



Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.

Quantitative structure–activity relationship-based computational approaches

*Virupaksha Bastikar*¹, *Alpana Bastikar*² and *Pramodkumar Gupta*³

¹Amity Institute of Biotechnology, Amity University, Mumbai, Maharashtra, India

²Navin Saxena Research and Technology Pvt. Ltd, Gandhidham, Gujarat, India

³School of Biotechnology and Bioinformatics, D Y Patil Deemed to be University, Navi Mumbai, Maharashtra, India

10.1 Introduction

Quantitative structure–activity relationship (QSAR) could be a methodology to associate the chemical arrangement of a molecule with its biochemical, physical, pharmaceutical, biological, etc., effect. The exploitation of QSAR developed strategies can be done significantly in chemo computing, drug discovery and to calculate the biological activity of chemical compounds, but also additionally for pharmacological medicine and ecotoxicological assessments of individual chemicals among the risk management. QSAR models are developed for computational drug design, activity prediction, and toxicology predictions. QSAR is outlined as the quantitative correlation of biological activities with chemical science properties (Puzyn & Leszczynski, 2012).

Biological activity = f (physicochemical parameter)

QSAR studies have a very important application in modern chemistry and biochemistry. QSAR helps in finding the compounds with desired properties using chemical information and its association with biological activity. The physicochemical properties such as partition coefficient and presence or absence of certain chemical features are taken into consideration. QSAR attempts to correlate structural, chemical, statistical, and physical properties with biological potency using various mathematical methods. The generated QSAR models are used to predict and classify the biological activities of new chemical compounds. QSAR guides the process of lead optimization and is also used as a screening

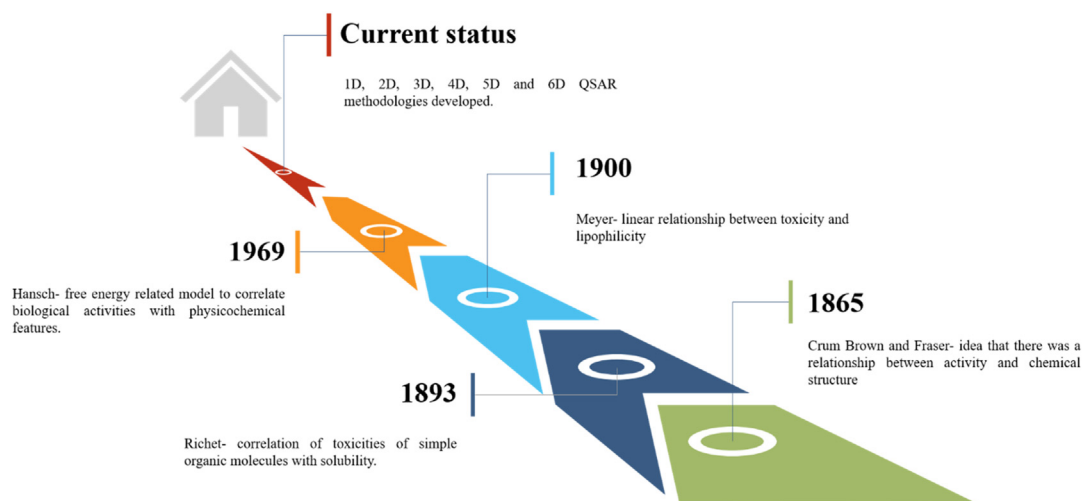


FIGURE 10.1 History of quantitative structure–activity relationship.

and enrichment tool to remove the compounds and molecules that do not possess drug-likeness properties or are predicted toxic (Gajewicz et al., 2012) (Fig. 10.1).

10.2 The importance of quantitative structure–activity relationship

The motivation behind developing in silico QSAR models examines, and incorporates the following points:

1. To foresee natural action of the compounds and understand physical-substance properties by mathematical methods. The natural activity of the compounds can be studied and predicted by the development of the QSAR models for many drug classes.
2. To comprehend and rationalize the mechanisms of action within a series of chemicals. By developing a QSAR model using these fixed mechanisms of actions for a series of molecules the activity of unknown molecules can be predicted. A group of similar molecules generally exhibit a similar type of activity and give activity in a particular fixed range. Thus any new molecule that is developed that belongs to a similar class, its activity can also be predicted and a QSAR mathematical model helps to improve its activity and design new molecules.
3. Savings in the expense of compound advancement (e.g., in the drug, pesticide) in terms of synthesis and manufacturing of the molecule as well as in vitro and in vivo testing of the molecule. Once mathematically it is proven that a given set of newly designed molecules gives a better activity, only those can be taken forward for synthesis, rejecting the others that are not predicted as having good activity. Thus the cost of synthesis and time required for the entire study is comparatively reduced as against the traditional method of drug design.

- Predictions could lessen the prerequisite for extensive and costly animal tests thereby avoiding ethical issues and concerns. Every time sacrificing an animal just to check whether a given novel molecule shows activity or not is overall not feasible, both in terms of cost, time, and also ethically. QSAR helps to avoid unnecessary testing of animals for the novel molecules.
- Other spaces of advancing green and greener science to expand productivity and eliminate waste by not following leads unlikely to be successful. Those molecules that are going to be harmful to the environment can be avoided to be synthesized based on the results of the QSAR (Aptula & Roberts, 2006).

10.3 Requirements to generate a good quantitative structure–activity relationship model

Based on the above (Fig. 10.2) a QSAR model requires the following tools-

- A set of molecules to be used for generating the QSAR model: A dataset consisting of molecules, structurally similar, whose QSAR model needs to be developed are to be prepared for the QSAR study. Depending upon the type of QSAR the molecules need to be minimized or cleaned.
- A set of molecular descriptors generated for the data set of molecules: Once the molecules are finalized, the parameters of the molecules known as the descriptors are calculated, which can be the overall structural properties of the molecules, two-dimensional properties of the molecules, three-dimensional properties of the molecule in space, or the different conformational properties of the molecules.
- Biological activity (IC₅₀, EC₅₀, etc.) of the set of molecules: The molecules whose QSAR model is to be developed should have a definite and known biological activity value that can be correlated with the molecular descriptors generated, to develop a good and reliable QSAR model.

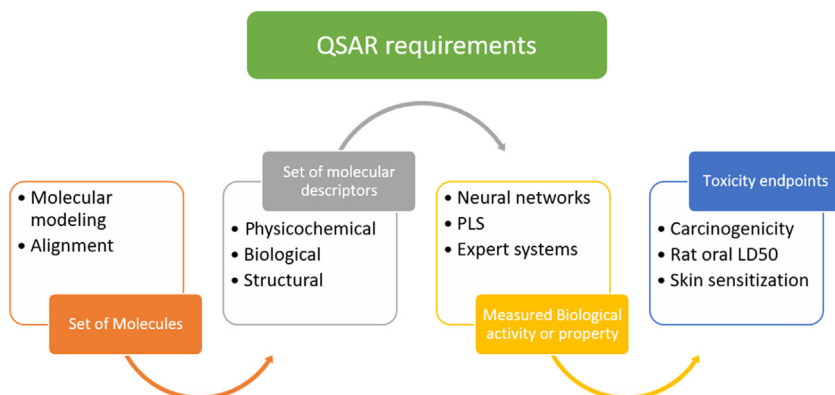


FIGURE 10.2 Quantitative structure–activity relationship.

4. Statistical methods to develop a QSAR model: Various statistical methods like clustering, partial least square, regression, principal component analysis (PCA), etc., can be used to develop a mathematical correlation between the biological activity and the descriptors calculated.
5. The QSAR model thus generated is validated and if found to be full-proof is used further to predict the activity of any unknown compound belonging to the same class of molecules as the data set in terms of the same disease, the same type of biological activity, same scaffold, same pharmacophore, etc.

10.4 Applications of quantitative structure–activity relationship in various fields

The capacity to foresee an organic movement is important in quite a few ventures. While some QSARs give off an impression of being minimal more than scholarly examinations, there are countless uses of these models inside the industry, the scholarly world, and (administrative) offices (OECD, 2007). Few potential uses are recorded beneath:

1. *Chemical*: One of the primary authentic applications is to anticipate limits. It is notable, for example, that inside a specific group of substance compounds, particularly of natural science, these are solid connections amongst the construction of the molecule and its noticed properties. A basic model is a connection between the quantity of carbon in alkanes and their limit. There is an unmistakable pattern in the increment of the edge of boiling over with an increment in the carbon, and this serves as a method for foreseeing the edge of boiling over of higher alkanes. Thus this chemical property can be exploited by generating a QSAR model of the said property and predicting alkanes based on their boiling points.
2. *Natural*: The organic action of a particle is normally estimated in order to set up the degree of the hindrance of specific signal transduction or metabolic pathway. Medication disclosure frequently includes the utilization of QSAR to recognize synthetic design that could have a great inhibitory impact on the said protein target. A set of organic molecules can be tested against a particular protein or enzyme target to study their effect on the metabolic pathway involved. A QSAR model developed is definitely useful to study the mechanism of action of the drugs on the metabolic pathway.
3. The QSAR model gives a sensible distinguishing proof of new leads with pharmacological, biocidal, or pesticide activities.
4. The QSAR model deals with the enhancement of pharmacological, biocidal, or pesticide activities.
5. The QSAR model allows distinguishing toxic compounds at the beginning phases of ligand improvement or the screening of various databases of existing compounds.
6. The QSAR model forecasts the poisonousness to natural species. The choice of mixtures with ideal pharmacokinetic properties, regardless of whether they be synthesized or accessible in organic frameworks can be given.
7. The forecast of an assortment of physicochemical properties of atoms (whether they be drugs, pesticides, individual items, fine synthetic substances).

Characteristic features of a good QSAR model (Todeschini & Consonni, 2000):

1. A defined endpoint: Every QSAR model should be developed for a specific endpoint, for example, biological activity, toxicity, skin Sensitization, mutagenetic, etc., which should be specified at the beginning of the model prediction.
2. An unambiguous algorithm: An algorithm or mathematical model which can predict the given defined endpoint and not give any other vague result.
3. A defined domain of applicability: The Physicochemical, structural or biological space, data, or information on which the training set of the model has been established, and for which it is applicable to make calculations for new compounds.
4. An appropriate measure of goodness of fit: The goodness of fit of a statistical mathematical model describes how well it fits a set of observations. Measures of goodness of fit classically encapsulate the inconsistency between observed values and the values expected under the model developed.

10.5 The different stages of advancement of quantitative structure–activity relationship

1. One-dimensional QSAR: This is the first type of QSAR model to be developed that correlates the pKa (dissociation constant) and log P (partition coefficient). This takes into account the overall structure and its pKa and logs P correlation.
2. Two-dimensional QSAR: The biological Activity correlates to the overall structure pattern of drug molecules. It takes into account the entire structure of the molecule in two-dimensional space. Various parameters of the structure of the molecule are calculated and correlated to the biological activity. For example, no hydrogen bonds, molecular refractivity, topological indices, dipole moment, etc.
3. Three-dimensional QSAR: The biological Activity correlates with the three-dimensional structure of the molecule and its properties. It takes into account the molecule in its three-dimensional space. The different parameters like a steric hindrance, h-bond acceptors, h-bond donors, hydrophobic interactions are a part of three-dimensional QSAR.
4. Four-dimensional QSAR: It is the same as three-dimensional QSAR along with multiple representations of ligand conformations. It takes into account the different conformations of the ligand molecule in space. It studies how the ligand can be placed in different conformations in the space and what are the changes in the three-dimensional parameters based on the conformational changes. Based on the changes in the parameter values different QSAR models are developed.
5. Five-dimensional QSAR: Same As that with the four-dimensional along with multiple representations of ligands in the docked complexes. It takes into account the ligand–receptor binding and the different conformations of ligand in the docked complex three-dimensional space. It studies the different conformations of the ligand however now it includes the receptor binding interactions of the ligands. The different conformations are based on the changes in the docked complexes of the ligands and receptors.
6. Six-dimensional QSAR: Same As with five-dimensional along with multiple representations of molecular dynamic studies of the receptor–ligand complexes. Along with the different

conformations of the ligands in the receptor–ligand complexes, this QSAR also takes into account the changes occurring in the stability of the complex during the molecular dynamics simulations. The energy calculated for different ligand conformations at different time intervals forms the basis of the development of this QSAR.

10.5.1 Steps and strategies for quantitative structure–activity relationship

QSAR modeling process consists of five main steps (Ekins, 2007):

1. Begins with the selection of molecules to be used: Preparation of dataset—it consists of a set of molecules against which the QSAR model is to be prepared.
2. Selection of descriptor; numerical represented of molecular feature (e.g., no. of carbon): Various parameters of the dataset are generated that can be correlated with the biological activity of the dataset molecules.
3. The original descriptor pool must be reduced in size: Screening of the generated descriptors to keep only the relevant directly linked to the biological activity.
4. Model building: Using statistical methods a mathematical model is built correlating the screened descriptors with the biological activity.
5. The reliability of the model should be tested: The prediction capacity of the model is checked on a given set of test compounds.

10.6 Molecular descriptors

Atomic descriptors are a mathematical portrayal of compound data present inside a particle. This numerical portrayal must be invariant to the particle's size and the number of atoms for building a model with measurable methodologies (Tropsha et al., 2003). The three significant kinds of boundaries and related descriptors are given in Fig. 10.3. The data about atomic descriptors relies upon two central points:

1. The molecular representation of compounds.
2. The algorithm used for the calculation of the descriptor.

10.7 Methods of quantitative structure–activity relationship

A wide range of ways to deal with QSAR has been created since Hansch's fundamental works. QSAR strategies can be investigated from two perspectives (Gramatica, 2007):

1. The sorts of underlying boundaries that are utilized to describe subatomic personalities begin from the various portrayals of particles, from basic synthetic equations to 3D conformities.
2. The numerical system is utilized to acquire the quantitative connection between these primary boundaries and organic action. The figure clarifies the technique of QSAR utilized for any broad QSAR type. Constructions are divided to build up their pertinent descriptor properties. With the assistance of different numerical investigation devices,

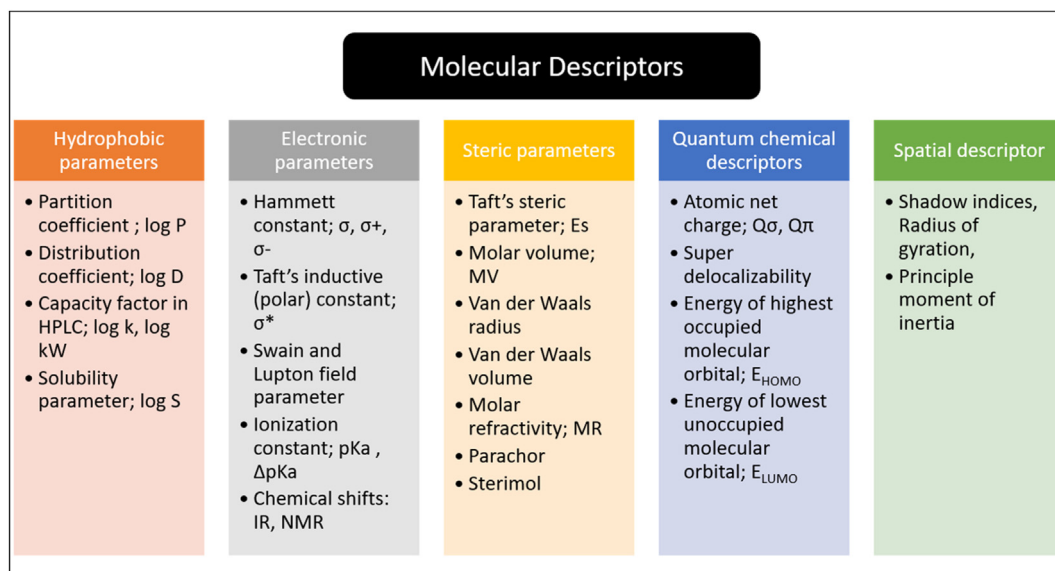


FIGURE 10.3 Molecular descriptors for quantitative structure–activity relationship.

the information is prepared to set up a numerical QSAR model, which will associate with the natural movement. The model created is approved by different approval techniques and tried for outside expectations. At long last, a powerful QSAR model is set up that considers the pertinent boundaries for the natural action for the given arrangement of mixtures.

A model medication applicant is required to have unmistakable properties, that is, compound properties, solvency, enzymatic soundness, penetration across natural layers, low leeway by the liver or kidney, strength, and wellbeing. Out of various accessible descriptors, the choice of the central atomic descriptors is the main test in a QSAR. Subsequently, to comprehend the QSAR model, to diminish overfitting, speed up preparation, and to improve the general model consistency, the decision of suitable and interpretable descriptors to set up QSAR models is a very pivotal advance.

10.8 Data analysis methods

10.8.1 Free Wilson analysis

It is a structure–activity evaluation technique that considers the contribution of diverse structural fragments to the general organic activity. Indicator variables outline the presence or absence of a specific structural characteristic in a molecule. This mathematical model considers the symmetry equation to limit linear dependency between variables (Fig. 10.4) (Puzyn & Leszczynski, 2012).

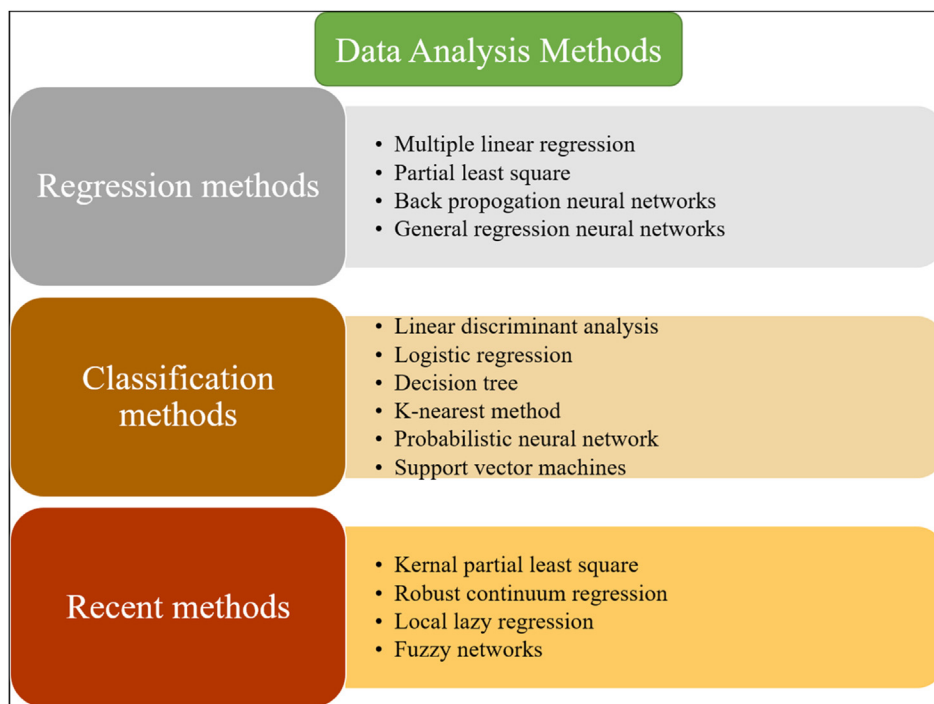


FIGURE 10.4 Data analysis methods.

10.8.2 Statistical methods

Statistical techniques offer the premise for the improvement of QSAR evaluation. The software of multivariate evaluation, data description, classification, and regression evaluation are used for interpretation and theoretical prediction of organic features for new compounds (Puzyn & Leszczynski, 2012).

10.8.3 Discriminant analysis

Discriminant evaluation is used to split molecules into their constituent classes. It reveals a linear mixture of things that high-quality discriminates among one-of-a-kind constituents classes. This approach is used for the evaluation in preference to a couple of linear regressions because the organic interest information isn't on a nonstop scale of interest however labeled as lively and inactive (Puzyn & Leszczynski, 2012). It is used to symbolize a quantitative courting among molecular descriptors and the organic property.

10.8.4 Cluster analysis

Clustering is the manner of dividing a set of devices into agencies in order that every cluster includes distinctly comparable gadgets, and items in a single cluster are dissimilar

gadgets of different clusters. When cluster evaluation is implemented on a compound dataset, the range of clusters affords records approximately the range of structural kinds found in a compound set. A numerous subset of compounds may be prepared with the aid of using taking one or extra compounds from every cluster (Puzyn & Leszczynski, 2012). It is implemented to pattern numerous subsets of compounds from a bigger compound dataset. Hierarchical clustering, k-way clustering, and nonhierarchical clustering are the techniques used for compound clustering.

10.8.5 Principal component analysis

The quantity of variables used to explain an item is referred to as dimensionality. PCA is used to lessen the dimensionality of the statistics set while a huge correlation exists among a few or all the variables (descriptors). PCA gives facts approximately the huge essential additives and represents most facts on impartial variables.

10.8.6 Quantum mechanical methods

Quantum mechanical strategies are used to understand correct molecular identities such as electrostatic capacity or polarizabilities, ionization capacity or electron affinities, etc. This approach is implemented to QSAR via way of means of the direct derivation of digital descriptors from the molecular wave function.

10.9 Quantitative structure–activity relationship model validation

After a QSAR model is developed it is necessary to validate it for its accuracy and predictively as well as precision (Fig. 10.5) (Veerasingam et al., 2011).

After the model validation, the model applicability domain needs to be checked, where the outliers will be thrown out, during model building (Gramatica, 2007) (Fig. 10.6).

10.10 Quantitative structure–activity relationship and Coronavirus disease-2019

The emission of COVID-19 has borne contrarily on populations' day-by-day lives. Indeed, it has undermined their wellbeing genuinely, intellectually, and mentally and hampered social and monetary improvement. Individuals during the time of isolation are experiencing plenty of burdensome manifestations because of numerous reasons among which the absence of actual work and dread are the most well-known ones. Researchers and analysts are dashing to bring a way forward and to discover immunizations or medications against COVID-19. By the by, there is no particular medication that has been accounted for in light of the fact that the creation of an effectual and solid medication requires quite a while of examination and clinical preliminaries. Subsequently, drug

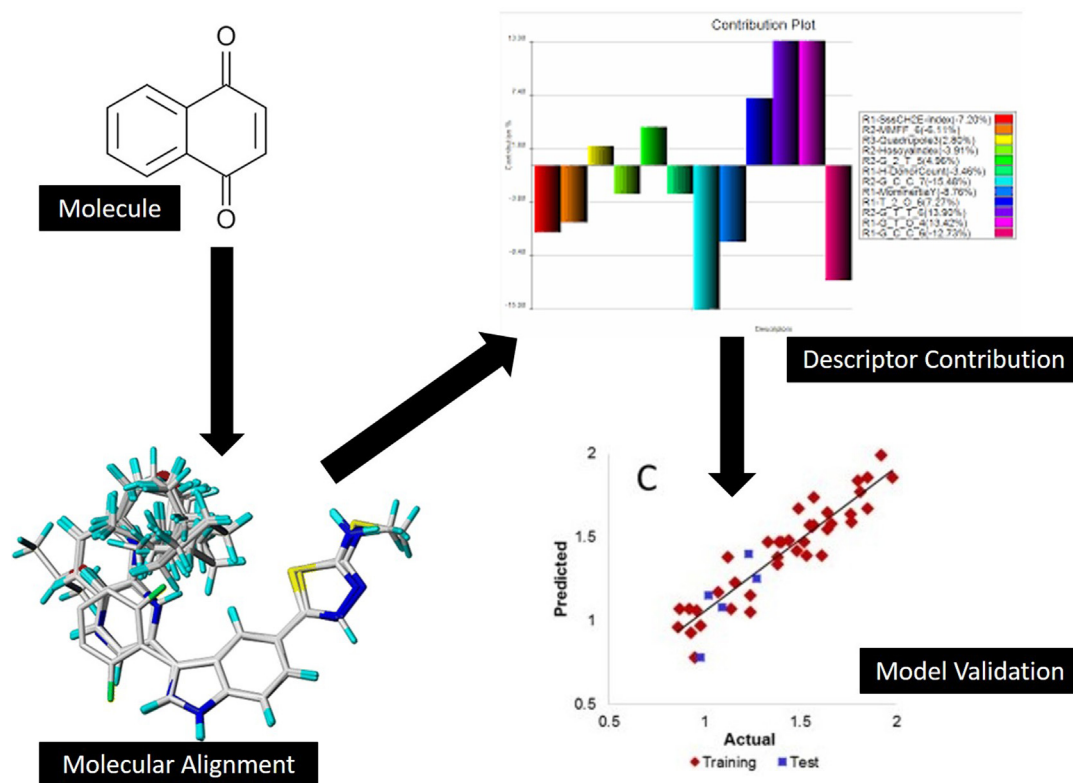


FIGURE 10.5 Model validation.

repositioning has been a methodology embraced by a majority of specialists worldwide to look for viable treatment in a brief timeframe (Tandon et al., 2019) (Fig. 10.7).

There have been various studies like docking analysis, molecular modeling, and simulations to develop new drugs against COVID-19. Many researchers are focusing on the repurposing of drugs as a potential treatment against COVID-19. To that effect, various computational techniques have been used to assist the development of molecules. Various QSAR studies have been reported that are used to develop leads and hits for COVID-19. Some studies have been reported below.

Sulfonamides are organically dynamic compounds since they are of essential significance. There are numerous sulfonamide drugs in the business sector for treating infections of various nature. Sulfonamide subsidiaries, for example, methazolamide, dichlorphenamide, ethoxzolamide, acetazolamide, and dorzolamide have been clinically wagered on for quite a long time as inhibitors of the zinc catalyst carbonic anhydrase. On account of their moderateness and minimal expense, they are intensely utilized as veterinary antimicrobials in many parts of the world, particularly in Asia, a few regions of Europe, and many rising nations. Sulfonamide subordinators are a significant moiety of various scopes of bioactive molecules and drug particles like antibacterial, anticancer, antitumor, and

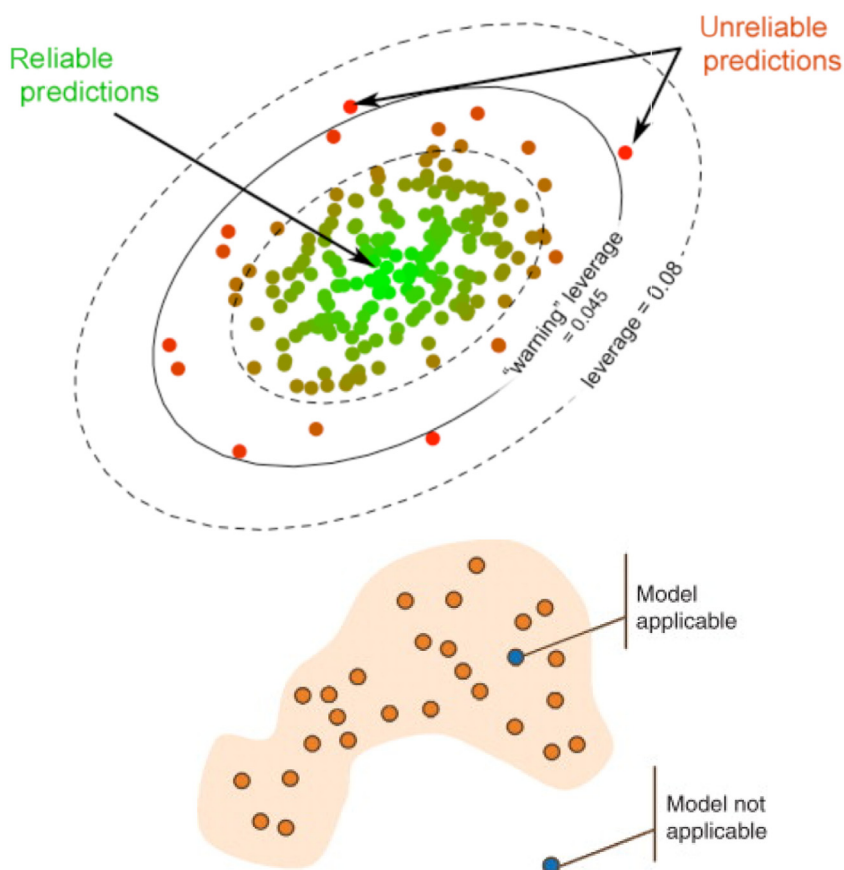


FIGURE 10.6 Model applicability domain.

antimalarial. Inferable from the general medical problem and absence of a powerful fix, numerous nations are settling on Chloroquine as an antimalarial drug for the therapy of COVID-19. Thusly, it has become critical to attempt to find new medications that can be more believable and compelling without having any destructive results than the Chloroquine used to fix the new pandemic. With that in mind, a bunch of eighteen carboxamides sulfonamide analogs, present antimalarial action were examined utilizing both CoMFA and CoMSIA approaches which are a type of three-dimensional QSAR modeling. Moreover, subatomic docking reproduction was accomplished to investigate the binding between SARS-CoV-2 primary protease and carboxamides sulfonamide compounds. In this examination, the antimalarial action and synthetic designs of 18 carboxamides sulfonamide subsidiaries were taken from the literature. These particles were considered to direct the three-dimensional QSAR examination by parting the information base into two datasets; a preparation set of 14 atoms to foster the quantitative model and a test set of four compounds to affirm the capability of the former model (Khaldan et al., 2021). The following figure demonstrates the SAR established with the help of the developed QSAR model (Fig. 10.8).

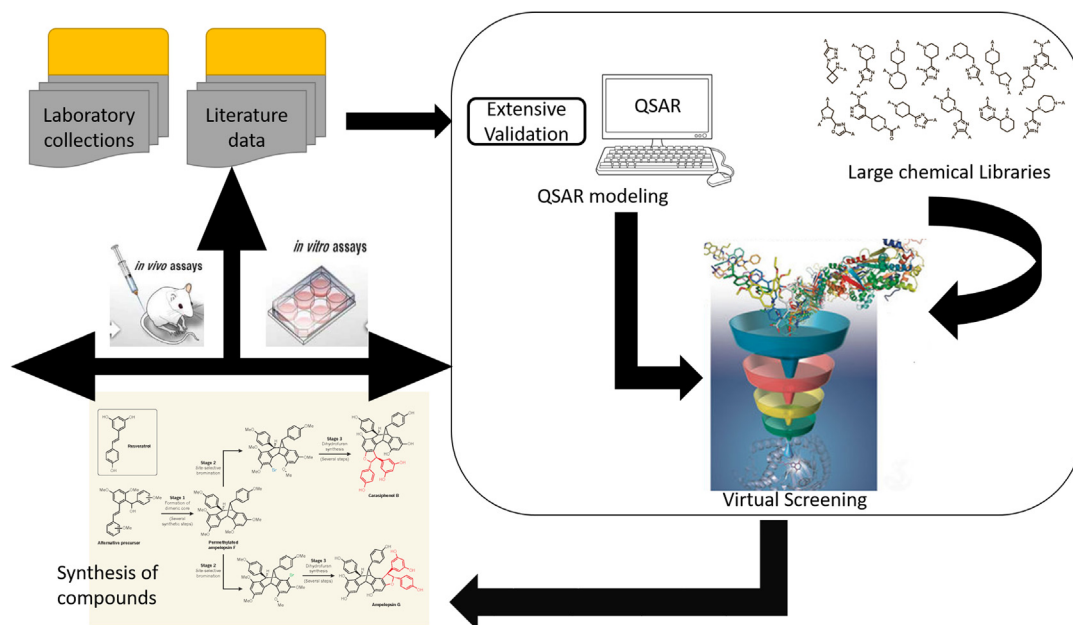


FIGURE 10.7 Quantitative structure–activity relationship and drug design.

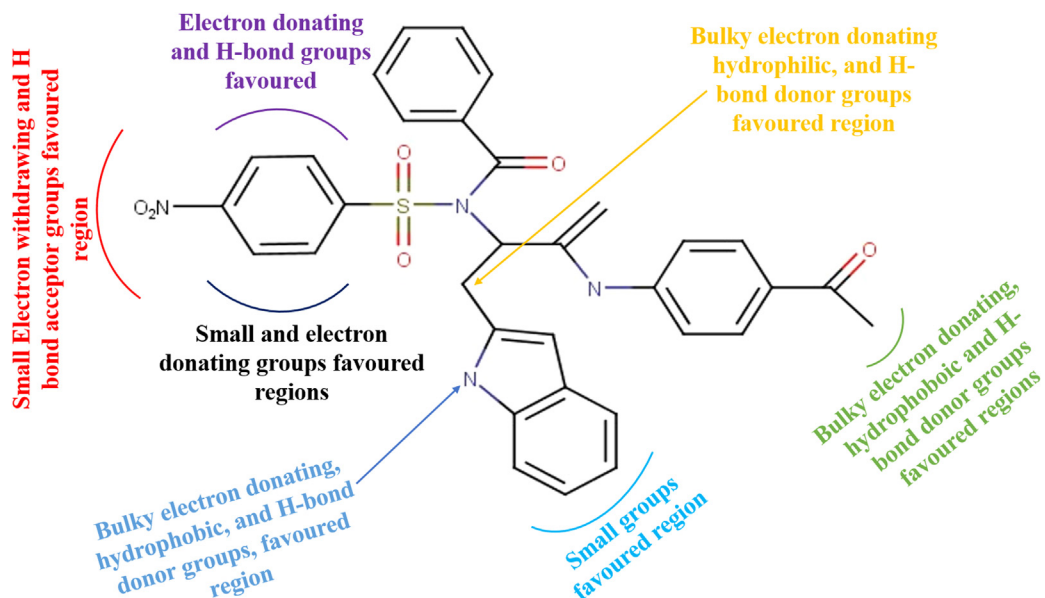


FIGURE 10.8 Severe acute respiratory syndrome from quantitative structure–activity relationship.

In the point of finding new powerful medications against COVID-19, the three-dimensional QSAR and subatomic docking considers were applied on a progression of eighteen carboxamides sulfonamide subordinates. The ideal CoMFA and CoMSIA models unveiled great factual results as far as a few thorough measurable keys, like Q2, R2, and R2test, thereupon, these models can be proficiently upheld to anticipate new molecules with significant activity. The shape maps created by CoMFA and CoMSIA models, uncover the significant destinations where steric, electrostatic, and hydrophobic collaborations may essentially be impacting (increment or lessening) the action of the particles. These form maps guided to propose eight atoms with significant inhibitory movement (Ivanov et al., 2020).

In one more examination, researchers curated more than 1000 inhibitors with structure – bioactivity information as preparing atoms for 3CLpro and RdRp protein targets. They gathered this information from the most current SARS-CoV-2 bioassay concentrates just as existing investigations with SARS-CoV-1, MERS-CoV, and other related infections in the CAS information assortment. Utilizing this information, they applied an assortment of AI calculations to assemble a few dozen QSAR models selecting from among these, the most grounded performing models one focusing on 3CLpro and one focusing on RdRp (Amin et al., 2020).

The subsequent models were utilized to screen 1087 FDA-endorsed drugs, almost 50,000 substances from the CAS COVID-19 Antiviral Candidate Compounds Dataset, a rundown of 113,000 substances with CAS-appointed pharmacological action or a helpful job filed in SARS, MERS, and COVID-19-related records distributed since 2003. Some anticipated atoms of these models were approved by distributed bioassay considers and clinical preliminaries as a positive sign of the prescient models. The model was then likewise applied to the CAS COVID-19 Antiviral Candidate Compounds Dataset, which contains 49,437 mixtures with potential antiviral movement recognized by CAS researchers. The model anticipated that 970 of these substance compounds are probably going to be dynamic against 3CLpro of the Covid. From every one of these applications, a couple of chosen atoms with the most elevated hindrance likelihood. True to form, the model recognized a few notable HIV-1 protease inhibitors (ritonavir and lopinavir) and distinguished substances (RNs 2243743–58-8, 1934276–50-2, and 2229818–46-4) that objective 3C protease/3CLpro and was appeared to hinder Enterovirus, MERS-CoV, and SARS-CoV-1 when tried in bioassays. These could address new lead applicants as helpful specialists for COVID-19 or other viral diseases. The model additionally recognized substances against have proteins engaged with cell measures, including diltiazem hydrochloride and leflunomide. Leflunomide is a dihydroorotate dehydrogenase inhibitor and is associated with nucleotide amalgamation (Rafi et al., 2020).

The investigation configuration was made out of two significant angles (1) Ligand-based methodologies: (A) grouping QSAR-based information mining of different SARS-CoV Papin-like protease (PLpro) inhibitors, (B) QSAR-based virtual screening (VS) to distinguish in-house particles that could be viable against putative objective SARS-CoV PLpro and (2) Structure-based methodologies: at long last approval of hits through receptor—ligand association examination. Subsequently, this investigation presented key ideas, set up for particle ID and QSAR-based screening of in-house atoms dynamic against putative SARS-CoV-2 PLpro chemical. Here, a model was developed which was an order-based QSAR model that could be utilized as a device for foreseeing new atoms and additionally VS. The model created by Monte Carlo advancement-based QSAR was trailed by VS of some in-house synthetic compounds. At that point, ADME information-driven screening was performed by SwissADME and distinguished

mixtures with great medication resemblance. At long last, atomic docking investigation of QSAR inferred virtual hits was performed to build the trust in the last theories. The subatomic docking study performed against putative objective SARS-CoV-2 PLpro recommended the probability of these researched in-house particles. Hence, it tends to be inferred that the in-house particles can possibly use as a seed for drug plan and enhancement against SARS-CoV-2 PLpro. After broad in vitro and in vivo considers, these in-house VS hits might arise as helpful alternatives for COVID-19. This investigation may likewise propel restorative physicists to plan comparative kinds of mixtures in desires to trigger natural power just as viability without gathering poison levels (Płonka et al., 2020; Tejera et al., 2020).

10.11 Conclusion

COVID-19 has been creating havoc throughout the world. Scientists and researchers are emerged in developing vaccines and medicines against the virus. Various techniques like drug repurposing and high throughput screening are used to develop medicines for the immediate treatment of SARS-COV-2. QSAR is a computational methodology that has been used for ages for the screening of molecules by developing mathematical models to predict the activity of unknown lead compounds. The same technique has been used for the development of mathematical models in the treatment of COVID-19 to develop hits for the treatment of patients suffering from COVID-19. This gives hope that by using computational techniques more molecules can be developed against the pandemic.

References

- Amin, S. A., Ghosh, K., Gayen, S., & Jha, T. (2020). Chemical-informatics approach to COVID-19 drug discovery: Monte Carlo based QSAR, virtual screening and molecular docking study of some in-house molecules as papain-like protease (PLpro) inhibitors. *Journal of Biomolecular Structure and Dynamics*, 39, 4764–4773. Available from <https://doi.org/10.1080/07391102.2020.1780946>.
- Aptula, A. O., & Roberts, D. W. (2006). Mechanistic applicability domains for nonanimal-based prediction of toxicological end points: General principles and application to reactive toxicity. *Chemical Research in Toxicology*, 19(8), 1097–1105. Available from <https://doi.org/10.1021/tx0601004>.
- Ekins, S. (2007). *Computational toxicology: Risk assessment for pharmaceutical and environmental chemicals*. New Jersey: Wiley-Interscience.
- Gajewicz, A., Rasulev, B., Dinadayalane, T. C., Urbaszek, P., Puzyn, T., Leszczynska, D., & Leszczynski, J. (2012). Advancing risk assessment of engineered nanomaterials: Application of computational approaches. *Advanced Drug Delivery Reviews*, 64(15), 1663–1693. Available from <https://doi.org/10.1016/j.addr.2012.05.014>.
- Gramatica, P. (2007). Principles of QSAR models validation: Internal and external. *QSAR and Combinatorial Science*, 26(5), 694–701. Available from <https://doi.org/10.1002/qsar.200610151>.
- Ivanov, J., Polshakov, D., Kato-Weinstein, J., Zhou, Q., Li, Y., Granet, R., Garner, L., Deng, Y., Liu, C., Albaiu, D., Wilson, J., & Aultman, C. (2020). Quantitative structure–activity relationship machine learning models and their applications for identifying viral 3CLpro- and RdRp-targeting compounds as potential therapeutics for COVID-19 and related viral infections. *ACS Omega*, 5(42), 27344–27358. Available from <https://doi.org/10.1021/acsomega.0c03682>.
- Khaldan, A., Bouamrane, S., En-Nahli, F., El-mernissi, R., El khatabi, K., Hmamouchi, R., Maghat, H., Ajana, M. A., Sbai, A., Bouachrine, M., & Lakhlifi, T. (2021). Prediction of potential inhibitors of SARS-CoV-2 using 3D-QSAR, molecular docking modeling and ADMET properties. *Heliyon*, 7(3), e06603. Available from <https://doi.org/10.1016/j.heliyon.2021.e06603>.

- OECD (2007). *Guidance document on the validation of (quantitative) structure–activity relationships models*. Organisation for Economic Co-operation and Development.
- Plonka, W., Paneth, A., & Paneth, P. (2020). Docking and QSAR of aminothioureas at the SARS-CoV-2 S-protein-human ACE2 receptor interface. *Molecules*, 25(20), 4645. Available from <https://doi.org/10.3390/molecules25204645>.
- Puzyn, T., & Leszczynski, J. (2012). *Towards efficient designing of safe nanomaterials: Innovative merge of computational approaches and experimental techniques*. The Royal Society of Chemistry.
- Rafi, M. O., Bhattacharje, G., Al-Khafaji, K., Taskin- Tok, T., Alfasane, M. A., Das, A. K., Parvez, M. A. K., & Rahman, M. S. (2020). Combination of QSAR, molecular docking, molecular dynamic simulation and MM-PBSA: analogues of lopinavir and favipiravir as potential drug candidates against COVID-19. *Journal of Biomolecular Structure and Dynamics*, 1–20. Available from <https://doi.org/10.1080/07391102.2020.1850355>.
- Tandon, H., Chakraborty, T., & Suhag, V. (2019). A concise review on the significance of QSAR in drug design. *Chemical and Biomolecular Engineering*, 45. Available from <https://doi.org/10.11648/j.cbe.20190404.11>.
- Tejera, E., Munteanu, C. R., López-Cortés, A., Cabrera-Andrade, A., & Pérez-Castillo, Y. (2020). Drugs repurposing using QSAR, docking and molecular dynamics for possible inhibitors of the SARS-CoV-2 Mpro protease. *Molecules*, 25(21), 5172. Available from <https://doi.org/10.3390/molecules25215172>.
- Todeschini, R., & Consonni, V. (2000). *Handbook of molecular descriptors*. Weinheim: Wiley-VCH.
- Tropsha, A., Gramatica, P., & Gombar, V.K. (2003). The importance of being earnest: Validation is the absolute essential for successful application and interpretation of QSPR models. In *QSAR and combinatorial science* (22, Issue 1, pp. 69–77). Wiley-VCH Verlag. <<https://doi.org/10.1002/qsar.200390007>>.
- Veerasamy, R., Harish, J., Avijeet, S., Shalini., Christopher, P., & Agrawal. (2011). Validation of QSAR models—Strategies and importance. *International Journal of Drug Design and Discovery*, 2, 511–519.