03368, 6(2):1941-1950
11. Ojha P and Roy K (2018) Development of a robust and validated 2D-QSPR model for sweetness potency of diverse functional organic molecules, Food and Chemical Toxicology, 10.1016/j.fct.2017.03.043, (112):551-562
12. Cheng Z, Yang B, Chen Q, Shen Z and Yuan T (2018) Quantitative relationships between molecular parameters and reaction rate of organic chemicals in Fenton process in temperature range of 15.8 °C–60 °C, Chemical Engineering Journal, 10.1016/j.cej.2017.12.105, (350):534-540
13. Gajewicz A (2018) How to judge whether QSAR/read-across predictions can be trusted: a novel approach for establishing a model's applicability domain, Environmental Science: Nano, 10.1039/C7EN00774D, 5(2):408-421
14. An P and Fukusaki E (2018) Metabolomics: State-of-the-Art Technologies and Applications on Drosophila melanogaster Drosophila Models for Human Diseases, 10.1007/978-981-13-0529 (14):257-276
15. Basant N and Gupta S (2018) Multi-target QSPR modeling for simultaneous prediction of multiple gas-phase kinetic rate constants of diverse chemicals, Atmospheric Environment, 10.1016/j.atmosenv.2017.11.028 (177):166-174
16. Yang M, Kohler M, Heyder T, Forsslund H, Garberg H, Karimi R, Grunewald J, Berven F, Nyrén S, Magnus Sköld C and Wheelock Å (2018) Proteomic profiling of lung immune cells reveals dysregulation of phagocytotic pathways in female-dominated molecular COPD phenotype, Respiratory Research, 10.1186/s12931-017-0699-2, 19:1
17. Satpathy R (2018) Quantitative Structure-Activity Modelling of Toxic Compounds Nanotechnology, Food Security and Water Treatment, 10.1007/978-3-319-70166 (10):313-331
18. Rasti B and Shahangian S (2018) Proteochemometric modeling of the origin of thymidylate synthase inhibition, Chemical Biology & Drug Design, 10.1111/cbdd.13163, 91(5):1007-1016
19. Kell D and Oliver S (2014) How drugs get into cells: tested and testable predictions to help discriminate between transporter-mediated uptake and lipoidal bilayer diffusion, Frontiers in Pharmacology, 10.3389/fphar.2014.00231, 5
20. Duchowicz P (2018) Linear Regression QSAR Models for Polo-Like Kinase-1 Inhibitors, Cells, 10.3390/ cells7020013, 7:2, (13)
21. Yang M, Li X, Li Z, Ou Z, Liu M, Liu S, Li X, Yang S and Bajic V (2013) Gene Features Selection for Three-Class Disease Classification via Multiple Orthogonal Partial Least Square Discriminant Analysis and S-Plot Using Microarray Data, PLoS ONE, 10.1371/journal.pone.0084253, 8:12, (e84253)
22. Simeon S, Anuwongcharoen N, Shoombuatong W, Malik A, Prachayasittikul V, Wikberg J and Nantasenamat C (2016) Probing the origins of human acetylcholinesterase inhibition via QSAR modeling and molecular docking, PeerJ, 10.7717/peerj.2322, 4, (e2322)
23. Molnar M, Komar M, Brahmbhatt H, Babić J, Jokić S and Rastija V (2017) Deep Eutectic Solvents as Convenient Media for Synthesis of Novel Coumarinyl Schiff Bases and Their QSAR Studies, Molecules, 10.3390/molecules 22091482, 22:9, (1482
24. Carlsen L (2009) The Interplay between QSAR/QSPR Studiesand Partial Order Ranking and Formal Concept Analyses, International Journal of Molecular Sciences, 10.3390/ijms10041628, 10(4):1628-1657

25. Singh S and Supuran C (2015) In silico modeling of β -carbonic anhydrase inhibitors from the fungus Malassezia globosa as antidandruff agents , Journal of Enzyme Inhibition and Medicinal Chemistry, 10.3109/ 14756366.2015.1031127, (1-8)
26. Rath E, Gill H, Bai Y and Tang H (2017) Identification of potential antimicrobials against Salmonella typhimurium and Listeria monocytogenes using Quantitative Structure Activity Relationmodeling, PLOSONE, 10.1371/journal. pone.0189580, 12:12, (e0189580)
27. Zhang Y, Xia Z, Yan L and Liu S (2015) Prediction of Placental Barrier Permeability: A Model Based on Partial Least Squares Variable Selection Procedure, Molecules, 10.3390/molecules20058270, 20(5):8270-8286
28. Lankadurai B, Furdui V, Reiner E, Simpson A and Simpson M (2013). 1H NMR-Based Metabolomic Analysis of Sub-Lethal Perfluorooctane Sulfonate Exposure to the Earthworm, Eisenia fetida, in Soil Metabolites, 10. 3390/metabo3030718, 3(3)718-740
29. Shao L, Wu L, Fang H, Tong W, Fan X and Rapallo F (2010) Does Applicability Domain Existin Microarray Based Genomic Research?, PLoSONE, 10.1371/journal.pone.0011055, 5:6, (e11055)

CITATION OF THIS ARTICLE

Phool Singh Yaduwanshi, Sheelu Singh, Indra Kumar Sanodiya, Shivkant Patel, Manglesh Kumar, Samisha Sharma. A Mini Review on Descriptors Used in Qsar Study. Bull. Env.Pharmacol. Life Sci., Vol 12 [12] November 2023: 350-357