Regular Article

PHYSICAL CHEMISTRY RESEARCH
Published by the
Iranian Chemical Society
www.physchemres.org
info@physchemres.org

Phys. Chem. Res., Vol. 10, No. 1, 31-44, March 2022
DOI: 10.22036/pcr.2021.272257.1884

Docking and 2D-Structure-activity Relationship and ADMET Studies of Acetylcholinesterase Inhibitors

F. Ansari[a], A. Niazi[b,*], Jahan B. Ghasemi[c,*] and A. Yazdanipour[b]

[a]*Department of Chemistry, Arak Branch, Islamic Azad University, Arak, Iran*
[b]*Department of Chemistry, Central Tehran Branch, Islamic Azad University, Tehran, Iran*
[c]*Drug Design in Silico Lab, School of Sciences, Chemistry Faculty, University of Tehran, Teheran, Iran*

(Received 7 February 2021, Accepted 17 August 2021)

In this work, a quantitative structure-activity relationship (QSAR) for some tacrine derivatives inhibitors of acetylcholinesterase was modeled using ligand-receptor interconnection interaction space. The descriptors were obtained by multivariate image analysis (MIA) of each molecule. Docking studies were performed to determine the best conformers of inhibitors. In the first step, the best pose of all the ligands was selected. Afterward, an MIA-QSAR model using ligand-receptor interconnection data was developed. The pool of descriptors was compressed by principal component analysis (PCA). Variable selection was carried out by genetic algorithm (GA) followed by model building using the support vector machine (SVM) regression method. The validation of the model's predictive ability was studied by a validation set containing 11 individual compounds. The Q^2, r^2, and Δr_m^2 test prediction values for PCA-GA-SVM model were 0.62, 0.89, and 0.145, respectively. After validating the results with all statistical data, three new molecules were designed by the MIA-QSAR model. Afterward, new molecules were docked in the AChE active site. Docking studies showed that the amino acids TYR70, TYR121, TYR334, TRP279, PHE288, PHE290, TRP84, TRP334, and SER286 are active amino acids in the complex. Finally, the ADMET parameters of the new compounds were calculated that were in acceptable ranges.

Keywords: Molecular docking, Multivariate image analysis-QSAR (MIA-QSAR), Ligand-receptor interaction, Acetylcholinesterase (AChE), ADMET

INTRODUCTION

One of the primary tasks of theoretical and computational chemistry is the development of existing computational methods to improve their utility in molecular modeling perspectives [1,2]. Image analysis in chemistry involves applying chemometrics methods to obtain information from simple images of chemical structure or chemical system, for example, images of a group of molecules [3,4]. There are many reports on this approach that discuss image analysis and its usage in chemistry.

Moreover, image analysis is important because it actively provides better quality control in chemical analysis. Images are mostly considered as sources of multivariate data used as input for various multivariate regression methods. In fact, the visual appearance of any particular image is less important [5,6]. Quantitative structure-activity relationship (QSAR) coupled with multivariate image analysis (MIA) was extensively used to predict the bio-reactivity of the drug-like compounds. The MIA-QSAR provides two main advantages: high accuracy, and ease ofobtaining molecular descriptors using 2D image analysis. In MIA-QSAR, pixels of a bi-dimensional picture of a molecule are chosen as raw input data. In this method, pixels are called descriptors that may provide useful information in chemistry [7].

On the other hand, descriptors do not have a direct

*Corresponding authors. E-mail: jahan.ghasemi@ut.ac.ir; ali.niazi@gmail.com

physicochemical meaning; they are binaries. According to a research, in QSAR, images (2D chemical structure) contain chemical information that shows the correlation between chemical structures and properties [8]. The basic argument here is that the biological properties depend on the functional group attached to the main scaffold. this means that the prediction ability of biological properties are affected by modifying the functional group [9-12]. Previous literature showed that employing different colors and sizes for different atoms provides more meaningful molecular descriptors [13]. These novel ideas improved the relationship between the functional group of the scaffold and the substance's bioreactivity. Therefore, it led to more accurate models in MIA-QSAR [14-16].

In this study, we focused exclusively on the acetylcholinesterase (AChE) inhibitors. The biological role of AChE is quite important. It mainly causes the quick hydrolysis of a neurotransmitter called acetylcholine (ACh) by cutting off the transmission at the cholinergic synapse [17,18]. Previously, some new information was reported on AChE and its inhibitors. These reports utilized computational chemistry to study the structure of the enzyme and enzyme-ligand interaction. Also, classical kinetic methods, such as QSAR modeling, were employed to investigate the site-directed mutagenesis [18-20]. In addition, AChE has a great specific activity. The main active site of the AChE is down below a narrow catalytic gorge. Therefore, this structure limits the products and substrate inside and outside of the active site. Historically, maybe the most significant feature of cholinesterase has been accommodating the physiological substrate cationic charge. The enzyme-ligand electrostatic interaction through the positive charge of the ligand is important; this implies the role of cholinesterase in the binding of the cationic substrate and ligands [21,22].

Recently, AChEs inhibitors were among the widely used treatment for several neuromuscular sicknesses. Also, it should be noted that AChEs inhibitors are the first generation of medications for Alzheimer's disease (AD). AD is a complex neurodegenerative process occurring in the central nervous system (CNS) [23,24]. Thus, the design of new inhibitors attracted great attention among researchers. AD patients should go through using limited types of medications, including galantamine, rivastigmine, N-methyl-D-aspartate receptor antagonist, donepezil, tacrine, and memantine [23-25]. This paper describes a developed MIA-QSAR model using the interactions of the best conformer with the AChE active site. Therefore, the docking method is used to anticipate whether the compound is a promising inhibitor for a particular protein or not. Then, the first pose of all ligands was chosen to apply in the MIA-QSAR method. Finally, principal component analysis (PCA) was conducted on the raw molecular images to project the input data on another coordinate system with fewer dimensions while data preserves its characteristics [29]. In order to properly choose the number of PC as input for support vector machine regression (SVR), a genetic algorithm (GA) was utilized. GA offers an intelligent feature selection process, which reduces the number of input variables prior to the modeling procedure.

Both PCA and GA are known as the methods of data reduction. GA can be employed to solve optimization problems in a stochastic manner. However, the GA applies the Darwin hypothesis evolution on variables to select the most important variables. Therefore, GA can be successfully used as a variable selection method [30]. Also, the modeling was performed by using PC-GA-SVM. The PC-GA-SVM was used to model the inhibitory activity of three tacrine derivatives. In this study, for the first time, images of the ligand-receptor complexes in the binding site from docking simulation were used for the MIA-QSAR modeling. Afterward, the pixels of aligned 2D images of the ligands were used as descriptors for making the QSAR model. The genetic algorithm was used to select pixel descriptors that are different for the various applied molecules. In fact, the constant or invariant parts of pixel descriptors are mainly removed by the genetic algorithm in the selection phase. In this way, the calibration method is very simple, and its statistical parameters are much better than the original version. As a result, the model's prediction ability in the prediction phase for predicting IC50 values is very close to actual values. Based on the obtained results, new molecules were designed. The *in silico* absorption, distribution, metabolism, excretion, and toxicity (ADMET) studies were done on the new molecules.

MATERIALS AND METHOD

Dataset

In this paper, three types of tacrine derivatives as AChE

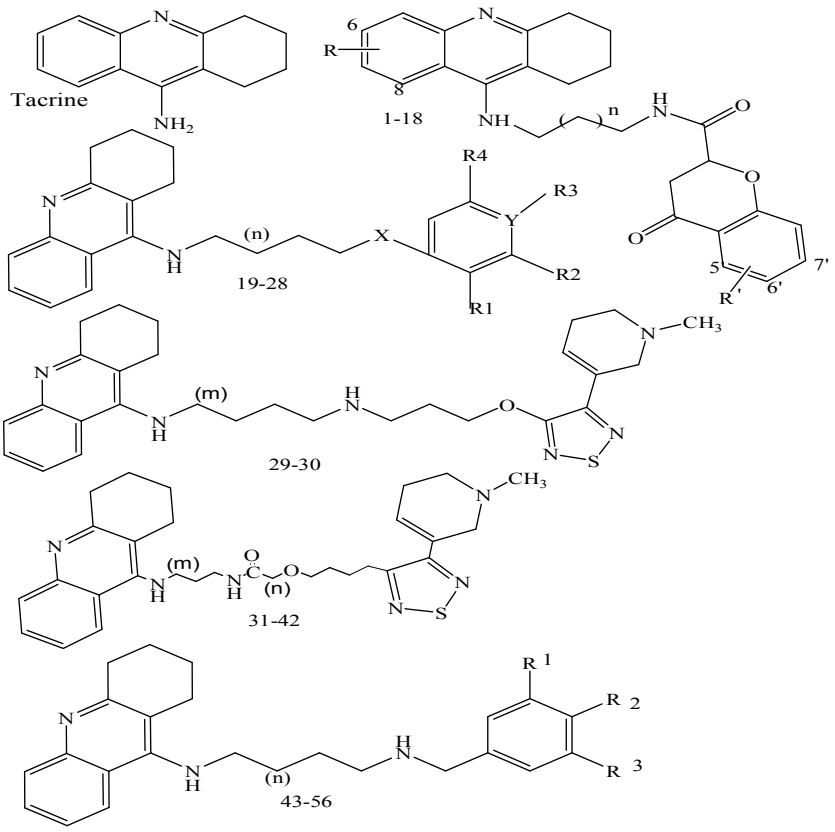

Fig. 1. Structures of the molecules (1-18) tacrine-4-Oxo-4H-chromene hybrids, (19-28 and 43-56) heterodimers of tacrine and substituted benzene derivatives and (29-42) tacrine-xanomeline dimers.

inhibitors, including Tacrine-4-Oxo-4H-chromene Hybrids, heterodimers of tacrine, substituted benzene derivatives, and tacrine-xanomeline dimers were extracted from the literature (Fig. 1) and used for MIA-QSAR and docking studies [31,32].

Moreover, inhibitory activity (IC_{50}) value has been converted to logarithmic scale pIC50 value, which was taken as the dependent variable for the MIA-QSAR study. The histogram of pIC_{50} of 56 compounds of tacrine derivatives showed an asymmetric distribution pattern (Fig. 2). This resulted in a rational description of data in MIA-QSAR modeling. Figure 3 shows the basic structures of inhibitors. First, the Kennard-Stone algorithm was performed on the data set to split the data into training and test set diversely. Then, data series diversity was checked using molecular diversity analysis.

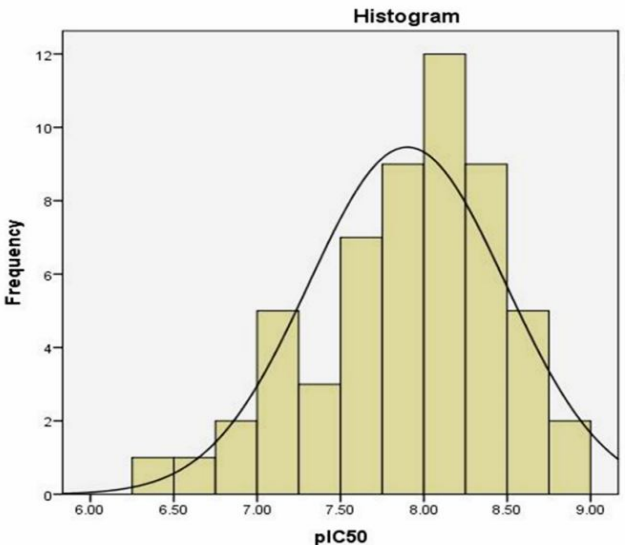

Fig. 2. The histogram of logIC50 of 56 compounds of tacrine derivatives.

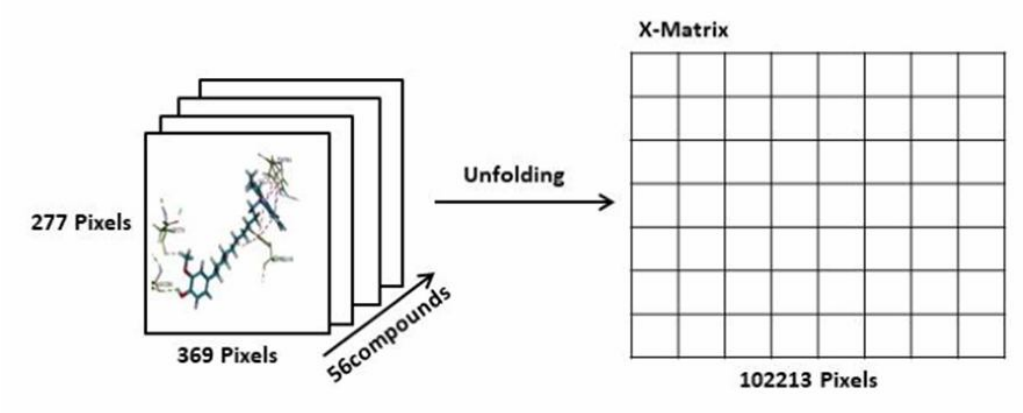

Fig. 3. Illustration of MIA-QSAR workflow.

Protein and Ligand Preparation

The functional groups of the molecular structures were drawn colorful in Marvin Sketch (ChemAxon) module and then were transferred to Discovery Studio 4 software. Also, the docking study was performed on the Ache monomeric units where the predestinate ligand was positioned. The alkylene-linked tacrine dimers with acetylcholinesterase enzyme X-ray crystallographic structure were achieved from the protein data bank (2CKM) 2.15 Å resolution [33] (http://www.pdb.org). Also, Discovery Studio 4.1 software was used to remove water molecules. Both protein and ligand files were prepared and saved in PDB format.

Molecular Docking Protocol

All structures of inhibitors were transferred into the Discovery Studio 4 software workspace. The structures were optimized using CHARMm force field [34]. Then, the partial charges of inhibitors were calculated by Momany-Rone [35,36] .The pH value of the protein was set to 7 andwater molecules were removed. A spherical graphics object with the specified center and radius around the active site was created. Other parameters were considered based on the default protocol settings. Then CDOCKER algorithm was used for molecular docking of inhibitors into the protein binding site. Afterward, the molecular docking was used in a specified space with a radius of 11.88 angstroms around the nonbonded ligand in the binding site. The radius must have a value to cover the ligand, and the involved residues have interactions with the ligand in the binding site. After performing CDOCKER individually for each of these ligands, all the ligands' first pose was chosen. Subsequently, screenshots were taken and transferred to a workspace in the windows paint program.

Descriptor Calculation with MIA-QSAR Method

All calculations were performed using a personal computer with the CPU Intel core i7 and 16 GB of RAM on the MATLAB software 2016b (MathWorks).

The pixels of ligand images that can be 2D or 3D were considered as MIA's input descriptors. These pixels were used for making QSAR models [37]. 2D images were saved as jpg file type and then overlaid as a 2D arrangement by taking a pixel on the specific coordinate. Then, each image was read in MATLAB software in 277 × 369 pixels window size, resulting in a 56 × 277 × 369 three-way dataset. The three-way dataset was extended to a 56 × 102213 matrix and the computer memory usage was minimized. Columns with a standard deviation of less than 20 were removed to generate the final input matrix (X) with 56 rows and 22829 columns.

Before using the SVM method on the dataset, PCA was carried out. Then, the PC scores were reduced via the GA procedure. Finally, 19 PC scores were chosen for model construction. The GA parameters were defined as the following: 1% the probability of mutation, 90 crossovers, 100 runs, and smoothing window size of 3. In addition, the preprocessing method (mean centering) was performed on the dataset to normalize them before performing SVM regression. After constructing the X matrix, in order to evaluate the performance of generated regression models,

about 20% of the molecules were selected as a test set using the Kennard-Stones algorithm (11 out of 56). Then, the SVM regression was used to correlate the X matrix to the activities, pIC50 (Y matrix). The modeling procedure, including Kennard-Stone, cross-validation, GA variable selection, and PCA, SVM, was performed on PLS Toolbox Version 6.5 (Eigenvector Research Inc.).

The main statistical parameters such as squared regression coefficients (r^2) for the training and leave one out cross-validation (Q^2 and LOO-CV) were calculated. Also, r^2_p and root mean square of external validation (RMSE) were calculated for the test set. Moreover, the Δr_m^2 K and mean absolute error (MAE) were also obtained.

Variable Selection and Modeling

A standard deviation higher than 20 was chosen in the preprocessing step. In addition, PCA, as a nonparametric method for orthogonal linear transformation, was performed on the bi-dimensional image descriptors [38,39]. This method projects the data on another coordinate system with fewer dimensions while preserving its characteristics [39]. The PCs are linear combinations of the original variables [40,41]. Each PC is orthogonal to the other [42]. Then, the PCs were reduced *via* the GA procedure. Finally, 19 PCs were selected for building the model. After that, the Kennard-Stones algorithm was performed on the input dataset to separate it into two individual sets, namely the training set (45 data) and the validation set (11 data). In the next step, the SVM method was employed to build the model.

Computational Theory of Support Vector Regression

SVM was first interduced in 1992 by Vapnik *et al.* as a promising classification and regression method [43]. In SVM regression, a nonlinear mapping function is applied to move the descriptors into a higher-dimensional space; in the next step, linear regression would be performed in the new space according to Eq. (1) [44,45]:

$$y = f(x) = < W, Q(x) > + b \tag{1}$$

where W is calculated using the following equation and b is the bias value;

$$W = \sum_{i=1}^{n} \alpha_i y_i x_i \tag{2}$$

$$\alpha_i = 2\gamma e_i \tag{3}$$

$$\varphi = \frac{1}{2} W^T w + c \frac{1}{I} \sum_{i=1}^{I} \varepsilon_i \tag{4}$$

In Eq. (1), Φ is the kernel function used for nonlinear mapping, α_i is the Lagrange multiplier, and w is the coefficients vector. Thus, the term of <w, Q(x)> defines the dot product of w and Φ(x).

The minimum value of the φ determines the optimal regression function [46,47]. The I is the number of training compounds, ε is the tolerance zone, and $1/2w^Tw$ is employed to determine the model complexity. C is the regularized constant, which determines the tradeoff between the model complexity and the empirical error. If C is a high value, it minimizes the empirical error for the training set that results in a model with the low generalization ability of chemicals in the test set. However, if C is too small, inadequate stress will be established on fitting the training data. The unknown data set noise determines the optimal value of ε. The Lagrange multipliers can provide the optimum value for ε, C, and kernel function parameters [46].

According to the definition of the error function, SVR is separated into two general regression categories. The error function can be a multinomial linear function called radial basis function (RBF). The RBF is the most commonly used function for QSAR modeling. The most general Kernel function is RBF which is defined as follows:

$$K(x_i, x_j) = e^{-\gamma \left\| x_i - x_{ij} \right\|^2} \tag{5}$$

where $K(x_i,x_j)$ is the kernel function, x_i is the input dataset, and γ is one of the kernel parameters [45].

Diversity Analysis

Molecular diversity analysis describes the behavior of target molecules to cover a certain structural space and triggers various tactics to select an appropriate compound. The diversity analysis can be utilized to make sure that both training and test sets are descriptive of the whole data set. In

diverse spaces, a precise method is performed to place molecules properly in a suitable parametrized way. An optimum metric space should be chosen for structural diversity demonstration, which is a key factor in the model efficiency. The Euclidean distance norm d_{ij} can be used to define a distance score for two individual compounds X_i and X_j in a database, including n compound generated from m significantly correlated chemical descriptors:

$$d_{ij} = \left\| x_i - x_j \right\| = \sqrt{\left(\sum_{k=1}^{m} (x_{ik} - x_{jk}) \right)^2} \quad (6)$$

where x_{ik} and x_{jk} are compound descriptors (Eq. (6)). Also, the mean distance from a target sample to the rest of the samples was calculated using the following method [48]:

$$d_i = \frac{\sum_{j=1}^{n} d_{ij}}{n-1} \qquad i = 1,\ldots,n \quad (7)$$

The normalized mean distances of chemical values were plotted against experimental $\log K_i$. These results indicate that the training set with a comprehensive illustration of the chemistry space was enough to ensure the stability of the model and appropriate distribution of the test set in the entire dataset. Therefore, the predictive ability of the model can be accurately determined (Fig. 4).

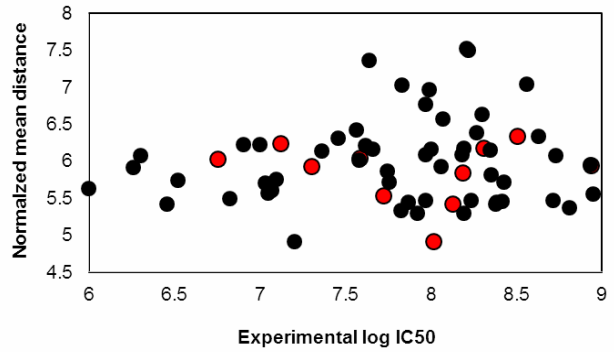

Fig. 4. The result of diversity analysis on AChE inhibitors.

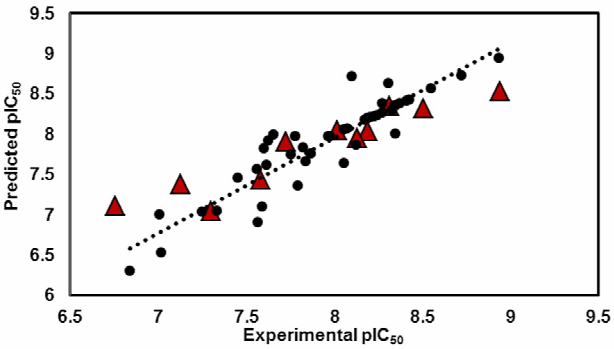

Fig. 5. Predictive *vs*. experimental pIC50 values derived from the SVR Model of the training (●) and test sets▲).

In silico ADME Profile Prediction

The ADMET properties were determined to predict the pharmacokinetic properties for selecting the effective bio compounds. This step precedes thesynthesis phase. There is a relationship between chemical structures and physiological parameters. Therefore, some chemical descriptors can be used to calculate pharmacokinetic properties. Calculations of important ADMET properties of tacrine derivatives including Atom-based logP98 (AlogP), 2D polar surface area (2D_PSA), plasma protein binding (PPB), hepatotoxicity, cytochrome P4502D6 (CYP2D6) binding, aqueous solubility, and blood-brain barrier (BBB) were performed by ADMET analysis using Discovery Studio 4.1 software.

RESULTS AND DISCUSSION

The SVR results for predicting the tacrine derivatives activities are listed in Table 1 (in Supplementary). It is shown that the RBF was used, and the capacity values (C), epsilon (ε), and the Kernel parameter (γ) were optimized. The optimum values for developing SVM model were obtained as ε = 0.01, and γ = 0.001. The model parameters including R_c^2, RMSE of calibration (RMSEC), R_p^2, and RMSEP were calculated (0.85, 0.25, 0.89, and 0.23, respectively). Also, q^2 value of 0.62 and RMSECV value of 0.36 were computed. Moreover, $bias^2$ (0.255), variance (0.005), and the mean square error (0.254) for prediction of $\log IC_{50}$ were obtained using bootstrap sampling (with 1000 bootstrap models) in bias-variance estimator software (Fig. 5). These results suggest a high predictive capability of the model.

Statistical Metrics to Examine the Quality of the Dveloped Model

After PCA, 56 scores were selected for building the

Table 1. Experimental and Redicted pIC50 Values for by SVR Model (Test Set)

Compound	Experimental pIC50	Predicted pIC50	Residual error	n; m; R; R'
M5Test	7.301	7.349	-0.048	n = 8; 6-Cl; 5,7-diOCH_3
M6 Test	8.187	8.03	0.157	n = 8; 6-Cl; 6,7-diOCH_3
M7Test	6.757	7.1046	-0.348	n = 8; 6,8-diCl
M14Test	7.125	7.377	-0.252	n = 8; 6,8-diCl; 5-OH
M19 Test	7.583	8.002	-0.419	n = 4 or 5, x = CH_2, Y = C, R1 = R4 = H, R2 = R3 = OMe
M24Test	8.939	8.632	0.307	n = 8, x = CH_2, Y = C, R1 = H, R4 = R2 = OMe, R3 = OH
M26 Test	8.506	8.31	0.196	n = 8, x = CH_2, Y = C, R1 = OH, R2 = R3 = R4 = H
M28 Test	8.016	8.048	-0.032	n = 8, x = C = O, Y = N
M44 Test	7.722	7.916	-0.194	R1-R2 = OCH_2, R3 = OMe, n = 3
M52 Test	8.312	8.319	-0.007	R1 = R2 = R3 = OMe, n = 7
M53 Test	8.127	7.951	0.176	R1 = R2 = OMe, R3 = H, n = 4

model. Then, model validation was achieved through LOO-CV and external validation. Golbraikh and Tropsha [49] have reported that for a QSAR model, the high value of Q^2 is essential for high predictive ability in a developed model. However, this condition is not an adequate reason for a built model to show high predictive ability. Thus, one can conclude that an internal validation, such as LOO-CV, should not be performed solely to define a model's predictive ability. Consequently, one should use a rougher external validation procedure by utilizing the molecules that are not present in the training set. One of the suitable external validation methods is the r_m^2 metrics which was introduced by Roy [50]. The r_m^2 determines the proximity between the observed and predicted activity. The r_m^2 metric can be measured for the both training and validation sets and includes the predicted training set and/or predicted validation set of the equivalent observed data. Currently, two various r_m^2 metrics can be calculated: r_m^2 , and Δr_m^2 [51,52]. They can be calculated using the following equations:

$$r_m^2 = \frac{r_m^2 + r_m'^2}{2} \tag{8}$$

$$\Delta r_m^2 = \left| r_m^2 + r_m'^2 \right| \tag{9}$$

$$r_m^2 = r^2\left(1 - \sqrt{\left(r^2 - r_0^2\right|)}\right) \text{ and } r_m'^2 = r'^2\left(1 - \sqrt{\left(r^2 - r_0'^2\right|)}\right) \tag{10}$$

Also, the predictive ability of the QSAR model was determined by calculating the "with intercept r^2" and "without intercept (r^2_0)" using Eq. (11).

$$q^2 > 0.5 \qquad r^2 > 0.6$$

$$\left[\frac{r^2 - r_0^2}{r^2}\right] < 0.1 \quad or \left[\frac{r^2 - r_0'^2}{r^2}\right] < 0.1 \tag{11}$$

The k or k' parameter indicates the slope of regression lines through the origin.

$$0.85 \le k \le 1.15 \text{ or } 0.85 \le K' \le 1.15 \tag{12}$$

It has been suggested that for a model, the value of r^2_m should be more than 0.5, and the Δr_m^2 value should be lower than 0.2. Moreover, RMSE and MAE are the most commonly used metrics for errors in QSAR. Here, the RMSEC, RMSECV, and RMSEP were used as statistical parameters, and also the mean average error (MAE) was calculated [52]. The better predictive ability of a model can be obtained at the lower MAE value. The following criteria indicate the prediction ability of a model:

$$MAE_{Test} \le 0.1 \times \text{training set range and } MAE_{Test} + 3 \times \sigma \le 0.2 \times \text{training set range} \tag{13}$$

Table 2. The Validation Parameters of the QSAR Models Estimated by BXternalValidationPlus 1.2 (https://sites.google.com/site/dtclabxvplus/)

Model biasness test	Systematic error result	Absent
	r^2Test(95% data)	0.8700
	r_0^2Test(95% data)	0.8622
Classical metric	$r_0'^2$Test(95% data)	0.6513
(after removing	Q^2_{F1}(95% data)	0.8714
5% data with	Q^2_{F2}(95% data)	0.8619
high residuals)	$\overline{r_m}^2$(95% data)	0.7428
	Δr_m^2(95% data)	0.1217
	CCC(95% data)	0.9193
	RMSEP(100% data)	0.2268
Error-based metrics	SD(100% data)	0.1142
(for 100% data)	SE(100% data)	0.0344
	MAE(100% data)	0.1989
	RMSEP(95% data)	0.2000
Error-based metric	SD(95% data)	0.0960
(after removing 5% data	SE(95% data)	0.0303
with high residuals)	MAE(95% data)	0.1781
	MAE+3*SD(95% data)	0.4660
BASIC DATA STRUCTURE INFORMATION		
RESULT (MAE-based criteria applied on 95% data)	Prediction quality	GOOD

MAE_{Test} is the prediction error among the 95% of test set compounds and where the σ value is related to the standard deviation of the absolute error values for the validation set. The data set and the model should be revised when the following condition applies, indicating the poor performance of a model [53,54]:

MAE > 0.15 × training set range or MAE + 3 × σ > 0.25 × training set range.

The MAE-based standard and other validation metrics were measured using an online tool, "BXternalValidationPlus" 1.2, presented by Roy et al. In addition, this program inspects whether a systematic error in the prediction has occurred.

The SVM model showed good predictive ability, and there was no detectable and systematic error in the prediction (Table S1 in the supplementary information). Finally, the calculated statistical parameters of the proposed MIA-QSAR model were:

$r_m^2{}_{train} = 0.63$, $\Delta r_m^2{}_{train} = 0.18$, $r_m^2{}_{test} = 0.69$, $\Delta r_m^2{}_{test} = 0.145$, $\Delta r_m'^2{}_{overll} = 0.63$, $\Delta r_m^2{}_{overll} = 0.19$, $r^2 - r_0^2/r^2 = 0.02$, $k = 1.006$.

Figure 6 shows the LOO-CV results corresponding to the SVM algorithm optimization. Each various colors show individual γ values representing the error levels. The brown and blue regions show higher and lower errors, respectively. The optimum conditions were C = 31.62, ε = 0.01, and γ = 0.001.

Molecular Design and Docking Studies

The SVM model was investigated to predict the inhibitory activity of three new tacrine derivatives. To the best of our knowledge, there is no report on any biological tests on these compounds. Figure 7 shows the molecular structures of the new compounds that their inhibitory

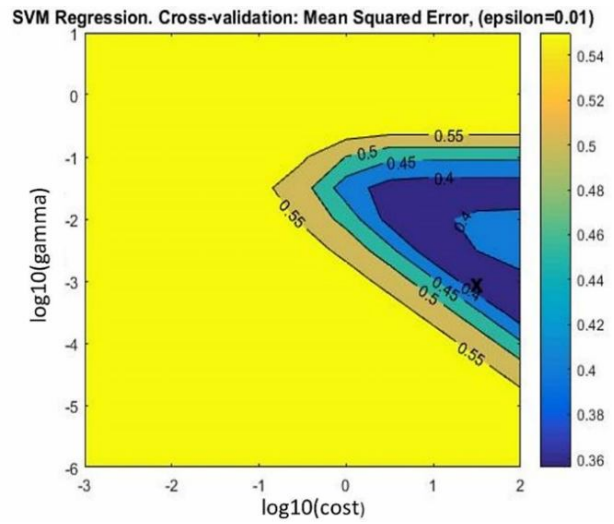

Fig. 6. Optimization of support vector machine parameters; a black star highlighted the optimized point.

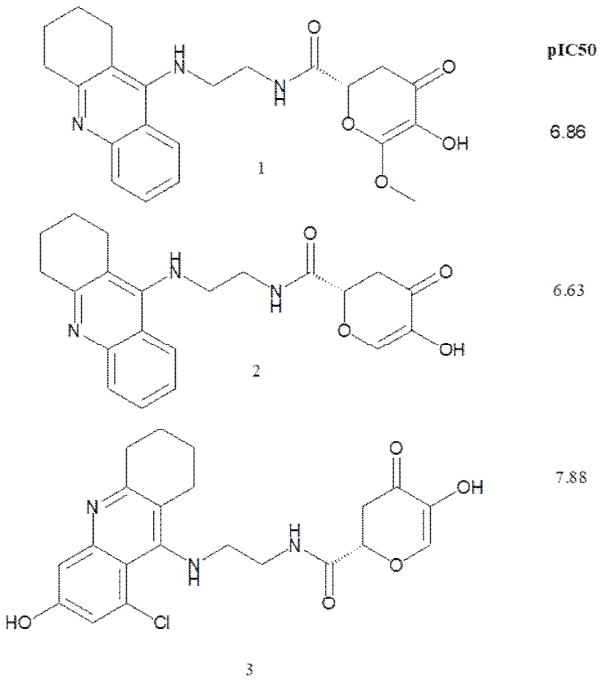

Fig. 7. Structures of the designed molecules.

activities were calculated.

In order to discover the precise conformation of all molecules and validate the main interaction in the enzyme active sites, docking simulations were used. The root means

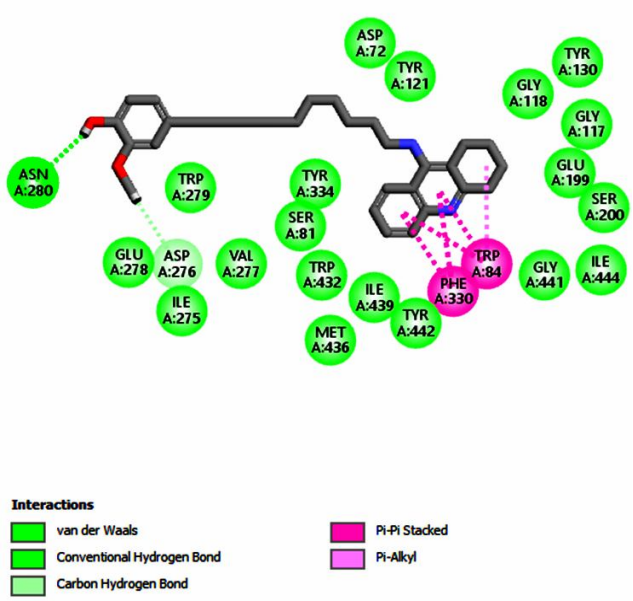

Fig. 8. Schematic docking interaction between compound 27 with AChE

square deviation (RMSD = 0.82 Å) of co-crystal ligand (2CKM), and the re-docked ligand was computed to evaluate the validity of the method. The RMSD value of 0.825 Å indicates great reliability of the method for rebuilding the empirical bonding mode for AChE inhibitors. It showed that the most active compound is compound 27, which was selected to obtain more information about the enzyme and inhibitors interactions. The stacked π-π interaction of component 27 and tryptophan and tyrosine is illustrated in Fig. 8. It is indicated that compound 27 has a decent interaction with the active site of the receptor. A hydrogen bond is formed between the hydrogen atom of the hydroxyl group and the carbonyl part of the asparagine group. It can be said that asparagine is among the charge-free polar R groups. The R groups of the non-polar amino acids compared to that of asparagine shows higher water solubility and hydrophilicity. The reason for that is their ability to form hydrogen bonds with water. The negative CDOCKER energy for compound 27 was 42.3 kcal mol^{-1} (Fig. 8).

The new molecules were computationally docked. The correct binding sites for each of the compounds were examined. Docking studies showed thatthe amino acids TYR70, TYR121, TYR334, TRP279, PHE288, PHE290,

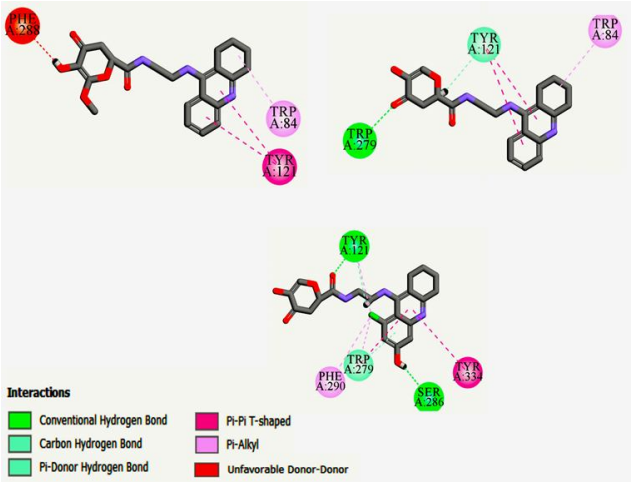

Fig. 9. The best-docked conformer of proposed molecules.

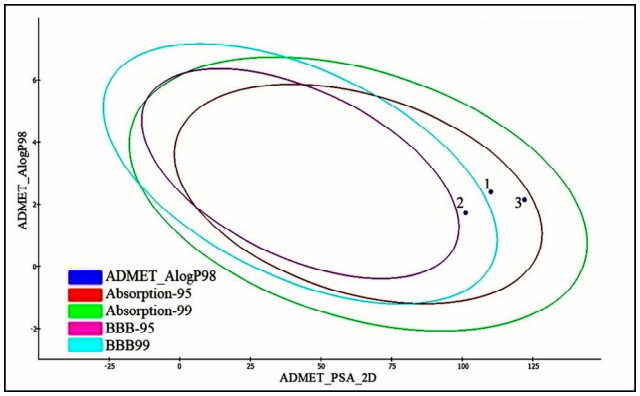

Fig. 10. ADMET Plot of the 2-D polar surface area (PSA_2D) *vs*. ALogP98 for three new molecules.

TRP84, TRP334, and SER286 are active amino acids in the complex (Fig. 9).

Tryptophan, which is among hydrophobic amino acids that engages in hydrophobic mutual reactions, appeared in the molecular docking of all new components. As shown in Fig. 7, tryptophan has formed a π-π bond with the benzene substitution ring and also the tacrine ring. Component 2 and component 3 demonstrate the stacked π-π bonds of tryptophan and tyrosine. Also, in component 2 and component 3, a carbon-hydrogen bond has been formed. These amino acids have aromatic R groups that are almost non-polar (hydrophobic), and they have aromatic side chains. All these groups can engage in mutual hydrophobic reactions. In component 3, serine amino acid has formed a hydrogen bond as a hydrophilic residue due to its hydroxyl groups. The negative CDOCKER energy amounts in these new compounds are 33.098, 35.32, and 25.33, respectively. According to the docking results, the main interactions between inhibitors and the protein's active site are hydrogen bonding, stacked π-π, and hydrophobic interactions.

ADMET Studies

It is reported that pharmacokinetic studies are one of the greatest challenging subjects in rational drug delivery. These properties depend on ADMET of selected drugs. The pharmacokinetics feature of the chemicals, including permeability, molecular weight, octanol-water coefficient, *etc*., can be determined by ADMET information [55,56].

In order to evaluate the drug-likeness, we studied the ADMET properties of new candidate molecules. The results are summarized in Table 3. Figure 10 shows acceptable values for the structures. According to these results, the new ligands are confirmed by drug‐likeness parameters. Moreover, the bipolar of AlogP versus 2D polar surface area (PSA) for new molecules is shown in Fig. 10. The blood-brain barrier penetration (BBP) was predicted using descriptors AlogP98 and 2D_PSA at 95% and 99% confidence ellipses. According to the summarized in silico ADMET parameters in Table 3 and biplot for studying drug-likeness of compounds, it was found that compound number 2 is a favorable compound for further assessment of *in vitro* and in vivo biological activity evaluation.

CONCLUSIONS

In this study, tacrine derivatives as AChE inhibitors were designed by using the structure-based MIA-QSAR modeling. A set of 56 compounds were used in order to construct the model. The 2D images were used as the raw input for the model. The ligand-protein interaction was studied by performing molecular docking, and the correct binding site for these sorts of compounds was examined. The docking's first pose images were transferred to the windows paint program for being applied in the MIA-QSAR method. Docking studies for the suggested compounds showed that hydrogen bonding and hydrophobic

Table 3. Prediction of ADME Properties of Designed Ligands

Molecular formula	Molecular weight	Solubility	BBB	CYP2D6	Hepatotoxic	PPB	AlogP98	PSA_2D
C22 H25 N3 O5	411.463	-4.117		-6.23902	1.65865	-5.87159	2.388	110.158
C21 H23 N3 O4	381.436	-3.399	-1.223	-5.82606	1.80017	-6.83591	1.723	101.228
C21 H22 Cl N3 O5	431.881	-4.083		-7.06706	2.08585	-5.29175	2.145	122.044

interactions are critical factors in the relationship between the inhibitors and the receptor. Docking studies showed that these compounds have a good interaction with the active receptor site. According to the results, PC-GA-SVM modeling coupled with MIA-QSAR showed versatility and has a high potential predictive ability which can be used to predict the pIC50 values. Furthermore, the *in silico* ADMET studies were performed on the new molecules. As a result, the computed values of ADMET descriptors showed that all properties are within the expected ranges.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the support of this work from Islamic Azad University, Arak Branch, and the University of Tehran.

REFERENCES

[1] Van de Waterbeemd, H.; Camenisch, G.; Folkers, G.; Raevsky, O. A., Estimation of Caco-2 cell Permeability using calculated molecular descriptors. *Quant. Struct.-Act. Relat.*, **1996**, *15*, 480. DOI: 10.1002/qsar.19960150604.

[2] Oprea, T. I.; Gottfries, J., Toward minimalistic modeling of oral drug absorption. *J. Mol. Graph. Mod.,* **1999**, *17*, 261. DOI: 10.1016/s1093-3263(99)00034-0.

[3] Stenberg, P.; Luthman, K.; Artursson, P., Prediction of membrane permeability to peptides from calculated dynamic molecular surface properties. *Pharm. Res.,* **1999**, *16*, 205. DOI: 10.1023/A:1018816122458.

[4] Zamora, I.; Ungell, A. L., Correlation between drug absorption and molecular surface descriptors: Comparison between different experimental models. *Eur. J. Pharm. Sci.,* **1996**, 85, 32-39. DOI: 10.1021/JS950285R.

[5] Oprea, T. I.; Zamora, I.; Ungell, A. U., Pharmacokinetic based mapping device for chemical space navigation. *J. Comb. Chem.,* **2002**, *4*. DOI: 10.1021/CC010093W.

[6] Muthukumaran, P.; Rajiniraja, M., Aug-MIA-QSAR based strategy in bioactivity prediction of a series of flavonoid derivatives as HIV-1 inhibitors. *J. Theor. Biol.,* **2019**, *469*, 18-24. DOI: 10.1016/j.jtbi.2019.02.019.

[7] Banjare, P.; Matore, B.; Singh, J.; Roy, P. P., *In Silico* local QSAR modeling of bioconcentration factor of organophosphate pesticides. *Silico Pharmacol.*, **2021**, *9*, 1-13. DOI: 10.1007/s40203-021-00087-w.

[8] amraei, A.; Niazi, A.; Alimoradi, M.; Delfan, B., Pixel selection by successive projections algorithm method in multivariate image analysis for a QSAR study of antimicrobial activity for cephalosporins and design new cephalosporins. *Iran. J. Pharm. Res.,* **2018**, *17*, 1240-1248. DOI: 10.22037/ijpr.2018.2286.

[9] Bitencourt, M.; Freitas, M. P., MIA-QSAR evaluation of a series of sulfonylurea herbicides. *Pest Manag. Sci.,* **2008**, *64*, 800-807. DOI: 10.1002/ps.1565.

[10] Freitas, M. P. D.; Duarte, M. H., Evolution of Multivariate Image Analysis in QSAR: The Case for a Neglected Disease. In Quantitative Structure-Activity Relationships in Drug Design, Predictive Toxicology, and Risk Assessment; IGI Global: Federal University of Lavras, Brazil, **2015**; pp. 84-122. DOI: 10.4018/978-1-4666-8136-1.ch003.

[11] Bitencourt, M.; Freitas, M. P.; Rittner, R., The MIA-QSAR method for the prediction of bioactivities of

possible acetylcholinesterase inhibitors. *Arch. Pharm. (Weinheim).,* **2012**, *345*, 723-728. DOI: 10.1002/ardp.201200079.

[12] Xu, M. L.; Wang, X.; Li, B. Q.; Zhai, H. L.; Lu, S. H., Application of image moments in MIA-QSAR. *J. Chemom.,* **2018**, *32*, e2958. DOI: 10.1002/cem.2958.

[13] Barigye, S. J.; Freitas, M. P., 2D-Discrete fourier transform: Generalization of the MIA-QSAR strategy in molecular modeling. *Chemom. Intell. Lab. Syst.,* **2015**, *143*, 79-84. DOI: 10.1016/j.chemolab.2015.02.020.

[14] Muthukumaran, P.; Rajiniraja, M., MIA-QSAR based model for bioactivity prediction of flavonoid derivatives as acetylcholinesterase inhibitors. *J. Theor. Biol.,* **2018**, *459*, 103-110. DOI: 10.1016/j.jtbi.2018.09.030.

[15] Daré, J. K.; Ramalho, T. C.; Freitas, M. P., 3D Perspective into MIA-QSAR: A Case for anti-HCV agents. *Chem. Biol. Drug Des.,* **2018**, *93*, 1096-1104, DOI: 10.1111/cbdd.13440.

[16] Daré, J. K.; Freitas, M. P., Different approaches to encode and model 3D information in a MIA-QSAR perspective. *Chemom. Intell. Lab. Syst.,* **2021**, *212*, 104286. DOI: 10.1016/j.chemolab.2021.104286.

[17] Dvir, H.; Silman, I.; Harel, M.; Rosenberry, T. L.; Sussman, J. L., Acetylcholinesterase: From 3D structure to function. *Chem. Biol. Interact.*, **2010**, *187*, 10-22. DOI: 10.1016/j.cbi.2010.01.042.

[18] Sanson, B.; Colletier, J. P.; Xu, Y.; Lang, P. T.; Jiang, H.; Silman, I.; Sussman, J. L.; Weik, M., Backdoor opening mechanism in acetylcholinesterase based on X-ray crystallography and molecular dynamics simulations. *Protein Sci.,* **2011**, *20*, 1114-1118. DOI: 10.1002/pro.661.

[19] Zhou, Y.; Wang, S.; Zhang, Y., Catalytic reaction mechanism of acetylcholinesterase determined by born-oppenheimer *ab initio* QM/MM molecular dynamics simulations. *J. Phys. Chem. B,* **2010**, *114*, 8817-8825. DOI: 10.1021/jp104258d.

[20] Akula, N.; Lecanu, L.; Greeson, J.; Papadopoulos, V., 3D QSAR studies of AChE inhibitors based on molecular docking scores and CoMFA. *Bioorg. Med. Chem. Lett.,* **2006**, *16*, 6277-6280. DOI: 10.1016/j.bmcl.2006.09.030.

[21] Ansari, F.; Ghasemi, J. B.; Niazi, A., Three dimensional quantitative structure activity relationship and pharmacophore modeling of tacrine derivatives as acetylcholinesterase inhibitors in alzheimer's treatment. *Med. Chem.,* **2019**. DOI: 10.2174/1573406415666190513100646.

[22] Barnard, E. A., Neuromuscular Transmission-Enzymatic Destruction of Acetylcholine. In The Peripheral Nervous System; Hubbard, J. I., Ed.; Springer US: Boston, MA, **1974**, pp. 201-224. DOI: 10.1007/978-1-4615-8699-9_9 LB-Barnard1974.

[23] Cheung, J.; Rudolph, M. J.; Burshteyn, F.; Cassidy, M. S.; Gary, E. N.; Love, J.; Franklin, M. C.; Height, J. J., Structures of human acetylcholinesterase in complex with pharmacologically important ligands. *J. Med. Chem.*, **2012**, *55*, 10282-10286. DOI: 10.1021/jm300871x.

[24] Carlier, P. R.; Han, Y. F.; Chow, E. S. H.; Li, C. P. L.; Wang, H.; Lieu, T. X.; Wong, H. S.; Pang, Y. P., Evaluation of short-tether bis-THA AChE inhibitors. A further test of the dual binding site hypothesis. *Bioorg. Med. Chem.,* **1999**, *7*, 351. DOI: 10.1016/S0968-0896(98)00213-2.

[25] Sirin, G. S.; Zhou, Y.; Lior-Hoffmann, L.; Wang, S.; Zhang, Y., Aging mechanism of soman inhibited acetylcholinesterase. *J. Phys. Chem. B,* **2012**, *116*, 12199-12207. DOI: 10.1021/jp307790v.

[26] Shao, D.; Zou, C.; Luo, C.; Tang, X.; Li, Y., Synthesis and evaluation of tacrine-E2020 hybrids as acetylcholinesterase inhibitors for the treatment of alzheimer's disease. *Bioorg. Med. Chem. Lett.,* **2004**, *14*, 4639-4642. DOI: 10.1016/j.bmcl.2004.07.005.

[27] Girek, M.; Szymański, P., Tacrine hybrids as multi-target-directed ligands in alzheimer's disease: Influence of chemical structures on biological activities. *Chem. Pap.,* **2019**, *73*, 269-289. DOI: 10.1007/s11696-018-0590-8 LB-Girek2019.

[28] Eagger, S. A.; Levy, R.; Sahakian, B. J., Tacrine in alzheimer's disease. *Lancet,* **1991**, *337*, 989-992. DOI: 10.1016/0140-6736(91)92656-M.

[29] Wold, S.; Esbensen, K.; Geladi, P., Principal component analysis. *Chemom. Intell. Lab. Syst.,* **1987**, *2*, 37-52. DOI: 10.1016/0169-7439(87)80084-9.

[30] Chen, J. J. F.; Visco Jr., D. P., Identifying novel factor

XIIa inhibitors with PCA-GA-SVM developed VHTS models. *Eur. J. Med. Chem.,* **2017**, *140*, 31-41. DOI: 10.1016/j.ejmech.2017.08.056.

[31] Fernández-Bachiller, M. I.; Pérez, C.; Monjas, L.; Rademann, J.; Rodríguez-Franco, M. I., New Tacrine-4-Oxo-4H-chromene hybrids as multifunctional agents for the treatment of alzheimer's disease, with cholinergic, antioxidant, and β-amyloid-reducing properties. *J. Med. Chem.,* **2012**, *55*, 1303-1317. DOI: 10.1021/jm201460y.

[32] Vitorovic-Todorovic, M. D.; Cvijetic, I. N.; Juranic, I. O.; Drakulic, B. J., The 3D-QSAR study of 110 diverse, dual binding, acetylcholinesterase inhibitors based on alignment independent descriptors (GRIND-2). The effects of conformation on predictive power and interpretability of the models. *J. Mol. Graph. Model,* **2012**, *38*, 194-210. DOI: 10.1016/j.jmgm.2012.08.001.

[33] Rydberg, E. H.; Brumshtein, B.; Greenblatt, H. M.; Wong, D. M.; Shaya, D.; Williams, L. D.; Carlier, P. R.; Pang, Y. -P.; Silman, I.; Sussman, J. L., Complexes of alkylene-linked tacrine dimers with torpedo californica acetylcholinesterase: Binding of bis(5)-tacrine produces a dramatic rearrangement in the active-site gorge. *J. Med. Chem.,* **2006**, *49*, 5491-5500. DOI: 10.1021/jm060164b.

[34] Makhouri, F. R.; Ghasemi, J. B., *In silico* studies in drug research against neurodegenerative diseases. *Curr. Neuropharmacol.,* **2018**, *16*, 664-725. DOI: 10.2174/ 1570159X15666170823095628.

[35] Williams, D. H.; Cox, J. P. L.; Doig, A. J.; Gardner, M.; Gerhard, U.; Kaye, P. T.; Lal, A. R.; Nicholls, I. A.; Salter, C. J.; Mitchell, R. C., Toward the semiquantitative estimation of binding constants. Guides for peptide-peptide binding in aqueous solution. *J. Am. Chem. Soc.,* **1991**, *113*, 7020. DOI: 10.1021/ja00018a047.

[36] Frank, A. M.; Rebecca, R., Validation of the general purpose QUANTA ®3.2/CHARMm® force field. *J. Comput. Chem.,* **1992**, *13*, 888-900. DOI: 10.1002/ jcc.540130714.

[37] Khorshidi, N.; Sarkhosh, M.; Niazi, A., QSPR study of maximum absorption wavelength of various flavones by multivariate image analysis and principal components-least squares support vector machine. *J. Sci. Innov. Res.,* **2014**, *3*, 189-202. DOI: 10.1016/j.chemolab.2021.104286.

[38] Tipping Michael, E.; Bishop Christopher, M., Probabilistic principal component analysis. J. R. *Stat. Soc. Ser. B (Statistical Methodol.,* **2002**, *61*, 611-622. DOI: 10.1111/1467-9868.00196.

[39] Thomaz, C. E.; Giraldi, G. A., A New ranking method for principal components analysis and its application to face image analysis. *Image Vis. Comput.,* **2010**, *28*, 902-913. DOI: 10.1016/j.imavis.2009.11.005.

[40] Martínez-Archundia, M.; Correa-Basurto, J.; Montaño, S.; Rosas-Trigueros, J. L., Studying the collective motions of the adenosine A2A receptor as a result of ligand binding using principal component analysis. *J. Biomol. Struct. Dyn.,* **2019**, 1-16. DOI: 10.1080/ 07391102.2018.1564700.

[41] Ghafouri, H.; Ranjbar, M.; Sakhteman, A., 3D-QSAR studies of some reversible acetyl cholinesterase inhibitors based on CoMFA and ligand protein interaction fingerprints using PC-LS-SVM and PLS-LS-SVM. *Comput. Biol. Chem.,* **2017**, *69*, 19-27. DOI: 10.1016/j.compbiolchem.2017.05.001.

[42] Giuliani, A., The application of principal component analysis to drug discovery and biomedical data. *Drug Discov. Today,* **2017**, *22*, 1069-1076. DOI: 10.1016/j.drudis.2017.01.005.

[43] Pirhadi, S.; Shiri, F.; Ghasemi, J. B., Pharmacophore elucidation and 3D-QSAR analysis of a new class of highly potent inhibitors of acid ceramidase based on maximum common substructure and field fit alignment methods. *J. Iran. Chem. Soc.,* **2014**, *11*, 1329-1336. DOI: 10.1007/s13738-013-0402-6 LB-Pirhadi2014.

[44] Samghani, K.; HosseinFatemi, M., Developing a support vector machine based QSPR model for prediction of half-life of some herbicides. *Ecotoxicol Env. Saf.,* **2016**, *129*, 10-15. DOI: 10.1016/ j.ecoenv.2016.03.002.

[45] Zhang, X.; Amin, E. A., Highly predictive support vector machine (SVM) models for anthrax toxin lethal factor (LF) inhibitors. *J. Mol. Graph. Model*, **2016**, *63*, 22-28. DOI: 10.1016/j.jmgm.2015.11.008.

[46] Huang, X.; Maier, A.; Hornegger, J.; Suykens, J. A.

K., Indefinite kernels in least squares support vector machines and principal component analysis. *Appl. Comput. Harmon. Anal.*, **2017**, *43*, 162-172. DOI: 10.1016/j.acha.2016.09.001.

[47] Guenther, N.; Schonlau, M., Support vector machines. *Stata J.,* **2016**, *16*, 917-937. DOI: 10.1177/1536867X1601600407.

[48] Gholami Rostami, E.; Fatemi, M. H., Molecular docking and receptor-based QASR studies on pyrimidine derivatives as potential phosphodiesterase 10A inhibitors. *Struct. Chem.,* **2019**. DOI: 10.1007/s11224-019-01353-6.

[49] Golbraikh, A.; Tropsha, A., Beware of Q2! *J. Mol. Graph. Model.,* **2002***, 20,* 269-276. *DOI:* 10.1016/S1093-3263(01)00123-1.

[50] Roy, K.; Chakraborty, P.; Mitra, I.; Ojha, P. K.; Kar, S.; Das, R. N., Some case studies on application of "Rm2" metrics for judging quality of quantitative structure-activity relationship predictions: Emphasis on scaling of response data. *J. Comput. Chem.,* **2013**, *34*, 1071-1082. DOI: 10.1002/jcc.23231.

[51] Roy, K.; Ghosh, G., QSTR with extended topochemical atom indices. 4. Modeling of the acute toxicity of phenylsulfonyl carboxylates to vibrio fischeri using principal component factor analysis and principal component regression analysis. *QSAR Comb. Sci.,* **2004**, *23*, 526-535. DOI: 10.1002/qsar.200430891.

[52] Roy, K.; Das, R. N.; Ambure, P.; Aher, R. B., Be aware of error measures. Further studies on validation of predictive QSAR models. *Chemom. Intell. Lab. Syst.,* **2016**, *152*, 18-33. DOI: 10.1016/j.chemolab.2016.01.008.

[53] Hariri, S.; Ghasemi, J. B.; Shirini, F.; Rasti, B., Probing the origin of dihydrofolate reductase inhibition *via* proteochemometric modeling. *J. Chemom*., 0, e3090. DOI: 10.1002/cem.3090.

[54] Bagheban Shahri, F.; Niazi, A.; Akrami, A., Application of wavelet and genetic algorithms for QSAR study on 5-lipoxygenase inhibitors and design new compounds. *J. Mex. Chem. Soc.*, **2015**, *59*, 203-210. DOI: 10.1080/1062936X.2021.1950832.

[55] Daoud, I.; Melkemi, N.; Salah, T.; Ghalem, S., Combined QSAR, molecular docking and molecular Dynamics Study on New Acetylcholinesterase and Butyrylcholinesterase Inhibitors. *Comput. Biol. Chem.,* **2018**, *74*, 304-326. DOI: 10.1016/j.compbiolchem.2018.03.021.

[56] Pirhadi, S.; Ghasemi, J. B., Pharmacophore identification, molecular docking, virtual screening, and *in silico* ADME studies of non-nucleoside reverse transcriptase inhibitors. *Mol. Inform.,* **2012**, *31*, 856-866. DOI: 10.1002/minf.201200018.