Molecular Modeling Studies of Anti-Alzheimer Agents Using QSAR, Molecular Docking and Molecular Dynamic Simulations Techniques

Running title: Molecular modeling studies of anti-Alzheimer agents

Keyword: Alzheimer's disease, CoMFA, CoMSIA, HQSAR, pyrimidinylthiourea derivatives, molecular docking

.

Title Page

Molecular modeling studies of anti-Alzheimer agents using QSAR, molecular docking and molecular dynamic simulations techniques

Rahman Abdizadeh[a], Farzin Hadizadeh[b,c], Tooba Abdizadeh[d,e*]

[a]Department of Medical Parasitology and Mycology, Faculty of Medicine, Shahrekord University of Medical Sciences, Shahrekord, Iran

[b]Biotechnology Research Center, Mashhad University of Medical Sciences, Mashhad, Iran

[c]Department of Medicinal Chemistry, School of Pharmacy, Mashhad University of Medical Sciences, Mashhad, Iran

[d]Clinical Biochemistry Research Center, Basic Health Sciences Institute, Sharekord University of Medicinal Sciences, shahrekord, Iran

[e]Department of Medicinal Chemistry, Damghan Branch, Islamic Azad University, Damghan, Iran

*Corresponding author: Clinical Biochemistry Research Center, Basic Health Sciences Institute, Sharekord University of Medicinal Sciences, shahrekord, Iran; phone: 983833346692, fax: 983833330709; t.abdizadeh@gmail.com

Abstract: ***Background:*** Acetylcholinesterase (AChE) is one of the most important targets in the treatment of Alzheimer's disease (AD). It was claimed that novel AChE inhibitors were optimized as potential drug candidates, designed for regional or systematic release, and created as significant inhibitors.

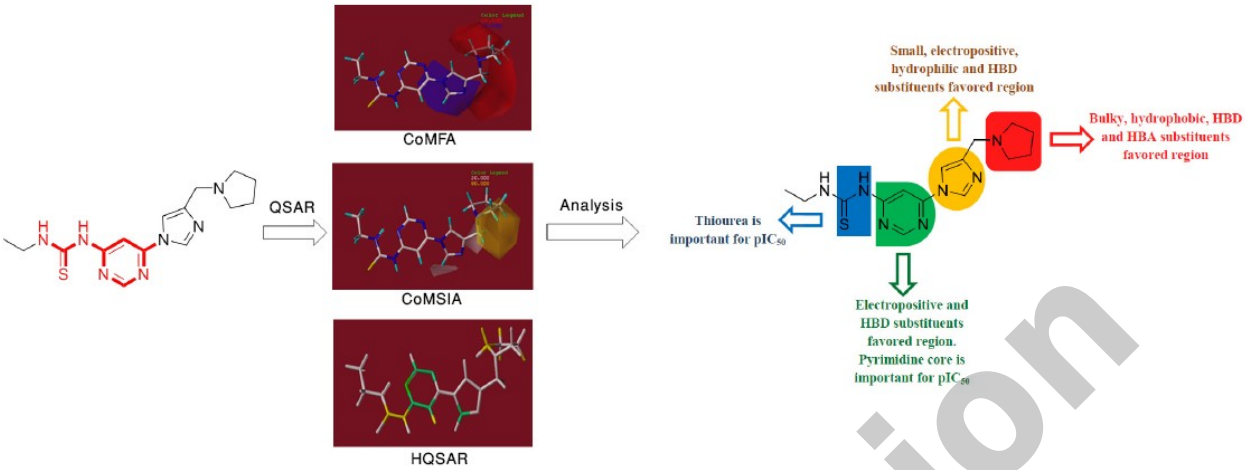
Objective: In this work, molecular modeling studies including CoMFA, CoMFA-RF, CoMSIA , HQSAR and molecular docking and molecular dynamic simulations were used to provide a theoretical basis for finding highly potent anti-Alzheimer drugs.

Methods: QSAR was used to generate models and predict the anti-Alzheimer activity using the Sybyl program (x1.2 version). pyrimidinylthiourea derivatives as AChE inhibitors were selected as our data set, which was split randomly into training and test sets. Docking and molecular dynamic simulation were carried out using the MOE software and the Sybyl program, respectively. Partial least square was used as QSAR model-generation method. The statistical qualities of generated models were justified by internal and external validation i.e., cross-validated correlation coefficient (q^2), non-cross-validated correlation coefficient () and predicted correlation coefficient (), respectively.

Results: The CoMFA (q^2, 0.775;, 0.901; 0.773), CoMFA-RF (q^2, 0.629;, 0.901; 0.824), CoMSIA (q^2, 0.754;, 0.919; 0.874) and HQSAR models (q^2, 0.622;, 0.949; 0.854) for training and test set yielded significant statistical results.

Conclusion: These QSAR models were excellent, robust and had good predictive capability. Contour maps of the QSAR models were generated and validated by molecular dynamic simulation-assisted molecular docking study. The final QSAR models could be useful for design and development of novel potent AChE inhibitors in Alzheimer's treatment.

Graphical abstract



Keyword: Alzheimer's disease; CoMFA; CoMSIA; HQSAR; pyrimidinylthiourea derivatives; molecular docking.

1. Introduction

Alzheimer's disease (AD), the most common form of dementia among elderly people, is a chronic, fatal, and neurodegenerative disease in the brain which related to loss of memory and behavioral function [1-3].

It is estimated that 44 million people suffer from AD worldwide and the number is increasing annually [4, 5]. With the growth of geriatric population, AD has become an urgent public health issue and had turned into a major social and financial burden. Thus, more efforts are needed to explore better ways for drug discovery to treat the AD disease and prevent its development [6-8]. Pathologically, β-Amyloid oligomerization, τ-protein aggregation, oxidative stress and low levels of acetylcholine (ACh) play key roles in the progress and development of AD [9-13]. Acetylcholine (ACh), a major neurotransmitter widely exiting in the brain, plays an important role in neural system and is synthesized and hydrolyzed by acetylcholinetransferase (AChT) and acetylcholinesterase (AChE), respectively [14-16]. Acetylcholinesterase (AChE) is a key target which is responsible for the metabolic breakdown of ACh. Accordingly, the inhibition of acetylcholinesterase (AChE) enzyme has been regarded as one of the most promising approaches for Alzheimer's treatment [17, 18].

The anti-AChE drugs such as donepezil, galanthamine, tacrine, rivastigmine and one N-methyl-D-aspartic acid (NMDA) antagonist memantine (Fig. 1) increase the ACh concentration in the synaptic cleft, but with limited success and efficiency in the current clinical therapy for AD [19-22].

Fig.1. Several AChE inhibitors that have been approved by FDA for the treatment of AD.

Therefore, it is necessary the design and development of new AChE inhibitors with potential application and low side effects for the treatment of AD.

The x-ray crystallographic structure of AChE contains two binding sites: the catalytic active site (CAS) at the bottom, and a peripheral anionic site (PAS) at the entrance of a long (20 Å) and narrow gorge [23-25]. It was proved that the dual binding site AChE inhibitors with action on PAS and CAS were potent inhibitors compared to the compounds that interact with only one site of the enzyme [26].

Quantitative structure- activity relationship (QSAR) is a technique that is used in computer-assisted rational drug design and predicts the protein-ligand interaction and to explore correlation between biological activity and molecular structure [27-30]. Three-dimensional QSAR (3D-QSAR) is a broad term encompassing all those QSAR methods which is utilized to calculate the highly specific interactions and a molecule, how far and with how much power can be connected to the active site of an enzyme or protein [31-33]. Recently, comparative molecular field analysis (CoMFA), CoMFA region focusing (CoMFA-RF), comparative molecular similarity index analysis (CoMSIA) and hologram QSAR (HQSAR) are especially effective methods of QSAR based on statistical techniques [34-36]. The CoMFA model proposed by Cramer et al. describes the molecular properties by steric (Lennard-Jones) and electrostatic (Coulomb) energy fields of important regions of a set of aligned compounds that predict their biological activity over a lattice of point [37, 38]. In CoMFA-RF model, steric and electrostatic fields are calculated for aligned fragments by creating specific grid space at the specific lattice points [39]. In CoMSIA model, proposed by Klebe et al., a probe atom is used to calculate similarity indices, at regularly placed grid points for the aligned molecules. Compared to CoMFA, CoMSIA uses a Gaussian-type distance-dependent function to assess five fields of different physicochemical properties (i.e., steric, electrostatic, hydrophobic, hydrogen binding donor and acceptor [40]. Also, CoMSIA is differentiated by Gaussian functions and no arbitrary definitions of cut off limits should be used.

HQSAR study is a comparatively new 2D-QSAR method which employs the fragment fingerprints of molecular holograms and other molecular descriptors to predict the biological activity of a series of molecules [41-43]. In these models, all regression analyses performed in two steps using the partial least squares (PLS) method [44-47].

In the present study, we performed a molecular modeling study by combined 2D- and 3D-QSAR, molecular docking and molecular dynamic (MD) simulations techniques. 2D-QSAR, using HQSAR method, and 3D-QSAR, using

CoMFA, CoMFA-RF and CoMSIA methods, were used to identify the key structural factors influence on inhibitory activity. Molecular docking was used to identify some key amino acid residues at the active site of AChE protein and investigate the binding modes between AChE and the selected inhibitors. Molecular dynamic (MD) simulations were employed to determine the detailed interactions in AChE protein and validate the rationality of docking results. The obtained results can apply to the further structural modification, design and development new and more potent anti-Alzheimer drugs.

2. Materials and Methods

2.1. Data Set

QSAR studies were performed on a set of 40 pyrimidinylthiourea derivatives as a new class of anti- Alzheimer agents with their biological activities (IC_{50} values) that recently reported by Li group [48].

These activity values (IC_{50} in μM) were converted to corresponding pIC_{50} (-log IC_{50}) values and used as a dependent variable in CoMFA, CoMFA-RF, CoMSIA and HQSAR models. The data set was randomly divided into a training set (30 compounds, 75%) for QSAR model generation and a test set (10 compounds, 25%) for external validation of the models (Fig. 2).

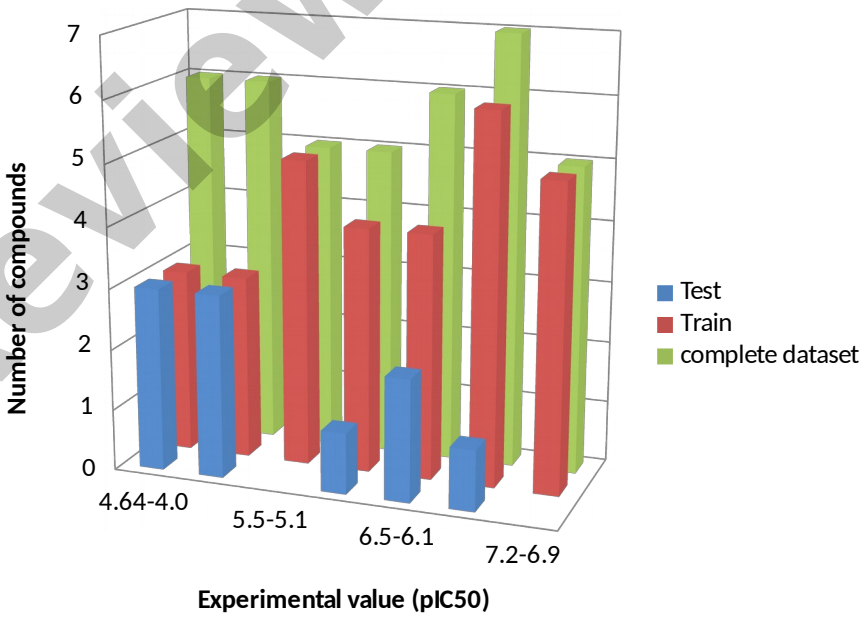


Fig. 2. Distribution of experimental inhibitory activities (pIC_{50}) for the training and test sets compounds in the QSAR models.

2.2. Molecular Modeling and Alignment

The QSAR models including CoMFA, CoMFA-RF, CoMSIA and HQSAR were performed using the SYBYL-X 1.2. molecular modeling software (Tripos, Inc, St. Louis, MO). Before modeling with these primary methods, the 3D structures of compounds were drawn using Chemoffice Bio 3D Ultra (version 12.0, Cambridge Soft Corporation, Cambridge, UK, 2010). All the compounds were energy minimized using the standard molecular mechanics force field with a distance dependent dielectric and the powell conjugate gradient algorithm with a convergence criterion of 0.05 kcal/molÅ using the maximum iteration set to 5000 [49]. Partial atomic charges of the compounds for electrostatic interactions were calculated by the Gasteiger-Hückel method. Structure alignment was one of the most important input variables in 3D-QSAR analysis and the accuracy of the prediction power of the models was reliability dependent on contour maps according to the structural alignment of the molecules. In this study, rigid body alignment of molecules in a Mol2 database was performed using maximum common substructures defined by Distill alignment. Compound **36** was selected as template because the most active compound of data set and other compounds were aligned according to the common structure. The structure of compound **36** with bold red common substructure and final super imposition of compounds are shown in Fig. 3a, 3b.

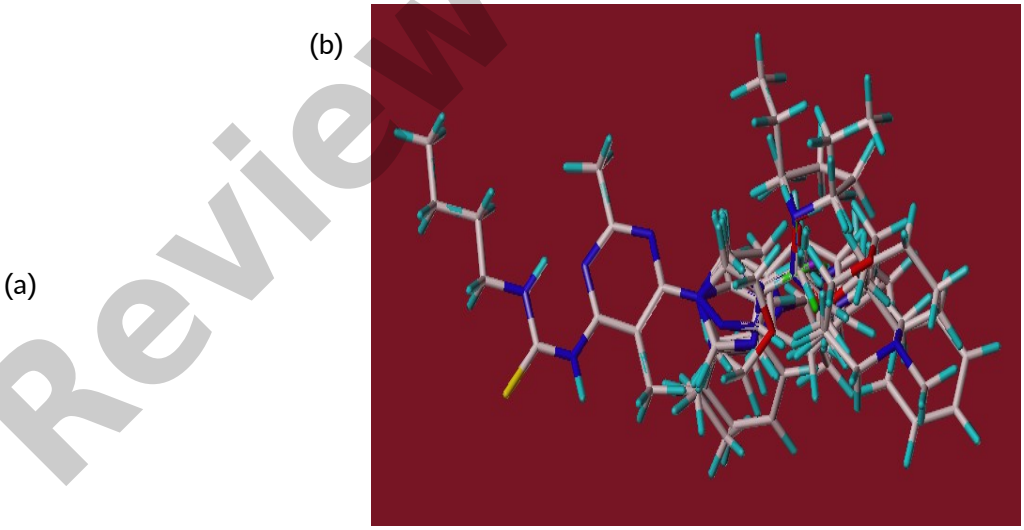


Fig. 3. Compound **36** used as the template molecule for database alignment and Common substructure in Distill alignment shown in the bold red (a) and aligned compounds in the training and test sets (b).

2.3. CoMFA and CoMSIA Analysis

In CoMFA method, the aligned molecules in optimal orientation were located in a 3D cubic lattice with grid spacing of 2 Å in the x, y, and z direction which extended 4.0 Å around the align molecules in all Cartesian directions. The CoMFA steric and electrostatic fields were calculated for each molecule using a hybridized sp^3 carbon probe atom with a Vander Waals radius of 1.52 Å and a charge of +1.0. The Coulomb and Lennard-Jones potential functions were used to estimate the electrostatic and steric interactions, respectively. The energy cut off values for both steric and electrostatic fields were set at 30 kcal/mol. In order to reduce noise and improve efficiency, column filtering was tested in the range of 0.0 to 2.0 kcal/mol and a threshold column filtering value of 2.0 kcal/mol. CoMFA-RF in the "Advanced CoMFA" module is a technique of application of weight to the lattice point in a CoMFA region to increase or decrease contribution of these points to subsequent analysis. "StDev*Coefficients" values as different weighting factors were employed in addition to grid spacing for getting the better models. This increases the resolution and predictive capability (q^2, cross validated r^2) of a followed PLS analysis.

The CoMSIA method calculates the similarity indices descriptors with the same lattice box used in CoMFA. Five physicochemical properties of steric, electrostatic, hydrophobic, hydrogen binding donor and acceptor fields were evaluated using a probe atom with charge +1.0, radius 1 Å, hydrophobicity +1.0, hydrogen binding donor +1.0, hydrogen binding acceptor +1.0, attenuation factor α of 0.3 and grid spacing 2.0 Å. A distance-dependent Gaussian type was used between the probe atom and each molecule atom [49, 50].

2.4. HQSAR Analysis

Hologram QSAR study is a 2D-QSAR technique which certain the relationship between the biological activity with the structural fragments. This method eliminates the need for 3D structure, the ability to achieve molecular alignment and conformational specification [51, 52] by transforming the chemical representation of a molecule into its corresponding molecular hologram. 2D chemical database storage and searching technologies rely on linear notations that define chemical structures [Wiswesser line-formula notation (WLN), simplified molecular input line entry system (SMILES); SLN-SYBYL line notation]. The process involves generation of fragments that are hashed into array is called molecular hologram and bin occupancies are the descriptor variable [53-55].

The HQSAR method employs different parameters for the molecular hologram generation such as hologram length (HL) values (53, 59, 61, 72, 83, 97, 151, 199, 257, 307, 353 and 401), fragment distinction (atom (A), bonds (B),

connections (C), hydrogen atoms (H), chirality (Ch), and donor and acceptor (DA), and the fragment size (2-5, 3-6, 4-7, 5-8, 6-9, 7-10).

2.5. Partial Least-Square (PLS) Analysis

In 3D-QSAR studies, PLS method [56] an extension of multiple regression analysis was used for the model building. Calculated CoMFA and CoMSIA descriptors as independent variables were used with the pIC_{50} values as dependent variables in the PLS regression analysis, respectively. Before the PLS analysis, the CoMFA and CoMSIA columns were filtered by using column filtering value equal to 2.0 kcal/mol. The predictive ability of the models was evaluated by leave-one-out (LOO) and leave-ten-out (L-10-O) methods. LOO cross-validation method was used as an internal validation to generate the optimal number of components (ONC) with the lowest standard error of prediction (SEP) and the highest cross-validated coefficient q^2 () that was calculated by Equation (1):

Whereas, Y_{pred}, Y_{obs} and Y_{mean} are predicted, observed, and mean activity values of the training set, respectively [5]. The is the predictive residual sum of squares (PRESS).

After cross validation, the final PLS analysis was carried out using the optimal number of components with no validation to generate final QSAR model. The non-cross-validated analysis performed by the conventional correlation coefficient (), standard error of estimation (SEE) and F values calculated with the same column filtering set. High q^2 and r^2 (q^2> 0.5, r^2> 0.6) values are regarded as a proof of high predictive ability of the built model and also for a good model should not be more than 0.3 [52].

Bootstrapping analysis was performed for 100 runs to assess the statistical confidence of the derived models [37, 56-58]. Contour maps were generated graphically after models were developed in CoMFA/CoMFA-RF and CoMSIA using the field type “StDev*Coeff” and the contour levels were set to default values.

In HQSAR, LOO cross-validation was applied to determine the number of components that yields a good predictive model. PLS then yields a mathematical equation that related the molecular hologram bin values to the inhibition activity of the compounds in the database.

4.5. Validation of the QSAR model

A good internal validation showed only a high q^2 in the training set of compounds, but it did not indicate the high predictive ability of the established models, therefore external validation was essential. The predictive ability of 3D-

QSAR models were validated by calculating biological activities of the compounds which were not included in the training set and used as a test set. Test set was marked with * in Table 1.

Table 1
Chemical Structure and the Corresponding Experimental and Predicted pIC_{50} Values by QSAR models.

Compd	R_1, R_2	Experimental pIC_{50}	**Predicted pIC_{50}**			
			CoMFA	CoMFA-RF	CoMSIA	HQSAR
1		5.089	5.601	5.592	5.062	5.043
2*		4.657	5.602	4.904	5.062	5.047
3		4.630	4.602	5.108	4.963	4.299
4		5.798	5.498	5.647	5.103	5.680
5*		4.841	5.684	4.522	5.628	5.091
6		5.978	5.663	5.712	5.628	6.057
7		5.105	5.366	5.139	5.356	5.116
8		4.397	4.401	4.312	5.043	4.435

9*	4.346	4.089	4.534	4.144	4.598
10*	4.634	5.153	5.253	4.241	4.027
11	5.486	5.457	5.356	5.613	5.526
12	4.292	4.560	4.219	4.150	4.231
13	5.059	5.448	5.610	5.075	5.020
14	4.693	4.112	4.403	4.528	4.360
15*	5.070	5.020	5.332	4.897	5.246
16	4.042	4.112	4.403	4.528	4.36
17	5.214	5.323	5.403	4.995	5.372
18	5.173	5.234	5.184	5.151	5.139
19	5.488	5.956	5.858	5.936	5.830

20*	6.211	5.697	5.730	5.464	5.791
21*	6.011	5.805	5.871	5.568	5.672
22	6.196	5.975	5.640	5.893	6.136
23	6.019	6.080	5.971	6.402	5.899
24*	6.769	6.795	6.519	6.806	6.120
25	6.779	6.197	6.475	6.348	6.596
26	6.459	6.440	6.539	6.380	6.506
27	6.002	6.190	6.205	6.204	6.048
28	6.600	6.497	6.297	6.509	6.557

29		6.466	6.634	6.410	6.590	6.489
30*		6.372	6.802	6.761	6.451	6.615
31		6.607	6.672	6.566	6.587	6.561
32		6.815	6.637	6.774	6.663	6.887
33		6.939	6.818	6.995	6.783	6.791
34		6.690	6.885	7.142	6.810	6.885
35		6.742	6.964	6.888	6.955	6.828

36	7.173	6.816	7.046	7.060	6.973
37	6.954	6.959	6.962	7.014	6.999
38	6.917	6.499	6.832	7.041	6.963
39	6.517	6.874	6.352	6.625	6.378
40	6.987	6.822	6.756	6.656	7.027

*Test set

The predictive correlation coefficient (> 0.6) [59], based on the test set was calculated using Equation (2):

SD is the sum of squared deviation between the biological activities of the test set molecules and the mean activity of the training set molecules. PRESS is the sum of squared derivations between the predicted and actual activities of the test set molecules.

Performance of the regression models constructed here was evaluated using the root mean squared error (RMSE), mean absolute error (MAE) (RMSE and MAE close to zero), Residual sum of squares (RSS) and concordance correlation coefficient (CCC; CCC ≥ 0.85) of the training and validation sets [60]. The RMSE and the MAE are calculated for the data set as Equations (3-6)

RMSE=

MAE=

RSS=

CCC=

To obtain the good predictive model for test set, additional validation of model, the following

Parameters [59] were used (Equation (7)):

(or Eq. (7

k' 0.85 k 0.85

and are squared correlation coefficients of determination for regression lines through the origin between predicted (y) and observed (x) activities and vice versa. The values of k and k' are the slopes of their models, respectively.

To further assess the models, another validation statistical parameter and were determined by following Equations (8), (9):

value more than 0.5 (>0.5) and <0.2 show good external predictability of the models.

2.6. Molecular Docking Study

Molecular docking as one of the most frequently methods in drug design was used to investigate the mode of interaction of small molecules with the appropriate target binding sites. The docking study was performed using Operation Environment (MOE) software (www.chemcomp.com) between the most and least active compounds with AChE enzyme. Fro the preparation of ligands prior to docking, the 2D structures of ligands were prepared by Chemoffice ultra (version 12.0, Cambridge Soft Corporation, Cambridge, UK, 2010) and converted to 3D format by Hyper Chem7 (Hyper cube Inc, USA) using AM1 semi-empirical method. The ligands in our data set were docked in the active site of AChE (PDB ID: 1eve) by MOE software. The docking was performed by triangle matcher placement algorithm in combination with London dG scoring function and force field as refinement method and the

conformation of compounds were further analyzed by LigX module in MOE software. The best docking pose of compound **36** was chosen for Molecular dynamic (MD) simulation.

2.7. Molecular Dynamic Simulation

The molecular dynamics (MD) simulations, based on Newton's second law or the equation of motion, were performed to investigate the interaction between the receptor and ligand in atomic details using the dynamics module of SYBYL [61]. The compound **36** was used as the template molecule to elucidate the MD simulations. Energy minimization of the docked ligand was performed with Tripos force field and Gasteiger-Huckel charge without water using Boltzmann initial velocity.

The simulations were executed using normal temperature and volume (NTV) ensemble 300 K with coupling 100 fs. The MD simulations were performed with a time step of 2 fs for 10000 fs and conformation snapshot at every 100 fs to calculate RMSD values. Further, time dependent change of temperature, potential energy, kinetic energy and total energy for inhibitor was determined and recorded as a plot.

3. Results and discussion

3.1. CoMFA and CoMFA-RF Statistical Results

The statistical results of CoMFA and CoMFA-RF models are summarized in Table 2. The CoMFA analysis was carried out with steric and electrostatic fields at column filtering of 2.0 kcal/mol.

Table 2
Statistical parameters of QSAR models.

Parameters	CoMFA	CoMFA-RF	CoMSIA	HQSAR
PLS analysis				
$\mathbf{q^2}$	0.629	0.775	0.754	0.823
(L-10-O)	0.619	0.793	0.767	0.662
ONC	2	3	4	4
SEP	0.586	0.448	0.487	0.431
	0.901	0.910	0.919	0.976
SEE	0.303	0.297	0.279	0.159
$\mathbf{R}_{\mathbf{pearson}}$	0.943	0.951	0.921	0.962
F	78.639	123.076	70.916	116.361
	0.995	0.993	0.992	0.990
$\mathbf{SEE}_{\mathbf{bs}}$	0.007	0.006	0.006	0.006
Contribution				
Steric	0.844	0.641	0.261	-
Electrostatic	0.156	0.359	0.115	-
Hydrophobic	-	-	0.267	-
Donor	-	-	0.190	-
Acceptor	-	-	0.166	-

q^2: cross-validated correlation coefficient after the leave-one-out procedure; ONC: optimal number of principal components; non-cross-validated correlation coefficient; SEE: standard error of estimate; F: the value of F statistic;: the average r^2 value from a bootstrapping analysis for 100 runs; SEE_{bs}: the average SEE value from a bootstrapping analysis for 100 runs; :(mean) the average r_{cv} from ten times tenfold cross-validation.

PLS analysis of CoMFA for training set including leave-one-out (LOO) and leave-ten-out (L-10-O) cross validation with ONC 2 showed q^2 value of 0.629, (L-10-O) value of 0.619 and SEP of 0.586. These statistical results showed that the model had a good predictive capability.

The non-cross-validated PLS analysis gave a of 0.901 with standard error of estimate (SEE) of 0.303, F value of 78.639, of 0.272 and $R_{pearson}$ of 0.943 which supported the statistical validity of the developed model. The contributions from steric and electrostatic field descriptors explained 0.844 and 0.156 of the total variance, respectively that indicated steric effect was more important than the electrostatic fraction.

After using region focusing, a new model of CoMFA-RF with improvement in the statistical parameters was created. The cross-validation and non-cross-validated PLS calculation results were found better in CoMFA-RF as compared to CoMFA. This approach showed an increase in the q^2 value from 0.629 to 0.775 with ONC of 3 and (L-10-O) from 0.619 to 0.793 and SEP of 0.448. The non-cross-validated PLS analysis resulted in high value of 0.910 with a low SEE value of 0.297, F value of 123.076, value of 0.135 and $R_{pearson}$ value of 0.951. The contribution of steric and electrostatic field descriptors was 0.641 and 0.359, respectively in CoMFA-RF.

The bootstrapped results were shown in and values of 0.995 and 0.007 (CoMFA) and 0.993 and 0.006 (CoMFA-RF), respectively that suggesting a good internal consistency and the absence of systematic errors of the models within the training data set.

3.2. CoMSIA statistical results

The CoMSIA technique deals with direct correlation of ligand affinities to changes in molecular properties [62]. The CoMSIA model was generated using combinations of five steric (S), electrostatic (E), hydrophobic (H), hydrogen binding acceptor (A) and hydrogen binding donor (D) fields. The statistical parameters of CoMSIA model were summarized in Table 2. In PLS analysis, the q^2 value of 0.754 with ONC of 4, SEP of 0.487 and (L-10-O) of 0.767 was obtained with column filtering of 2.0 kcal/mol. The non-cross-validated PLS analysis gave a value of 0.919 with SEE value of 0.279, F value of 70.916, value of 0.165 and $R_{pearson}$ value of 0.921.

A high bootstrapped r^2 value of 0.992 and of 0.006 suggest a high degree of confidence in the analysis. For CoMSIA, the contribution of the steric, electrostatic, hydrophobic, hydrogen bond donor and hydrogen bond acceptor field descriptors were 0.261, 0.115, 0.267, 0.190, and 0.166, respectively. These molecular fields were not completely independent of each other and could form 31 combinations (Fig. 4).

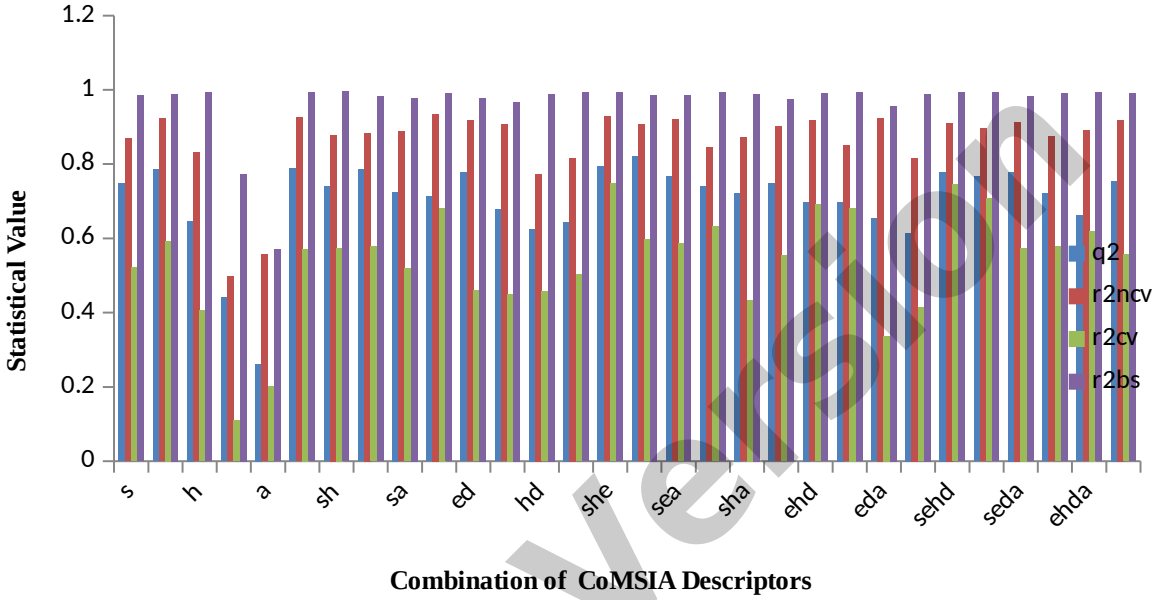


Fig. 4. Results of the distribution of q^2,,,values that were obtained from 31 combinations of CoMSIA fields. s, steric; e, electrostatic; h, hydrophobic; d, H-bond donor; a, H-bond acceptor.

Among these, combination of steric, electrostatic and hydrogen bond donor (SED) was found to be the best. CoMSIA (SED) combination gave q^2 value of 0.820, of 0.907, of 0.597 and of 0.986. In the model CoMSIA, this combination shared the large part and indicated that internal prediction of SED combination was good.

3.3. HQSAR Statistical Results

The HQSAR is a technique of QSAR analysis that is useful in exploring the combination of each molecule under study to the biological activity and eliminates the need of alignment, generation of 3D structures and putative binding conformation. The performance of the HQSAR model was affected by three parameters, including the fragment size, the fragment type (fragment distinction) and hologram length. The HQSAR models with statistical parameters are showed in Table 2.

The best statistical results of HQSAR model were obtained with q^2 value of 0.823, ONC of 4, SEP of 0.431 and (L-10-O) of 0.662, of 0.976 with SEE of 0.159, F value of 116.361, of 0.153, of 0.990 with of 0.006 and $R_{pearson}$ of 0.962 using a relevant hologram length (HL) of 307, fragment distinction (atom (A), chirality (Ch), and donor and acceptor (DA), and the fragment size of 4-7 (Table 3-4). All the results demonstrated that the HQSAR model was also highly predictive.

Table 3
HQSAR analysis for various fragment distinctions on the key statistical parameters using fragment size (4-7)

N	HL	SEE	$\mathbf{r^2}$	SEP	$\mathbf{q^2}$	Fragment distinction	model
3	97	0.278	0.917	0.468	0.763	A	1-1
5	353	0.189	0.964	0.463	0.786	A/B	1-2
5	257	0.159	0.975	0.520	0.731	A/C	1-3
5	257	0.292	0.915	0.547	0.702	A/H	1-4
5	97	0.187	0.965	0.468	0.781	A/Ch	1-5
6	307	0.163	0.975	0.447	0.810	A/DA	1-6
4	401	0.181	0.966	0.466	0.775	A/B/C	1-7
6	257	0.203	0.961	0.512	0.749	A/B/H	1-8
3	151	0.280	0.916	0.497	0.734	A/B/Ch	1-9
5	97	0.231	0.947	0.492	0.759	A/B/DA	1-10
6	59	0.203	0.961	0.555	0.706	A/C/H	1-11
5	257	0.159	0.975	0.520	0.731	A/C/Ch	1-12
6	353	0.130	0.984	0.465	0.794	A/C/DA	1-13
4	97	0.349	0.874	0.540	0.698	A/H/Ch	1-14
5	151	0.287	0.918	0.576	0.670	A/H/DA	1-15
4	**307**	**0.159**	**0.976**	**0.431**	**0.823**	**A/Ch/DA**	**1-16**
6	71	0.245	0.943	0.547	0.715	A/B/C/H	1-17
4	401	0.181	0.966	0.466	0.775	A/B/C/Ch	1-18
4	401	0.166	0.973	0.491	0.760	A/B/C/DA	1-19
6	97	0.248	0.941	0.491	0.770	A/B/H/Ch	1-20
6	53	0.286	0.922	0.531	0.731	A /B/H/DA	1-21
4	199	0.233	0.943	0.490	0.750	A/B/Ch/DA	1-22
6	59	0.203	0.961	0.555	0.706	A/C/H/Ch	1-23
5	151	0.293	0.914	0.556	0.692	A/H/Ch/DA	1-24
5	401	0.299	0.911	0.581	0.664	A/C/H/DA	1-25
4	199	0.357	0.867	0.561	0.674	A/C/H/Ch/DA	1-26
4	53	0.324	0.891	0.521	0.718	A/B/H/Ch/DA	1-27
6	199	0.159	0.976	0.519	0.743	A/B/C/Ch/DA	1-28
6	61	0.225	0.952	0.524	0.738	A/B/C/H/DA	1-29
6	71	0.245	0.943	0.547	0.715	A/B/C/H/Ch	1-30
6	307	0.212	0.957	0.566	0.694	A/B/C/H/Ch/DA	1-31

q^2, cross-validated correlation coefficient; r^2, non-cross-validated correlation coefficient; SEE, standard estimated error; HL, hologram length; N, optimal number of components. Fragment distinction: A, atom; B, bond; C, connections; H, hydrogen atom; Ch, chirality; D, donor and acceptor.
The model chosen for analysis is highlighted in bold fonts.

Table 4
Statistical results of QSAR model using the model 1-16 (including fragments A/Ch/DA) with different fragment sizes

N	HL	SEE	$\mathbf{r^2}$	SEP	$\mathbf{q^2}$	Fragment size	model
6	257	0.252	0.940	0.517	0.745	1-4	2-1

6	307	0.170	0.972	0.445	0.811	2-5	2-2
6	307	0.171	0.972	0.431	0.822	3-6	2-3
4	307	0.159	0.976	0.431	0.823	4-7	2-4
6	61	0.188	0.966	0.412	0.838	5-8	2-5
4	401	0.299	0.946	0.457	0.783	6-9	2-6
5	151	0.194	0.963	0.436	0.811	7-10	2-7
6	**61**	**0.172**	**0.972**	**0.407**	**0.842**	**8-11**	**2-8**
3	59	0.303	0.901	0.482	0.749	9-12	2-9

q^2, cross-validated correlation coefficient; r^2, non-cross-validated correlation coefficient; SEE, standard estimated error; HL, hologram length; N, optimal number of components. Fragment distinction: A, atom; B, bond; C, connections; H, hydrogen atom; Ch, chirality; D, donor and acceptor.
The model chosen for analysis is highlighted in bold fonts.

3.4. Validation of QSAR models

The predictive abilities of the QSAR models were externally validated using the independent test set that was not used for the model generation [59]. q^2 and r^2 parameters, obtained from internal validation, were used for confirm the stability and the predictive ability of the models. The QSAR models for the whole test set including compounds10 gave the and values of 0.773 and 0.560 (CoMFA), 0.824 and 0.580 (CoMFA-RF), 0.874 and 0.606 (CoMSIA), and 0.854 and 0.658 (HQSAR) and high slope regression lines with k and k´ values of 1.026 and 0.967 (CoMFA), 0.990 and 1.004 (CoMFA-RF), 1.012 and 0.983, and 0.976 and 1.018 (HQSAR, respectively. and values of 0.667 and 0.704 (CoMFA), 0.719 and 0.735 (CoMFA-RF), 0.755 and 0.823 (CoMSIA), and 0.766 and 0.794 (HQSAR), respectively were used to calculate the relationship between , and that (r^2-)/r^2 and (r^2-)/r^2 values of 0.065 and 0.013 (CoMFA), 0.038 and 0.015 (CoMFA-RF), 0.086 and 0.014 (CoMSIA), and 0.037 and 0.003 (HQSAR), respectively were obtained.

The QSAR models yielded RMSE, MAE and CCC values of 0.244, 0.170, 0.948; 0.247, 0.095, 851 (CoMFA); 0.245, 0.165, 0.948; 0.217, 0.090, 0.861 (CoMFA-RF); 0.224, 0.148, 0.958; 0.183, 0.084, 0.892 (CoMSIA); 0.140, 0.079, 0.988; 0.197, 0.091, 0.882 (HQSAR) for Training and test set, respectively.

From the values of the performance criteria parameters yielded by the QSARs in training and test data (Table 5), it is evident that all of the models yielded considerably low RMSE and MAE values and high CCC values which show that models built by training set could be used for the prediction of these chemotypes.

The results of external validation parameters are listed in Table 5. These results confirm that the QSAR models could be used to predict the biological activities of new compounds and their derivatives.

Table 5

Statistical Parameters of Validation Method for QSAR Models.

HQSAR	CoMSIA	CoMFA-RF	CoMFA	Parameters
0.854	0.874	0.824	0.773	
0.766	0.755	0.719	0.667	
0.794	0.723	0.735	0.704	
0.037	0.086	0.038	0.065	$(r^2\text{-})/r^2$
0.003	0.014	0.015	0.013	$(r^2\text{-})/r^2$
0.976	1.012	0.990	1.026	**k**
1.018	0.983	1.004	0.967	**k´**
0.658	0.606	0.580	0.560	
0.510	0.642	0.607	0.645	
0.148	0.036	0.027	0.085	
0.140	0.224	0.245	0.244	$\mathbf{RMSE}_{train}$
0.197	0.183	0.217	0.247	$\mathbf{RMSE}_{test}$
0.079	0.148	0.165	0.170	$\mathbf{MAE}_{train}$
0.091	0.084	0.090	0.095	$\mathbf{MAE}_{test}$
0.584	1.813	2.392	2.383	$\mathbf{RSS}_{train}$
1.560	1.341	1.88	2.432	$\mathbf{RSS}_{test}$
0.988	0.958	0.948	0.948	$\mathbf{CCC}_{train}$
0.882	0.892	0.861	0.851	$\mathbf{CCC}_{test}$

predicted correlation coefficient for the test set of compounds;
: correlation coefficient for regression through the origin for predicted versus observed activities (test set);
: correlation coefficient for regression through origin for observed versus predicted activities (test set);
: modified squared correlation coefficient (test set); RMSE: root mean squared error; MAE: mean absolute error ; RSS: residual sum of squares; CCC: concordance correlation coefficient.

The correlation plots between the predicted and experimental activities are shown in Fig. 5. Most of the compounds were located on or near to the trend line in the QSAR models and these results confirm that these models had good predictive ability for new compounds.

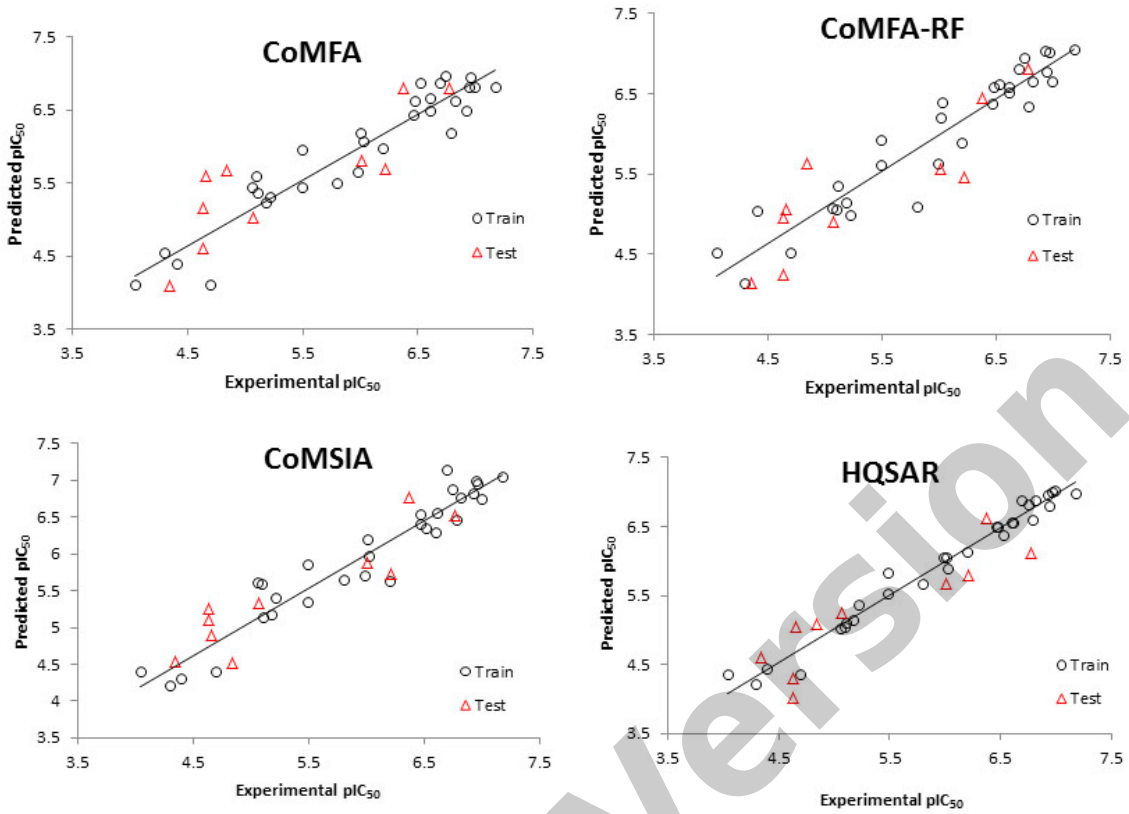


Fig. 5. The plot of predicted pIC_{50} versus experimental pIC_{50} values for training and test sets compounds by QSAR models.

The residual values of the QSAR models are shown in Fig. 6. The CoMSIA and HQSAR models showed smaller residuals than the CoMFA and CoMFA-RF models and were the better models.

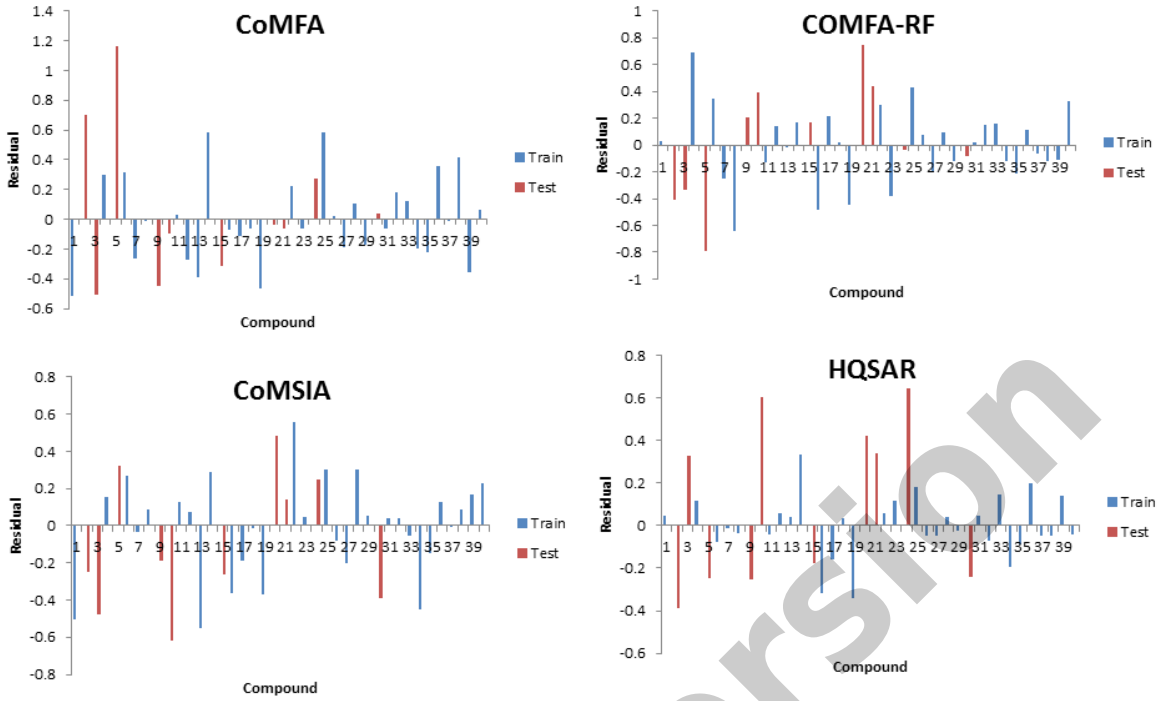


Fig. 6. Residual plots between experimental and predicted values foe QSAR models.

3.5. Evaluation of the Y -randomization test and application Domain (AD) of model

The QSAR models were further validated by applying the Y-randomization test to assess the robustness of the models and to avoid chance correlation [63, 64]. Thus, foe every original model, several random shuffles of the dependent variable (biological activity) were performed and a new QSAR model was developed using the original independent variable matrix and the results are shown in Table 6. The low q^2 and values (q^2< 0.5 and < 0.6) show that the good results obtained in the formulation of the final models were not by chance. after several Y-tests.

Table 6
q^2 and values randomization

CoMSIA		CoMFA		
	q^2		q^2	Y-random Iteration
0.218	0.346-	0.364	0.204-	1
0.334	0.212-	0.432	0.316-	2
0.295	0.156-	0.325	0.182-	3
0.323	0.136-	0.317	0.173-	4
0.341	0.175-	0.347	0.081-	5
0.163	0.141-	0.311	0.112-	6
0.388	0.715-	0.321	0.108-	7
0.275	0.206-	0.375	0.237-	8
0.320	0.163-	0.336	0.195-	9
0.169	0.097-	0.326	0.100-	10
0.355	0.241-	0.361	0.243-	11
0.298	0.188-	0.374	0.255-	12
0.338	0.195-	0.324	0.188-	13
0.343	0.179-	0.305	0.178-	14
0.361	0.241-	0.343	0.192-	15
0.388	0.230-	0.367	0.271-	16
0.337	0.215-	0.353	0.234-	17
0.373	0.253-	0.362	0.247-	18
0.386	0.222-	0.327	0.211-	19
0.321	0.105-	0.361	0.253-	20
0.919	0.754	0.901	0.629	Non-Random

For a new compound with no experimental data, a predicted value of QSAR models without an idea of reliability of the value is not useful. Therefore, for evaluating new compounds, a very important step in QSAR model development is the definition of the applicability domain of regression or classification models [65].

The Williams plot, the plot of the standardized residuals (δ) vs. leverage values (h_i), was used to illustrate the prediction and express the applicability domain of the models for each chemical compound [66-68].

:[The standardized residuals (δ) value is calculated by Equation (10) [69

Where are the observed and predicted values for *i-th* compound, respectively, n is the number of compounds and A is the number of descriptors. Also, the leverage value (h_i) is defined by Equation (11):

$$(h_i=(X)^{-1} \quad (i=1,\ldots, n) \qquad \text{Eq. (11}$$

where x_i is the descriptor-row vector of the *i-th* compound, is the transpose of x_i, X is the descriptor matrix of the training set compounds and is the transpose of X.

The warning leverage value (h^*), as a prediction tool, is expressed as:

where k is the number of model descriptors and n is the number of training compounds.

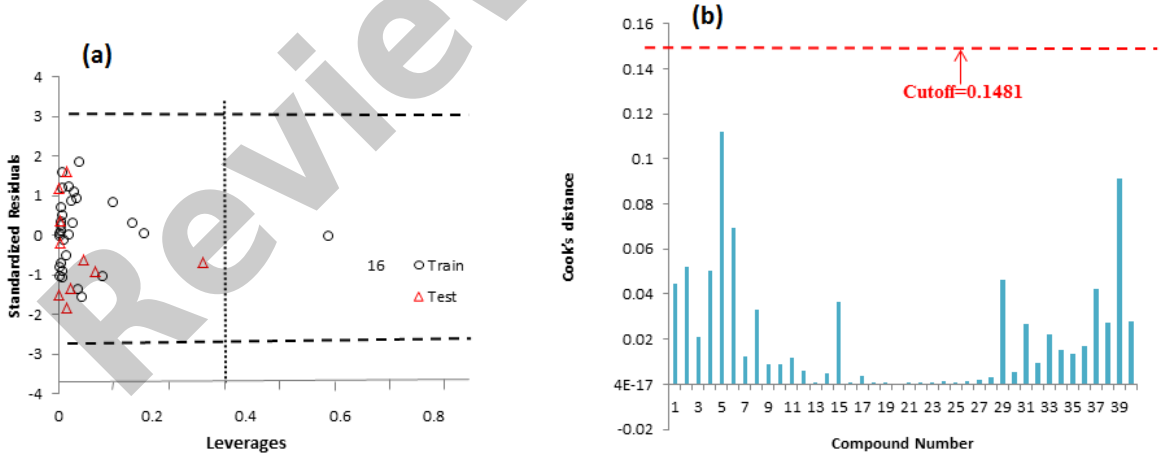
The Williams plot illustrates the distribution of data and its restricting rang termed cutoff lines which all data should be between ±3 units (horizontal dotted line) for standardized residuals and the leverage value (h_i) should be less than warning leverage ($h_i < h^*$). The Williams plot for the training set is used to identify molecules with the greatest structural influence ($h_i < h^*$) in developing the QSAR models. Molecules with $h_i > h^*$ are evaluated to be unreliably predicted by the models due to substantial extrapolation.

Cook's distance is used to estimate the influence of a single observation to the model [70], and is defined by Equation (12):

Where is the standard residual of the *i-th* compound, p is the number of descriptors, and is the leverage value of the *i-th* compound. The cutoff of the Cook's distance is defined as , and the compounds with Cook's distance higher than the cutoff value are marked as highly influential points of the model.

In this work, for CoMFA and CoMSIA models, most of the compounds fall into their corresponding application domain. These results indicated that our QSAR models had achieved a reliable activity prediction for the compounds.

As shown in the Williams plot of CoMFA model for data set (Fig. 7a), only one compound (**16**) of training set had greater value than the warning leverage (h^*) value of 0.3. This compound had low standard residual value and could be considered as influential in fitting the model performance but not necessarily outlier to be deleted from the training set. The test compounds were within the applicability domain (AD) indicating that their predicted activity values were reliable. Also, at the Cook's plot of CoMFA model (Fig. 7b); there was not any highly influential compound for training and test set. In addition, the histogram of the residuals distribution was confirmed with histogram plot as shown in Fig. 7c.



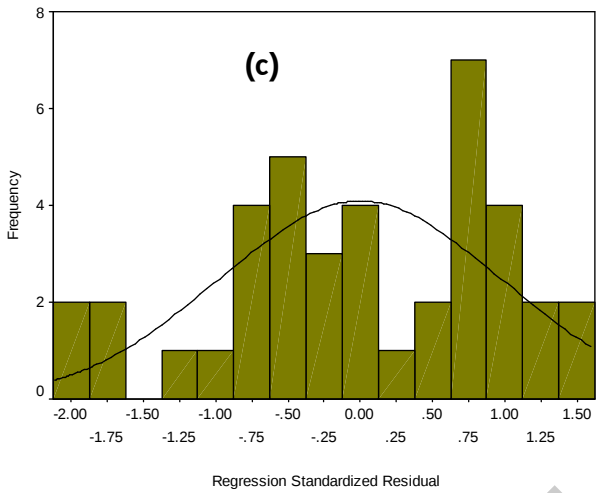
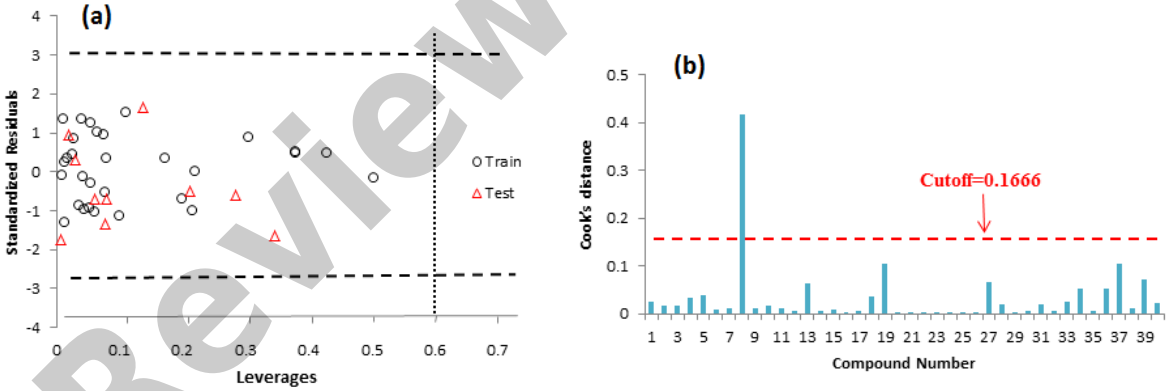


Fig. 7. Williams plot describing the applicability domain of the CoMFA model for the training and test sets (h* = 0.30) (a); Cook's distance plot (b) and Histogram of model CoMFA residuals (c).

Also, at the Williams plot of CoMSIA model for data set (Fig. 8a); there was not any outlier compound for training and test set. Otherwise, according to the Cook's distances (cutoff=0.166) of the compounds in the data set, one highly influential compound may slightly distort the regression (Fig. 8b), also, the histogram of the residuals distribution was confirmed with histogram plot as shown in Fig. 8cand prediction of CoMSIA model is reliable.



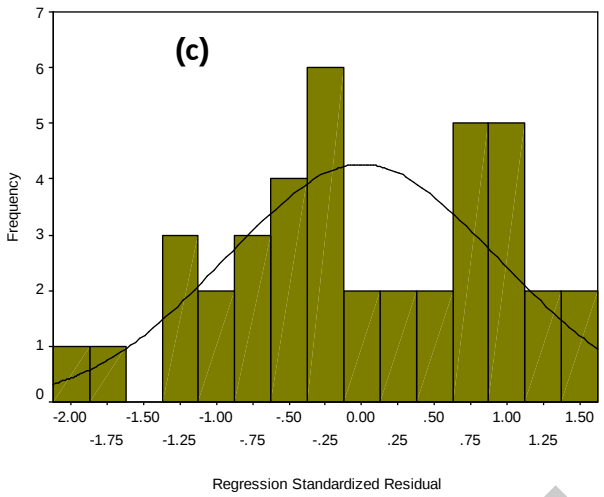
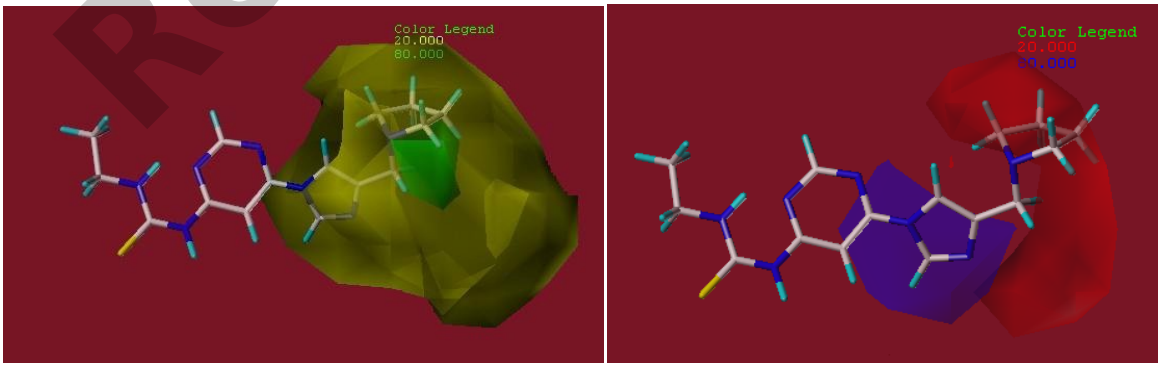


Fig. 8. Williams plot describing the applicability domain of the CoMSIA model for the training and test sets (h* = 0.60) (a), Cook's distance plot (b) and Histogram of model CoMSIA residuals (c).

3.6. Interpretation of CoMFA and CoMSIA contour maps

The QSAR contour maps were used as an informative tool to visualize the effects of the different fields on the target compound 3D grid orientation of models. The CoMFA and CoMSIA results were graphically interpreted by field contribution maps using the standard deviation (StDev) at each grid point and the coefficient from the PLS analysis (StDev*Coefficients).

The CoMFA contour maps of the steric and electrostatic fields for the best anti-Alzheimer agent (compound **36**) are shown in Fig. 9a, 9b. The field steric is shown by favorable groups (80% contribution) in green color and unfavorable ones (20% contribution) in yellow where the introduction of bulky groups may enhance or diminish the activity.



(a) (b)

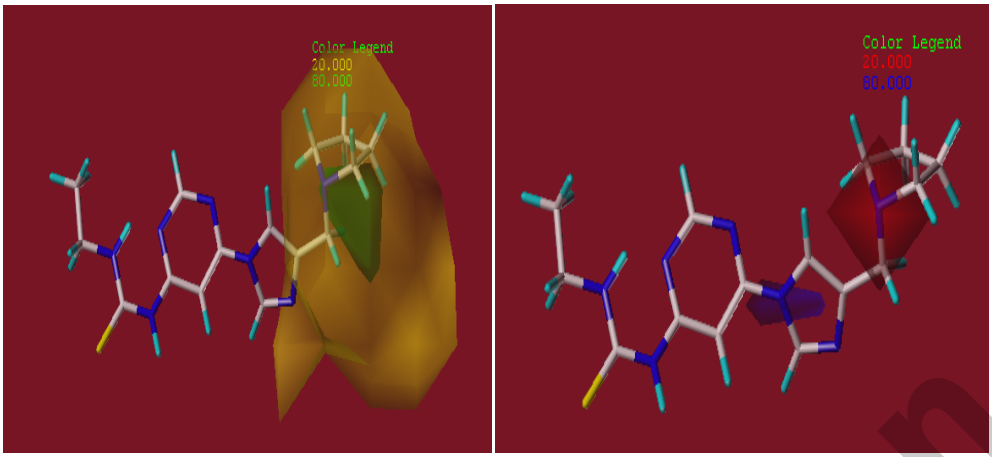
Fig. 9. CoMFAStDev*coeff. Contour plots with the combination of compound **36**. (a) *Steric contour maps*: Green contours indicate regions where bulky groups increase activity and yellow contours indicate regions where bulky groups decrease activity. (b) *Electrostatic contour maps*: Blue contours indicate regions where positive charges increase activity and red contours indicate regions where negative charges increase activity.

In the CoMFA steric maps, there was a green contour covering the pyrrolidine group at C-4 position of imidazole ring. The bulky groups at this position of compound improved anti-Alzheimer activity and had the highest activity.

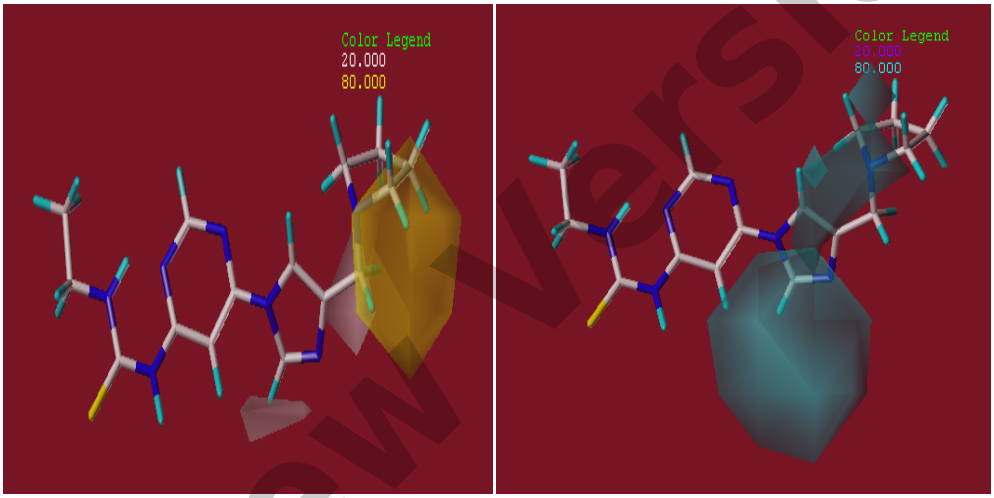
The compounds **21**-**35** and **37**-**40** with bulky substituents (e.g. , , , , Cl and Br) at this region exhibited more potency, while compounds **3**, **13** and **17**-**18** due to the absence of this group had relatively low activity. Substituting the morpholine and benzyl amine at 6 position of pyrimidine ring in compounds **10** and **12** decreased activity because these substituents were located at disfavored yellow contours. Therefore, this position of pyrimidine ring should be occupied by the steric moderate and low crowed substituents.

In CoMFA electrostatic contour maps (Fig. 9b), the blue region (80% contribution) are favorable for electropositive groups and red regions (20% contribution) are favorable for electronegative groups. The blue contour on the 6 position of pyrimidine ring of compound **36** indicated the introduction of electropositive groups in this position could improve the biological activity. Besides, a red contour in the C-4 position of imidazole ring showed that the electronegative substituent was beneficial to activity (compounds **19**<**27**<**21**<**23**<**22**<**20**<**24**<**21**).

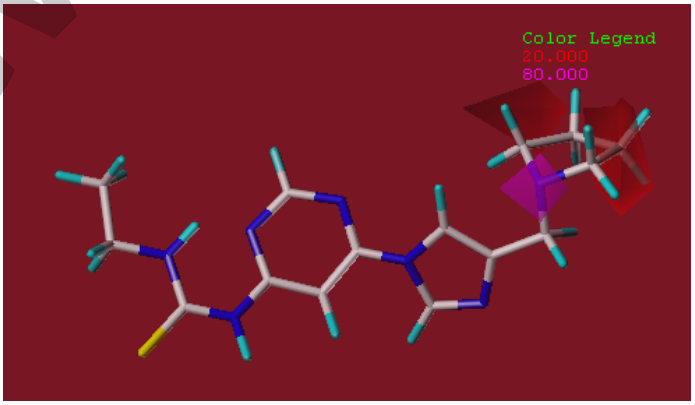
In CoMSIA model, the steric and electrostatic, hydrophobic, hydrogen binding (H-bond) donor and acceptor contour maps of compound **36** are shown in Fig. 10. The CoMSIA steric and electrostatic contours were nearly similar to that of CoMFA contours, so the hydrophobic interaction and hydrogen bond fields were described here. In the hydrophobic contour map, the yellow region is favorable (80% contribution) for hydrophobic group while white region (20% contribution) is favorable for the hydrophilic group.



(a) (b)



(c) (d)



(e)

Fig. 10. CoMSIAStDev*Coeff contour plots with the combination of compound **36**. (a) *Steric contour maps*: Green contours indicate regions where bulky groups increase activity; yellow contours indicate regions where bulky groups decrease activity; (b)*Electrostatic contour maps*: Blue contours indicate regions where positive charges increase activity; red contours indicate regions where negative charges increase activity. (c) *Hydrophobic contour maps*: yellow contours indicate regions where hydrophobic substituents enhance activity; white contours indicate regions where hydrophobic groups decrease activity. (d) *Hydrogen bond donor contour maps*: Cyan contours indicate regions where H-bond donor groups increase activity and purple contours indicate the unfavorable regions for hydrogen bond donor substituents. (e) *H-bond acceptor contour maps*: Magenta contours indicate regions where H-bond acceptor substituents increase activity; red contours indicate the disfavor regions for H-bond acceptor groups.

A white region near C-2 and C-4 positions of imidazole ring showed that the introduction of hydrophilic groups into these positions might be beneficial for inhibitory activity (Fig. 10c). The yellow contour in C-4 position of imidazole ring indicated that hydrophobic groups such as aryl and heterocyclic in this region could be increased the activity of

compounds. The compounds **31**, **35**, **37**-**38** and **40** with hydrophobic substituent (e.g. , ,

) at this region exhibited more potency, while compounds **8**-**16** due to the absence of this hydrophobic group, had relatively lower activity.

These results confirm that the yellow contour of hydrophobic map was in agreement with green contour of steric map.

The CoMSIA H-bond donor and acceptor contour maps correlated with hydrogen bind interactions of ligand with target. The cyan and purple contour maps of H-bond donor indicated favorable (80% contribution) and unfavorable (20% contribution) interactions and the magenta and red contour maps indicated favorable (80% contribution) and unfavorable (20% contribution) H-bond acceptor groups (Fig. 10d). However, no unfavorable purple contour was observed. There were two cyan contours near to C-5 and C-6 positions of pyrimidine ring and C-4 position of imidazole ring that the H-bond donor might improve anti-Alzheimer activity (compounds **28**-**30**). Changing the position of N atom in the pyrimidine ring and introducing substituents on this ring (compounds **13**-**14**) had relatively lower activity relative to compound **36**.

There was a magenta contour in N atom of pyrrolidine ring in C-4 position of imidazole ring which was favorable for H-bond acceptor (compounds **31**-**33** and **37**-**40**). However, the red contours over carbon positions of pyrrolidine ring were unfavorable for H-bond acceptor interactions (Fig. 10e).

3.7. Interpretation of HQSAR contribution map

HQSAR calculations are based on the contributions of molecular fragments to the biological activity for each molecule. Results of the HQSAR contribution maps can be graphically shown as a color-coded structure diagram which the color of each atom reflects its contribution to the molecule's overall activity. The red end of the spectrum (red, red orange and orange) reflects negative contribution to the activity, while the green end (yellow, blue, green-blue and green) represents positive effect and intermediate contributions are colored in white. The individual atomic contributions of the most active anti-Alzheimer analogues (compound **36**) were displayed in Fig. 11.

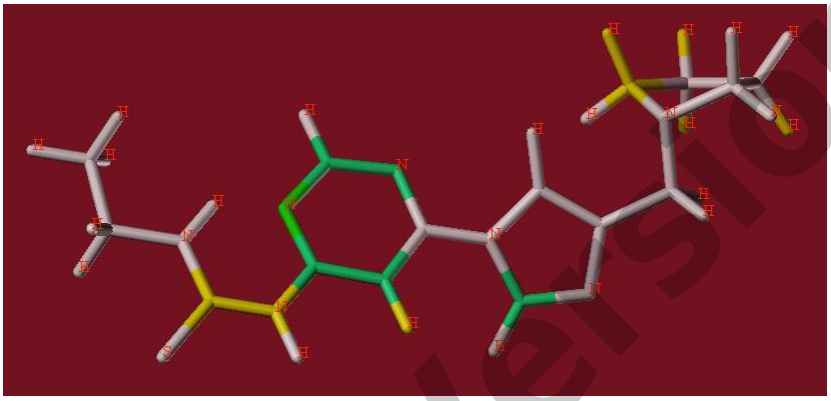


Fig. 11. The HQSAR contribution map of the most active compound (**30**). The colors in yellow, blue, green-blue or green indicate positive contributions, while colors with red, red-orange or orange represent negative contributions and intermediate contributions are colored in white.

The pyrimidine scaffold as maximal common structural fragment represented by green and yellow color codes because it was common fragment to all molecules and contributed in the same way to all inhibitors. The amino methyl derivatives in C-4 position of imidazole ring were highlighted in yellow color, indicating the importance of these fragments to biological activity. C-2 position of imidazole ring was colored in green that positive contribution to inhibitory activity. The atoms of thiourea moiety was colored in yellow and made a positive contribution to increase activity. But, replacement of the thiourea group with a similar urea group was not tolerated and a decrease in potency of anti-Alzheimer compounds was observed. Finally, the structure-activity relationship and binding features obtained by present QSAR models and molecular docking analysis are summarized in Fig. 12.

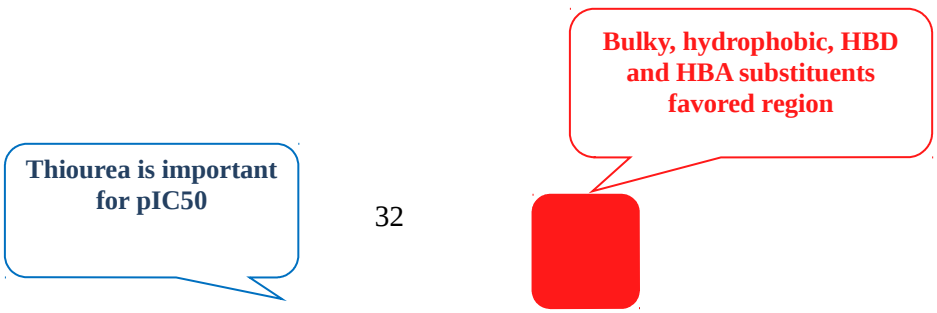


Fig. 12. Structure-activity relationship revealed by 3D- and 2D-QSARand docking studies.

3.8. Molecular Docking Studies

In this study, the MOE program was run to explore the possible binding modes of the anti-Alzheimer agents. To confirm the validity of used docking parameters, the co-crystallized ligand E2020 was re-docked into the active site of AChE enzyme. The re-docking result and the cognate ligand (green) were almost completely superimposed and the RMSD value (0.86 A°) guaranteed the reliability of the docking procedure (Fig. 13).

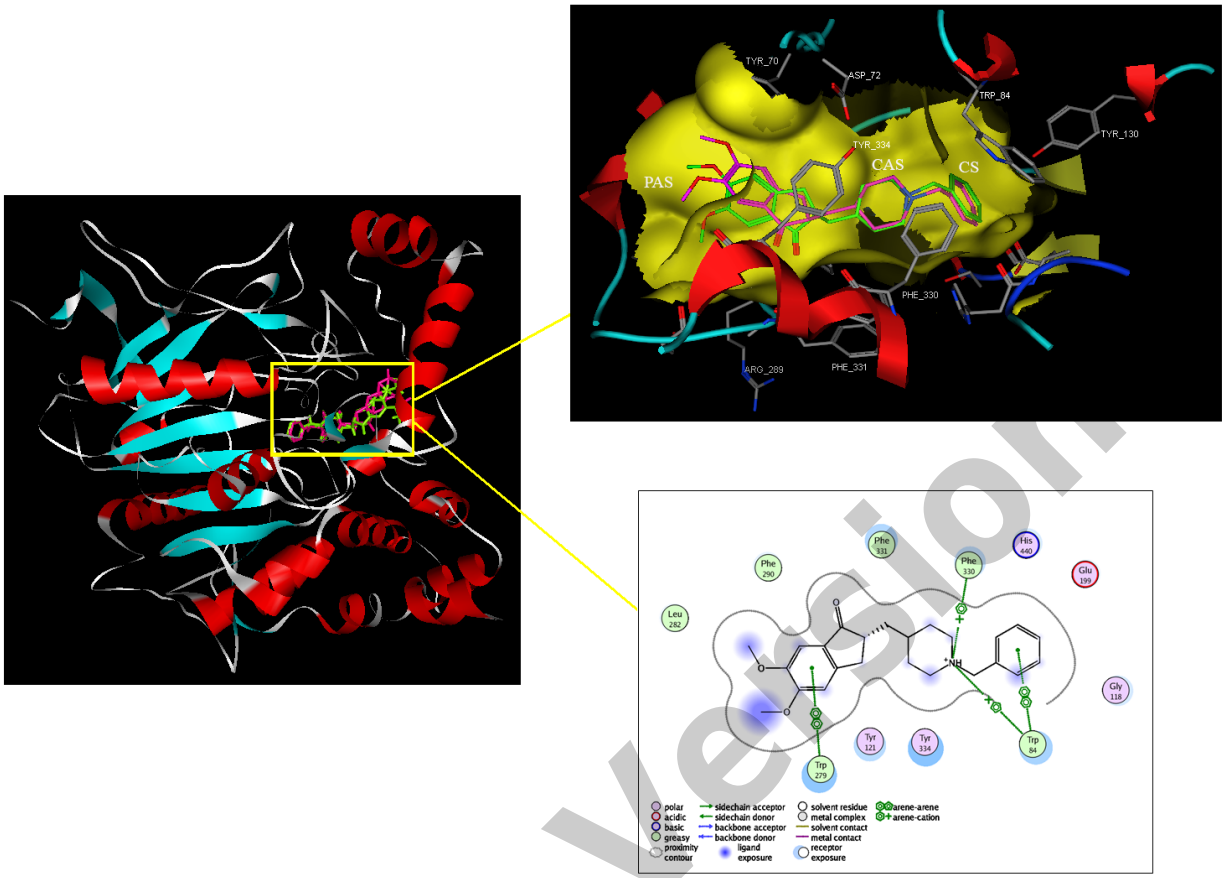
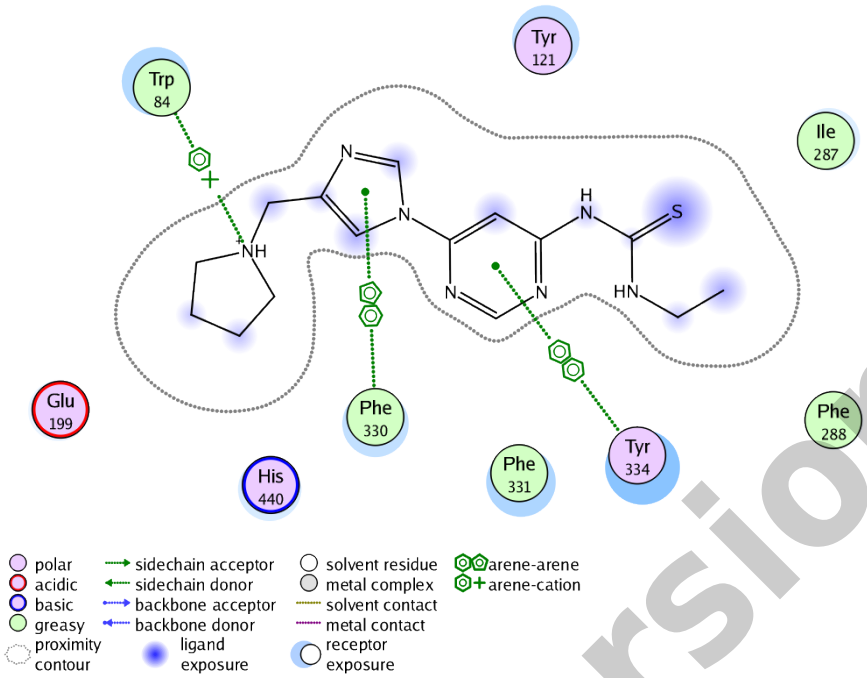


Fig. 13. Re-docking result of the co-crystallized ligand E2020 (green) into the binding of AChE and the cognate ligand (red) from the protein (PDB: 1eve). The catalytic site (CS), catalytic anionic subsite (CAS) and peripheral anionic site (PAS) were shown in the receptor.

In order to gain functional and structural insight into the binding mode of the most potent (compound **36**) and lest potent (compound **16**) inhibitors and AChE enzyme and also, to validate the results of QSAR contour maps, docking studies were carried out using MOE software (Fig. 13a, 13b).

(a)



(b)

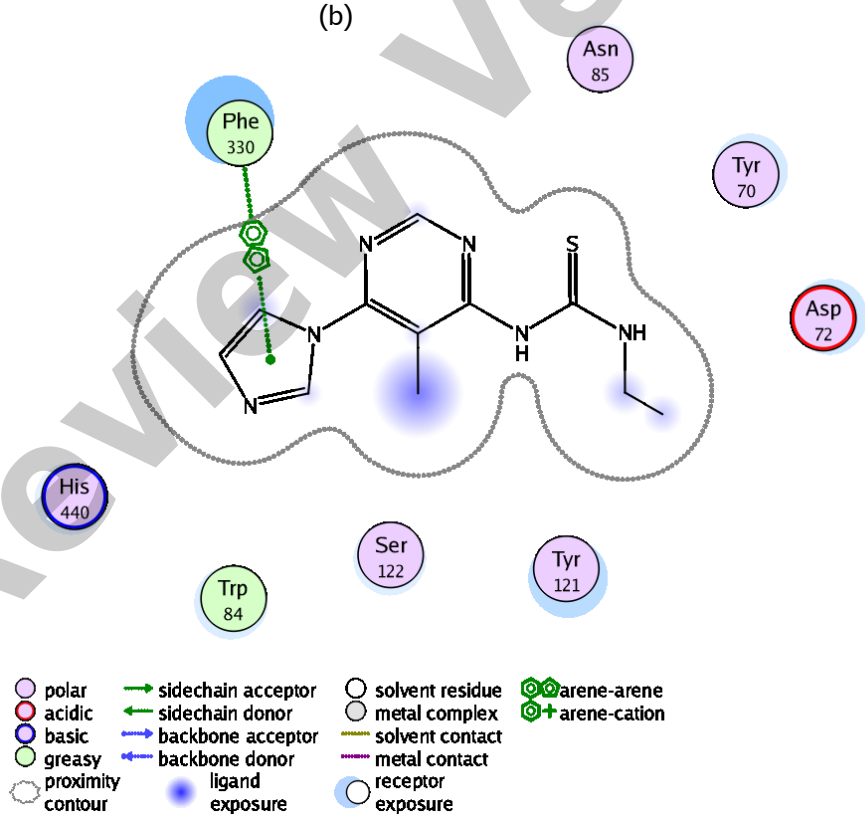
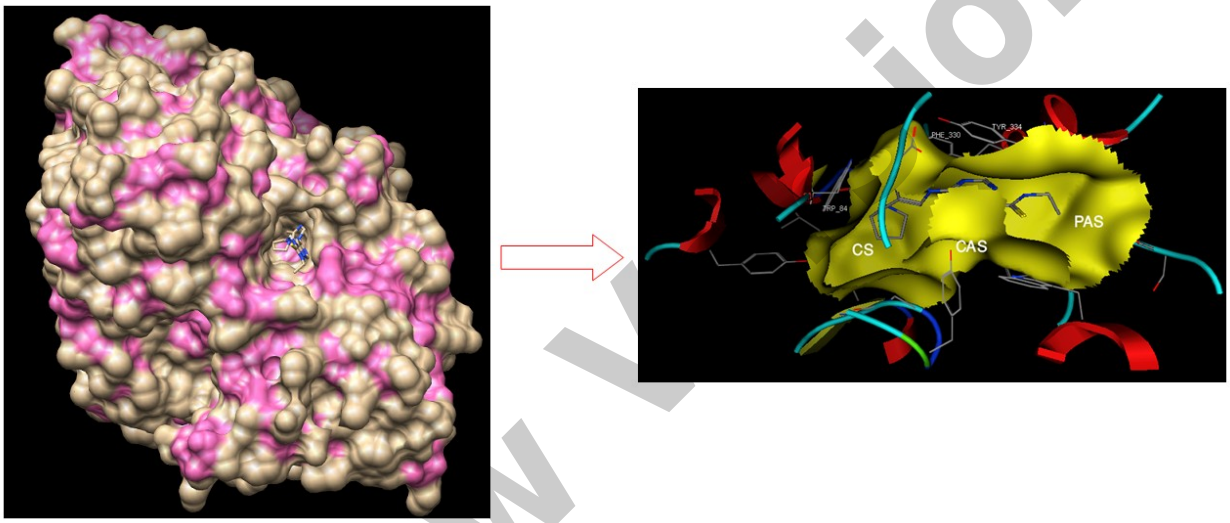


Fig. 13. (a) The 2D representation of the interaction between compound **36 (**the most active compound) **(**b**)**The 2D representation of the interaction between compound **16** (the least active compound) in the crystal structure of AChE (PDB ID: 1eve) using LigX in MOE.

Analysis of docking results revealed a high docking score (-13.30 kcal/mol) for the most active compound **36** in comparison to that of the least active compound **16** (-10.66 kcal/mol). The compound **16** bound with less numbering of active site residues of enzyme and had less interaction with AChE compared to compound **36**.

The compound **36** was well stabilized in the active site of AChE and had significant interactions with the key amino acid residues of CAS and PAS (Fig. 14).



.**Fig. 14.** 3D representation of docked ligand **36** into binding site of AChE

Regarding the docking studies, three types of interactions, hydrophobic π-π interaction, π-cation interaction and hydrophobic were involved in the attachment of compound **36** to active site of receptor. With few exceptions, the binding mode of best-scored ligands with AChE by LigX of MOE suggested that the compounds were oriented towards the gorge of the enzyme. The ethyl thiourea moiety was caged into the PAS, while pyrimidine and substituted imidazole rings were embedded into CAS. In the PAS region, the ethyl thiourea moiety was well fitted in the hydrophobic pocket composed by Tyr 121, Phe 288, Trp 279, and Ike 287. The pyrimidine ring was placed in the mid gorge of AChE active site in parallel with phenyl rings of Tyr 334 and Phe 331 and made two π-π stacking interactions with these residues. In addition to, imidazole ring could form another π-π stacking interaction with Phe 330 in CAS and the quaternary nitrogen of the pyrrolidine ring on imidazole moiety was caged into catalytic site (CS) at the bottom of gorge through a π-cation interaction with six member ring of Trp 84. Therefore, the higher

potency of the compounds with substituted imidazole moiety could be due to the more favorable interactions of compounds with target. These docking results validated the contour maps of QSAR models.

3.9. Molecular Dynamic Simulation

The molecular dynamic simulation has been done to elucidate behavior of AChE enzyme upon binding to ligand and stability and interaction of ligand-protein throughout simulation. This ligand-protein complex was simulated from 0 to 10,000 fs with respect to temperature, potential energy, kinetic energy and total energy, respectively (Fig. 15). The compound **36** formed a stable complex with AChE enzyme when simulated up to or beyond 1 ns with respect to temperature (TEMP) (at or above 298.3145 K), potential energy (PE) (at or above 1808.372 kcal/mol), kinetic energy (KE) (at or above 3820.903 kcal/mol) and total energy (TOT-ENG) (at or above 5629.275).

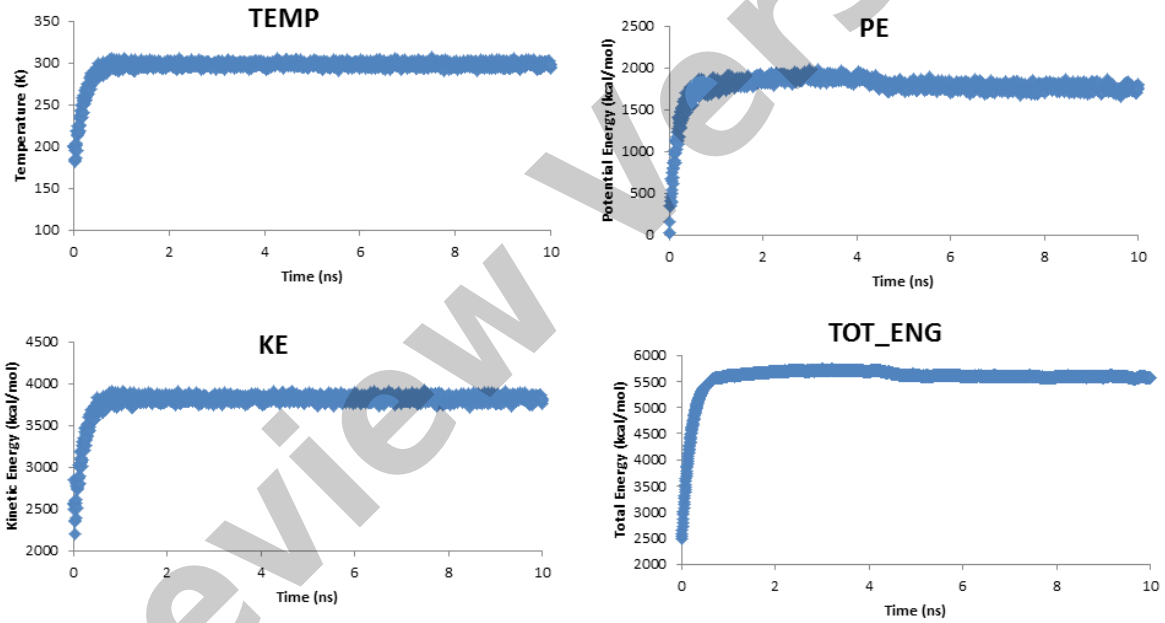
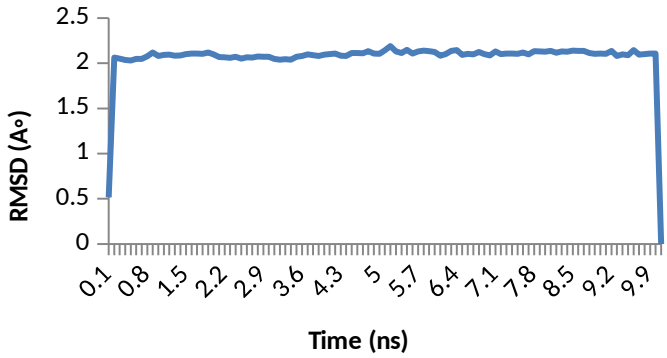


Fig. 15. Graph representing change in temperature, potential energy, kinetic energy, and total energy of AChE-ligand complex as a function of molecular dynamic simulation time.

The overall simulation convergence and ligand-protein equilibration were determined with root-mean-square deviation (RMSD) of backbone atoms (Cα, C, and N) that is a measure of the stability of the structures. The RMSD vs. Time is shown in Fig. 16.



.**Fig. 16.** RMSD between AChE with ligand

This plot indicated that The RMSD of complex with average 2.08 Å was stable after 0.2 ns of simulation. Also, this conformation of compound **36** with AChE showed a new hydrogen bonding interaction with amino acid residue of Asp 72.The simulation results showed that the final structure and initial docked structure were in the same binding pocket and ligand-protein conformation was stable. The 2D representation of the interaction between compound **36** after 10 ns simulation has been depicted in Fig. 17.

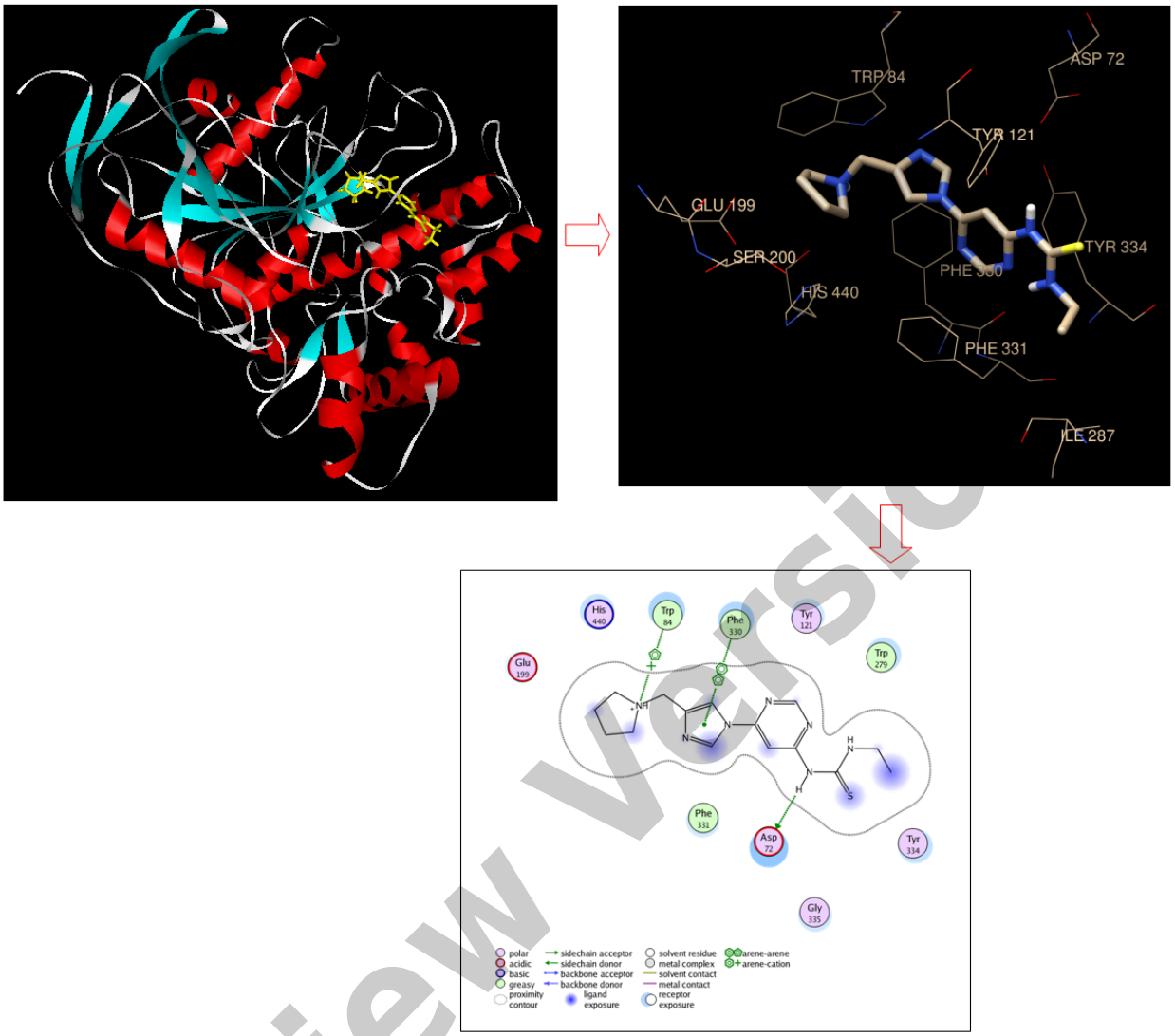


Fig. 17. A diagrammatic representation of interactions between compound **36** and AChE at the active site after 10ns molecular dynamic simulations.

4. Conclusion

The 2D- (HQSAR) and 3D-QSAR (CoMFA, CoMFA-RF and CoMSIA) methods were employed to study a series of pyrimidinylthiourea derivatives as anti-Alzheimer agents. The CoMFA, CoMFA-RF, CoMSIA and HQSAR models provided statistically significant results for internal and external validations including q^2 Values of 0.629, 0.775, 0.754, and 0.622, values of 0.910, 0.910, 0.919, and 0.949, values of 0.773, 0.824, 0.874, and 0.854 and values of 0.560, 0.580, 0.606, and 0.658, respectively. The CoMFA andCoMSIAcontour maps and the HQSAR fragment contribution map were explained structure-activity relationship of this series of anti-Alzheimer agents. Also,

molecular docking and molecular dynamic simulations studies were carried out to confirm the rationality of the derived models. The pyrimidine group as scaffold and the bulk groups in the heterocyclic moiety as a hydrophobic part were key factors to improve inhibitory activity of AChE. These results showed good predictive models for the rational design of novel acetylcholinesterase inhibitors for the treatment of Alzheimer's disease.

Abbreviations: AChE, Acetylcholinesterase; AD, Alzheimer's disease; AChT, Acetylcholinetransferase; AChE, Acetylcholinesterase; CAS, catalytic active site; PAS, peripheral anionic site; QSAR, Quantitative structure-activity relationship; 3D-QSAR, Three-dimensional QSAR; CoMFA, Comparative molecular field analysis; CoMFA-RF, CoMFA region focusing; CoMSIA, Comparative molecular similarity index analysis; HQSAR, Hologram QSAR;MD, Molecular dynamic; AD, Aplication Domain.

Consent for Publication

Not applicable.

Conflict of Interest

The authors declare no conflict of interest, financial or otherwise.

Acknowledgements

We are grateful to Clinical Biochemistry Research Center, Basic Health Sciences Institute, Sharekord University of Medicinal Sciences for financial support of this research.

References

[1] Bukhari, S.N.A.; Jantan, I.; Masand, V.H.; Mahajan, D.T.; Sher, M.; Naeem-ul-Hassan, M.; Amjad, M.W. Synthesis of α, β-unsaturated carbonyl based compounds as acetylcholinesterase and butyrylcholinesterase inhibitors: Characterization, molecular modeling, QSAR studies and effect against amyloid β-induced cytotoxicity. *Eur. J. Med. Chem.*, **2014**, 83, 355-365.

[2] Bautista-Aguilera, O.M.; Esteban, G.; Bolea, I.; Nikolic, K.; Agbaba, D.; Moraleda, I.; Iriepa, I.; Samadi, A.; Soriano, E.; Unzeta, M.; Marco-Contelles, J. Design, synthesis, pharmacological evaluation, QSAR analysis, molecular modeling and ADMET of novel donepezil–indolyl hybrids as multipotent cholinesterase/monoamine oxidase inhibitors for the potential treatment of Alzheimer's disease. *Eur. J. Med. Chem.*, **2014**, *75*, 82-95.

[3] Vitorović-Todorović, M.D.; Cvijetić, I. N.; Juranić, I.O.; Drakulić, 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.

[4] Dighe, S,N.; Deora, G.S.; De, l.a.; Mora, E.; Nachon, F.; Chan, S.; Parat, M.-O.; Brazzolotto, X.; Ross, B.P. Discovery and structure–activity relationships of a highly selective butyrylcholinesterase inhibitor by structure-based virtual screening. *J. Med. Chem*., **2016**, *59*, 7683-7689.

[5] Tam, C.; Wong, J.H.; Ng, T.B.; Tsui, S.K.W.; Zuo, T. Drugs for targeted therapies of Alzheimer's disease. *Current. Med. Chem.*, 2018**,** *25*, 1-25**.**

[6] Puiatti, M.; Borioni, J.L.; Vallejo, M.G.; Cabrera, J.L.; Agnese, A.M.; Ortega, M.G.; Pierini, A.B. Study of the interaction of Huperzia saururus Lycopodium alkaloids with the acetylcholinesterase enzyme. *J. Mol. Graph. Model*., **2013**, *44*, 136-144.

[7] Mount, C.; Downton, C. Alzheimer disease: progress or profit? *Nature. Medicine*, **2006**, *12*, 780-784.

[8] Lalut. J.; Santoni, G.; Karila, D.; Lecoutey, C.; Davis, A.; Nachon, F.; Silman, I.; Sussman, J.; Weik, M.; Maurice, T.; Dallemagne, P.; Rochais, C. . Novel multitarget-directed ligands targeting acetylcholinesterase and σ1 receptors as lead compounds for treatment of Alzheimer's disease: Synthesis, evaluation, and structural characterization of their complexes with acetylcholinesterase. *Eur. J. Med. Chem*., **2019**, *162*, 234-248.

[9] Grundke-Iqbal, I.; Iqbal, K.; Tung, Y-C.; Quinlan, M.; Wisniewski, H.M.; Binder, L.I. Abnormal phosphorylation of the microtubule-associated protein tau (tau) in Alzheimer cytoskeletal pathology. *Proc. Natl. Acad. Sci*., **1986**, *83*, 4913-4917.

[10] Gella, A.; Durany, N. Oxidative stress in Alzheimer disease. *Cell. Adh. Migr*., **2009**, *3*, 88-93.

[11] Talesa VN. Acetylcholinesterase in Alzheimer's disease. *Mech. Ageing. Dev*., **2001**, *122*, 1961-1969.

[12] Goedert, M.; Spillantini, M.G. A century of Alzheimer's disease. *science*., **2006**, *314*, 777-781.

[13] Birks, J.; Grimley Evans, J.; Iakovidou, V.; Tsolaki, M. Rivastigmine for Alzheimer's disease. *The Cochrane database of systematic reviews*. **2000** (4):CD001191.

[14] Perry, E.; Walker, M.;, Grace, J.; Perry, R. Acetylcholine in mind: a neurotransmitter correlate of consciousness? *Trends. Neurosci*., **1999**, *22*, 273-280.

[15] Raevsky, O.A.; Mukhametov, A.; Grigorev, V.Y.; Ustyugov, A.; Tsay, S.C.; Jih-Ru Hwu, R.; Yarla, N.S.; Tarasov, V.V.; Aliev, G.; Bachurin, S.O. Applications of multi-target computer-aided methodologies in molecular design of CNS drugs**.** *Current. Med. Chem.*, **2018**, *25*, 5293-5314.

[16] Vitorović-Todorović, M,D.; Juranić, I.O.; Mandić, L.M.; Drakulić, B.J. 4-Aryl-4-oxo-N-phenyl-2-aminylbutyramides as acetyl- and butyrylcholinesterase inhibitors. Preparation, anticholinesterase activity, docking study, and 3D structure–activity relationship based on molecular interaction fields. *Bioorg. Med. Chem*., **2010**, *18*, 1181-1193.

[17] Scarpini, E.; Scheltens, P.; Feldman, H. Treatment of Alzheimer's disease: current status and new perspectives. *Lancet. Neurol*., **2003**, *2*, 539-547.

[18] Cheng, Z.Q.; Zhu, K.K.; Zhang, J.; Song, J.L.; Muehlmann, L.A.; Jiang, C.S.; Liu, C.L.; Zhang, H. Molecular-docking-guided design and synthesis of new IAA-tacrine hybrids as multifunctional AChE/BChE inhibitors. Bioorg. Chem. 2019, 83, 277-288.

[19] Korabecny, J.; Dolezal, R.; Cabelova, P.; Horova, A.; Hruba, E.; Ricny, J.; Sedlacek, L.; Nepovimova, E.; Spilovska, K.; Andrs, M.; Musilek, K.; Opletalova, V.; Sepsova, V.; Ripova, D.; Kuca, K. 7-MEOTA–donepezil like compounds as cholinesterase inhibitors: Synthesis, pharmacological evaluation, molecular modeling and QSAR studies. *Eur. J. Med. Chem*., **2014**, *82*, 426-438.

[20] Sun, X.; Jin, L.; Ling, P. Review of drugs for Alzheimer's disease. *Drug. Discov. Ther*., **2012**, *6*, 285-290.

[21] Racchi, M.; Mazzucchelli, M.; Porrello, E.; Lanni, C.; Govoni, S. Acetylcholinesterase inhibitors: novel activities of old molecules. *Pharmacological. Research*., **2004**, *50*, 441-451.

[22] Leon, R.; Garcia, A.G.; Marco-Contelles, J. Recent advances in the multitarget-directed ligands approach for the treatment of Alzheimer's disease. *Med. Res. Rev*., **2013**, *33*, 139-189.

[23] Sussman, J.L.; Harel, M.; Frolow, F.; Oefner, C.; Goldman, A.; Toker, L.; Silman, I. Atomic structure of acetylcholinesterase from Torpedo californica: a prototypic acetylcholine-binding *protein*. *Science*, **1991**, 253, 872-879.

[24] Munoz-Ruiz, P.; Rubio, L.; García-Palomero, E.; Dorronsoro, I.; del Monte-Millán, M.; Valenzuela, R.; Usán, P.; de Austria, C.; Bartolini, M.; Andrisano, V. Design, synthesis, and biological evaluation of dual binding site acetylcholinesterase inhibitors: new disease-modifying agents for Alzheimer's disease. *J. Med. Chem*., **2005**, *48*, 7223-7233.

[25] Chaudhaery, S.S.; Roy, K.K.; Saxena, A.K. Consensus superiority of the pharmacophore-based alignment, over maximum common substructure (MCS): 3D-QSAR studies on carbamates as acetylcholinesterase inhibitors. *J. Chem. Inf. Mod*., **2009**, 49, 1590-1601.

[26] Leonetti, F.; Catto, M.; Nicolotti, O.; Pisani, L.; Pisani, L.; Cappa, A.; Stefanachi, A.; Carotti, A. Homo-and hetero-bivalent edrophonium-like ammonium salts as highly potent, dual binding site AChE inhibitors. *Bioorg. Med. Chem*., **2008** *16*, 7450-7456.

[27] Kubinyi, H. QSAR and 3D QSAR in drug design Part 1: methodology. *Drug. Discov. Today*., **1997**, *2*, 457-467.

[28] , H. QSAR and 3D QSAR in drug design Part 2: applications and problems. *Drug. Discov. Today*., **1997**, *2*, 538-546.

[29] Zheng, X.; He, M.; Tan, X.; Zheng, J.; Wang, F.; Liu, S. 3D-quantitative structure–activity relationship and docking studies of coumarin derivatives as tissue kallikrein 7 inhibitors. *J. Pharm. Pharmacol*., **2017**, *69*, 1136-1144.

[30] Fang, C.; Xiao, Z. Receptor-based 3D-QSAR in drug design: methods and applications in Kinase studies. *Curr. Top. Med. Chem.*, **2016**, *16, 1463-1477*.

[31] Akamatsu, M. Current state and perspectives of 3D-QSAR. *Curr. Top. Med. Chem*., **2002**, *2*, 1381-1394.

[32] Abdizadeh, T.; Ghodsi, R,.; Hadizadeh, F. 3D-QSAR (CoMFA, CoMSIA) and molecular docking studies on histone deacetylase 1 selective inhibitors. *Recent. Pat. Anticancer. Drug. Discov*., **2017**, *12*, 365-383.

[33] Verma, J.; Khedkar, V.M.; Coutinho, E.C. 3D-QSAR in drug design-a review. *Curr. Top. Med. Chem*., **2010**, *10*, 95-115.

[34] Wu. S.; Qi, W.; Su, R.; Li, T.; Lu, D.; He, Z. CoMFA and CoMSIA analysis of ACE-inhibitory, antimicrobial and bitter-tasting peptides. *Eur. J. Med . Chem*., **2014**, *84*, 100-106.

[35] Gupta, N.; Vyas, V.K.; Patel, B.; Ghate, M. Predictive 3D-QSAR and HQSAR model generation of isocitrate lyase (ICL) inhibitors by various alignment methods combined with docking study. *Med. Chem. Res*., **2014**, *23*, 2757-2768.

[36] Sharma, R.; Dhingra, N.; Patil, S. CoMFA, CoMSIA, HQSAR and molecular docking analysis of ionone-based chalcone derivatives as antiprostate cancer activity. *Indian.J. Pharm. Sci*., **2016**, *78*, 54.

[37] Cramer III, R.D.; Bunce, J.D.; Patterson, D.E.; Frank, I.E. Crossvalidation, bootstrapping, and partial least squares compared with multiple regression in conventional QSAR studies. *Quant. Struct.Act. Relat*., **1988**, *7*, 18-25.

[38] Kellogg, G.E.; Semus, S.F.; Abraham, D.J. HINT: a new method of empirical hydrophobic field calculation for CoMFA. *J. Comput. Aid. Mol. Des*., **1991**,*5*, 545-552.

39. Borisa, A.; Bhatt, H. 3D-QSAR (CoMFA, CoMFA-RG, CoMSIA) and molecular docking study of thienopyrimidine and thienopyridine derivatives to explore structural requirements for aurora-B kinase inhibition. *Eur. J. Pharm. Sci*., **2015**, *79*, 1-12.

[40] Klebe, G.; Abraham, U.; Mietzner, T. Molecular similarity indices in a comparative analysis (CoMSIA) of drug molecules to correlate and predict their biological activity. *J. Med. Chem*., **1994**, *37*, 4130-4146.

[41] Moda, T.L.; Montanari, C.A.; Andricopulo, A.D. Hologram QSAR model for the prediction of human oral bioavailability. *Bioorg. Med. Chem*., **2007**, *15*, 7738-7745.

[42] Lowis, D.R. HQSAR: a new, highly predictive QSAR technique. *Tripos. Technical. Notes*., **1997**,*1*, 17.

[43] Castilho, M.S.; Postigo, M.P.; de Paula, C.B.; Montanari, C.A.; Oliva, G.; Andricopulo, A.D. Two-and three-dimensional quantitative structure–activity relationships for a series of purine nucleoside phosphorylase inhibitors. *Bioorg .Med. Chem*., **2006**,*14*, 516-527.

[44] Dunn Iii, W.; Wold, S.; Edlund, U.; Hellberg, S.; Gasteiger, J. Multivariate structure-activity relationships between data from a battery of biological tests and an ensemble of structure descriptors: The PLS method. *Quant. Struct.Act .Relat*., **1984**, *3*, 131-137.

[45] Geladi, P. Notes on the history and nature of partial least squares (PLS) modelling. *J. Chemom*., **1988**, *2*, 231-246.

[46] Kubinyi, H.; Martin, YC.; Folkers, G. 3D QSAR in drug design: volume 1: theory methods and applications: *Springer Science & Business Media*; **1993**.

[47] Bush, B.L.; Nachbar, R.B. Sample-distance partial least squares: PLS optimized for many variables, with application to CoMFA. *J. Comput. Aid. Mol. Des*., **1993**, *7*, 587-619.

[48] Li, X.; Wang, H.; Lu, Z.; Zheng, X.; Ni, W.; Zhu, J.; Fu, Y.; Lian, F.; Zhang, N.; Li, J.; Zhang, H.; Mao, F. Development of multifunctional pyrimidinylthiourea derivatives as potential anti-Alzheimer agents. *J. Med. Chem*., **2016**, *59*, 8326-8344.

[49] Clark, M.; Cramer, R.D.; Van, O.N. Validation of the general purpose Tripos 5.2 force field. *J. Comput. Chem*., **1989**, *10*, 982-1012.

[50] Politi, A.; Durdagi, S.; Moutevelis-Minakakis, P.; Kokotos, G.; Papadopoulos, M.G.; Mavromoustakos, T. Application of 3D QSAR CoMFA/CoMSIA and in silico docking studies on novel renin inhibitors against cardiovascular diseases. *Eur. J. Med. Chem*., **2009**, *44*, 3703-3711.

[51] Sainy, J.; Sharma, R. QSAR analysis of thiolactone derivatives using HQSAR, CoMFA and CoMSIA. *SAR. QSAR. Environ. Res*., **2015**, *26*, 873-892.

[52] Waller, C.L. A comparative QSAR study using CoMFA, HQSAR, and FRED/SKEYS paradigms for estrogen receptor binding affinities of structurally diverse compounds. *J. Chem. Inf. Compu. Sci*., **2004**, *44*, 758-765.

[53] Zhang, H.; Li, H.; Liu, C. CoMFA, CoMSIA, and molecular hologram QSAR studies of novel neuronal nAChRs ligands-open ring analogues of 3-pyridyl ether. *J. Chem. Inf. Model*., **2005**, *45*,440-448.

[54] Jiao, L.; Zhang, X.; Qin, Y.; Wang, X.; Li, H. Hologram QSAR study on the electrophoretic mobility of aromatic acids. *Chemom. Intell. Lab. Sys*., **2016**,*157*, 202-207.

[55] Sun, J.; Mei, H. QSAR and molecular mechanism analysis of N-substituted oseltamivir derivatives as potent avian influenza H5N1 neuraminidase inhibitors. *Chemom. Intell. Lab. Sys*., **2015**, *146*, 485-493.

[56] Ståhle, L.; Wold, S. Partial least squares analysis with cross-validation for the two-class problem: A Monte Carlo study. *J. Chemom*., **1987**, *1*, 185-196.

[57] Wold, S. Cross-validatory estimation of the number of components in factor and principal components models. *Technometrics*., **1978**, *20*, 397-405.

[58] Kearns, M.; Ron, D. Algorithmic stability and sanity-check bounds for leave-one-out cross-validation. *Neural. Comput*., **1999**, *11*, 1427-1453.

[59] Golbraikh, A.; Tropsha, A. Beware of q2! *J. Mol. Graph. Mod*., **2002**, *20*, 269-276.

[60] Rácz, A.; Bajusz, D.; Héberger, K. Consistency of QSAR models: correct split of training and test sets, ranking of models and performance parameters. *SAR. QSAR. Environ. Res*., **2015**, *26*, 683-700.

[61] Wang, Z.; Cheng, L.; Kai, Z.; Wu, F.; Liu, Z.; Cai, M. Molecular modeling studies of atorvastatin analogues as HMGR inhibitors using 3D-QSAR, molecular docking and molecular dynamics simulations. *Bioorg. Med. Chem. Lett*., **2014**, *24*, 3869-3876.

[62] Zhang, S.; Lin, Z.; Pu, Y.; Zhang, Y.; Zhang, L.; Zuo, Z. Comparative QSAR studies using HQSAR, CoMFA, and CoMSIA methods on cyclic sulfone hydroxyethylamines as BACE1 inhibitors. *Comput. Biol. Chem*., **2017**, *67*, 38-47.

[63] Lorca, M.; Morales-Verdejo, C.; Vásquez-Velásquez, D.; Andrades-Lagos, J.; Campanini-Salinas, J.; Soto-Delgado, J.;, Recabarren-Gajardo, G.; Mella, J. . Structure-activity relationships Based on 3D-QSAR CoMFA/CoMSIA and design of aryloxypropanol-amine agonists with selectivity for theh β3-adrenergic receptor and anti-obesity and anti-diabetic profiles. *Molecules*., **2018**, *23*, 1191.

[64] Rücker, C.; Rücker, G.; Meringer, M. y-Randomization and its variants in QSPR/QSAR. *J. Chem. Inf. Model*., **2007**, *47*, 2345-2357.

[65] Dhingra, R.; Malhotra, M.;Sharma, V.; Bhardwaj, T. R.; Dhingra, N. Exploration of novel 5α-Reductase inhibitors for benign prostatic hyperplasia by 2D/3D QSAR, cytotoxicity Pre-ADME and docking studies. *Curr. Top. Med. Chem*, DOI: 10.2174/1568026619666190119145959

[66] Weaver, S.; Gleeson, M.P. The importance of the domain of applicability in QSAR modeling. J. Mol. Graph. Mod., 2008, 26, 1315-1326.

[67] Kaneko, H.; Funatsu, K. Applicability Domain Based on Ensemble Learning in Classification and Regression Analyses. *J. Chem. Inf. Model*., **2014**, *54*, 2469-2482.

[68] Veerasamy, R.; DRajak, H.; Jain, A.; Sivadasan, S.; Varghese, C. P.; Agrawal, R.K. Validation of QSAR models-strategies and importance. Inter. *J. Drug. Des. Discov*., **2011**, *2*, 511-519.

[69] Yang, X.; Liu, H.; Yang, Q.; Liu, J.; Chen, J.; Shi, L. Predicting anti-androgenic activity of bisphenols using molecular docking and quantitative structure-activity relationships. *Chemosphere*, **2016**, *163*, 373-381.

[70] Lei, T.; Chen, F.; Liu, H.; Sun, H., Kang, Y.; Li, D.; Li, Y.; Hou, T. ADMET Evaluation in Drug Discovery. Part 17: Development of Quantitative and Qualitative Prediction Models for Chemical-Induced Respiratory Toxicity. *Mol. Pharm*., **2017**, *14*, 2407-2421.