Erzincan Üniversitesi
Fen Bilimleri Enstitüsü Dergisi
2022 15(1), 168-182
ISSN: 1307-9085, e-ISSN: 2149-4584
Araştırma Makalesi

Erzincan University
Journal of Science and Technology
2022, 15(1), 168-182
DOI: 10.18185/erzifbed.973326
Research Article

4D-QSAR Analysis of Some New Pyrrole Derivatives Using Klopman Index Descriptor

Dilek Şeyma KIZILCAN[1], Burçin TÜRKMENOĞLU[2], Yahya GÜZEL[1*], Ertuğrul ASLAN[1]

[1]Erciyes University, Faculty of Science, Department of Chemistry, Kayseri, Turkey

[2]Erzincan Binali Yıldırım University, Faculty of Pharmacy, Department of Basic Pharmaceutical Sciences, Erzincan, Turkey

Geliş / Received: 19/07/2021, Kabul / Accepted: 15/10/2021

Abstract

We applied the Klopman Index, the Local Reactive Descriptor (LRD), for 3-Dimensional (3D) interactions between the Ligand-Receptor (L-R), for some new pyrrole derivatives as antituberculosis agents for 4-Dimensional Quantitative Structure-Activity Relationship (4D-QSAR) investigations. The parameters of the receptor side belonging to the Pharmacophore (Pha), which consists of the interaction points suggested between L-R, were calculated in the Molecular Conformer Electron Topological (MCET) method we developed based on the LRD values on the ligand side. After the 4D-QSAR model was established with the Leave One Out Cross-Validation (LOO-CV) technique on the molecules in the training set, the model was confirmed on the molecules in the external test set. Statistical results obtained from both sets, Q^2 = 0.875 and R_{ext}^2 = 0.918, respectively, were evaluated as satisfactory.

Keywords: 4D-QSAR, Klopman Index, LRD, MCET.

Klopman İndeksi Tanımlayıcısını Kullanarak Bazı Yeni Pirol Türevlerinin 4D-QSAR Analizi

Öz

Ligand-Reseptör (L-R) arasındaki 3 Boyutlu (3D) etkileşimler için yeni olarak kullanmaya başladığımız Yerel Reaktif Tanımlayıcı (LRD) olan Klopman İndeksi'ni antitüberküloz ajanlar olarak bazı yeni pirol türevlerinde 4D-QSAR incelemeleri için uyguladık. L-R arasında ileri sürülen etkileşim noktalarından oluşan farmakofora (Pha) ait reseptör tarafının parametreleri, ligand tarafındaki LRD değerlerine bağlı olarak geliştirdiğimiz MCET yönteminde hesaplandı. Eğitim setindeki moleküller üzerinde "Birini Dışarıda Bırak-Çapraz Doğrula" (LOO-CV) tekniği ile 4D-QSAR modeli kurulduktan sonra harici test setindeki moleküller üzerinde model doğrulandı. Her iki setten elde edilen istatistik sonuçları, sırasıyla, Q^2=0.875 ve R_{ext}^2=0.918 değerleri tatmin edici olarak değerlendirildi.

Anahtar Kelimeler: 4D-QSAR, Klopman İndeks, LRD, MCET

*Corresponding Author: yguzel@erciyes.edu.tr

1. Introduction

Tuberculosis (TB), a lung infection caused by Mycobacterium Tuberculosis (MTB), is an infectious disease that causes the death of millions of people, mostly in developing countries (Duncan and Barry, 2004). According to World Health Organization (WHO) reports, closed to 30,000 people get this disease every day and nearly 4500 people die due to TB (WHO, 2019). Currently used TB drugs were discovered mainly in the period of 1945-1965. Since then, numerous new drugs have been developed with the emergence of drug-resistant mycobacterium strains. Drug resistant species such as Isoniazid and Rifampicin have been the most dangerous forms of TB. This case has made it necessary to develop new therapeutic agents, as this causes greater difficulties for the treatment of TB. Among many organic compounds, pyrrole derivatives are heterocyclic compounds commonly discovered in plants and animals. In previous studies, it has been observed that pyrrole and pyrrole derivatives have antitumor Kamal et al. (2021), analgesic Ahmadi et al. (2011), antituberculosis (Biava et al., 2010; Sbardella et al., 2004) and anti-inflammatory Mohamed et al. (2011) activities. Because of these properties, it has been used quite widely in drug discovery (Ragno et al., 2000).

Examining the relationship between the molecular structures of chemical compounds and their biological activity and physicochemical properties according to their descriptor types, QSAR (Quantitative structure-activity relationship) is of great importance in drug discovery (Leszczynskı, 2010; Nantasenamat et al., 2009; Nantasenamat et al. 2010). The identifiers used emerge from some structural features such as polarizability, electronic properties, and steric parameters. However, 3D geometric properties of molecules that cause some difficulties in L-R interactions are not considered in classical QSAR approaches. Better and more general results can be obtained by applying the 3D-QSAR methodology to find out the 3D interaction map between L-R (Cramer et al., 1988; Klebe et al., 1994; Sippl, 2010). With primary descriptors, while the electronic properties such as the wave function of the molecules and the atomic charge / coefficient are calculated theoretically, secondary identifiers; physicochemical properties such as dipole moment, dispersion constant, equilibrium constant are determined experimentally. Various identifiers have been developed with new and different methods and real identifiers have been used in the calculation method to reveal the correct analysis. For a molecule, “There are no restrictions in the design of structural invariants; the limiting factor is one's own imagination”, this sentence emphasizing the descriptive diversity (Randic, 1996). Although more than 1000 descriptors have been found in studies conducted so far, how to find the "best" descriptor is an important problem (Consonni and Todeschini, 2010; Kubinyi, 1988). However, using the chemical reactivity of atoms in three dimensions, regardless of their physicochemical properties, may be the most realistic descriptor form. The Electron Topological Matrix (ETM) is an ideal descriptor for 3D interactions, both by evaluating atoms in their positions depending on their molecular structure and by using the most appropriate electronic properties of atoms (Bersuker et al., 1987; Dimoglo et al., 1989; Dimoglo et al., 1995; Saripinar et al., 1996). In the ETM, the position of the atoms is found as non-diagonal elements, while the electronic properties of the atoms are shown in the diagonal situation and these form the LRDs of those atoms (Bersuker

et al., 1991). In ETM used in MCET method developed by our group; New LRD types such as atomic charges (Natural, Mulliken and Electrostatic), Interaction Index, HOMO / LUMO frontier orbital coefficients, Fukui Indices and Klopman Index have been added (Yilmaz et al., 2011). The MCET method, such as CoMFA and CoMSIA, which is an internationally used ligand-based approach for 3D / 4D-QSAR investigations, also uses LRDs of atoms in 3D structure. The interaction pharmacophore elements classification scheme in CoMFA and CoMSIA defines five / seven different interaction types for each pharmacophore region of a 4D-QSAR model (Hong and Hopfinger, 2003). In MCET, a type of electronic property of atoms is selected, and the parameters of the receptor side belonging to the pharmacophore region are determined (Turkmenoglu et al., 2017; Yilmaz et al., 2011; Yilmaz et al., 2014). Among these, the "Klopman Index", which we have newly introduced to the literature, is a quite important descriptor (Guzel et al., 2018; Kizilcan et al., 2020; Türkmenoğlu and Güzel, 2021; Türkmenoğlu and Güzel, 2021; Tokat et al., 2019; Turkmenoglu and Guzel, 2018). The difference and superiority of this descriptor from others; it uses both atomic charge and frontier orbital (HOMO / LUMO) coefficients for each atom at the same time. Using the Levenberg – Marquardt approach in the MCET method, this descriptor was determined at each point of the receptor not as a single parameter, but as two parameters (simultaneous) for both atomic charge and atomic coefficients. As with CoMFA / CoMSIA, unlike mapping using probe types separately, mapping interaction with a single descriptor is easier and more realistic (Klopman, 1968; Klopman, 1984a and 1984b) Klopman index descriptor alone includes prop types in CoMFA / CoMSIA such as P +, P-, non-bonding, H-donor, H-acceptor.

The purpose of this article is to compare the MCET method, which is belong 3D-QSAR studies of biological activity of some new pyrrole derivatives as antituberculosis agents with other methods. And it is to show once again that the "Klopman Index" is better than the other descriptors with the results in the suggested Pha model.

2. Material and Methods

- For molecular modeling, three-dimensional structures of molecules were drawn with Spartan'10 package program and conformer distribution were performed with Molecular Mechanics Force Field (MMFF).
- Quantum chemical calculations were made in HF 6-31G * method for each conformer. Among the conformers, those with high energy were omitted. Because, according to the Boltzmann distribution, the molecule with high energy is the least likely to be found in number.
- To hold the geometric and quantum information of the remaining conformers, ETM.txt and cartCoord.txt files were created with the ETM program that we prepared (Bersuker et al., 1989; Bersuker et al., 1991).

There are geometrical and electronic properties of atoms for each conformer discussed in the ETM.txt file. Our homemade program MCET method is briefly explained below, to make 3D / 4D-QSAR calculations.

The lowest energy conformer of molecule, a simplest, most active and with the least amount of conformer is the most compatible with the receptor and the template is selected.

- ✓ All conformers are aligned on the core structure created by making a combination between the atoms in the template (containing at least one functional atom).
- ✓ Of these conformers, the most overlapping with the template is selected to represent its own molecule and superimpose occurs. In the next statements, when the term "Molecule" or "Conformer" is said, the conformer structure chosen should be understood.
- ✓ According to the geometrically and electronically derived tolerance values, clusters are formed from overlapping atoms and M ' pieces space vectors are determined.
- ✓ Interaction areas from M pieces sub-cluster formed after made acceptance / rejection with GA between clusters are handled stochastic.
- ✓ The equation is created based on the nonlinear interaction between L-R and the parameters on the receptor side are calculated with the Levenberg-Marquardt algorithm using LRDs at the interaction points of the ligand sides (Guzel et al., 2018). The model with the best statistical result is predicted by using the LRDs in the training set and the calculated parameters, with the LOO-CV method.
- ✓ The model is validated using the parameters calculated with the LRDs of the molecules in the external test set corresponding to the subset discussed (validation).

The 3D interaction points formed by the geometric structure of the subset between L-R are indicated by the pharmacophore. It is the most realistic approach to use as descriptor the electronic values of the atoms on the ligand side corresponding these points. Therefore, it is very important to use as arguments the independent variables of the LRD values on the ligand side and the parameters of the interaction points on the receptor side in the model. Even so, it is not clear which electronic properties of ligand atoms will be used. Different parameter values can be determined from each of them and therefore different statistical results are obtained depending on different models. Pha is recommended for a model according to the proposed correlation coefficient (R^2_{pred}, Q^2) belong to the training set and the best values of the correlation coefficient (R^2_{ex}) in the external test set.

a. New Local Reactivity Descriptor: Klopman Index

Since the best of among different LRDs used in our previous studies (Guzel et al., 2018; Kizilcan et al., 2020; Tokat et al., 2019; Turkmenoglu and Guzel, 2018; Turkmenoglu et al., 2020; Türkmenoğlu and Güzel, 2021) stands out as the Klopman index, was used as a descriptor type also in this study. The main feature of this descriptor is that it treats the interaction of hard acid and hard base together in an equation, not separately. According to the Klopman-Salem equation, the interaction between the corresponding atoms of the two studied compounds can be calculated by both electrostatic and covalent bonds. This interaction can be separately, or by using both together in this equation, energy gain (or loss) can be found (Klopman, 1968; Klopman, 1984a and 1984b; Salem, 1968). The equation is shown in detail below.

$$\Delta E = -\sum_{a\epsilon A, b\epsilon B}(q_a + q_b)\beta_{ab}S_{ab} + \sum_{a\epsilon A, b\epsilon B}\left(\frac{Q_a Q_b}{\varepsilon R_{ab}}\right) + \{\sum_{r\epsilon A}^{occ.}\sum_{s\epsilon B}^{unocc.} - \sum_{s\epsilon B}^{occ.}\sum_{r\epsilon A}^{unocc.}\}\frac{2\,(\sum_{a\epsilon A, b\epsilon B}^{n} c_{ra}c_{sb}\beta_{ab})^2}{E_r - E_s} \quad (1)$$

A and B : Interacting molecules

q_α : Electron population in the orbital of the atom on the a_m ligand side, m = 1, 2, 3… M_n

β and S : Resonance and Coincidence integrals

Q_a : Total load on a_m atom

ε : Local dielectric constant

R_{ab} : Distance between a_m and b_m atoms

c_{ra} : Coefficient of a_m atomic orbital in molecular orbital r

E_r : Energy of the molecular orbital r

Electrophiles and nucleophiles prefer to be bonded to each other covalently or electrostatically, depending on their hardness / softness similarity. Thus, they react faster (LoPachin and Gavin 2016). For simplicity, the first term in Equation 1 is arranged as Equation 2, neglecting the closed shell repulsions arising from the electron density of the filled orbitals of the molecules. Instead of the reaction of A and B molecules, the interaction between the Ligand and the Receptor (**L** and **R**) is discussed and Equation 2 is explained as pertaining to these interaction points.

$$\Delta E \approx \sum_{\mathrm{m}}^{M}\left(\frac{Q_{nuc,m}Q_{elect,m}}{\varepsilon R_m}\right) - 2\sum_{\mathrm{m}}^{M}\frac{\left(c_{nuc,m}c_{elect,m}\beta_m\right)^2}{E_{LUMO(elec)} - E_{HOMO(nuc)}} \quad (2)$$

Here;

M :Total points of interaction on the receptor side

M_n :The total number of interactions of the nth molecule.

M : Points of interaction on the ligand; m = 1, ..., M_n for the nth molecule

L and R **:** Interacting structures

$E_{LUMO(elec)}$, $E_{HOMO(nuc)}$:Frontier orbital energies of L and R.

$c_{nuc,m}$:Frontier orbital coefficient of atoms at point m

On the receptor side, there are M pieces points of interaction belong the pharmacophore. The pharmacophore structure consists of M clusters, which are different from each other, stochastically designed from M' clusters after the ligands are aligned and superimposed. If the

nth ligand (n = 1, 2, ... N) with the M_n interaction point from the series of N molecules examined contains atoms corresponding to all the interaction points of the receptor side, then M_n=M, otherwise M_n<M. According to this, the contribution of each ligand at the mth interaction point is equal to the electronic value of the atom it carries. These electronic values are considered in Equation 2 as much as both the atomic charge and the coefficients in the frontier orbital. In the first term, when the atomic charges in the electrostatic interaction are positive / negative and the atomic coefficients in the frontier orbitals shown in the second term are the same / different phases, pushing / pulling forces occur. According to this, The Klopman Index is a descriptor in which the ionic bond in the first term and the non-bonding and hydrogen bonding forces in the second term are used simultaneously.

For the interaction of point m between L and R; The first term refers to electrostatic energy. This energy results from the interaction of partial charges ($Q_{nuc,m}$ and $Q_{elect,m}$) with respect to the dielectric constant (ε) and the distance between the atoms (R). The second term represents the HOMO and LUMO energy differences as well as the coefficients (c_m) on them and a covalent energy arising from the resonance integral (β_m). The interaction energy in the second term depends significantly on the frontier orbital energy level differences. When this difference is large, the contribution in the second term will decrease, while the contribution of the first term will be larger, and vice versa when the difference decreases. From the first term hard acid-base effects and from the second term soft acid-base effects arise. Accordingly, the ionic species, which are P^+ and P^-, among the interactions in the first term, are important, while in the second term, the non-polar and Hydrogen bonding causative species are more dominant. In the Klopman Index, both ionic characters and covalent characters take place together at each point of interaction. The above equation has been simplified to use HOMO of a nucleophile and LUMO of an electrophile.

In Equation 2, the smallest E_{LUMO}-E_{HOMO} energy differences should be determined to decide whether the nucleophile (electron donor side HOMO) and the electrophile (electron acceptor side LUMO) are ligand-receptor or receptor-ligand, respectively. In other words, the side with higher E_{HOMO} acts as the nucleophile and the side with lower E_{LUMO} acts as the electrophile. However, other parameters are as important as the E_{LUMO}-E_{HOMO} energies in the second term. Due to the complexity of the entire structure of the receptor, the frontier orbital values cannot be known. Therefore, the average E_{LUMO} and E_{HOMO} of the ligands are used for the receptor. However, both nucleophile-electrophile and electrophile-nucleophile, respectively, are taken into consideration for the ligand and the receptor and is decided according to the best result for all molecules. Thus, nucleophile and electrophile sides are revealed by using the parameters of the first term with c_m and β_m values, as well as E_{LUMO} and E_{HOMO}.

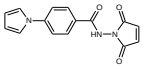
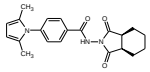
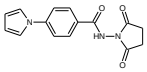
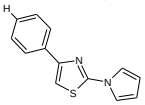
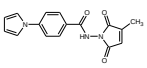
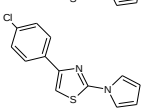
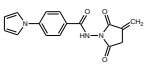
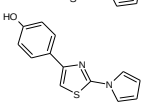
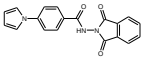
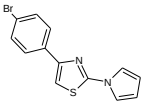
3. Results and Discussion

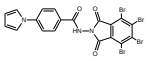
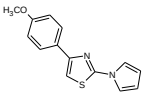
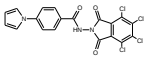
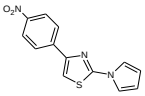
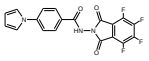
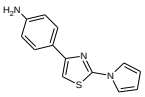
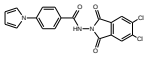
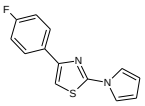
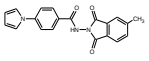
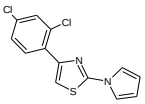
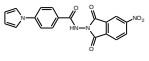
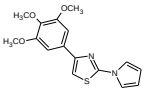
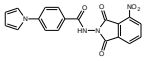
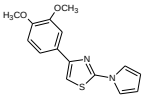
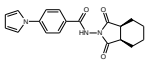
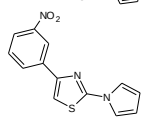
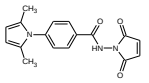
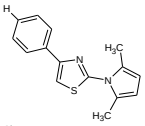
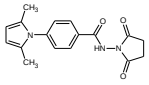
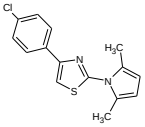
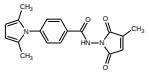
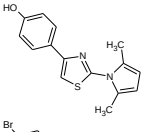
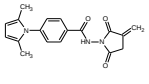
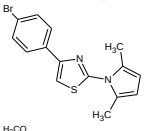
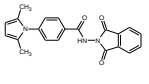
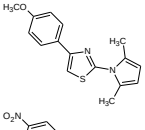
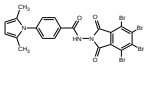
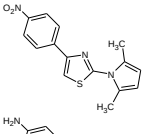
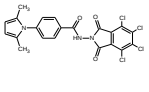
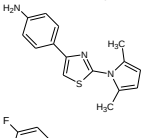
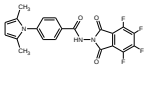
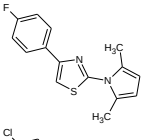
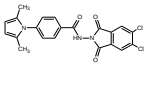
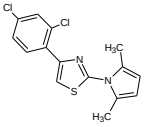
Pyrrole derivatives as anti-tuberculosis agents were taken from the literature, their activities were re-calculated by MCET method and the best model showing L-R interaction was derived using MCET method in line with the best statistical results (Joshi et al., 2014). The activity values of this model were calculated using the observed activity values for both training and

test sets. Table 1 shows the structures of the molecules and their theoretical and experimental activity values.

Using the LOO-CV technique, the activities of the compounds in the training set are verified, and the activities of the excluded test set are predicted by the generated model. The best model is found with the Q^2, R^2 and R_m^2 values calculated using the LRD of the group selected by GA among the clusters. These values are obtained based on the activities calculated for estimation of the receptor parameters. The parameters at these points specified are constants that have been proven to be accurate using the LRD values of the molecules in the training set. Using these constants, the activities of the molecules in the test set are calculated and validated. The parameter values on the receptor side are constant throughout the process, but there are different LRD values for each ligand corresponding to the active sites of the receptor. The force that will occur between two different ligands interacting at any point of the receptor depends on the atomic charges of the ligands. When the atomic charge of one is negative and that of the other is positive, electrostatically attractive force is formed, while if both have the same charge, the repulsive force is formed. In other words, for the κ parameter of the negative value at the point of electrostatic interaction, if the corresponding LRD value of the ligand is positive, electrostatic effect occurs, while if the LRD value is negative, a steric effect occurs. Considering this case for the frontier orbital interaction, if there is a positive ξ coefficient, covalent bonding occurs when the orbital coefficient of the ligand is positive at this point, while the negative orbital coefficient causes anti bond formation. The push or pull force at each point where the receptor and ligand interact; the partial charge of the ligand and the frontier orbital coefficients are determined by the κ and ξ parameter values and LRD values of the receptor. Because both electrostatic and covalent interactions are used simultaneously in the Klopman Index.

Table 1. The skeletons of the investigated molecules, the observed and predicted IC_{50} activities.

Molecule	Molecular Structure	Obs.	Pred.	Molecule	Molecular Structure	Obs.	Pred.
1		5.200	5.201	**26**		5.200	5.233
2		5.200	5.185	**27**		6.400	6.442
3		6.100	5.855	**28**		6.400	6.347
4		5.510	5.800	**29**		4.300	4.756
5		6.400	6.467	**30**		6.100	6.282

6		6.100	6.258	**31**		4.900	5.100
7		6.400	6.080	**32**		6.400	6.546
8		6.700	6.686	**33**		4.600	4.636
9		6.100	6.273	**34**		6.700	6.088
10		5.510	5.244	**35**		6.400	6.186
11		5.330	5.841	**36**		4.300	4.593
12		5.510	5.464	**37**		4.300	4.427
13		5.510	5.238	**38**		6.100	6.277
14		4.900	5.115	**39**		6.400	6.350
15		5.100	5.232	**40**		6.100	6.208
16		5.510	5.759	**41**		4.90	4.965
17		5.510	5.814	**42**		6.40	6.192
18		6.400	5.895	**43**		5.20	5.157
19		6.400	5.814	**44**		6.40	6.370
20		6.400	6.029	**45**		4.90	4.625
21		6.700	6.614	**46**		6.70	7.003
22		6.10	6.239	**47**		6.40	6.383

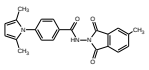
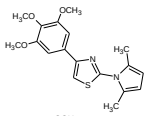
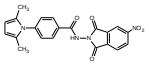
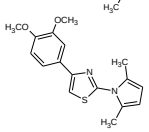
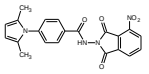
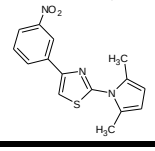
23		5.20	5.244	**48**		4.90	4.461
24		5.20	5.463	**49**		4.30	4.347
25		5.36	5.450	**50**		6.40	6.243

Table 2. The values of Q^2 and R^2 according to different descriptors

Descriptors	Q^2	R^2
E_LKlopman	0.554	0.430
M_LKlopman	0.709	0.910
N_Charge	0.741	0.869
N_LKlopman	0.705	0.743
N_Fukui (f^+)	0.718	0.821
P_Fukui (f^-)	0.711	0.661
M_HKlopman	**0.875**	**0.918**
E_HKlopman	0.730	0.603
M_Charge	0.541	0.805
HOMO	0.707	0.814
LUMO	0.741	0.699
N_HKlopman	0.779	0.960

M_Charge, is using Mulliken charge. N_Fukui (f+), is using the Fukui index against nucleophiles. E_HKlopman, is using Klopman index consist of electrostatic partial charge and HOMO coefficient. E_LKlopman is using Klopman index consist of electrostatic partial charge and LUMO coefficient. (Turkmenoglu and Guzel, 2018).

As LRD type, natural charge, Mulliken charge, electrostatic charge and Fukui index were compared with Klopman's index. For each LRD type, the combination with the best Q^2, R^2 results was chosen. The best statistical results for different LRDs are shown in Table 2. The combination that produces the best results among these creates a 3D-PhaM. In this study, the Klopman index is the best LRD type and the 3D-PhaM configuration is shown in Table 3.

The best model is proposed by using the Klopman Index descriptor, one of the LRD types known as independent variables. The names, positions, Cartesian coordinates and κ / ξ parameter values of the molecule and atom in the model proposed using this descriptor are shown in Table 3.

Table 3. Cartesian coordinate values of the receptor side and calculated κ and ξ parameters

Molecule Number	Atom Number	x, y, z Cartessian coord.			Position letters	κ Value	ξ Value
		x	y	z			
n01	**N3**	**0**	**0**	**0**	**a**	**-0.825**	**-11.607**
n01	**C15**	**2.295**	**0**	**0**	**b**	**0.005**	**7.714**
n01	**C17**	**1.088**	**0.871**	**0**	**c**	**0.849**	**-10.636**
n29	C2	2.339	2.761	-0.011	d	-0.059	0.119

n29	**C3**	**1.160**	**2.047**	**-0.008**	**e**	**0.268**	**-8.583**
n36	S1	6.726	-1.725	0.310	f	-0.113	-9.389
n36	C7	4.864	-0.042	-0.019	g	0.164	5.399
n37	C9	-1.020	0.770	0.008	h	0.025	5.126
n37	C12	-3.473	1.063	0.067	i	-0.133	-8.056
n21	C16	-4.819	3.350	0.766	j	0.496	6.263
n21	C19	-3.810	2.699	1.463	k	-0.855	-4.753
n21	C15	-2.888	1.988	0.746	l	0.098	-2.746
n21	C12	-3.951	2.554	-1.330	m	0.204	-2.873

The positions of the atoms at the points of interaction of the ligands given in Table 3 are marked with the letters a, b, and c. The occurrence of repulsive or pulling forces at each point of the parameters on the receptor side corresponding to points in this position depends on whether they are positive or negative. These formed forces act in the direction of increasing or decreasing activity. The meeting of the positive and negative sides, which is the reason for the pulling force, increases the activity, while the encounter of the same charged parties decreases the activity. If there is not any atom of the ligand at the point of interaction of the receptor, it will not contribute.

The magnitude of the repulsive or pulling forces occurring at these points is established by looking at the values of both the ligands and the receptors. Activity estimates are made using the Klopman equation. Both interactions are discussed here, as the Klopman equation also includes both electrostatic and frontier orbital interactions. In this study, for the frontier orbital interactions, the LUMO of the receptor was used as the electrophile and the HOMO of the ligand as the nucleophile.

The pharmacophore depends on the type of descriptor derived from the electronic properties of the atoms corresponding to the points of interaction on the active side of the receptor (Tokat et al., 2019). Thanks to Pha, the activity belonging to any ligand can be determined and analysis can be made about the 3D orientation of the atoms at the interaction points. To do this, it is checked whether a molecule under consideration has a nucleus structure that forms the basis of the Pha structure. After the nucleus structure of the template overlaps with the molecule, among the remaining atoms of the molecule, the ones that are similar in 3D Cartesian coordinate values of Pha are determined. By using the atomic charges and coefficients used in the Klopman index for each atom in the positions of the Pha structure, the interaction energy between L-R and its activity accordingly are calculated.

In this study, the pharmacophore model was formed with atoms positioned according to the 3D Cartesian coordinates of conformers using the Klopman index.

The 3D pharmacophore structure showing the interaction points marked as positions a, b, c is given in Figure 1. Some ligand atoms interacting with the receptor at these points indicated may be within the tolerance range deemed appropriate.

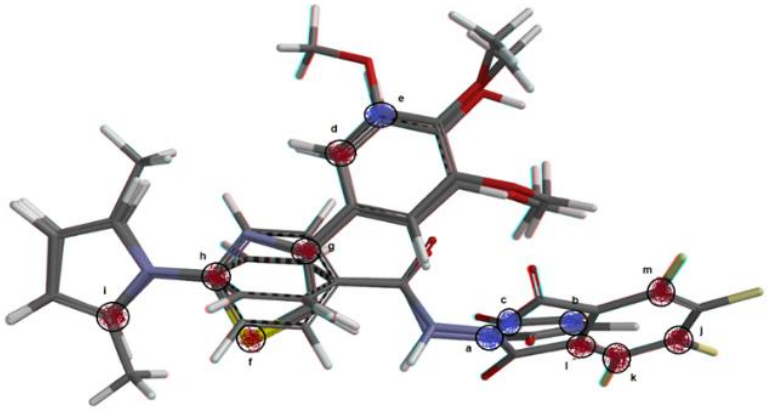


Figure 1. Point of interaction of the aligned and superimposed molecules with the receptor is indicated by regions marked a, b, c, etc.

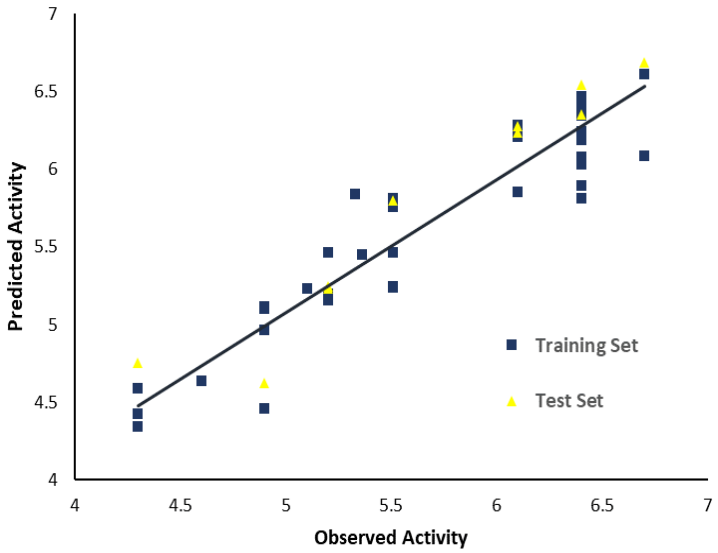


Figure 2. The observed and estimated activity values

Looking at the pharmacophore model proposed here, it is seen that the Klopman Index, which is a universal descriptor, has a strong performance. The accuracy of a QSAR model is usually understood by checking it with cross calculations with the ability to externally predict using test sets. Since the structure and activities of the molecules in the training and test sets are like each other, the activities of the molecules in the test set were calculated using the model set taken from the training set. The observed and predicted activities of molecules in both training and test sets are shown by the correlation curve in Figure 2.

Based on the observed and predicted activities of the compounds in the training and test sets, the pharmacophore model produced in the MCET method can be compared with the model previously specified by CoMFA for the same molecule series. For the training set (compound

40) and test set (compound 10), internal validation (R^2 excluding one, Q^2 =0.875) and external validation (R^2_{pred} =0.918) were calculated. These results are promising as they are better than the Q^2 = 0.815 and R^2 = 0.973 values found by the CoMFA method. In this article, the reliability of the results has been accepted using known validation parameters.

In the pharmacophore model, the activities can be calculated again using both Kappa, Xi parameter values and the LRD values of the ligands corresponding to these points (Tokat et al., 2019).

In MCET method, the data obtained after estimating the parameter values on the receptor side can be stored in an Excel file to calculate and display the activity of any molecule analytically. Because of this feature, the MCET method is of great importance.

4. Conclusion

For the 4D-QSAR analysis of the pyrrole derivatives series, the Klopman Index was used in our home-made program MCET method. The fact that the results obtained for the training and test sets were better than those calculated with the international software COMFA have given confidence both for the MCET method and the Klopman Index descriptor used as LRD. It was based on the energy values resulting from a 3D physicochemical interaction, depending on the Klopman index, for which the pharmacophore model obtained to estimate the IC_{50} values of the compounds is quite realistic. Thus, the model obtained using electrostatic and covalent interactions offers quite clear and sufficient information for the structure-activity relationship. Accordingly, a series of new series of molecules that affect activity and have specified structural properties can be designed.

Ethics in Publishing

There are no ethical issues regarding the publication of this study.

References

Ahmadi, A., Solati, J., Hajikhani, R., Pakzad, S. 2011. "Synthesis and Analgesic Effects of New pyrrole derivatives of phencyclidine in mice", *Arzneimittelforsch*, 61, 296-300.

Alp Tokat, T., Turkmenoglu, B., Guzel, Y., Kizilcan, D. S. 2019. "Investigation of 3D Pharmacophore Of N-benzyl Benzamide Molecules of Melanogenesis Inhibitors Using a New Descriptor Klopman Index: Uncertainties in Model", *J Mol Model*, 25(8), 247.

Bersuker, I. B., Dimoglo, A. S., Gorbachov, M. Y. 1987. "The Electron-Topological Approach in Investigation of Structure - Activity Relation - Inhibition of Thymidine Phosphorylase by Uracil Derivatives", *Bioorg Khim+*, 13, 38-44.

Bersuker, I. B., Dimoglo, A. S., Gorbachov, M. Y., Greni, A. I., Vysotskaya, L. E., Mikhailova, T. V. 1989 "Study Of The Electronic And Structural-Properties Of The Chemical-Compounds İn Garlic Aroma", *Nahrung*, 33, 405-411.

Bersuker, I. B., Dimoglo, A. S., Gorbachov, M. Y., Vlad, P. F. 1991. "Origin of Musk Fragrance Activity - the Electron-Topologic Approach", *New J Chem*, 15, 307-320.

Biava, M., Porretta, G. C., Poce, G., Battilocchio, C., Alfonso, S., De Logu, A., Serra, N., Manetti, F., Botta, M. 2010. "Identification Of A Novel Pyrrole Derivative Endowed With Antimycobacterial Activity And Protection Index Comparable To That Of The Current Antitubercular Drugs Streptomycin And Rifampin", *Bioorg Med Chem*, 18, 8076-8084.

Consonni, V., Todeschini, R. 2010. "Molecular Descriptors", *Chall Adv Comput Che*, 8, 29-102.

Cramer, R. D., Patterson, D. E., Bunce, J. D. 1988. "Comparative Molecular-Field Analysis (Comfa) .1. Effect of Shape on Binding of Steroids to Carrier Proteins", *J Am Chem Soc*, 110, 5959-5967.

Dimoglo, A. S., Beda, A. A., Shvets, N. M., Gorbachov, M. Y., Kheifits, L. A., Aulchenko, I. S. 1995. "Investigation of the Relationship between Sandalwood Odor and Chemical-Structure - Electron-Topological Approach", *New J Chem*, 19, 149-154.

Dimoglo, A. S., Bersuker, I. B., Popa, D. P., Kuchkova, K. I. 1989. "The Electron-Topological Study of the Growth-Regulating Activity in a Series of Abscisic-Acid Analogs", *Teor Eksp Khim*, 25, 590-594.

Duncan, K., Barry, C. E. 2004. "Prospects For New Antitubercular Drugs", *Curr Opin Microbiol*, 7, 460-465.

Guzel, Y., Aslan, E., Turkmenoglu, B., Su, E. M. 2018. "4D-QSAR Studies Using a New Descriptor of the Klopman Index: Antibacterial Activities of Sulfone Derivatives Containing 1, 3, 4-Oxadiazole Moiety Based on MCET Model", *Curr Comput-Aid Drug*, 14, 207-220.

Hong, X., Hopfinger, A. J. 2003. "3D-Pharmacophores of Flavonoid Binding at the Benzodiazepine GABA (A) Receptor Site Using 4D-QSAR Analysis", *J Chem Inf Comp Sci*, 43, 324-336.

https://www.who.int/campaigns/world-tb-day/world-tb-day-2019, in.

Joshi, S. D., More, U. A., Dixit, S. R., Korat, H. H., Aminabhavi, T. M., Badiger, A. M. 2014. "Synthesis, Characterization, Biological Activity, and 3D-QSAR Studies on Some Novel Class of Pyrrole Derivatives as Antitubercular Agents", *Med Chem Res*, 23, 1123-1147.

Kamal, A., Ramakrishna, G., Nayak, V. L., Raju, P., Rao, A. V. S., Viswanath, A., Vishnuvardhan, M. V. P. S., Ramakrishna, S., Srinivas, G. 2012. "Design and Synthesis of benzo[c,d]indolone-pyrrolobenzodiazepine Conjugates as Potential Anticancer Agents", *Bioorgan Med Chem*, 20, 789-800.

Kizilcan, D. S., Turkmenoglu, B., Guzel, Y. 2020. "The Use of the Klopman İndex as A New Descriptor for Pharmacophore Analysis on Strong Aromatase İnhibitor Flavonoids against Estrogen-Dependent Breast Cancer", *Struct Chem*, 31, 1339-1351.

Klebe, G., Abraham, U., Mietzner, T. 1994. "Molecular Similarity Indexes in a Comparative-Analysis (Comsia) of Drug Molecules to Correlate and Predict Their Biological-Activity", *J Med Chem*, 37, 4130-4146.

Klopman, G. (1968). Chemical reactivity and the concept of charge- and frontier-controlled reactions. *Journal of the American Chemical Society*, 90(2), 223-234.

Klopman, G. (1984a). Citation Classic - Chemical-Reactivity and the Concept of Charge-Controlled and Frontier-Controlled Reactions. *Current Contents/Engineering Technology & Applied Sciences*(20), 16-16. Retrieved from <Go to ISI>://WOS:A1984SP75100001

Klopman, G. (1984b). Citation Classic - Chemical-Reactivity and the Concept of Charge-Controlled and Frontier-Controlled Reactions. *Current Contents/Physical Chemical & Earth Sciences*(20), 16-16. Retrieved from <Go to ISI>://WOS:A1984SP74700001

Kubinyi, H. 1988. "Free Wilson Analysis - Theory, Applications and Its Relationship to Hansch Analysis", *Quant Struct-Act Rel*, 7, 121-133.

Leszczynskı, J. (2010), "Recent Advances in QSAR Studies: Methods and Applications, Chall Adv Comput Che, 8", *Springer Science+Business Media,* Springer Dordrecht Heidelberg London New York, 1-423.

LoPachin, R. M., Gavin, T. 2016. "Reactions Of Electrophiles With Nucleophilic Thiolate Sites: Relevance to Pathophysiological Mechanisms and Remediation", *Free Radical Res*, 50, 195-205.

Mohamed, M. S., Kamel, R., Fathallah, S. S. 2011. "Synthesis of New Pyrroles of Potential Anti-Inflammatory Activity", *Arch Pharm*, 344, 830-839.

Nantasenamat, C., Isarankura-Na-Ayudhya, C., Naenna, T., Prachayasittikul, V. 2009. "A Practical Overview of Quantitative Structure-Activity Relationship", *Excli J*, 8, 74-88.

Nantasenamat, C., Isarankura-Na-Ayudhya, C., Prachayasittikul, V. 2010. "Advances in computational methods to predict the biological activity of compounds", *Expert Opin Drug Dis*, 5 (2010) 633-654.

Ragno, R., Marshall, G. R., Di Santo, R., Costi, R., Massa, S., Rompei, R., Artico, M. 2000. "Antimycobacterial Pyrroles: Synthesis, Anti-Mycobacterium Tuberculosis Activity and QSAR Studies", *Bioorgan Med Chem*, 8, 1423-1432.

Randic, M. 1996. "Molecular bonding profiles", *J Math Chem*, 19, 375-392.

Salem, L. 1968. "Intermolecular Orbital Theory of the İnteraction between Conjugated Systems. I. General Theory", *J Am Chem Soc*, 90, 543-552.

Saripinar, E., Guzel, Y., Patat, S., Yildirim, I., Akcamur, Y., Dimoglo, A. S. 1996. "Electron-Topological Investigation of the Structure-Antitubercular Activity Relationship of Thiosemicarbazone Derivatives", *Arzneimittel-Forschung/Drug Research,* 46, 824-828.

Sbardella, G., Mai, A., Artico, M., Loddo, R., Setzu, M. G., La Colla, P. 2004. "Synthesis and in Vitro Antimycobacterial Activity of Novel 3-(1H-pyrrol-1-yl)-2-oxazolidinone Analogues of PNU-100480", *Bioorg Med Chem Lett*, 14, 1537-1541.

Sippl, W, 2010. "3D-QSAR-Applications, Recent Advances, and Limitations", Chall Adv Comput Che, 8, (ed., Puzyn, T., Leszczynski, J., Cronin, M. T. D.), *Springer Science+Business Media*, Springer Dordrecht Heidelberg London New York, 103-125.

Türkmenoğlu, B., and Güzel, Y. 2021. "Confirmation by molecular docking of the pharmacophore defined as 4D-QSAR using the MCET method for the 2-hydroxydiarylamide derivatives", *International Journal of Chemistry and Technology*, 5(1), 11-25.

Türkmenoğlu, B., and Güzel, Y. 2021. "Klopman Index: A New Descriptor Used In MCET For 3D/4D-QSAR Modeling", Current Academic Studies in Science and Mathematics Sciences, *Livre de Lyon*, Lyon, 1-16.

Turkmenoglu, B., Guzel, Y. 2018. "Molecular Docking And 4D-QSAR Studies of Metastatic Cancer Inhibitor Thiazoles", *Comput Biol Chem*, 76, 327-337.

Turkmenoglu, B., Guzel, Y., Su, E. M., Kizilcan, D. S. 2020. "Investigation Of İnhibitory Activity Of Monoamine Oxidase A With 4D-QSAR Using Fukui İndices İdentifier", *Mater Today Commun*, 25, 1-6.

Türkmenoglu, B., Yilmaz, H., Su, E. M., Alp Tokat, T., Güzel, Y. 2017. "4D-QSAR Study of Flavonoid Derivatives With MCET Method, *International Journal of Chemistry and Technology*, 1, 14-23.

Yilmaz, H., Boz, M., Turkmenoglu, B., Guzel, Y. 2014. "Pharmacophore and Functional Group Identification of 4, 4'-dihydroxydiphenylmethane as Bisphenol-A (BSA) Derivative", *Trop J Pharm Res*, 13, 117-126.

Yilmaz, H., Guzel, Y., Onal, Z., Altiparmak, G., Kocakaya, S. O. 2011. "4D-QSAR Study of p56 (lck) Protein Tyrosine Kinase Inhibitory Activity of Flavonoid Derivatives Using MCET Method", *B Korean Chem Soc*, 32, 4352-4360.