http://dx.doi.org/10.5012/bkcs.2011.32.12.4352

4D-QSAR Study of p56^{lck} Protein Tyrosine Kinase Inhibitory Activity of Flavonoid Derivatives Using MCET Method

Hayriye Yılmaz,[*] Yahya Güzel,[†] Zulbiye Önal,[†] Gökçe Altıparmak,[†] and Safak Ozhan Kocakaya[‡]

Erciyes University, Faculty of Pharmacy, Kayseri 38039, Turkey. [*]E-mail: hayriyey@erciyes.edu.tr
[†]Erciyes University, Faculty of Science, Department of Chemistry, Kayseri 38039, Turkey
[‡]University of Dicle, Faculty of Science, Department of Chemistry, Diyarbakir 21280, Turkey
Received July 27, 2011, Accepted October 17, 2011

A four dimensional quantitative structure activity relationship analysis was applied to a series of 50 flavonoid inhibitors of p56^{lck} protein tyrosine kinase by the molecular comparative electron topological method. It was found that the -log (IC50) values of the compounds were highly dependent on the topology, size and electrostatic character of the substituents at seven positions of the flavonoid scaffold in this study. Depending on the negative or positive charge of the groups correctly embedded in these substituents, three-dimensional bio-structure to increase or decrease -log (IC50) values in the training set of 39 compounds was predicted. The test set of 11 compounds was used to evaluate the predictivity of the model. To generate 4D-QSAR model, the defined function groups and pharmacophore used as topological descriptors in the calculation of activity were of sufficient statistical quality ($R^2 = 0.72$ and $Q^2 = 0.69$). Ligand docking approach by using Dock 6.0. These compounds include many flavonoid analogs, They were docked onto human families of p56lck PTKs retrieved from the Protein Data Bank, 1lkl.pdb.

Key Words : Drug design, 4D-QSAR, Flavonoid, ETM, Protein tyrosine kinase

Introduction

The protein tyrosine kinases (PTKs) are emerging as an exciting class of targets for drug discovery.[1] It plays a very important role in control of cellular signaling and is involved in multiplication, differentiation and metabolism.[2] Abnormal PTKs signaling has been identified in a wide range of diseases including cancer,[3] inflammation,[4] neurological and neurodegenerative diseases, cardiovascular diseases, allergies and asthma or hormone-related diseases. Therefore, considerable effort has been made to identify PTK inhibitors that are effective as therapeutic agents against these diseases. p56^{lck} is a lymphoid-specific protein tyrosine kinase that is principally expressed in T lymphocytes.[5] The association of p56^{lck} with the cytoplasmic tail of various cell surface receptors, as well as its associations with intracellular targets of phosphorylation, suggests that this tyrosine kinase plays a central role in coordinating early signal transduction events.[6] Based on this knowledge, it is clear that substances which can simulate the activity of PTKs might be potentially effective therapeutic agents.[7] A variety of compounds can inhibit the function of PTKs in a manner which is competitive with respect to nucleotide binding. Among such competitive inhibitors are flavonoids, a group of low molecular weight plant natural products that include one of the largest classes of naturally occurring polyphenolic compounds.[8,9] The enzyme inhibitory activity of a series of flavonoids was measured by using p56^{lck}, a lymphoid cell lineage-specific PTK of the src family which is over expressed in several lymphomas.[10] A number of compounds have been studied to find their inhibitory activities against these PTKs.

The quantitative structure-activity relationship (QSAR) model seeks to discover and uses mathematical relationships between chemical structure and biological activity. The QSAR models help in understanding the nature of the drug-target interactions for the various biological targets.[11] Once the structure is known, any molecular descriptor can be calculated, whether or not the compound is synthesized. The topological index, geometrical descriptors, and other descriptors are included in QSAR studies.[12]

Three-dimensional (3D) QSAR, analysis is a major applications methodology in computer-assisted molecular design, CAMD. Several schemes to do 3D-QSAR analysis have been developed and are discussed in recent reviews.[13-15] Probably the most popular, and one of the first 3D-QSAR schemes, is comparative molecular field analysis, CoMFA.[16]

Basically, four-dimensional (4D) QSAR investigates the conformational space of the molecular objects.[17] We attempt to find consistent relationships between values of molecular properties and biological activity, and try to rationalize 4D-QSAR analyses using the molecular comparative electron topological (MCET) method developed by us. Our main research includes choosing the atoms responsible for activity with the 3D-space of the conformers. This method progresses into the development of 4D-QSAR using the atomic characteristic values in the data set according to the topological descriptors, which consist of distance, angle and dihedral-angle.

4D-QSAR analysis has three inherent problems to overcome. First is the identification of the common and functional groups of conformations/molecular shapes of flexible compounds in the set. In the most straightforward interpret-

ation, particularly for in Vitro activity, the bio-structure of a ligand corresponds to the receptor-bound shape. Our working definition of the bio-structure is that it is the one which optimizes the quantitative 4D-QSAR model together with the selected conformers of each ligand. The second problem to be overcome is the specification of the basis for comparing conformers in constructing a 4D-QSAR which is referred to as the molecular alignment.[18] Finally, each molecule in the training set must be partitioned with respect to intermolecular (receptor) interactions. That is, different parts of the conformers of each molecule can be expected to have different types of interactions with sites on a common receptor and/or in a common medium. This partitioned form of the molecule is called the bio-structure. The 4D-QSAR formalism has been developed to deal with each of these problems in constructing a 4D-QSAR model.

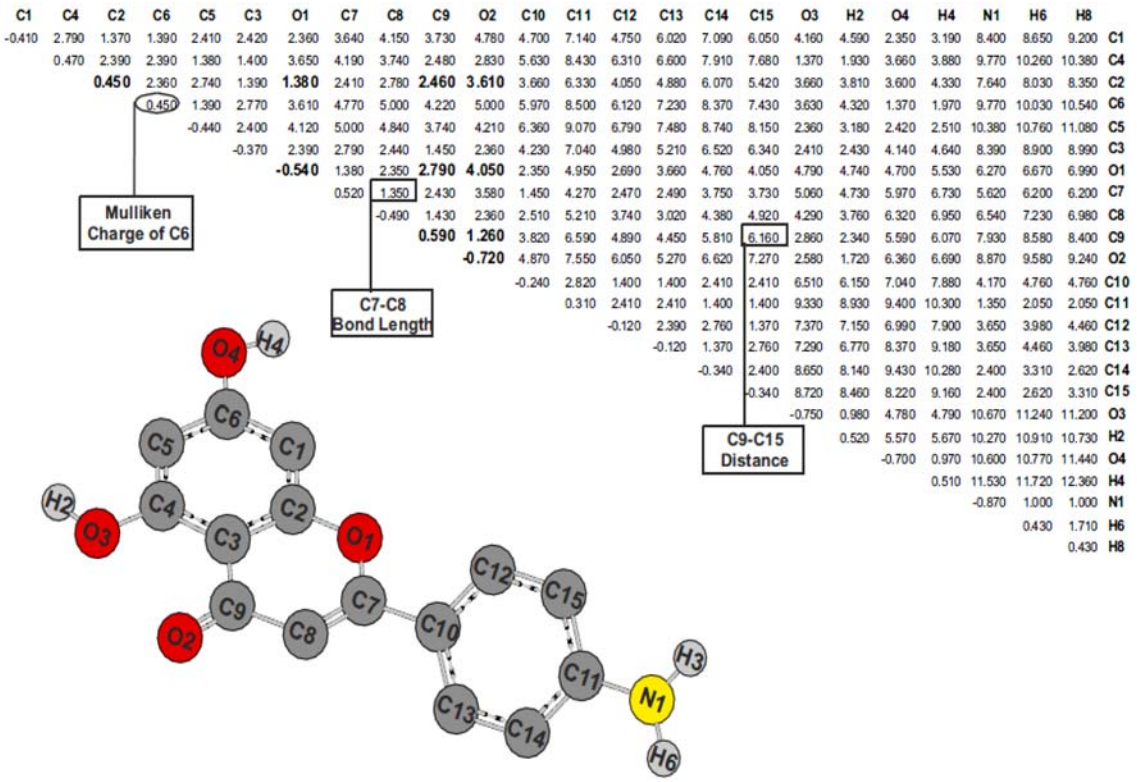
Materials and Methods

A set of flavonoid derivatives was collected from the literature along with their activity data.[7] Both molecular mechanics and ab initio quantum chemical calculations were used for all compounds, respectively to obtain the most stable conformers and to find the optimum 3D geometries of conformers, the heats of formation, the Mulliken charge on each atom and so on. The conformational analysis was performed applying the systematic search tool available in the PC Spartan'08 using the MMFF force field.[19] Default options were used, including the maximum number of 100 conformations and an energy cutoff of 10 kcal mol^{-1} from the minimum energy conformation found. The conformations generated were optimized using the Hartree-Fock/3-21G(*) level (in water). We excluded the most similar conformations for each compound according to all atoms superposition. The acceptable conformers for each compound were selected according to heat of formation with the lowest energy. Output files obtained from Spartan'08 parallel software were transformed into Electron Topological Matrix (ETM) as a different representation of the electronic and geometric properties of the conformer in a numerical form. All significantly populated conformations should be subjected to the electronic structure calculations and construction of the ETM. After all the electron topological parameters were auto loaded from the ETM database, upper triangular part of ETM, which was a symmetric matrix, was used in the computational process. Each of the ETMs given in the Figure 1 represents only one conformer in an electronic and geometric way. The ETM was taken as a language for conformer description (one matrix = one conformer).

The ETM formation proceeds in the following way[20];

1) Let A_i be an atom of a conformer being described, the corresponding diagonal matrix element a_{ii} is one of m1 which is the number of local atomic properties (for example, charge, polarizability, HOMO- and LUMO-orbital coefficients, etc.)

2) If A_i and A_j are any two atoms of the conformer then



C1	C4	C2	C6	C5	C3	O1	C7	C8	C9	O2	C10	C11	C12	C13	C14	C15	O3	H2	O4	H4	N1	H6	H8	
-0.410	2.790	1.370	1.390	2.410	2.420	2.360	3.640	4.150	3.730	4.780	4.700	7.140	4.750	6.020	7.090	6.050	4.160	4.590	2.350	3.190	8.400	8.650	9.200	C1
	0.470	2.390	2.390	1.380	1.400	3.650	4.190	3.740	2.480	2.830	5.630	8.430	6.310	6.600	7.910	7.680	1.370	1.930	3.660	3.880	9.770	10.260	10.380	C4
		0.450	2.360	2.740	1.390	**1.380**	2.410	2.780	**2.460**	**3.610**	3.660	6.330	4.050	4.880	6.070	5.420	3.660	3.810	3.600	4.330	7.640	8.030	8.350	C2
			0.450	1.390	2.770	3.610	4.770	5.000	4.220	5.000	5.970	8.500	6.120	7.230	8.370	7.430	3.630	4.320	1.370	1.970	9.770	10.030	10.540	C6
				-0.440	2.400	4.120	5.000	4.840	3.740	4.210	6.360	9.070	6.790	7.480	8.740	8.150	2.360	3.180	2.420	2.510	10.380	10.760	11.080	C5
					-0.370	2.390	2.790	2.440	1.450	2.360	4.230	7.040	4.980	5.210	6.520	6.340	2.410	2.430	4.140	4.640	8.390	8.900	8.990	C3
						-0.540	1.380	2.350	**2.790**	**4.050**	2.350	4.950	2.690	3.660	4.760	4.050	4.790	4.740	4.700	5.530	6.270	6.670	6.990	O1
							0.520	1.350	2.430	3.580	1.450	4.270	2.470	2.490	3.750	3.730	5.060	4.730	5.970	6.730	5.620	6.200	6.200	C7
								-0.490	1.430	2.360	2.510	5.210	3.740	3.020	4.380	4.920	4.290	3.760	6.320	6.950	6.540	7.230	6.980	C8
									0.590	**1.260**	3.820	6.590	4.890	4.450	5.810	6.160	2.860	2.340	5.590	6.070	7.930	8.580	8.400	C9
										-0.720	4.870	7.550	6.050	5.270	6.620	7.270	2.580	1.720	6.360	6.690	8.870	9.580	9.240	O2
											-0.240	2.820	1.400	1.400	2.410	2.410	6.510	6.150	7.040	7.880	4.170	4.760	4.760	C10
												0.310	2.410	2.410	1.400	1.400	9.330	8.930	9.400	10.300	1.350	2.050	2.050	C11
													-0.120	2.390	2.760	1.370	7.370	7.150	6.990	7.900	3.650	3.980	4.460	C12
														-0.120	1.370	2.760	7.290	6.770	8.370	9.180	3.650	4.460	3.980	C13
															-0.340	2.400	8.650	8.140	9.430	10.280	2.400	3.310	2.620	C14
																-0.340	8.720	8.460	8.220	9.160	2.400	2.620	3.310	C15
																	-0.750	0.980	4.780	4.790	10.670	11.240	11.200	O3
																		0.520	5.570	5.670	10.270	10.910	10.730	H2
																			-0.700	0.970	10.600	10.770	11.440	O4
																				0.510	11.530	11.720	12.360	H4
																					-0.870	1.000	1.000	N1
																						0.430	1.710	H6
																							0.430	H8

Figure 1. The illustrated ETM and three dimensional structure of reference compound n01. Diagonal elements show atomic charges while non-diagonal elements show bond length, which is distance between chemically bonded atoms, or distance between atoms. The sub-matrix which belongs to the Pha is given in bold letter.

two cases may occur:

a. A_i and A_j are not chemically bonded. In this case the distance between these two atoms represents the corresponding non-diagonal matrix element, a_{ij}.

b. A_i and A_j are chemically bonded, and a_{ij} describes this bond by means of electronic properties (bond length, bond energy, etc.). Suppose that we have m2 which is the number of such characteristics. As a result, the number of all ordinary n*n matrices is m = m1*m2 (m-dimensional ETM). The matrices, based on the data known from quantum chemical calculations, can be formed as well as the number of atoms in the corresponding conformer. It has been assumed that a single electronic matrix in the ETM represents a configuration of molecule. The matrix, which is containing charge on the diagonal, bond length and distance on the non-diagonal, is one-dimensional as shown in Figure 1.

Both in the electron topological (ET)[21-24] and in the electron conformer (EC) methods,[25] the ETM have been employed as a tool for exploring bio-structure. The ET method accounts to only one conformer with the lowest energy of each ligand, and emphasis upon only the pharmacophore (Pha) in three dimensions, therefore it is an investigation of 3D-QSAR. Likewise, the EC method, which uses the auxiliary group and anti-pharmacophore shielding group besides Pha, each of which influence the activities, is an analysis of 4D-QSAR, because it takes into account all the acceptable conformers more than one conformer.[25] Since the formulation of 4D-QSAR analyses used in the EC method has been previously described in detail by Bersuker and *et al.*,[26] it was only summarized as following in this study. While the Pha is responsible for the existence or non-existence of the activity, the APS and AG, which are functional groups of the conformers having the Pha, are characteristic variable, and affect of the activity values.

A criterion that is commonly used in structural methods for evaluating the probable Pha occurrence in a series under study is given by the following formulas:

$$P_\alpha = (\nu 1 + 1)/(\nu 1 + \nu 3 + 2);$$

$$\alpha_a = (\nu 1 * \nu 4 - \nu 2 * \nu 3)/(\mu 1 * \mu 2 * \mu 3 * \mu 4)^{1/2}$$

where $\nu 1$ and $\nu 2$ are the numbers of molecules possessing and not possessing, respectively, the feature of activity in the class of molecules; $\nu 3$ and $\nu 4$ have analogous meaning in the low active molecules; and $\mu 1$ and $\mu 2$ are the numbers of molecules in the class of active and low active molecules $\mu 3 = \nu 1 + \nu 3$; $\mu 4 = \nu 2 + \nu 4$. In this way, P_α evaluates the deposit of only active molecules, while α_a reflects the deposit of both active and low active molecules in the feature of the activity found. Then, without setting any constraints on tolerance values, maximum tolerance values are defined for all active molecules.[25]

To reveal the presence (or absence) of Pha, MCET method is processed with comparing ETMs. The choice of Pha is based upon matching of the acceptable conformers of the other compounds with the reference compound's conformer, preferably ones with the lowest energy. Once reasonable Pha responsible for the activity is found, the new Cartesian coordinate values of conformer is placed according to its three ordered atoms. After Pha of each conformer is aligned using the tolerance intervals, the functional group around of its is seek. To form bio-structure, those functional groups together with Pha can be estimated in the Eq. (1) as independent variables.

$$S_{ni} = \sum_{j=1}^{J} \kappa_j \cdot a_{ni}^{(j)} \tag{1}$$

where j and J are the number and the total number of chosen parameters, κ_j is adjustable constant in the j^{th} position and is found with Newton Raphson approximation, a_{ni}^{j} and is the independent variable described for the j^{th} position in the i^{th} conformation of the n^{th} compound.

$$A_n = A_l \frac{\sum_{i=1}^{m_l} e^{-E_{li}/kT} \sum_{i=1}^{m_n^{Pha}} e^{-S_{ni}} e^{-E_{ni}/kT}}{\sum_{i=1}^{m_n} e^{-E_{ni}/kT} \sum_{i=1}^{m_l^{Pha}} e^{-S_{li}} e^{-E_{li}/kT}} \tag{2}$$

In Eq. (2),[26]

A_l/A_n : the activity of the reference l^{th}/n^{th} compound
T : the room temperature
k : Boltzman constant
m_n/m_l : the number of the acceptable conformers of the n^{th}/l^{th} compound
m_n^{Pha}/m_l^{Pha}, : the number of conformers, containing Pha, of the n^{th}/l^{th} compound
E_{ni} : the energy of i^{th} conformer of the n^{th} compound

A brief summary of the basic algorithm proposed in the MCET method is given in Figure 2. The method in this study is modified using C# programming language by us in such a similar form of the EC, but the applied procedure and the computer algorithm is different from that described. Some of the most important progresses and differences in the method are following:

1) After the Pha is found by comparing the acceptable conformers not just for the selected molecules but for all in training set with the lowest energy conformation of reference, the conformers containing Pha are aligned with respect to Pha's atoms, and then the spatial orientation of the rest atoms are again determined.

2) In the EC method, it is pointed out that the estimation of the variables in Eq. (1) is especially time-consuming because so far, an algorithm to extract them automatically from the ETM has not been yet found, so it has been done individually for each conformation.[26] However, such an algorithm has been constructed and employed in our method. To do this independent variables with the a_{ni}^{j} are estimated, and the values of the corresponding κ_j as parameter have been simultaneously predicted through Newton-Raphson approximation.[27]

For many κ_j-problem, we would need values on the all other κ_j to solve one κ_j for j = 1, 2,. J. Such a calculation requires a self-consistent approach as outlined below.

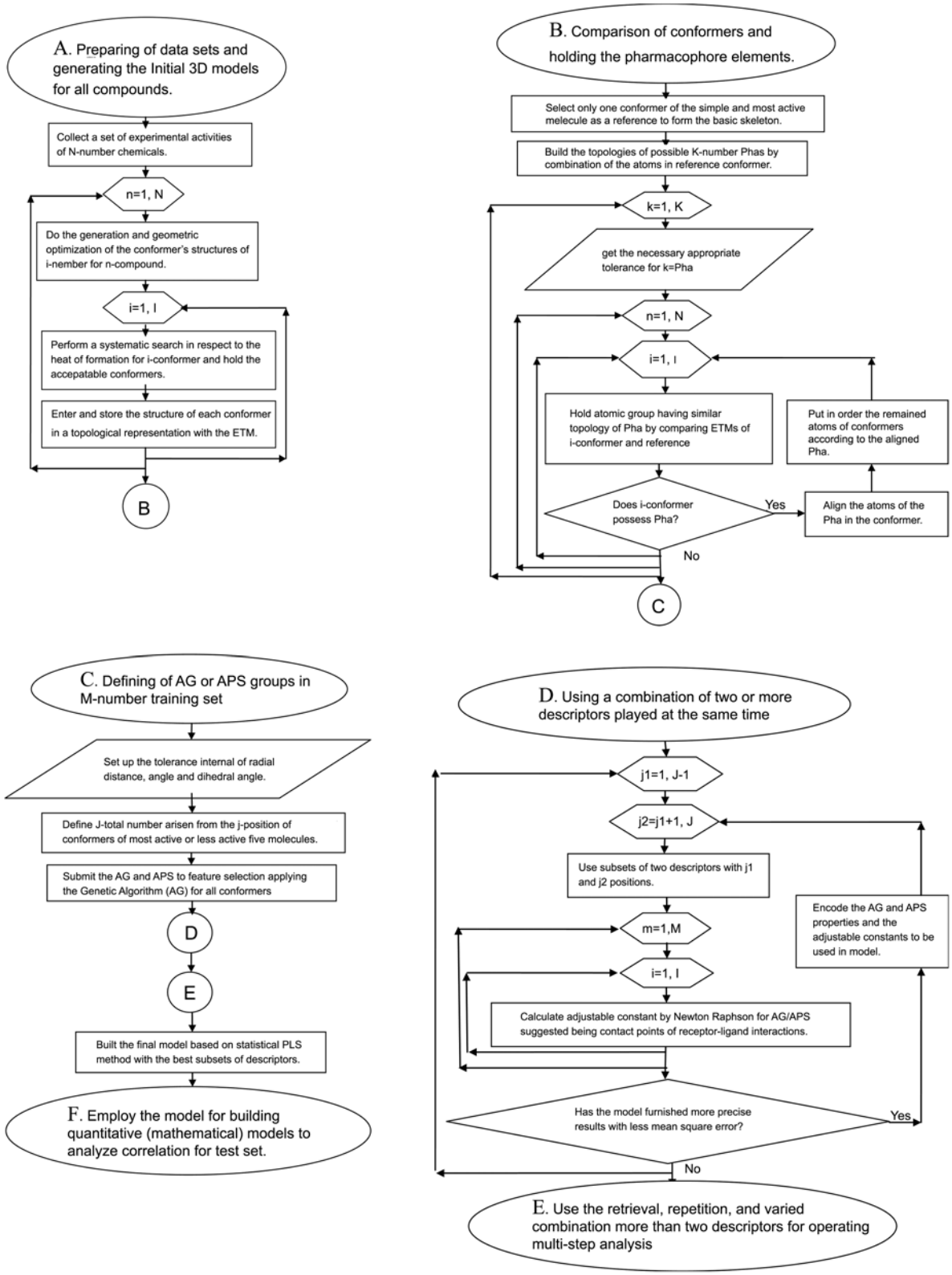


Figure 2. Algorithm of the MCET method.

a. Start with trial κ_j as number J-1 (except one κ_1).
b. To calculate one κ_j, use all remained κ_j.
c. Solve the Newton-Raphson Equation derived from Eq. (2) for one κ_j.
d. Define the new value of one κ_j.
e. Repeat the calculation for another κ_j coming after the defined ones.
f. Continue iteration until self-consistent results for all the

κ_j are obtained, i.e., the activities of all molecules don't change from one iteration to the next.

According to the independent variables present in the selected positions, all the κ_j are calculated to form the 4D-QSAR model in Eq. (2).

3) To find the Pha, both EC and MCET methods use the atomic properties, bond length and distance in ETM. In the EC method, although the various topological properties such as distance, angle, dihedral angle etc. are employed as independent variables, in the present investigation those topological properties are used to determine only the positions of the functional groups. Since the principal aim of this study is to search the bio-structure in the interaction points of the receptor, the functional groups with the positive and negative charges stated in related regions of the ligand. To pick out these groups distributed in various regions around the ligand, while the comparative molecular field analysis (CoMFA) uses the standard probe atoms,[16] our method uses the functional groups oriented around the Pha.

4) The selected positions in Eq. (1) through a genetic algorithm (GA) control variable, which specifies the number of independent variables in the 4D-QSAR models, are varied in order to determine the optimal number of descriptors.[28,29] GA can not only automatically select the optimum number of descriptors in regression analysis, but also construct multiple nonlinear regression (MNR) models. Therefore, GA is employed with MNR analysis for deriving and validating the model arised from Eq. (2) which includes nonlinear exponential functions, and as a result, it uses in the model the optimum calculated values of the κ_j, which are in good accordance with given values.

The bio-structure and the elevated model showed the topology of the simultaneously interactions of the ligand with the receptor. The action and nature of interactions between ligand and target protein were represented with Pha and functional groups of each conformer. The atomic charges of the functional groups together with Pha placed in the model should act like the electrostatic potential across the side of the receptor, and were used to interpret the results and to construct the 4D-QSAR model. For the calculation of the activity of each compound, Eq. (2) was employed taking into account the heat of formation of the conformers and their functional groups. κ-values related to positions in the model were simultaneously calculated from the activity of the compounds in the training set, and then used to predict the activity of those in test set. During the selection of the independent variables basing on partial least-squares (PLS),[30] leave-one-out cross-validation (LOOCV) was performed to identify an optimal number of explanatory variables or components which yielded an optimal model.[31] The most highly weighted PLS descriptors were used to form the trial basis set among the positioned atoms for GA model optimization.

Since the different positions of one conformer to another occurred due to the interchange of conformers, the variables at the defined positions might be present or absent for different conformers of any compound in Table 3. According to the selected positions in Table 2, which were the most appropriate functional groups affecting activity, adjustable constants (κ_j $j = 1, 2 \ldots J$) of corresponding interactions points were evaluated with Eq. (2). The security of the results was based upon the magnitude of the corresponding PLS regres- sion coefficients. For construction of a statistical model, which explains the observed biological activity, a procedure which involves iterative selection of structures, can be repeated until R^2 is reaches a high value.

According to Eq. (2) if the multiplication of κ-value by atomic charge had the negative result, because of the negative sign on the front of S_{ni}, the atom under consideration supported the ligand binding and acted as auxiliary group, otherwise a positive result of multiplication impeded and acted as anti-pharmacophore shielding group. According to Boltzmann distribution, conformers of higher energy have little effect on activity.

The most important reasons of the changing of the activity are that 1) the conformers with the lowest energy have the largest effect on the activity, 2) when the functional groups at the defined positions are present as auxiliary group, the activity increases and vice versa and 3) as another reason of the changing activity, the magnitude of different atomic charge should be taken into consideration.

Molecular Docking. In this study, dock calculations were performed to accommodate the flavonoid inhibitors within the p56 lck protein tyrosine kinases (PTKs). The docking studies of the guests were carried out using the protein tyrosine kinases (PTKs) and flavonoid parameters obtained for the minimized structures of the complexes between the protein and guests. The initial coordinates of the host and guest molecules for the docking studies were obtained via MD studies for 1 ns for the minimized the guests. Docking of the guests was caried out using program DOCK 6.0.[32] Docking was performed with default settings to obtain a population of possible conformations and orientations for the guests at the binding site. Spheres around the centre of the binding pocket was defined as binding pocket for the docking runs.

Molecular Dynamics Simulations. Molecular dynamics (MD) simulations were carried out in a Linux-Cluster system to determine conformations for the each flavonoid inhibitor molecules studied. All simulations were conducted by using AMBER (version 9.0)[33] suit of programs. AM1-Bcc (Austian model with Bond and charge correction)[34] atomic partial charges for the receptor and guests were determined by antechamber module of AMBER (v9) package. The General AMBER Force Field (GAFF)[35] was adopted in simulation because it handles small organic molecules. The individual contribution of Van der Waals electrostatic terms (in kcal mol^{-1}) to the free energy of binding (Dock score energy) of the enzyme to the guests calculated by Docking are given in Table 4.

Results and Discussion

4D-QSAR analysis of $p56^{lck}$ protein tyrosine kinase for the inhibitory activity of 50 flavonoid derivatives were

Table 1. The atoms of Pha in the reference molecule, n01, and its sub matrix

C2	O1	C9	O2
0.45	1.376	2.459	3.606
	–0.538	2.792	4.049
		0.591	1.257
			–0.719

Table 2. The functional groups and their positions defined within the tolerance values (Δ_{DA} = 15.0°, Δ_{An} = 15.0° and Δ_{Dis} = 1.0 Å) of dihedral-angle, angle and distance

Reference compound No:	Atom No:	Dihedral Angle	Angle	Distance[a]	Position Type
n01	C1	–0.01	30.87	2.36	a
n50	O6	22.37	169.35	6.72	b
n50	O4	–12.32	61.64	3.99	c
n48	C5	–34.00	0.56	4.13	d
n49	C16	21.97	34.25	5.57	e
n46	C17	–21.80	153.06	7.00	f
n47	O3	–52.54	13.58	6.16	g

[a]Distance between one of the pha atoms and the oriented atom.

studied. The values and positions of the Pha, shown as a matrix in Table 1, were unique for the molecule under consideration according to the principles of quantum mechanics. The lowest energy conformations at the most or least active compounds (n01, n46-n50) in Table 2 were selected as template structure containing effect positions, defined through distance, angle and dihedral angle. The functional groups at the positions around Pha were generated from substituent of these conformers, and the groups of remained conformers were matched and located by comparing with these positions within the tolerance values. The activity was calculated much more accurately via the functional groups of the conformers.

In Table 3, the functional groups for each conformer were given by the letters a, b, c, etc. together with heat of formation. Also the chemical structures of the flavonoids and their binding affinities were represented by the pIC_{50} scale (-log IC_{50}). One fifth of data set was used as the test set, and the remained compounds served as the training set. Thus, a series of 39 chosen compounds for the training set was used to construct model, and then the remaining 11 compounds, which were applied to the test set and marked with '§', were employed for the model validation.

The bio-structure forming the resulted model was shown in orange and yellow circles in Figure 3. According to Boltzmann distribution, each acceptable conformer's contributions were taken into account in Eq. (2), and were not again given in the following detailed explanation of Table 3. Also, although the effects of the functional groups were quantitatively studied for the activity of compounds, the following conclusion was qualitatively made and the illustration might be given as a result of the approach in Table 3. For compounds n08, n11, n12, n16, n17, n20, n22, n29 and n34,

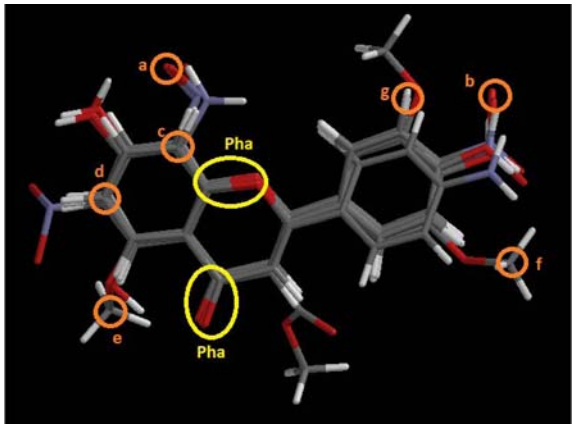


Figure 3. Bio-structure consists of Pha (yellow circles) and the functional groups (orange circles).

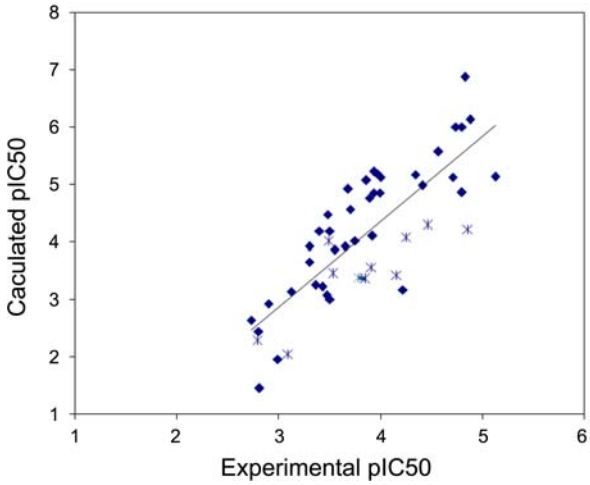
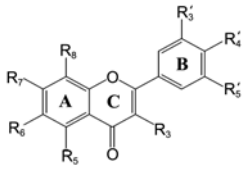


Figure 4. Correlation between experimental and calculated data sets on estrogenic activity given in training set and test set*.

since their conformers have only a-position, the activities were close to each other, and any minor differences among the activities were arisen from the different values of the atomic charges at this position. As for the compounds (n01-07, n09 and n23) whose conformers include both a- and b-positions, the activity values were bigger than those before, and so the group in b-position was acted as an auxiliary group (AG). Since the c-, d- and e-positions were mixed together into the conformers of the remained compounds; no comment about the contribution of any one of them could be made. But it was obvious that the activities of the compounds having one or more of them were dramatically decreased. Therefore, the functional groups in these positions will act as an anti-pharmacophore shielding (APS). Similarly, the activities of the compounds with e-, f- and g-positions had small values. The most consistent interpretation of all these results was that while a- and b-positions acted as an auxiliary group, and the other positions acted as an anti-pharmacophore shielding group.

The model prepared with the MCET method was used as a vehicle to find the correct quantitative 3D bio-structure. The

Table 3. The observed[7] and calculated activities of the compounds under study and the functional groups marked a, b, etc. for each conformer



Mol	Exp.	Calc.	R	The heat of formation (kJ/mol) and positions of conformers				
n1	5.13	5.13	5,7-OH,4'-NH_2	−2423870.07a	−2423866.32a b			
n2	4.88	6.13	3,5,7,3',4'-OH	−2866539.90a	−2866537.72ab	−2866537.53ab	−2866535.82a	
n03[§]	4.86	4.21	3,7,3',4'-OH	−2671050.76a	−2671049.19ab	−2671047.32ab	−2671046.84ac	
n4	4.83	6.87	5,7,4'-OH	−2475626.65ab	−2475626.13ab	−2475622.57a	−2475622.51ab	
n5	4.80	6.00	5,4'-OH	−2280159.29ab	−2280159.00a			
n6	4.80	4.86	6,3'-OH	−2280118.31a	−2280117.10a	−2280116.70ab	−2280116.46a	
n7	4.74	6.00	6-OH, 5,7,4'-NH_2	−2515786.66a	−2515786.66ab	−2515786.65a		
n8	4.71	5.12	5,7-OH	−2280165.55a	−2280161.87a			
n9	4.57	5.57	4'-OH,3',5'-OCH_3	−2679360.23a	−2679358.70ab			
n10[§]	4.46	4.28	5,7,3',4'-OH	−2671086.07ab	−2671083.90a	−2671083.81ac	−2671083.54a	−2671082.35a
n11	4.41	4.98	7,3'-OH	−2280125.87a	−2280125.40a	−2280125.05a	−2280124.30a	−2280124.02a
n12	4.34	5.17	6'-OH,5,7,3'-NH_2	−2515773.04a	−2515773.01a	−2515773.01a		
n13[§]	4.25	4.08	6-OMe,8,3'-NH_2	−2473892.43ac	−2473892.30abc			
n14	4.22	3.17	6-OH, 3',4',5'-OCH_3	−2976694.73abf	−2976694.61a			
n15[§]	4.16	3.43	3,5,7,4'-OH, 3',5'-OCH_3	−3265770.29a	−3265770.27a c			
n16	4.00	5.12	3,5,7,3',5'-OH	−2866530.08a				
n17	3.99	4.85	6,4'-NH_2	−2176595.11a				
n18	3.97	5.18	6,8,4'-NH_2	−2320284.67abc				
n19	3.93	5.22	6-OH,8,4'-NH_2	−2372057.42abc				
n20	3.93	4.85	6,4'-OH	−2176595.11a				
n21	3.92	4.10	7,8,4'-OH,3',5'-OCH_3	−3070246.40ae	−3070245.63a	−3070245.61abe	−3070243.84ab	−3070243.81ab
n22[§]	3.91	3.56	8,4'-NH_2	−2176594.93abc				
n23	3.89	4.76	6,4'-OH,3,5'- OCH_3	−2874814.74a	−2874812.89ab	−2874812.87ab		
n24	3.86	5.07	7-OH,4'-NH_2	−2228373.43a	−2228373.35a			
n25[§]	3.85	3.36	7-OH,6,4'-NH_2	−2372066.93a	−2372066.91a c			
n26[§]	3.78	3.38	7,4'-OH	−2280131.66a	−2280131.51a	−2280131.43ac	−2280131.26a	
n27	3.75	4.02	7,8,3'-OH	−2475578.24a	−2475577.38a	−2475577.12ac	−2475576.75ab	
n28	3.7	4.56	6,3'-NH_2	−2176580.53a	−2176578.88ab			
n29	3.68	4.92	4'-NH_2	−2032911.67a				
n30	3.65	3.92	5-OH,6,4'-NH_2	−2372082.23a	−2372082.22a			
n31[§]	3.53	3.45	3,5,7-OH	−2475616.77a	−2475614.27a			
n32	3.55	3.86	5,4'-OH,7-OCH_3	−2577519.76a	−2577519.25ab	−2577516.51a		
n33	3.5	3.00	5,3'-OH	−2280152.25a	−2280151.18a	−2280150.65a	−2280150.51a	
n34	3.5	4.18	7,8-OH	−2280124.18a				
n35[§]	3.49	4.02	5-OH,8,4'-NH_2	−2372072.72ac				
n36	3.48	4.47	7-OH,8,4'-NH_2	−2372065.04a	−2372052.77abc			
n37	3.47	3.07	7-OH	−2084671.24ac	−2084671.04a			
n38	3.43	3.22	6-OCH_3,8,4'-NH_2	−2473949.65abc				
n39	3.4	4.18	7,8-OH,3',4',5'-OCH_3		−3172154.91a	−3172154.60a c		
n40	3.36	3.25	3-$COOCH_3$,4'-OH	−2676375.59abd	−2676373.81a	−2676373.18a		
n41	3.3	3.92	4'-OH	−2084670.00a	−2084669.85ac			
n42	3.3	3.65	7-OH,6,3'-NH_2	−2372052.10a	−2372050.04ab	−2372049.99ab		
n43	3.12	3.14	7-OH,6,8,4'-NH_2	−2515742.75a	−2515742.52ace			
n44[§]	3.09	2.05	3-$COOCH_3$,4'-NH_2	−2624617.10abd				
n45	2.99	1.96	3-$COOCH_3$,7-OCH_3,4'-OH		−2973732.10abd	−2973730.09abd	−2973729.76a	−2973729.67abd
n46	2.9	2.93	7,4'-OH,3',5'-OCH_3	−2874822.33af	−2874821.19af	−2874820.94a	−2874819.81a	−2874819.80a
n47	2.81	1.45	7-OH,6,8,4'-NO_2	−3687566.90acg	−3687566.88abcg	−3687566.51abc	−3687566.43ac	
n48	2.8	2.44	3-COOH,4'-OH	−2676375.59abd	−2676373.81abd	−2676373.18a		
n49[§]	2.79	2.29	5-OCH_3,8,4'-NH_2	−2473917.4abce	−2473917.35abc			
n50	2.73	2.64	7-OH,8,4'-NO_2	−3153276.91abc	−3153276.68abc	−3153276.67abc	−3153276.57abc	

[§]Test set compounds

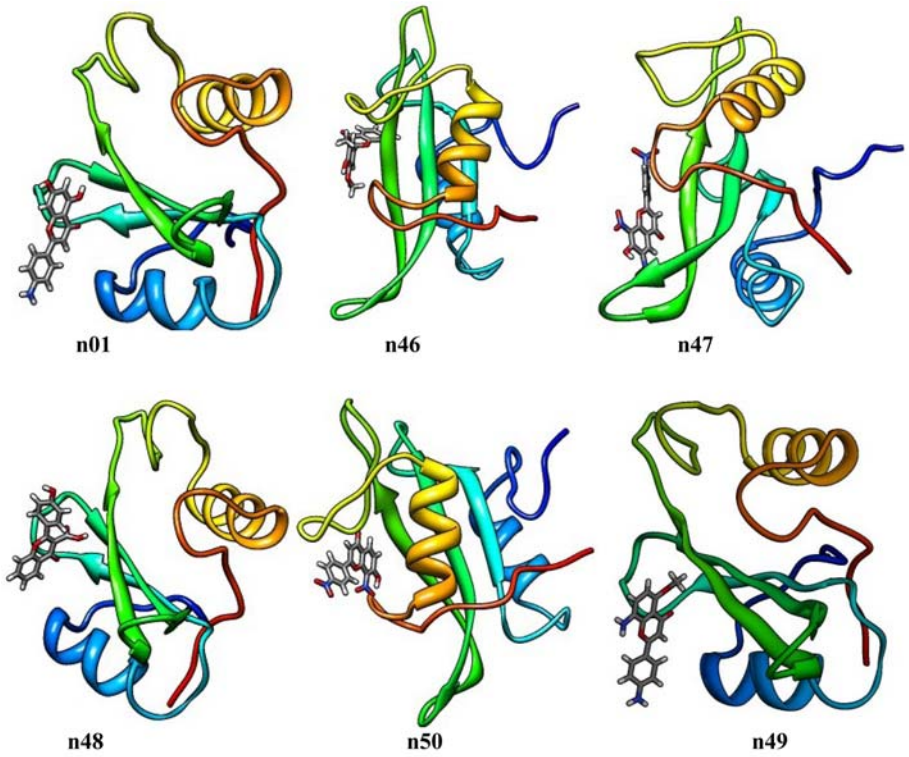


Figure 5. Graphic representation of the complexes of the p56Lck enzyme with the flavonoid derivatives obtained by Docking calculation.

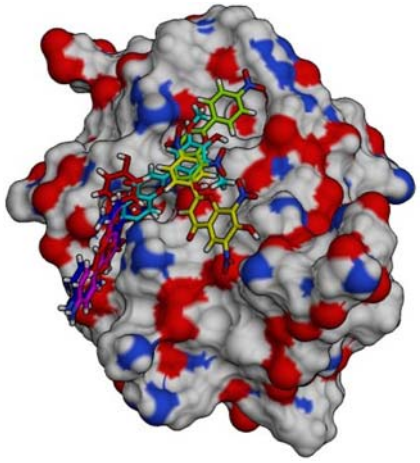


Figure 6. Superimposse graphic representation of the complexes of (1, 46, 47, 48, 49, 50) obtained by Docking calculation. Pink with 1, ciyan with 46, yellow with 47, red with 48, blue with 49, green with 50.

capability of making the model using the accepted conformers through 4D-QSAR analyses, in turn, permitted us to propose a bio-structure for the flavonoid binding affinities. By overview of inhibitors, it was proven that substitutions at the defined position of inhibitors were important for their activity. Then the a- and b-positions on the ligand were very important to produce potent biological action. It was shown that the defined topological properties of the compound affected the biological activity, which was concurrent with the model already deduced. The model generated from the functional groups together with the Pha showed 69% variance in biological activity for the test set.

Table 4. The individual contribution of Van der Waals[a] electrostatic terms[b] (in kcal mol^{-1}) to the free energy of binding (Dock score energy[c]) of the enzyme to the guests calculated by Docking

Ligand	VdW[a]	Ele[b]	Dock score[c]
1	−19.371	−6.219	−25.591
46	−28.892	−1.605	−30.498
47	−27.181	−3.458	−30.639
48	−25.201	−5.990	−31.192
49	−21.790	−5.126	−26.916
50	−26.338	−4.051	−30.389

The R2, Q2 and se for training, test set values and standart error of the optimal QSAR model were 0.72, 0.69 and 0.34, respectively.

Docking results suggested that several of these derivatives are active receptor with a significant preference for binding modes. During the docking process, the receptor was treated as fixed while the ligands are flexible. All torsion angles in each compound were allowed to rotate freely. Graphic representation of the complexes of the p56Lck enzyme with the flavonoid derivatives obtained by Docking calculation were given in Figure 5. Superimposse graphic representation of the complexes were given in Figure 6.

Conclusions

In this work, we applied the MCET method to examine the activities of 50 flavonoid inhibitors of $p56^{lck}$ and their structural descriptors. The method showed the potent and reliable

properties of 4D-QSAR model developed according to the values given in Table 3. Because the only topological properties of ligand as independent variables in model were used, the model performed for all compounds was significant, and the predicted activity were even more reliable and accurate, although small (Q2 = 0.69). Docking results suggested that several of these derivatives are active receptor with a significant preference for binding modes. It was important that the interaction points of the receptor were evaluated in 4D-QSAR analysis. In the Figure 4, the results of calculation demonstrated that the model established could be a useful and reliable tool in identifying flavonoid inhibitors.

Acknowledgments. This work was financially supported by Erciyes University Scientific Research Projects (BAP) of Turkey (Grant No; FBD-10-2983).

References

1. Cohen, P. *Nature Reviews Drug Discovery 1* 2002; p 309.
2. Manning, G.; Whyte, D. B.; Martinez, R.; Hunter T.; Sudarsanam, S. *Science 6* **2002**, *298*, 1912.
3. Dancey, J.; Sausville, E. A. *Nature Reviews in Drug Discovery and Development* **2003**, *2*, 296.
4. Adams, J. L.; Badger, A. M.; Kumar, S.; Lee, J. C. *Prog. Med. Chem.* **2001**, *38*, 1.
5. Weil, R.; Veillette, A. *Current Topics Micro. Immunol.* **1996**, *205*, 63.
6. Anderson, S. J.; Levin, S. D.; Perlmutter, R. M. *Adv. Immunol*. **1994**, *56*, 151.
7. Fasihsi, A.; Sabet, R. *Int. J. Mol. Sci.* **2008**, *9*, 1876.
8. Cushman, M.; Nagarathnam, D.; Burg, D. L.; Geahlen, R. L. *J. Med. Chem.* **1991**, *34*, 798.
9. Cushman, M.; Zhu, H.; Geahlen, R. L.; Kraker, A. J. *J. Med. Chem.* **1994**, *37*, 3353.
10. Novic, M.; Nikolovska-Coleska, Z.; Solmajer, T. *J. Chem. Inf. Comput. Sci.* **1997**, *37*, 990.
11. Prasanna1, S.; Doerksen, R. J. *Current Medicinal Chemistry* **2009**, *16*, 21.
12. Cheng, Z.; Zhang, Y.; Zhang, W. *Med. Chem. Res.* **2009**, *16*, 2479.
13. Kubinyi, H. *Ed.ESCOM Science Publishers*, Leiden, The Netherlands, 1993.
14. Gren, S. M.; Marshall, G. R. *Trends Pharmacol. Sci.* **1995**, *16*, 285.
15. Hopfinger, A. J.; Tokarski, J. S., Charifson, P. S. *Ed. Marcel Dekker*: New York, 1997, p 106.
16. Crame, R. D.; Patterson, D. E.; Bunce, J. D. *J. Am. Chem. Soc*. **1988**, *110*, 5959.
17. Hopfinger, A. J.; Wang S.; Tokarski, J. S.; Jin, B.; Albuquerque, M.; Madhav, P. J.; Duraiswami, C. *J. Am. Chem. Soc*. **1997**, *119*, 10509.
18. Hopfinger, A. J.; Burke, B. J.; Dunn, W. J. *J. Med. Chem.* **1994**, *37*, 3768.
19. SPARTAN '08 for Windows; Wavefunction, Inc.: Irvine, CA, USA, **2008**.
20. Guzel, Y.; Ozturk, E. *Bioorganic & Medicinal Chem.* **2003**, *11*, 4383.
21. Bersuker, I. B.; Dimoglo, A. S. *Reviews in Computational Chemistry*, Lipkowitz, K. B., Boyd, D. B., Eds; VCH: New-York, 1991.
22. Dimoglo, A. S.; Beda, A. A.; Shvets, N. M.; Gorbachov, M. Y.; Kheifits, L. A; Aulchenko, I. S. *New. J. Chem.* **1995**, *19*, 149.
23. Dimoglo, A. S.; Gorbachov, M. Y.; Bersuker, I. B.; Greni, A. I.; Vysotskaya, L. E.; Stepanova, O. V.; Lukash, E. Yu. *Die Nahrung* **1988**, *32*, 461.
24. Guzel, Y. *J. Molecular Structure* **1996**, *366*, 131.
25. Bersuker, I. B.; Bahçeci, S.; Boggs, J. E.; Pearlman, R. S. *SAR and QSAR in Env.* **1999**, *10*, 157.
26. Bersuker, I. B.; Bahceci, S.; Boggs, J. E.; Pearlman, R. S. *J. Comput.-Aided Mol. Des.* **1999**, *13*, 419.
27. Tjalling, J. Y. *SIAM Review* **1995**, *37*, 531.
28. Melanie, M. *MIT Press*, London. 1999.
29. Cho, D. H.; Lee, S. K.; Kim, B. T.; No, K. T. *Bull. Korean Chem. Soc.* **2001**, *22*, 388.
30. Höskuldsson, A. *J. Chemometrics* **1988**, *2*, 211.
31. Efron, B. *J. Am. Stat. Assoc*. **1983**, *78*, 316.
32. Lang, P. T.; Moustakas, D.; Brozell, S.; Carrascal, N.; Mukherjee, S.; Pegg, S.; Raha, K.; Shivakumar, D.; Rizzo, R.; Case, D.; Shoichet, B.; Kuntz, I. **2007**. DOCK 6.1, University of California, San Francisco. http://dock.compbio.ucsf.edu/
33. Case, D. A.; Cheatham III, T. E.; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M.; Onufriev, A. Jr.; Simmerling, C.; Wang, B.; Woods, R. The Amber biomolecular simulation programs. *J. Computat. Chem.* **2005**, *26*, 1668.
34. Lee, M. R.; Duan, Y.; Kollman, P. A. *Proteins: Struct. Funct., Genet.* **2000**, *39*, 309.
35. Srinavasan, J.; Cheatham, T. E.; Cieplak, P.; Kollman, P. A.; Case, D. A. *J. Am. Chem. Soc.* **1998**, *120*, 9401.