CHAPTER 2

Quantitative Structure-activity Relationship (QSAR) in Studying the Biologically Active Molecules

Serap ÇETINKAYA[1], Burak TÜZÜN[2,*] and **Emin SARIPINAR[3]**

[1] *Department of Molecular Biology and Genetics, Science Faculty, Sivas Cumhuriyet University, Sivas, Turkey*

[2] *Plant and Animal Production Department, Technical Sciences Vocational School of Sivas, Sivas Cumhuriyet University, Sivas, Turkey*

[3] *Department of Chemistry, Faculty of Science, Erciyes University, Melikgazi, 38039, Kayseri, Turkey*

Abstract: Recently, many new methods have been used in the research and development of a new drug. In this article, QSAR, which is one of the usable areas of artificial intelligence during molecule research, and the analysis and formulation studies related to the suitability of this area are discussed. It is explained how a model to be created is prepared and calculation formulas for how to verify this model are shown. Examples of the most recent 4D-QSAR calculations are given.

Keywords: Molecular Modelling, Pharmacophore, QSAR, Quantitative Structure-activity Relationship, Validation.

INTRODUCTION

Quantitative structure-activity relationship (QSAR) analysis uses the molecular structure of a compound or ligand to predict its biological activity. It presupposes that similar biological activities are retained in similar molecular structures [1]. It also uses known biological activity data to predict unknown activities. This approach has been adapted to diverse but related scientific disciplines [2-5], including the design of new chemical entities (NCEs) [5, 6] with high biological potentials.

* **Corresponding author Burak TÜZÜN:** Plant and Animal Production Department, Technical Sciences Vocational School of Sivas, Sivas Cumhuriyet University, Sivas, Turkey; E-mail: theburaktuzun@yahoo.com

Igor José dos Santos Nascimento (Ed.)

QSAR is a systematic multi-step process (Fig. **1**), made up of dataset preparation, selection, and generation of molecular descriptors; derivation of mathematical or statistical models; model training and validation using a training dataset; and model testing on a test dataset [7 - 10].

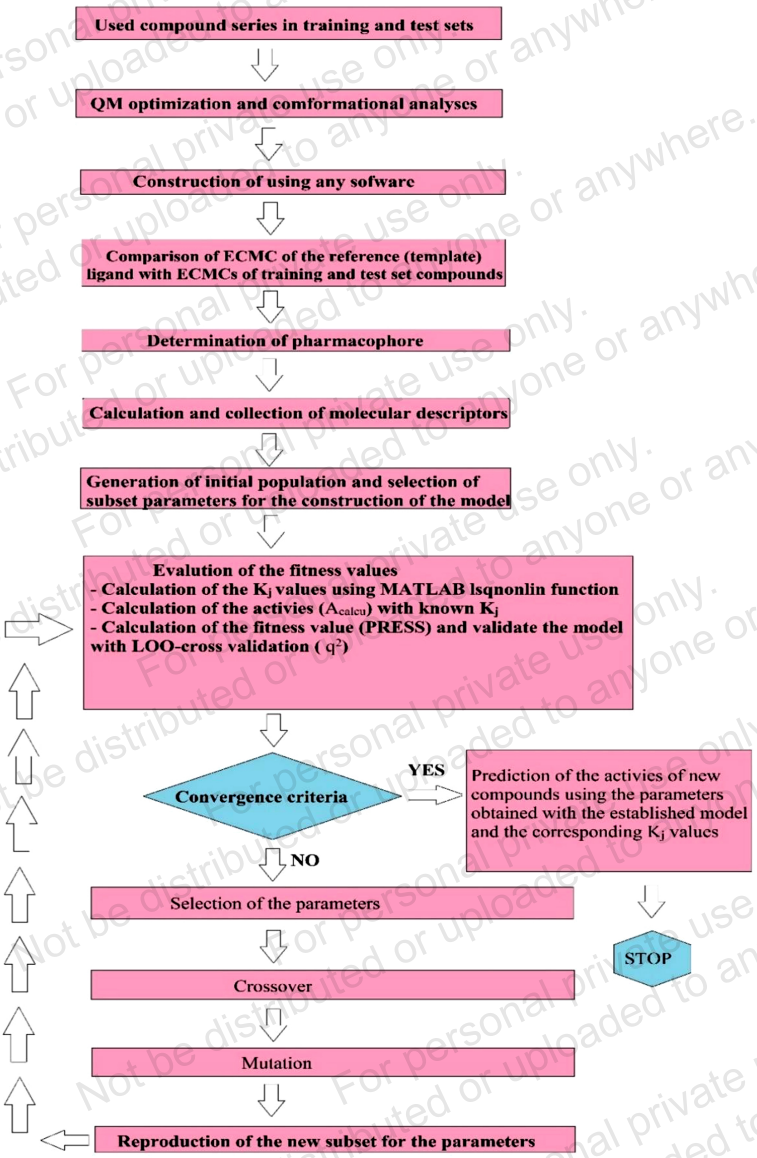

Fig. (1). Outlines of a QSAR model development.

In order to create a consistent QSAR model, it is central to utilize high-quality data that have been derived from bioassays, and to use an adequate number of compounds. Biological data are preferred to have been produced in a single laboratory [5, 11].

Selection and generation of molecular identifiers form the second step. Here the selection of appropriate descriptors, describing structural variations, is important. Various methods, such as machine learning techniques (*e.g.*, forward selection) and evolutionary algorithms (*e.g.*, genetic algorithm [11]), are utilized for descriptor/variable/feature selection.

A suitable mathematical or statistical model must be chosen to define the correlation between relevant descriptors and biological activities. The model can be linear partial least squares (PLS) [12], multiple linear regression (MLR) [13] or nonlinear. The selected model is then trained on a randomly chosen dataset, and the rest is used as test compounds. Model training often involves validation procedures, for example, exclusion cross validation (LOOCV) [14]. The training process is reiterated in order to reach an acceptable performance. The final step involves the testing process [11].

The concept of QSAR was first envisioned by Free, Wilson, Hansch and Fujita in 1964 [15, 16]. Subsequently, a new 3D-QSAR method, named comparative molecular field analysis (CoMFA) [17], has been worked out to overwhelm general 3D-QSAR problems. It has provided the basis for the development of multidimensional (nD) QSARs.

QSAR's Use

QSAR should not be seen as an academic tool that allows for the subsequent rationalization of data. It aims to derive molecular structure relationships between biology and chemistry for a valid reason. Models can be developed from these relationships and are thought to be predictive with common sense, luck, and expertise. A QSAR model can have many practical commitments [18, 19]:

- Rational estimation of biological activity and physicochemical properties.
- Understand and rationalize the action mechanisms of a wide variety of chemicals.
- Cost-effective product development.
- Minimization of the production time.
- Elimination of the ethical concerns.
- Spurring of “green” chemistry.

The ability to predict biological activity is of great importance in industry. Some QSARs have many applications in industry, academia and government (regulatory bodies). Some potential application areas can be seen below [18]:

- Rational compound identification with significant pharmacological, bactericidal or insecticidal activity.
- Optimization of the available biological activities.
- Design of versatile products to be used in diverse fields.
- Determination of toxicity in general and side effects in pharmaceuticals.
- Determination of compound stability.
- Rationalization and estimation of the combinatorial effects.

QSAR Model Development

A prerequisite in the creation of a QSAR model is the accuracy and reliability of the experimental data. If the experimental data are not accurate and precise, it would not be possible to obtain and develop a meaningful model. To develop an appropriate QSAR model, these data must include a wide range of chemically similar compounds [20]. Various biologically active compounds should be determined experimentally, and predetermined chemical structures should be selected [21].

While classical Hansch-like applications predominantly use the properties of substituents, atomic properties can also be taken into consideration, including charge, logP, orbital energy, and molar refractivity. The diversity of experimentally viable molecular variables contributes most to the strength of a given model. Software for any chemical structure determination variables are commercially available [22, 23], and these consider geometric, structural, topological, electrostatic, thermodynamic, and quantum mechanical properties of the products. In CODESSA, generally GAUSSIAN [24] and AMPAC [25] output files are used for the production of variables. DRAGON program on the other hand utilizes Molfiles, Multiple SD, Sybyl, SMILES, HyperChem, and MacroModel output files [26]. These programs compute compound parameters at the lowest energy fitness and exclude complex components. Therefore, the programs CODESSA and DRAGON only formulate those parameters related to one conformer of a compound. Spartan software [27] can calculate the Monte Carlo structure of compounds and calculate their relative energies, providing an overwhelming solution for the study of complex compounds. Another program, EMRE, uses the output files of SPARTAN 02, SPARTAN 04, SPARTAN 06, SPARTAN 08, SPARTAN 10, SPARTAN 14. An advantage of EMRE is that after determining both the matrix and the pharmacophore group, it also provides

information on the perpendicular distance of the farthest atom in the molecule to the pharmacophore plane, the angle between the pharmacophore plane and the two atoms, the angle between atoms, the torsion angle, and the van der Waals radius [28 - 30].

As mentioned above, a statistically robust model relies on the quality of the selection of variables that can help explain the evolution of the structure-activity relationship. The QSAR choice, on the other hand, is crucial to develop an appropriate structure-activity correlation. Today, least squares (PLS) and multiple linear regression (MLR) are used in QSAR modelling. The commonest problem is chanced upon in the biological activity calculations if the number of variables is larger than that of the compounds. It is important in statistical calculations that the number of parameters produced at the end of the calculation should not exceed the number of compounds (Table **1**) [21].

Table 1. QSAR techniques and used parameters.

QSAR techniques	Used Parameters
1D-QSAR	Physicochemical parameters such as pKa, logP
2D-QSAR	1D-QSAR + Structural, geometric, electrostatic, thermodynamic parameters
3D-QSAR	2D-QSAR + Electrostatic, steric, hydrophobic parameters
4D-QSAR	3D-QSAR + Parameters related to conformations, protonation and stereoisomers
5D-QSAR	4D-QSAR + Parameters related to conformational changes in ligand-protein binding
6D-QSAR	5D-QSAR + Parameters related to dissolution

2D-QSAR Analysis

The 2D-QSAR analysis uses linear regression to attain quantitative relationships. Statistical data for most 2D-QSAR techniques are questionable [31]. It focuses on the initial interaction between ligand and receptor. Thus, 2D QSAR analysis is not as precise as 3D-QSAR. It typically performs structure-based drug identification and focuses on reaction information, structural characterization predictions, and chemical information [32].

There are generally three types of molecular parameters in 2D-QSAR:

- Two-dimensional molecular coupling parameters
- Three-dimensional molecular surface area parameters
- LogP

2D-QSAR involves structural-, geometric-, topological-, quantum-, electrostatic-,

and thermodynamic parameters. In its simplest definition, a structure parameter reflects a molecular structure, excluding its geometric or electronic structure. Topology parameters contain the bonding information between two pairs of atoms in a molecule. Geometric parameters consider molecular space, density, and volume. Quantum chemical parameters are descriptive symbols such as HOMO and LUMO that provide information about the electronic properties of a molecule. Thermodynamic parameters cover heat formation, hydrophobicity, and molar refractivity [33, 34].

Fragment-Based 2D-QSAR Methods

As time and technology progressed, modelling calculations of many ligands or compounds can be performed, including 2D methods, 3D crystal receptors or where target structures are not available [35, 36].

Hologram-QSAR (HQSAR)

The hologram-QSAR method is fragment-based HQSAR (Hologram QSAR), originally developed by Tripos [37, 38]. The first step in the HQSAR method is to create molecular holograms that contain structural, geometric, electrostatic, and thermodynamic parameters of a large number of molecules and can be correlated with 2D molecular structures. Calculations are continued by deriving a mathematical regression equation to correlate these prepared hologram box values or components with the corresponding biological activities.

Fragment-Based QSAR (FB-QSAR)

In 2009, Du *et al.* [39] showed that a 2D-QSAR method based on molecular fragments uses two equations, the first being the Hansch-Fujita [40] linear free energy equation and the second the Free-Wilson [41] equation.

Fragment-Similarity Based QSAR (FS-QSAR)

A fragment similarity based QSAR (FS-QSAR) method [42] was developed using the original Free-Wilson method by introducing the concept of fragment similarity in the linear regression equation in 2010. It was applied for the first time to develop the traditional Free-Wilson equation instead of using physicochemical properties that often produce non-unique solutions. In this approach, part similarity calculations are made on the basis of similarity, using the lowest or highest eigenvalues calculated from the BCUT matrices [43, 44]. The

matrices contain the partial charges of the individual atoms as well as the atomic connectivity information contained in each individual part.

3D-QSAR

Even when computational studies did not exist, chemists were able to predict structure-activity relationships in a wide variety of molecules, and earlier QSAR approaches have often been effectively used to elucidate drug-receptor interactions. The most important limitation in their use is the total absence of numerical parameters which are essential in the production of substituents.

The main 3D-QSAR approaches, emerged in the 1980s, are the Active Analog Approach [35], Comparative Molecular Field Analysis [45, 46], and Molecular Shape Analysis [38]. Many others were developed in the 1990s, and have evolved considerably over the past decade [47]. Because the three-dimensional structure determines the biological activity of a compound, 3D-QSAR helps understand how structural modifications change biological properties. The 3D-QSAR uses informatics to reveal structure-activity relationships [33].

The general rules in the 3D-QSAR modelling are as follows [48]:

1. Conformational data are to be used in the selection of minimum energy conformers.

2. Distance between two atoms, optimization of the ring system, donor-acceptor hydrogen bond or charged functional groups should be determined.

3. Active Cluster Analysis is to be used in the selection of the linker.

4. Atomic properties such as dipole moment, polarity, point charges, molecular electrostatic potential, HOMO and LUMO orbitals, shape and volume that affect biological activity are to be determined.

5. Pharmacophore properties are to be selected [49].

Thanks to the use of the correlation between biological activity and the compound structure, 3D-QSAR has become the most popular approach in recent years. Parameters describing the chemical structure are used for the activity calculation. The 3D-QSAR method is divided into two basic classes according to the parameter type. One involves volume-based parameters while the other employs surface-based parameters. As a volume-based method, CoMFA is the most preferred 3D-QSAR method [34].

4D-QSAR

Hopfinger added another dimension to 3D-QSAR, in 1997, which is the mean of the matching conformers, and invented the term 4D-QSAR. Later on, Vedani has developed the 4D-QSAR approach. The latter used the concepts of 4D-QSAR by combining the effects of multidirectional conformations. This method is particularly useful in estimating the ligand's free binding energy to its receptor when the receptor structure is unknown [50]. The most active conformer of a given compound is defined at the minimum energy state (2 Kcal/mol). In this method, the active conformation is not a minimum energy conformer and instead it uses multi-temperature molecular dynamics (MDS). Hence, it is useful to estimate the potential energy of all low-energy conformers on a 3-D surface. Best results are produced when low energy conformers are calculated individually. The molecular coincidence problem is overwhelmed by similar sampling and computation. The conformers of each compound are placed on a predetermined rectangular surface to find the best agreement [51, 52]. The main difference of this method from CoMFA is that it combines both molecular conformers and overlapping molecules within a set of known molecular structures [33].

5D- and 6D-QSARs

Vedani [53] has added another dimension to 4D-QSAR, giving rise to 5D-QSAR. The latter examines the conformational changes inflicted on a protein by the binding of its ligand. The 6D-QSAR, also developed by Vedani, is able to map various dissolution patterns formed between solvent and solute molecules [54, 55].

Molecular Modelling and QSAR

In simple terms, molecular modelling can be thought of as a computer technique that can either predict molecular and biological properties or analyze molecules and molecular systems, based on experimental knowledge and theoretical chemical methods. Available techniques help describe properties that are responsible for biological activity, such as hydrophobicity, electronic properties of atoms and molecules, and geometry of the molecule. All these features are important in understanding the structure-activity relationship in drug design. SYBYL, AMBER, DOCK, MODELER and RasMol are the most used programs in molecular modelling [56].

Molecular modelling employs computer chemistry and graphical design techniques to explore structural properties of a molecule [57, 58]. Here three of the outstanding approaches are molecular dynamics, molecular mechanics, and quantum mechanics [59]. Models are generated by QM and MM calculations. The

models first allow the energy regulation of molecules and atoms in a calculated system, and they help understand how alterations in the position of atoms and molecules change the energy of a system. Three main steps in a molecular modelling can be stated as follows [60]:

1. Geometric work involves the construction of bond angle and bond lengths.

2. Fragments with appropriate geometry are demonstrated.

3. Structure is constructed using the experimental information, produced by X-ray crystallography, neutron diffraction or NMR.

A fourth step involves Monte Carlo Simulation or molecular dynamics. Finally, the calculations are analyzed.

Importance of the Validation of QSAR Models

QSAR has largely been matured, but it needs further improvement. Estimating the precision of calculations is for instance a serious problem [61]. Thus, the need for the proof of QSAR models has been brought into consideration [62]. Four common tools have so far been set for the assessment of the proof of QSAR models [63]: (i) cross validation, (ii) bootstrapping, (iii) randomization of response data, and (iv) external validation.

Some proposals have also been made for evaluating the rationality of QSAR models at an International workshop held in Setubal (Portugal) that have been revised later on at the OECD QSAR Work Program in 2004 [64, 65].

Means of Proof for QSAR Models

The proof is needed both to confirm the prediction strength of a model and to conclude the intricacy of an equation (Fig. **2**).

Least squares fit (R^2), cross-validation (Q^2) [66, 67], adjusted R^2 (R^2adj), chi-square test (χ^2) [68], root mean square error (RMSE) [69], bootstrapping, and shuffling (Y-Randomization) [70, 71] are internal methods for validating a model. However, external statistical methods serve best to make sure that the models created are robust and impartial [64].

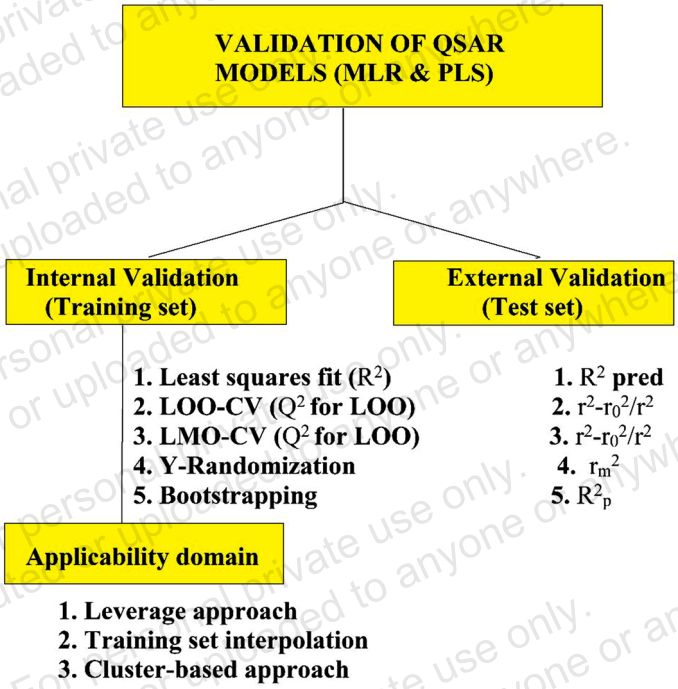

Fig. (2). A diagram of QSAR model validation.

Internal Validation

Least Squares Fit

It is an internal validation method, similar to linear regression. A solid straight line fit is produced to compute R^2. A substitute of this method excludes compounds in the training set (outliers) to validate the QSAR model. A difference between R^2 and R^2adj value of less than 0.3 shows that the number of descriptors included in the QSAR model is satisfactory [64].

$$R^2 = \left[\frac{N\Sigma XY - (\Sigma x)(\Sigma Y)}{\sqrt{([N\Sigma X^2 - (\Sigma X)^2][N\Sigma Y^2 - (\Sigma Y)^2])}}\right] \quad \textbf{(1)}$$

Fit of the Model

The prediction power of a QSAR model, reflected in R^2, can also be validated by chi-square (χ^2) and mean square error (RMSE) tests. *The chi-square* value is the difference between experimental and theoretical bioactivity scores. RMSE on the

other hand finds the error between the mean of the experimental- and the predicted figures [64].

$$\chi^2 = \sum_{i=1}^{n} \left(\frac{(y_i - \hat{y}_i)^2}{\hat{y}_i} \right) \quad \textbf{(2)}$$

$$RMSE = Sqrt \left(\sum_{i=1}^{n} \frac{(\hat{y}_i - y_m)^2}{n-1} \right) \quad \textbf{(3)}$$

y and $\hat{y}$, the experimental and predicted bioactivity for a single compound in the training set; *ym*, the mean of the experimental bioactivities; and n, the number of molecules in the studied data set.

High *chi*-square or RMSE scores (≥0.5 and 1.0, respectively) imply that the model built is of poor quality despite its high R^2 scores (≥0.7). An appropriate model should have *chi*- and RMSE values of 0.5 and <0.3, respectively [72] but they are insufficient for the complete validity test [21].

Cross-validation

Cross-validation is another internal validating test (CV, Q^2, q^2 or jack-knifing). The CV process reiterates the regression on the data subsets. It is often used to determine the suitability of a model to a given data set. In the test, the molecules are sequentially scrutinized and the missing molecule is used in the computation of R.

Cross-validation is exploited to understand the predictive capability of a model and to find out whether the model overfits, that is, a predictive model can only define the relationship between predictors and response in the existing compounds. This phenomenon comes to the fore when the difference between R^2 and Q^2 is higher than 0.3) [21, 66, 67].

Cross-validation test can be cross-checked using majority-out (LMO). This method is mostly handy when the training set is small (≤20 composites) or does not exist.

Equation for CV:

$$Q^2 = 1 - \frac{PRESS}{\sum_{i=1}^{N}(y_i - y_m)^2} \quad \textbf{(4)}$$

$$PRESS = \sum_{i=1}^{N}(y_{pred,i} - y_i)^2 \quad \textbf{(5)}$$

y_i, the data values not used to build the CV model; PRESS [65], the estimated remaining sum of squares.

Bootstrapping

Bootstrapping [73, 74] is another internal validation test which analyzes randomized subsamples rather than analyzing subsets. Here groups are indicated with *K* and object numbers with *n*. The test performed with randomly selected *n* objects is employed to assess target features for excepted samples. A high average Q^2 indicates the robustness of a models.

Randomization Test (Scrambling Model)

In this test, activity values are casually reallocated and the whole procedure is reiterated. Scrambled Model controls the identifiers to warrant that the chosen descriptors are proper. It is created using the original identifiers. For this test R^2 and Q^2 values are also computed. Lower scores established by repetitive testing indicate the robustness of the model built and there should not be a strong correlation between R^2 (R^2>0.50) [58] and Q^2 [75, 76].

External Validation

It has sometimes been argued that the best way to test the prediction capacity of a QSAR model is to compare the predicted and observed activities of compounds which are not included in the model creation process [77, 78]. The steps of the test have been featured as follows: computing the correlation coefficient, R and R^2 [79]; and defining the k and k' regression slopes passing through the origin.

$$R^2{}_{pred} > 0.6 \quad \textbf{(6)}$$

$$r^2 - r_0^2/r^2 < 0.1, \qquad r^2 - r_0^2/r^2 < 0.1 \quad \textbf{(7)}$$

$$0.85 \leq k \leq 1.15 \; or \; 0.85 \leq k' \leq 1.15 \quad \textbf{(8)}$$

The prediction power of the selected model should also be verified by the external R^2 predicted R^2 values. R2 >0.6 is a sufficient indicator of proper external predictability.

$$R^2_{Pred} = 1 - \frac{\sum_{i=1}^{test}(y_{exp} - y_{pred})^2}{\sum_{i=1}^{test}(y_{exp} - \bar{y}_{tr})^2} \quad \textbf{(9)}$$

y_{tr}, the mean value of the dependent variable for the training set.

The final tool available to check the external predictability of the selected model is r^2_m, which has been put forward by Roy and Paul [80]:

$$r^2_m = r^2(1 - \sqrt{r^2 - r_0^2} \quad \textbf{(10)}$$

r^2, the square of the correlation coefficient between the experimental and predicted values; and r^2_0, the square of the correlation coefficient between the experimental and predicted values with the intercept set to zero. An r^2_m value greater than 0.5 can be taken as an indicator of good external predictability [81].

The r^2_m (overall) statistic can be beneficial when the test set size is quite small and the regression-based external validation [82, 83] parameter is less dependent and highly reliant on test set observations. The r^2_m (overall) statistic can be utilized to choose the best analytical models among comparable models, because some models produce better internal validation parameters while others are good at generating external validation parameters.

Another proof parameter, R^2_p, checks the fitness of the selected model [82]. It penalizes the R^2 model (Eq. 11).

$$R^2_p = R^2 * \left(1 - \sqrt{|R^2 - R^2_r|}\right) \quad \textbf{(11)}$$

The r^2_m (overall) value defines whether the range of activity values predicted for the entire dataset is actually close to the experimental results. Therefore, a QSAR model can be satisfactory if the r^2_m (total) and R^2_p values are equal to or greater than 0.5 [52].

QSAR calculations are an effective theoretical method used to design more efficient and more active molecules (Fig. **3**). These calculations can be used to design new high activity molecules [84, 85].

Easily Reproducible QSAR Protocol

Spartan 10's software [23] is typically used for optimizations, conformational analysis of related molecules, and the creation of molecular structures at the chosen basis set level. Water was deemed to be the best solvent for cancer in general. When extremely sophisticated approaches like DFT (Density Functional Theory) are combined with enormous basis sets, it's a crucial technique for getting more precise findings. However, creating high basis sets usually takes long time, since molecular conformers are also taken into account. A relatively simple method and a larger basis set should be used to produce trustworthy results within acceptable time schedules.

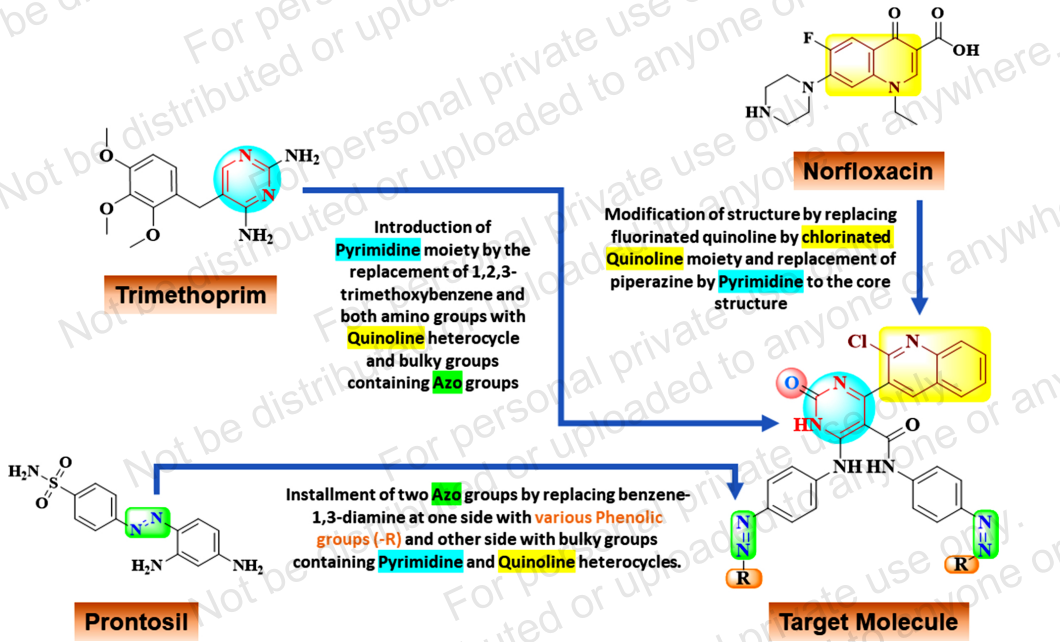

Fig. (3). Molecule design.

Low-energy conformers are known to be densely positioned at room temperature because conformers are ordered according to Boltzmann weights, which rely on the energies of the conformers relative to the energy of the conformer. Then, using the Monte Carlo random search approach, conformational analysis is carried out to look for these conformers. When the conformational search is complete, conformers with lower energies are retained and those with a Boltzmann distribution of less than 1/10000 are removed. Mulliken charges in diagonal elements, and bond configurations/atomic distances in non-diagonal elements were then carried out using EMRE software, and computations of electron conformational conformity matrices (ECMC) with residual conformers [2, 22, 24]. The conformers of each molecule were used to generate the ECMC based on the findings of the quantum chemical calculations. We defined a template compound that generates a predictive model for use in the comparison process of ECMCs [86]. Typically, the template was chosen from the atoms in the backbone of the lowest energy molecule. The molecules in the examined series were ranked according to their activities in order to distinguish between active and inactive compounds. The molecules were split into two after this sorting, by taking into account the order of the activities. The row's top individuals were labeled as active molecules. The lowest group was labeled as having inactive molecules. The pharmacophore, also known as the electron conformational activity sub-matrix (ECSA), was then defined by comparing the model compound's ECMC with all other ECMCs that fell within the specified tolerance range. Then, using the MATLAB application and a variety of produced quantum chemical parameters, QSAR calculations were performed. These computations were used to build a model.

QSAR studies are important in the designing of new and more effective molecules, based on the existing group of molecules. Activity estimations are made using many different programs. These programs are made by some researchers using programs developed by them and paid advanced programs. There are many studies conducted in the light of this information. Some of them are as follows:

In a study by Sahin and co-workers [87], a theoretical comparison of the activities of isatin derivatives was made against the BCL-2 inhibitor with the EMRE program using the Electron-Conformational Genetic Algorithm (EC-GA) hybrid method. In their calculations, it is seen that the molecules are prepared in a total of 801 parameters under four different categories ((i) electronic (ii) quantum-chemical (iii) geometrical and (iv) thermodynamic). It is seen that a good model is obtained by using the best seven parameters among these 801 parameters (Table **2**).

Table 2. κj values and description of the optimum parameters.

$a_{ni}^{(j)}$	Molecular Descriptors	κj values
1	C2 Nucleophilic Atom Boundary Electron Density (eV)	−15.43
2	H5 Electrophilic Atom Boundary Electron Density(eV)	−22.80
3	C7 Nucleophilic Atom Boundary Electron Density (eV)	−4.11
4	C4 Fukui Atomic Electrophilic Reactivity Index (eV)	152.31
5	log P	−0.03
6	PSA (u)	0.00
7	max.el. pot.(u)	0.00

In the model, it was seen that activity calculations were made for 27 molecules whose activity was unknown. It can be seen that very high results are obtained for the model when more than one internal and external validation is made. If these are the results of the validation; q^2_{ext1}=0.79, q^2_{ext2}=0.79, q^2_{ext3}=0.83, CCC_1=0.97, CCC_2=0.90 (Table **2**).

Another study has identified 80 methanone derivatives and biological comparison of molecules [52]. Electron Conformational Compatibility Matrices were obtained with EMRE software by using the geometric, thermodynamic and topological properties of 80 molecules studied. A model was created using 804 parameters. (Table **3**).

Table 3. κj values and description of the optimum parameters.

$a_{ni}^{(j)}$	Molecular Descriptors	κj values
1	The distance of O2–F1	0.027
2	Bond degree of C10–N2 (Lowdin)	−3.720
3	Lowdin values of the N2 atom (e−)	1.295
4	Orthogonal distance of the C11 atom to the N3 C20 O1 plane	−0.129
5	Bond degree of N2–C16 (Lowdin)	1.801
6	Orthogonal distance of the C11 atom to the C12 N1 O1 plane	−0.069
7	Bond degree of C6–C1 (Lowdin)	−0.435
8	H6 C6 C1 angle	−0.752

The numerical values of the model were were $R^2_{training}$ =0.834, q^2 =0.768 and $SE_{training}$ =0.075, q^2_{ext1} =0.875, q^2_{ext2} =0.839, q^2_{ext3} =0.764, ccc_{tr} =0.908, ccc_{test} =0.929 and ccc_{all} =0.920.

A chemical property-based pharmacophore model was developed for tetrahydrodibenzazosine derivatives with the EMRE package program [88]. All QSAR models were built with 40 compounds (training set), and then a test set was created with an additional nine compounds (test set) to create a consistent model (Table **4**).

Table 4. κj values and description of the optimum parameters.

$a_{ni}^{(j)}$	**Molecular Descriptors**	**κj values**
1	Orthogonal distance from the C2 atoms to C23–C14–C10 plane	−0.202
2	Orthogonal distance from the C2 atoms to C23–N1–C10 plane + van der Waals radius of C2 atom	−0.255
3	Orthogonal distance from the C12 atoms to C14–N1–C10 plane + van der Waals radius of C12 atom	0.135
4	The angle (degree) between C15–N1–C4atoms	−0.130
5	Fukui Lumo of the C5 atom	4.386
6	Fukui Lumo of the C13 atom	−5.912
7	Mulliken charge of C9 atom	−5.434
8	E (kcal/Mol)	0.002
9	Rel Eaq (kcal/Mol)	0.003
10	PSA (u)	0.008

In the end, a statistically valid 4D-QSAR (R^2_{training} =0:856, R^2_{test} =0:851 and q^2 = 0.650) model was obtained with a good good external set estimation.

The activity comparison of 86 alkynylphenoxyacetic acid derivatives has been made for the CRTh2 receptor [85]. A model was then created using the genetic algorithm and nonlinear least squares regression methods (Table **5**).

Table 5. κj values and description of the optimum parameters.

$a_{ni}^{(j)}$	**Molecular Descriptors**	**κj values**
1	C22-C18 distance (Å)	0.0216
2	C22-C18 distance (Å)+ vdW radius	0.0564
3	Orthogonal distance from the C22 atom to the C1-O1-O3 plane (Å)	0.0243
4	Orthogonal distance from the O2 atom to the C1-H6-C12 plane (Å) + vdW radius of O2 atom	–0.0139
5	Orthogonal distance from the C6 atom to the C1-O1-C12 plane (Å) + vdW radius of C6 atom	–0.0713
6	The angle of between Cl1-O3-C18 atoms	0.0108

(Table 5) cont.....

$a_{ni}^{(j)}$	Molecular Descriptors	κj values
7	The angle between the O3-H6-C12 plane and the O2-O4 lines	0.2155
8	The angle between the C6-O3-H6 plane and the Cl1-C18 lines	0.1040
9	The dihedral angle between the O2-C5-C22-C18 plane	–0.0109

A theoretical comparison of the anti-HIV-1 activity of 52 tetrahy- droimidazo [4,5,1-jk] 1,4] benzodiazepinone derivatives has been reported (Table **6**) [89].

Table 6. κj values and description of the optimum parameters.

$a_{ni}^{(j)}$	Molecular Descriptors	κj values
1	Distance between S1- the farthest atom bonded to N3 (C14)	0.127
2	Angle between line of N3-M* atoms and H25-S1-C9 plane	−0.003
3	Angle between S1-C1-N1 atoms	−0.466
4	Dihedral angle between H25-N1-C12- the farthest atom bonded to C12	−0.024
5	Angle (radian) between S1-C1-C12- the farthest atom bonded to C12 atoms	−95.346
6	Distance between N1 and C9+ van der Waals radius of C9 atom	5.124
7	Distance between N2 and C11+ van der Waals radius of the farthest atom bonded to C11	−0.118

CONCLUSION

QSAR calculations are an important method used to determine the parameter that contributes to the activities of molecules by comparing the numerical values of these properties by examining many properties of the molecules made before the experimental processes in order to design more effective and more active molecules.

REFERENCES

[1] Esposito EX, Hopfinger AJ, Madura JD. Methods for applying the quantitative structure-activity relationship paradigm. In Chemoinformatics Humana Press 2014; pp. 131-213.

[2] Bradbury SP. Quantitative structure-activity relationships and ecological risk assessment: An overview of predictive aquatic toxicology research. Toxicol Lett 1995; 79(1-3): 229-37. [http://dx.doi.org/10.1016/0378-4274(95)03374-T] [PMID: 7570660]

[3] Chen JZ, Han XW, Liu Q, Makriyannis A, Wang J, Xie XQ. 3D-QSAR studies of arylpyrazole antagonists of cannabinoid receptor subtypes CB1 and CB2. A combined NMR and CoMFA approach. J Med Chem 2006; 49(2): 625-36. [http://dx.doi.org/10.1021/jm050655g] [PMID: 16420048]

[4] Hansen C, Telzer BR, Zhang L. Comparative QSAR in toxicology: Examples from teratology and cancer chemotherapy of aniline mustards. Crit Rev Toxicol 1995; 25(1): 67-89. [http://dx.doi.org/10.3109/10408449509089887] [PMID: 7734060]

[5] Perkins R, Fang H, Tong W, Welsh WJ. Quantitative structure-activity relationship methods:

perspectives on drug discovery and toxicology. Environ Toxicol Chem 2003; 22(8): 1666-79. [http://dx.doi.org/10.1897/01-171] [PMID: 12924569]

[6] Salum LB, Andricopulo AD. Fragment-based QSAR: Perspectives in drug design. Mol Divers 2009; 13(3): 277-85. [http://dx.doi.org/10.1007/s11030-009-9112-5] [PMID: 19184499]

[7] Myint KZ. QSAR methods development, virtual and experimental screening for cannabinoid ligand discovery. Dissertation 2012.

[8] Akyüz L, Sarıpınar E. Conformation depends on 4D-QSAR analysis using EC–GA method: Pharmacophore identification and bioactivity prediction of TIBOs as non-nucleoside reverse transcriptase inhibitors. J Enzyme Inhib Med Chem 2013; 28(4): 776-91. [http://dx.doi.org/10.3109/14756366.2012.684051] [PMID: 22591319]

[9] Sahin K, Saripinar E. A novel hybrid method named electron conformational genetic algorithm as a 4D QSAR investigation to calculate the biological activity of the tetrahydrodibenzazosines. J Comput Chem 2020; 41(11): 1091-104. [http://dx.doi.org/10.1002/jcc.26154] [PMID: 32058616]

[10] Akyüz L, Sarıpınar E, Kaya E, Yanmaz E. 4D-QSAR study of HEPT derivatives by electron conformational–genetic algorithm method. SAR QSAR Environ Res 2012; 23(5-6): 409-33. [http://dx.doi.org/10.1080/1062936X.2012.665082] [PMID: 22452710]

[11] Keskin H. 4D-QSAR analysis of benzotienopyrimidin compounds by electron conformation-genetic algorithm (EC-GA) method. Master thesis. 2019.

[12] Köprü S. 4D-QSAR analysis of phenylpyrazine, alkinylpenoxyacetic acid, phthalazine derivatives by electron conformational-genetic algorithm (EC-GA) method. Master thesis. 2018.

[13] Yao XJ, Panaye A, Doucet JP, *et al.* Comparative study of QSAR/QSPR correlations using support vector machines, radial basis function neural networks, and multiple linear regression. J Chem Inf Comput Sci 2004; 44(4): 1257-66. [http://dx.doi.org/10.1021/ci049965i] [PMID: 15272833]

[14] Bahar R. Pharmacophore modelling and 4D-QSAR analysis of phenylpyrazole glutamic acid piperazine by electron conformational-genetic algorithm method. Master thesis. 2018.

[15] Free SM Jr, Wilson JW. A mathematical contribution to structure-activity studies. J Med Chem 1964; 7(4): 395-9. [http://dx.doi.org/10.1021/jm00334a001] [PMID: 14221113]

[16] Hansch C, Fujita T. ρ–σ–π Analysis. A method for the correlation of biological activity and chemical structure. J Am Chem Soc 1964; 86(8): 1616-26. [http://dx.doi.org/10.1021/ja01062a035]

[17] Kulak L. QSAR studies with CoMCET method on 2-amino-6-arylsulfonyl benzonitrile compounds as HIV-1 inhibitor. Master thesis. 2018.

[18] Puzyn T, Leszczynski J, Cronin MT, Eds. Recent advances in QSAR studies: methods and applications. Springer 2010. [http://dx.doi.org/10.1007/978-1-4020-9783-6]

[19] Aydın S. Qsar studies with electron conformation-genetic algorithm (EC-GA) method on 3-(1, 1-dioxo-2h-(1,2,4)-benzothiadiazine-3-yl)-4-hydroxy-2(1h)-quinoline compounds as hepatitis c inhibitor. Master thesis. 2018.

[20] Ravi M, Hopfinger AJ, Hormann RE, Dinan L. 4D-QSAR analysis of a set of ecdysteroids and a comparison to CoMFA modeling. J Chem Inf Comput Sci 2001; 41(6): 1587-604. [http://dx.doi.org/10.1021/ci010076u] [PMID: 11749586]

[21] Tüzün B. 4D-QSAR study of some pyridine carboxylic acid, oxadiazole, pyrimidine and oxazol derivatives by electron conformational-genetic algorithm method. Phd thesis. 2018.

[22] Karelson M, Maran U, Wang Y, Katritzky AR. QSPR and QSAR models derived using large molecular descriptor spaces. A review of CODESSA applications. Collect Czech Chem Commun 1999; 64(10): 1551-71.
[http://dx.doi.org/10.1135/cccc19991551]

[23] Polishchuk PG, Kuz'min VE, Artemenko AG, Muratov EN. Universal approach for structural interpretation of QSAR/QSPR models. Mole Informa 2013; 32(9-10): 843-53.
[http://dx.doi.org/10.1002/minf.201300029]

[24] Frisch MJ, Trucks GW, Schlegel HB, *et al.* revision D01. Wallingford: Gaussian Inc 2009.

[25] Holder AJ, Ye L, Yourtee DM, Agarwal A, Eick JD, Chappelow CC. An application of the QM-QSAR method to predict and rationalize lipophilicity of simple monomers. Dent Mater 2005; 21(7): 591-8.
[http://dx.doi.org/10.1016/j.dental.2004.08.004] [PMID: 15978267]

[26] Altaf R, Nadeem H, Iqbal MN, *et al.* Synthesis, biological evaluation, 2D-QSAR, and molecular simulation studies of dihydropyrimidinone derivatives as alkaline phosphatase inhibitors. ACS Omega 2022; 7(8): 7139-54.
[http://dx.doi.org/10.1021/acsomega.1c06833] [PMID: 35252705]

[27] Wavefunction inc. 18401 Von Karman Avenue. USA: Suite 370, Irvine 2006.

[28] Maria de Assis T, Ramalho TC, Ferreira da Cunha EF. 4D-QSAR models applied to the study of TGF-β1 receptor inhibitors. Curr Top Med Chem 2021; 21(13): 1157-66.
[http://dx.doi.org/10.2174/1568026621666210727161431] [PMID: 34315368]

[29] Liu J, Pan D, Tseng Y, Hopfinger AJ. 4D-QSAR analysis of a series of antifungal p450 inhibitors and 3D-pharmacophore comparisons as a function of alignment. J Chem Inf Comput Sci 2003; 43(6): 2170-9.
[http://dx.doi.org/10.1021/ci034142z] [PMID: 14632469]

[30] de Souza APM, Costa MCA, de Aguiar AR, *et al.* Leishmanicidal and cytotoxic activities and 4D-QSAR of 2-arylidene indan-1,3-diones. Arch Pharm 2021; 354(10): 2100081.
[http://dx.doi.org/10.1002/ardp.202100081] [PMID: 34323311]

[31] Altaf R, Nadeem H, Iqbal MN, *et al.* "Synthesis, biological evaluation, 2D-QSAR, and molecular simulation studies of dihydropyrimidinone derivatives as alkaline phosphatase Inhibitors". ACS Omega 2022; 7(8): 7139-54.
[http://dx.doi.org/10.1021/acsomega.1c06833] [PMID: 35252705]

[32] Yu X. Prediction of inhibitory constants of compounds against SARS-CoV 3CLpro enzyme with 2D-QSAR model. J Saudi Chem Soc 2021; 25(7): 101262.
[http://dx.doi.org/10.1016/j.jscs.2021.101262]

[33] Kocakaplan I. Investigation of Q2/R2 rating according to the training and test sets cutting and molecular distribution for 4d QSAR model of acetyl cholinesterase inhibitor activity Master thesis 2021.

[34] Kaushal T, Khan S, Fatima K, Luqman S, Khan F, Negi AS. Synthesis, molecular docking, and 2D-QSAR modeling of quinoxaline derivatives as potent anticancer agents against triple-negative breast cancer. Curr Top Med Chem 2022; 22(10): 855-67.
[http://dx.doi.org/10.2174/1568026622666220324151808] [PMID: 35331094]

[35] Myint KZ, Xie XQ. Recent advances in fragment-based QSAR and multi-dimensional QSAR methods. Int J Mol Sci 2010; 11(10): 3846-66.
[http://dx.doi.org/10.3390/ijms11103846] [PMID: 21152304]

[36] SYBYL80. Discovery Software for Computational Chemistry and Molecular Modeling. St. Louis, MO, USA: Tripos 2008.

[37] Salum LB, Andricopulo AD. Fragment-based QSAR: Perspectives in drug design. Mol Divers 2009;

13(3): 277-85.
[http://dx.doi.org/10.1007/s11030-009-9112-5] [PMID: 19184499]

[38] Lowis D. HQSAR: A new, highly predictive QSAR technique, tripos technique notes. St. Louis, MO, USA: Tripos 1997.

[39] Du QS, Huang RB, Wei YT, Pang ZW, Du LQ, Chou KC. Fragment-based quantitative structure-activity relationship (FB-QSAR) for fragment-based drug design. J Comput Chem 2009; 30(2): 295-304.
[http://dx.doi.org/10.1002/jcc.21056] [PMID: 18613071]

[40] Hansch C, Fujita T. ρ–σ–π Analysis. A method for the correlation of biological activity and chemical structure. J Am Chem Soc 1964; 86(8): 1616-26.
[http://dx.doi.org/10.1021/ja01062a035]

[41] Free SM Jr, Wilson JW. A mathematical contribution to structure-activity studies. J Med Chem 1964; 7(4): 395-9.
[http://dx.doi.org/10.1021/jm00334a001] [PMID: 14221113]

[42] Myint KZ, Ma C, Wang L, Xie XQ. The Fragment-similarity-based QSAR (FS-QSAR): A Novel 2D-QSAR method to predict biological activities of triaryl Bis-sulfone and COX2 analogs. 2010.

[43] Burden FR. Molecular identification number for substructure searches. J Chem Inf Comput Sci 1989; 29(3): 225-7.
[http://dx.doi.org/10.1021/ci00063a011]

[44] Xie XQ, Chen JZ. Data mining a small molecule drug screening representative subset from NIH PubChem database. J Chem Inf Model 2008; 48(3): 465-75.
[http://dx.doi.org/10.1021/ci700193u] [PMID: 18302356]

[45] DePriest SA, Mayer D, Naylor CB, Marshall GR. 3D-QSAR of angiotensin-converting enzyme and thermolysin inhibitors: A comparison of CoMFA models based on deduced and experimentally determined active site geometries. J Am Chem Soc 1993; 115(13): 5372-84.
[http://dx.doi.org/10.1021/ja00066a004]

[46] Cramer RD, Depriest SA, Patterson DE, Hecht P (1993) The developing practice of comparative molecular field analysis. In: Kubinyi H, editor. 3D QSAR in drug design: theory, method and applications, pp 443-85. ESCOM, Leiden.

[47] Kubinyi H. Comparative molecular field analysis (CoMFA). The encyclopedia of computational chemistry, 1998; 1: 448-60.

[48] Mezey PG. The degree of similarity of three-dimensional bodies: Application to molecular shape analysis. J Math Chem 1991; 7(1): 39-49.
[http://dx.doi.org/10.1007/BF01200814]

[49] Kubinyi H, Folkers G, Martin YC. 3D QSAR in Drug Design: Recent Advances. Springer Dordrecht 1998.

[50] Kubinyi H. QSAR and 3D QSAR in drug design Part 1: Methodology. Drug Discov Today 1997; 2(11): 457-67.
[http://dx.doi.org/10.1016/S1359-6446(97)01079-9]

[51] Kieser KJ, Kim DW, Carlson KE, Katzenellenbogen BS, Katzenellenbogen JA. Characterization of the pharmacophore properties of novel selective estrogen receptor downregulators (SERDs). J Med Chem 2010; 53(8): 3320-9.
[http://dx.doi.org/10.1021/jm100047k] [PMID: 20334372]

[52] Debnath AK. Quantitative structure-activity relationship (QSAR) paradigm-Hansch era to new millennium. Mini Rev Med Chem 2001; 1(2): 187-95.
[http://dx.doi.org/10.2174/1389557013407061] [PMID: 12369983]

[53] Çatalkaya S, Sabancı N, Yavuz SÇ, Sarıpınar E. The effect of stereoisomerism on the 4D-QSAR study

of some dipeptidyl boron derivatives. Comput Biol Chem 2020; 84: 107190.
[http://dx.doi.org/10.1016/j.compbiolchem.2019.107190] [PMID: 31918171]

[54] Tüzün B, Saripinar E. Molecular docking and 4D-QSAR model of methanone derivatives by electron conformational-genetic algorithm method. J Iranian Chem Soc 2020; 17(5): 985-1000.

[55] Vedani A, Briem H, Dobler M, Dollinger H, McMasters DR. Multiple-conformation and protonation-state representation in 4D-QSAR: The neurokinin-1 receptor system. J Med Chem 2000; 43(23): 4416-27.
[http://dx.doi.org/10.1021/jm000986n] [PMID: 11087566]

[56] Qin D, Zeng X, Zhao T, Cai B, Yang B, Tu G. 5D-QSAR studies of 1H-pyrazole derivatives as EGFR inhibitors. J Mol Model 2022; 28(12): 379.
[http://dx.doi.org/10.1007/s00894-022-05370-x] [PMID: 36342554]

[57] Yanmaz E, Sarıpınar E, Şahin K, Geçen N, Çopur F. 4D-QSAR analysis and pharmacophore modeling: Electron conformational-genetic algorithm approach for penicillins. Bioorg Med Chem 2011; 19(7): 2199-210.
[http://dx.doi.org/10.1016/j.bmc.2011.02.035] [PMID: 21419636]

[58] Gund T. Molecular modelling of small molecules. In: Guideb on Molec Model Drug Desi, Cohen 1996; 55-92.

[59] Hopfinger AJ, Wang S, Tokarski JS, *et al.* Construction of 3D-QSAR models using the 4D-QSAR analysis formalism. J Am Chem Soc 1997; 119(43): 10509-24.
[http://dx.doi.org/10.1021/ja9718937]

[60] Cai Z, Zafferani M, Akande OM, Hargrove AE. Quantitative structure–activity relationship (QSAR) study predicts small-molecule binding to RNA structure. J Med Chem 2022; 65(10): 7262-77.
[http://dx.doi.org/10.1021/acs.jmedchem.2c00254] [PMID: 35522972]

[61] Smith C. Molecular modeling -seeing the whole picture with modeling software packages. Scientist 1998; 12: 17-31.

[62] Esmel A. Insight into binding mode of nitrile inhibitors of Plasmodium falciparum Falcipain-3, QSAR and Pharmacophore models, virtual design of new analogues with favorable pharmacokinetic profiles. SDRP J Comput Chem Mole Model 2017; 2(1): 1-22.

[63] Tong W, Hong H, Xie Q, Shi L, Fang H, Perkins R. Assessing QSAR limitations-A regulatory perspective. Curr Computer-Aided Drug Des 2005; 1(2): 195-205.
[http://dx.doi.org/10.2174/1573409053585663]

[64] Worth AP, Van Leeuwen CJ, Hartung T. The prospects for using (Q)SARs in a changing political environment-high expectations and a key role for the european commission's joint research centre. SAR QSAR Environ Res 2004; 15(5-6): 331-43.
[http://dx.doi.org/10.1080/10629360412331297371] [PMID: 15669693]

[65] Worth AP, Hartung T, Van Leeuwen CJ. The role of the European centre for the validation of alternative methods (ECVAM) in the validation of (Q)SARs. SAR QSAR Environ Res 2004; 15(5-6): 345-58.
[http://dx.doi.org/10.1080/10629360412331297362] [PMID: 15669694]

[66] Veerasamy R, Rajak H, Jain A, Sivadasan S, Varghese CP, Agrawal RK. Validation of QSAR models-strategies and importance. Int J Drug Des Dis 2011; 3: 511-9.

[67] Von der Ohe PC, Kühne R, Ebert RU, Schüürmann G. Comment on "Discriminating toxicant classes by mode of action: 3. Substructure indicators" (M. Nendza and M. Müller, SAR QSAR Environ. Res. 18 155 (2007)). SAR QSAR Environ Res 2007; 18(7-8): 621-4.
[http://dx.doi.org/10.1080/10629360701698571] [PMID: 18038362]

[68] Leach AR. Molecular modeling: Principles and applications. Harlow, England: Pearson Education Ltd. 2001.

[69] Shao J. Linear model selection by cross-validation. J Am Stat Assoc 1993; 88(422): 486-94. [http://dx.doi.org/10.1080/01621459.1993.10476299]

[70] Tallarida RJ, Murray RB. Chi-square test. Manual of pharmacologic calculations. New York, NY: Springer 1987; pp. 140-2. [http://dx.doi.org/10.1007/978-1-4612-4974-0_43]

[71] Chai T, Draxler RR. Root mean square error *(RMSE)* or mean absolute error (MAE). Geosci Model Dev Discuss 2014; 7(1): 1525-34.

[72] Wold S, Ericksson L. Partial least squares projections to latent structures (PLS) in chemistry. In: Encyclopedia of computationalchemistry. Ragu , Schleyer P. Chichester: John Wiley & Sons 1998; 3: pp. 2006-21.

[73] Yasri A, Hartsough D. Toward an optimal procedure for variable selection and QSAR model building. J Chem Inf Comput Sci 2001; 41(5): 1218-27. [http://dx.doi.org/10.1021/ci010291a] [PMID: 11604021]

[74] Schneider G, Wrede P. Artificial neural networks for computer-based molecular design. Prog Biophys Mol Biol 1998; 70(3): 175-222. [http://dx.doi.org/10.1016/S0079-6107(98)00026-1] [PMID: 9830312]

[75] Thakur M, Thakur A, Khadikar PV. QSAR studies on psychotomimetic phenylalkylamines. Bioorg Med Chem 2004; 12(4): 825-31. [http://dx.doi.org/10.1016/j.bmc.2003.10.027] [PMID: 14759743]

[76] Zhai X, Chen M, Lu W. Predicting the toxicities of metal oxide nanoparticles based on support vector regression with a residual bootstrapping method. Toxicol Mech Methods 2018; 28(6): 440-9. [http://dx.doi.org/10.1080/15376516.2018.1449278] [PMID: 29644916]

[77] Choudhary G, Karthikeyan C, Hari Narayana Moorthy NS, Sharma SK, Trivedi P. QSAR analysis of some cytotoxic thiadiazinoacridines. Inter Elect J Mol Des 2005; 4: 793-802.

[78] Litang QIN, Shushen LIU, Qianfen XIAO, Qingsheng WU. Internal and external validtions of QSAR model. Environ Chem 2013; 32(7): 1205-11.

[79] Rücker C, Rücker G, Meringer M. y-Randomization and its variants in QSPR/QSAR. J Chem Inf Model 2007; 47(6): 2345-57. [http://dx.doi.org/10.1021/ci700157b] [PMID: 17880194]

[80] Golbraikh A, Tropsha A. Beware of q2!. J Mol Graph Model 2002; 20(4): 269-76. [http://dx.doi.org/10.1016/S1093-3263(01)00123-1] [PMID: 11858635]

[81] Guha R, Jurs PC. Determining the validity of a QSAR model-a classification approach. J Chem Inf Model 2005; 45(1): 65-73. [http://dx.doi.org/10.1021/ci0497511] [PMID: 15667130]

[82] Sachs L. Applied Statistics: A Handbook of Techniques. Berlin, Germany: Springer-Verlag 1984. [http://dx.doi.org/10.1007/978-1-4612-5246-7]

[83] Roy K, Paul S. Exploring 2D and 3D QSARs of 2, 4-diphenyl-1, 3-oxazolines for ovicidal activity against Tetranychus urticae. QSAR Comb Sci 2009; 28(4): 406-25. [http://dx.doi.org/10.1002/qsar.200810130]

[84] Pratim Roy P, Paul S, Mitra I, Roy K. On two novel parameters for validation of predictive QSAR models. Molecules 2009; 14(5): 1660-701. [http://dx.doi.org/10.3390/molecules14051660] [PMID: 19471190]

[85] Roy K. On some aspects of validation of predictive quantitative structure–activity relationship models. Expert Opin Drug Discov 2007; 2(12): 1567-77. [http://dx.doi.org/10.1517/17460441.2.12.1567] [PMID: 23488901]

[86] Özalp A, Yavuz SÇ, Sabancı N, Çopur F, Kökbudak Z, Sarıpınar E. 4D-QSAR investigation and

pharmacophore identification of pyrrolo[2,1-c][1,4]benzodiazepines using electron conformational–genetic algorithm method. SAR QSAR Environ Res 2016; 27(4): 317-42. [http://dx.doi.org/10.1080/1062936X.2016.1174152] [PMID: 27121415]

[87] Sahin K, Saripinar E, Durdagi S. Combined 4D-QSAR and target-based approaches for the determination of bioactive Isatin derivatives. SAR QSAR Environ Res 2021; 32(10): 769-92. [http://dx.doi.org/10.1080/1062936X.2021.1971760] [PMID: 34530651]

[88] Köprü S, Saripinar E. 4D-QSAR analysis and pharmacophore modeling for alkynylphenoxyacetic acids as CRTh2 (DP2) receptor antagonists. Turk J Chem 2018; 42(6): 1577-97. [http://dx.doi.org/10.3906/kim-1801-86]

[89] Akyüz L, Sarıpınar E. Conformation depends on 4D-QSAR analysis using EC–GA method: pharmacophore identification and bioactivity prediction of TIBOs as non-nucleoside reverse transcriptase inhibitors. J Enzyme Inhib Med Chem 2013; 28(4): 776-91. [http://dx.doi.org/10.3109/14756366.2012.684051] [PMID: 22591319]