



HAL
open science

QSAR Modeling and Drug-Likeness Screening for Antioxidant Activity of Benzofuran Derivatives

S. Boudergua, M. Alloui, S. Belaidi, M. Mogren Al Mogren, U. A. Abd Ellatif Ibrahim, M. Hochlaf

► **To cite this version:**

S. Boudergua, M. Alloui, S. Belaidi, M. Mogren Al Mogren, U. A. Abd Ellatif Ibrahim, et al.. QSAR Modeling and Drug-Likeness Screening for Antioxidant Activity of Benzofuran Derivatives. *Journal of Molecular Structure*, 2019, 1189, pp.307 - 314. 10.1016/j.molstruc.2019.04.004 . hal-03484965

HAL Id: hal-03484965

<https://hal.science/hal-03484965>

Submitted on 20 Dec 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

QSAR Modeling and Drug-Likeness Screening for Antioxidant Activity of Benzofuran Derivatives

S. Boudergua,^a M. Alloui,^{a,b} S. Belaidi,^{a,*} M. Mogren Al Mogren,^{c,*} U. A. Abd Ellatif Ibrahim,^c M. Hochlaf^{b,*}

^a University of Biskra, Faculty of Sciences, Department of Chemistry, Group of Computational and Pharmaceutical Chemistry, LMCE Laboratory, 07000 Biskra, Algeria.

^b Université Paris-Est, Laboratoire Modélisation et Simulation Multi Echelle, MSME UMR 8208 CNRS, 5 bd Descartes, 77454 Marne-la-Vallée, France.

^c Chemistry Department, Faculty of Science, King Saud University, PO Box 2455, Riyadh 11451, Kingdom of Saudi Arabia.

* Corresponding authors:

E-mail addresses: prof.belaidi@gmail.com (S. Belaidi), mmogren@KSU.EDU.SA (M. Mogren Al-Mogren), hochlaf@univ-mlv.fr (M. Hochlaf).

Abstract

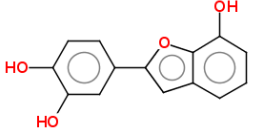
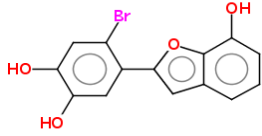
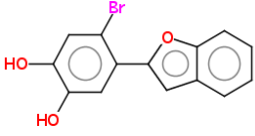
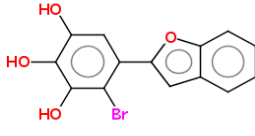
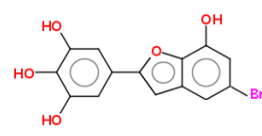
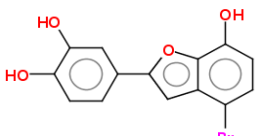
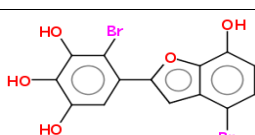
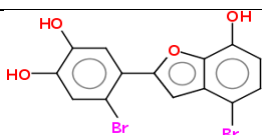
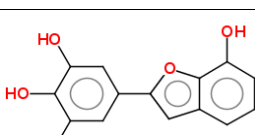
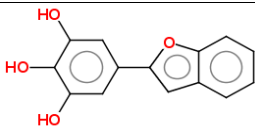
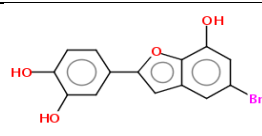
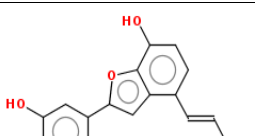
In order to explore the relationship between the antioxidant activity and structure of fifteen benzofuran derivatives, we carried out a QSAR study using multiple linear regression (MLR) and artificial neural network (ANN) methods. Six descriptors were used as input data (molar weight, surface area, octanol-water partition coefficient, hydration energy, highest occupied molecular orbital energy and lowest unoccupied molecular orbital energy). The electronic properties were derived at the B3LYP/6-31 G(d,p) level. Benchmarks on furan and benzofuran subunits and their comparison to the experiment showed that this level of theory is good enough. The output data correspond to the antioxidant activity as given by IC₅₀. The predicted properties are in agreement with experimental values. Our study shows that 80 % of studied molecules are in accordance with the Lipinski and Veber rules and reach the optimal lipophilicity indices. In addition, statistical analysis reveals that ANN technique with (6-2-1) architecture is more significant than MLR model.

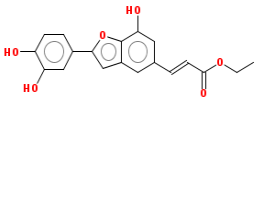
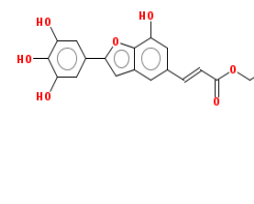
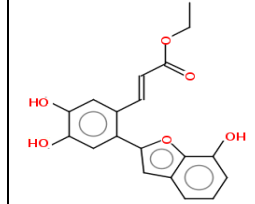
Keywords: benzofuran derivatives, antioxidant, QSAR, MLR, ANN, drug-likeness.

1. Introduction

Damage to cells caused by free radicals is believed to play a central role in disease progression [1] including cancer [2], cardiovascular [3] and Alzheimer diseases [4]. The need for antioxidants becomes even more critical with increased exposure to free radicals caused by pollution, cigarette smoke, drugs, illness and stress [1]. Scientists commonly agree that a combination of antioxidants, rather than unique entities, may be more efficient in the end. Antioxidants can be very helpful in improving the quality of life, for instance, by preventing or delaying the onset of degenerative diseases [5]. They are able to stabilize or disable free radicals before they attack the cells. In addition, they are essential for maintaining optimal cellular and system health and well-being [1]. In this context, oxygen heterocycles exhibit diverse biological and pharmacological activities due in part to the similarities with many natural and synthetic molecules with known biological activity. Among these compounds, furan [6] and benzofuran derivatives show efficient antioxidant activity. These derivatives exhibit also high potentialities for use as pharmacological agents. Indeed, these ring systems have emerged as powerful scaffolds for many biological evaluations and play an important role in the design and discovery of new physiological/pharmacologically active molecules. They can arrange to yield potent and selective drugs. We refer to the recent review by Khanam and Shamsuzzaman [7] for a wide presentation of the progress in the numerous pharmacological activities of benzofuran derivatives and of their applications in medicine.

Table 1: Chemical structures and experimental activity of the molecules under study [6].

N	Structure	IC ₅₀	N	Structure	IC ₅₀	N	Structure	IC ₅₀
1		18.5	2		20.6	3		33.8
4		18.0	5		14.6	6		18.4
7		19.4	8		7.8	9		12.8
10		10.8	11		28.6	12		16.8

13		25.4	14		11.7	15		26.2
----	---	------	----	---	------	----	---	------

The present theoretical contribution concerns the study of the relationship between antioxidant activity and the structure of fifteen benzofuran derivatives as recently determined by Hsieh et al. [6]. The structures of these compounds are specified in Table 1 as constructed using the Molinspiration Database [8]. Their corresponding antioxidant activity using 1, 1-diphenyl-2-picrylhydrazyl (DPPH) is compiled from the literature (Table 1). All these molecules show a free radical scavenging activity with IC_{50} values ranged between 7.8 ± 1.6 and 33.8 ± 2.8 μ M [6].

Many different chemometric methods, such as multiple linear regression (MLR) [9-11], partial least squares regression (PLS) [12], different types of artificial neural networks (ANN) [13-16], genetic algorithms (GA) [17], and support vector machine (SVM) can be employed to deduce correlation models between the molecular structure and properties [9]. At present, we derive a quantitative structure–activity relationship (QSAR) model using multiple linear regression (MLR) as well as artificial neural network (ANN) methods for the series of benzofuran derivatives of Table 1. Afterwards, Lipinski and Veber rules, and lipophilicity indices are applied to identify “drug-like” compounds. For electronic structure computations, we used Becke’s three-parameter Lee-Yang-Parr (B3LYP) [18] density functional theory (DFT) in conjunction with the 6-31G (d,p) basis set. This level of theory is viewed to be accurate enough after benchmarks on the structure and the vibrational spectroscopy of furan and benzofuran subunits using post Hartree-Fock (MP2, Coupled Clusters) methods and their comparison to experimental (IR, μ w) data.

2. Benchmarks on furan and benzofuran

2.1. Equilibrium geometries of furan and of benzofuran

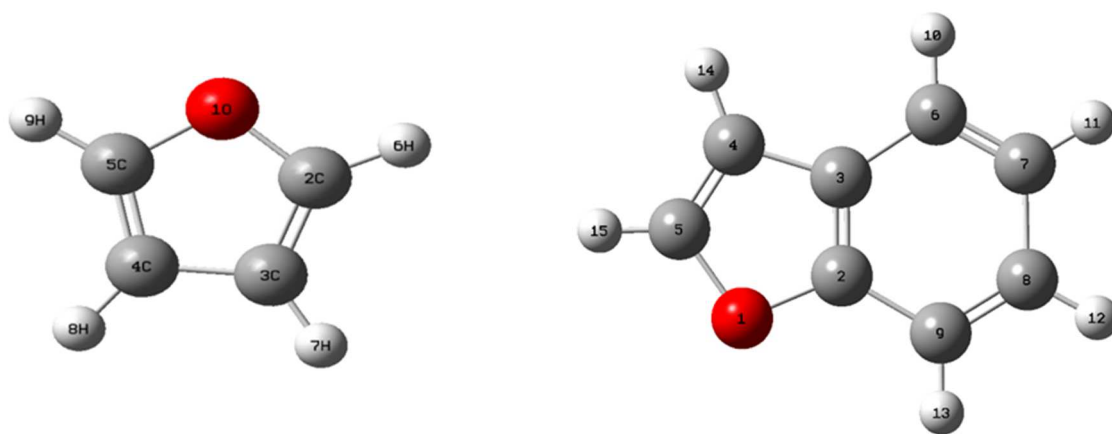


Figure 1: Structures of furan (left) and of benzofuran (right). We give also the atom numbering.

The most efficient theoretical approach to be used for the larger benzofuran derivatives of interest in the present study may be selected by comparison with experimental results. We started hence our investigations by optimizing the furan and benzofuran equilibrium structures in order to select the most reliable predictive method comparatively to experiment and with reduced computational cost. Indeed, a compromise between accuracy and computational cost (both CPU and disk occupancy) should be found in order to be able to compute the properties of our series of the benzofuran derivatives.

Table 2: Bond lengths (in Å) and valence angles (in degree, °) of furan and of benzofuran. See Figure 1 for the numbering of the atoms. Experimental data for furan and for benzofuran (See NOTE 1) are from Ref. [19] and Ref. [20], respectively.

Furan								
Parameters	Exp.	B3LYP		MP2			CCSD	
		6-31G(d,p)	cc-pVDZ	6-31G(d,p)	cc-pVDZ	aug-cc-pVTZ	6-31G(d,p)	cc-pVDZ
O1-C2	1.3621	1.364	1.363	1.366	1.363	1.360	1.367	1.365
C2-C3	1.3609	1.360	1.364	1.366	1.378	1.365	1.358	1.369
C3-C4	1.4309	1.436	1.438	1.427	1.436	1.426	1.440	1.450
C2-H6	1.0750	1.079	1.086	1.075	1.088	1.075	1.075	1.089
C3-H7	1.0768	1.080	1.087	1.076	1.089	1.076	1.076	1.090
O1-C2-C3	110.68	110.5	110.6	110.5	110.8	110.4	110.8	111.1
C2-C3-C4	106.50	106.1	106.0	106.2	105.8	106.1	106.0	105.6
C2-O1-C5	106.55	106.8	106.8	106.6	106.8	106.9	106.3	106.5
O1-C2-H6	115.92	115.8	115.7	115.7	115.8	115.8	115.7	115.8
C2-C3-H7	127.95	126.5	126.5	126.2	126.3	126.1	126.5	126.6
Benzofuran								
O1-C2	-	1.370	1.370	1.372	1.369	1.365	1.372	1.370
C2-C3	1.409	1.408	1.410	1.407	1.417	1.405	1.400	1.409
C3-C4	1.470	1.444	1.446	1.438	1.446	1.435	1.450	1.458
C4-C5	1.350	1.355	1.359	1.361	1.372	1.360	1.353	1.364
O1-C5	1.362	1.374	1.374	1.375	1.373	1.369	1.377	1.376
C2-C9	1.452	1.390	1.392	1.393	1.401	1.389	1.393	1.402
C3-C6	1.407	1.403	1.406	1.405	1.413	1.402	1.404	1.413
C6-C7	1.397	1.392	1.349	1.390	1.340	1.388	1.390	1.399
C7-C8	1.409	1.408	1.410	1.410	1.419	1.407	1.410	1.419
C8-C9	1.396	1.393	1.396	1.392	1.401	1.390	1.391	1.400
C2-O1-C5	106.7	106.0	105.9	106.0	105.6	105.9	105.8	105.5

O1-C2-C3	110.4	110.3	110.4	110.5	110.8	110.5	110.7	110.9
C2-C3-C4	105.2	105.3	105.3	105.9	105.2	105.4	105.3	105.1
C3-C4-C5	105.4	106.0	106.0	105.6	105.8	105.9	105.6	105.8
O1-C5-C4	112.3	112.4	112.4	112.4	112.6	112.3	112.5	112.8
C2-C3-C6	119.8	118.7	118.9	118.8	119.0	118.8	119.0	119.2
C3-C6-C7	117.6	118.4	118.4	118.3	118.3	118.2	118.3	118.3
C6-C7-C8	121.3	121.4	121.4	121.5	121.4	121.5	121.3	121.2
C7-C8-C9	121.4	121.3	121.2	121.4	121.3	121.4	121.4	121.3

Table 2 gives the equilibrium structures of furan and of benzofuran as computed at the B3LYP/6-31G (d,p), B3LYP/cc-pVDZ, MP2/6-31G (d,p), MP2/cc-pVDZ, MP2/aug-cc-pVTZ and CCSD/6-31G (d,p) and CCSD/cc-pVDZ levels of theory. These computations are carried out using the GAUSSIAN 09 program package [21]. Table 2 lists also the corresponding experimental geometrical parameters as measured using microwave spectroscopy [19,20]. Since furan and benzofuran are planar, the calculated dihedral angle values are either 0° or 180°.

Close inspection of Table 2 shows that CCSD method in conjunction with either the 6-31G (d,p) or the cc-pVDZ basis sets performs quite well since the differences between the computed and the measured equilibrium distances amount only to few thousandths of Å. The in-plane angles are also close to the experimental ones (within less than 1°). As established in the literature, CCSD approach can be used as a reference method to compare with. Interestingly, the less computationally demanding approaches i.e. B3LYP/6-31G(d,p), B3LYP/cc-pVDZ, MP2/6-31G (d,p), MP2/cc-pVDZ, MP2/aug-cc-pVTZ, perform also quite well compared to CCSD and to experiments since they lead to equilibrium structural parameters which differ by less than few hundredths of Å for the distances and less than 1° for the angles. Especially, a good agreement can be observed between the predicted B3LYP/6-31G (d,p) and experimental data. From that, we can conclude that the DFT/B3LYP method is appropriate enough for studies of benzofuran derivatives.

2.2. Vibrational analysis

Table 3: Anharmonic vibrational frequencies (in cm⁻¹) of furan and of benzofuran as computed using B3LYP/6-31 G(d,p). We give also the experimental data.

furan				Benzofuran			
Mode N°	Sym.	B3LYP / 6-31(d,p)	Exp. [25]	Mode N°	Sym.	B3LYP / 6-31(d,p)	Exp. [26]
1	a ₁	3165	3169	1	a'	3164.6	3158.0
2	a ₁	3143	3140	2	a'	3126.4	3124.5
3	a ₁	1494	1491	3	a'	3085.2	3094.0
4	a ₁	1392	1385	4	a'	3056.9	3077.0

5	a ₁	1150	1140	5	a'	3064.6	3067.0
6	a ₁	1075	1067	6	a'	3028.6	3047.0
7	a ₁	1005	995	7	a'	1631.2	1617.0
8	a ₁	874	870	8	a'	1601.4	1595.0
9	a ₂	865	864	9	a'	1558.1	1542.8
10	a ₂	713	722	10	a'	1487.1	1478.5
11	a ₂	606	600	11	a'	1461.0	1456.8
12	b ₁	3156	3161	12	a'	1373.0	1346.2
13	b ₁	3133	3130	13	a'	1343.5	1329.0
14	b ₁	1577	1558	14	a'	1274.2	1264.2
15	b ₁	1270	1267	15	a'	1259.6	1252.9
16	b ₁	1181	1181	16	a'	1181.6	1179.9
17	b ₁	1048	1043	17	a'	1164.0	1161.0
18	b ₁	878	873	18	a'	1138.4	1130.8
19	b ₂	827	838	19	a'	1112.4	1107.0
20	b ₂	747	745	20	a'	1044.0	1035.8
21	b ₂	612	603	21	a'	1018.6	1007.0
				22	a'	901.7	900.1
				23	a'	855.4	855.2
				24	a'	770.5	767.6
				25	a'	613.6	611.6
				26	a'	543.7	538.7
				27	a'	404.3	400.3
				28	a''	965.9	970.0
				29	a''	927.6	929.0
				30	a''	861.4	861.1
				31	a''	848.0	848.0
				32	a''	770.4	763.0
				33	a''	751.6	745.6
				34	a''	737.2	731.5
				35	a''	589.6	584.0
				36	a''	575.6	569.0
				37	a''	425.2	417.9
				38	a''	251.3	246.3
				39	a''	217.6	211.2

Accurate calculations of anharmonic frequencies of molecular systems allow correct identification of molecules through the assignment of their experimental RAMAN, IR or Far-IR spectra. Such computations are also needed to establish the reliability of the derivatives of the potentials, i.e. force fields, in semi empirical approaches. Obviously, this is crucial for accurate predictions of thermal contributions to enthalpies and entropies and of the reactive properties of these compounds using such methods. Here, we used second order perturbation theory (VPT2) approach as implemented in GAUSSIAN and MOLPRO [22] to evaluate the anharmonic vibrational frequencies of isolated furan and benzofuran. These data were calculated by means of B3LYP/6-31 G(d,p). The results are listed in Table 3, where they are compared to the respective FT-IR experimental data.

The benzofuran is a planar asymmetric rotor with C_s symmetry. The vibrations separate into 27 a' (in-plane) and 12 a'' (out-of-plane) fundamentals [23]. Furan is also a planar, near oblate symmetric top rotor. It possesses 8 a_1 , 3 a_2 , 7 b_1 , and 3 b_2 vibrations [24] (specified in Table 3). In brief, the comparison of the B3LYP/6-31G(d,p) computed and measured fundamentals for furan and benzofuran shows differences of less than 10 cm^{-1} between both sets of data. Again, this confirms the good performance of B3LYP/6-31G(d,p) chosen theoretical level for the derivation of properties of larger derivatives.

2.3.3D molecular electrostatic potential surface maps (3D MESP) of furan and of benzofuran

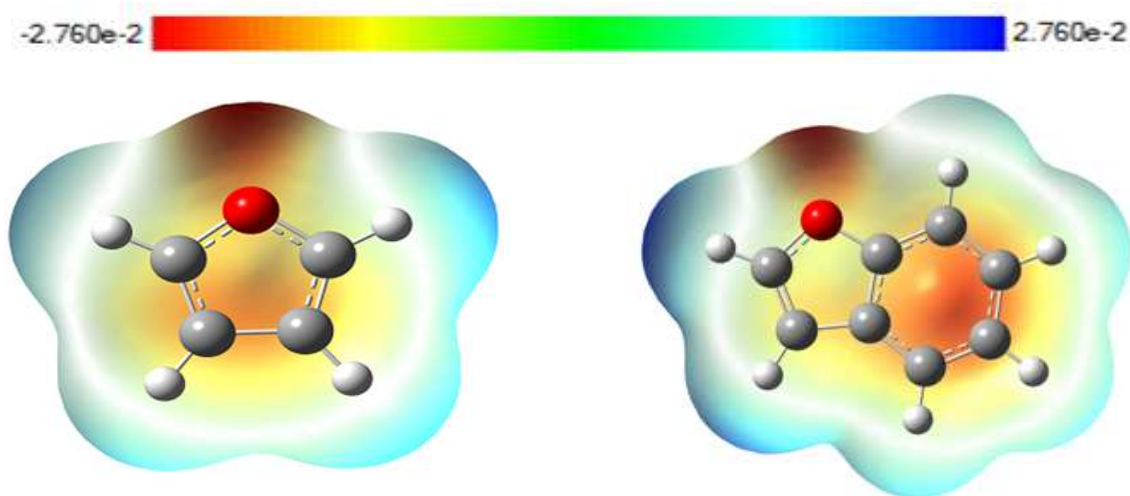


Figure 2: 3D MESP for furan (left) and for benzofuran (right). The results are shown by color, from red (most negative) to blue (most positive).

Figure 2 shows the 3D molecular electrostatic potential surface maps (3D MESP) of furan and of benzofuran. MESP allows understanding various physical and chemical phenomena such as molecular reactivity behavior, intermolecular interactions, molecular recognition, electrophilic reactions, substituent effects and the interactions induced by the reagents, for example between a drug and its cellular receptor [11].

As can be seen in Figure 2, furan and benzofuran present negative electrostatic potentials (red zone) around the oxygen atom due to its high electronegativity. An orange region can be observed around the carbon atoms of the five-ring of furan and the benzene ring of benzofuran. Thus, these parts may be subject to electrophilic attacks. We can see also positive electrostatic potentials (blue zone) around the atoms of hydrogens, which explain that these atomic sites are susceptible for nucleophilic attacks. In sum, furan and benzofuran exhibit common characteristics, which may be helpful for a qualitative understanding of the electrostatic interactions that may take place between reagents or enzyme active sites and the benzofuran derivatives under study.

3. Quantitative Structure Activity Relationships (QSAR) study

When chemical or physical properties and molecular structures are derived from numbers, it is often possible to propose mathematical relations connecting them, which allow making quantitative predictions. The obtained mathematical expressions can then be used as a predictive means of the biological response for similar structures. They are widely used in the pharmaceutical industry to identify promising compounds, especially at early stages of drug discovery [27].

Relationships between the physicochemical properties of chemical substances and their biological activities can be derived using QSAR (Quantitative Structure-Activity Relationships) concept. These models can also be used to predict the activities of new chemical entities and for their design [28]. Therefore, the biological activity is quantitatively expressed as the concentration of substance necessary to obtain a certain biological response. For that purpose, multiple linear regression, MLR, and artificial neural networks (ANNs) are used. The accuracy of such models is mainly evaluated by the correlation coefficient R^2 [9] and the average relative error ARE [29] (see NOTE 2). The MLR and ANN models were generated using the software JMP 8.0.2 [30].

The equilibrium geometries and the highest occupied molecular orbital energy (E_{HOMO}) and lowest unoccupied molecular orbital energy (E_{LUMO}) of benzofuran derivatives were determined at the B3LYP/6-31G (d,p) level of theory. We list in Table S1 of the supplementary material the Cartesian coordinates of the optimized benzofuran derivatives equilibrium structures. Then, the QSAR properties module from Hyper Chem 8.08 [33] was used to calculate: molar weight (MW), surface area (S), volume (V), refractivity (R), polarizability (P), octanol-water partition coefficient (log P) and hydration energy (HE). Finally, MarvinSketch17.1.2 Software [34] was employed to compute: hydrogen bond donors (HBD), hydrogen bond acceptors (HBA), number of rotatable bonds (NRB) and polar surface area (PSA).

3.1. Multiple linear regression (MLR)

MLR is one of the earliest and still one of the most commonly used methods for constructing QSAR mathematical models [31] because of its simplicity, transparency, reproducibility, and easy interpretability [32]. At present, QSAR was carried out on a series of fifteen furan derivatives (Table 1). In this work, nine descriptors were chosen to describe the structure of the benzofuran derivatives. Table 4 shows the values of the calculated descriptors used in order to establish the quantitative structure - antioxidant activity. The data

set was randomly divided into two sets: a training set (twelve compounds) and a testing set (three compounds: 6, 12 and 15) at a ratio of 80:20. Correlation matrix between parameters was performed on all nine descriptors. Nevertheless, the analysis revealed six independent descriptors for the development of the model.

Table 4: Values of molecular descriptors used in the QSAR study. We give the molar weight (MW, amu), surface area (S, Å²), volume (V, Å³), refractivity (R, Å³), polarizability (P, Å³), octanol-water partition coefficient (log P), hydration energy (HE, kcal/mol), energies of the HOMO (E_{HOMO}, eV) and LUMO (E_{LUMO}, eV) and IC₅₀ (in μM) (as given in the Table 1).

Compound	MW	S	V	R	P	log P	HE	E _{HOMO}	E _{LUMO}	IC ₅₀
1	242.23	419.57	685.41	74.61	25.35	-2.44	-23.21	-0.238	0.025	18.5
2	321.13	440.83	732.74	82.15	27.97	-2.39	-22.50	-0.240	0.008	20.6
3	305.13	433.63	711.92	80.54	27.33	-1.36	-16.04	-0.236	0.012	33.8
4	321.13	439.52	729.18	82.15	27.97	-2.39	-20.06	-0.235	0.013	18.0
5	337.13	458.90	765.50	83.75	28.61	-3.41	-27.98	-0.254	0.015	14.6
6*	321.13	447.86	741.16	82.15	27.97	-2.39	-22.68	-0.250	0.018	18.4
7	416.02	474.28	802.32	91.29	31.23	-3.36	-25.56	-0.257	0.004	19.4
8	400.02	471.11	790.38	89.68	30.60	-2.33	-22.03	-0.258	0.005	7.80
9	258.23	426.19	703.13	76.22	25.98	-3.46	-28.55	-0.237	0.023	12.8
10	242.23	418.17	682.36	74.61	25.35	-2.44	-22.56	-0.234	0.026	10.8
11	321.13	450.12	745.24	82.15	27.97	-2.39	-22.53	-0.251	0.017	28.6
12*	340.33	569.21	958.60	100.37	35.05	-2.27	-22.81	-0.237	0.017	16.8
13	340.33	582.08	965.45	100.37	35.05	-2.27	-22.70	-0.239	0.017	25.4
14	356.33	590.10	984.78	101.98	35.69	-3.29	-28.71	-0.240	0.018	11.7
15*	340.33	542.93	941.88	100.37	35.05	-2.27	-22.09	-0.241	0.016	26.2

The resulting MLR QSAR model is represented by the following equation:

$$IC_{50} = -212.620 - 0.785 MW + 0.397 S - 44.368 \log P + 9.274 HE - 1815.594 E_{HOMO} - 2746.546 E_{LUMO} \quad (1)$$

where, IC₅₀ is the response or dependent variable. MW, S, log P, HE, E_{HOMO} and E_{LUMO} are descriptors (features or independent variables). Within the regression, the coefficients in front of these descriptors are optimized [32].

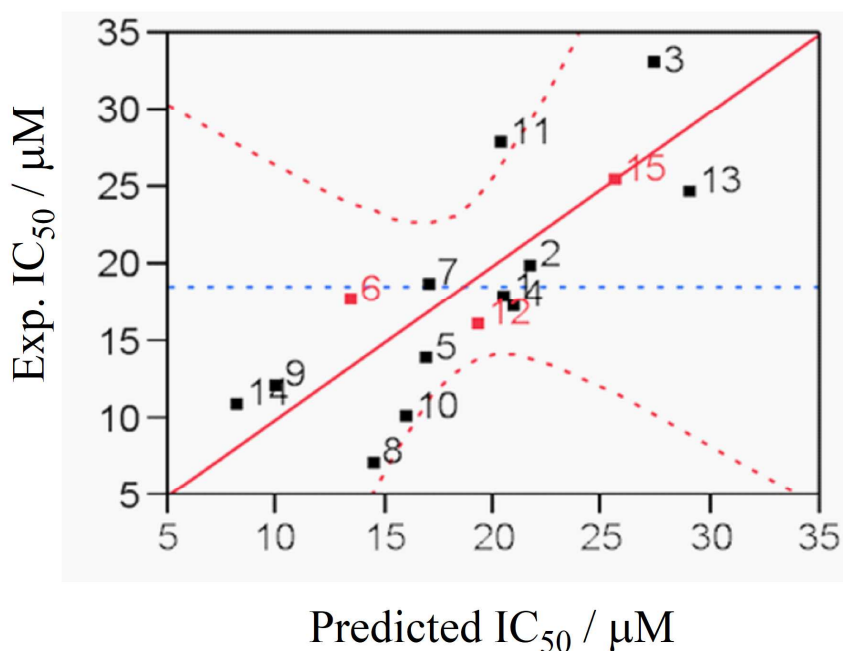


Figure 3: Correlation of experimental and predicted IC₅₀ (in μM) as derived using MLR.

For validation of the model, we plot in Figure 3 the experimental activities against the predicted values as determined by equation (1). We can observe that the predicted IC₅₀ values are in an acceptable agreement and regular distribution with experimental ones with R²= 0.64 and ARE of 0.74.

The positive values of S and HE coefficients indicate their promoter effect on the response, whereas the other factors (MW, log P, E_{HOMO} and E_{LUMO}) exhibit an opposite effect on IC₅₀. QSAR models aim also to define the importance of factor's effects. In our case, the surface is presented as the most influencing parameter on IC₅₀, followed by the molar weight and the hydration energy. The least important factor is log P. As the antioxidant activity is required to be improved, i.e. IC₅₀ decreased, it will be recommended to increase MW and reduce S and HE.

3.2. Artificial neural networks

Artificial neural networks (ANNs) models are non-linear models useful to predict the biological activity of large data sets of molecules [15]. In contrast to classical statistical methods such as regression analysis or partial least squares analysis, ANNs enable the investigation of complex and nonlinear relationships. Neural networks are therefore ideally suited for use in drug design and QSAR [35]. They are applied for simulating various non-linear complex systems of pharmaceutical, engineering, psychology and medicinal chemistry domains [36]. For instance, ANN was successfully used for the prediction and synthesis of new organic chemical compounds [16].

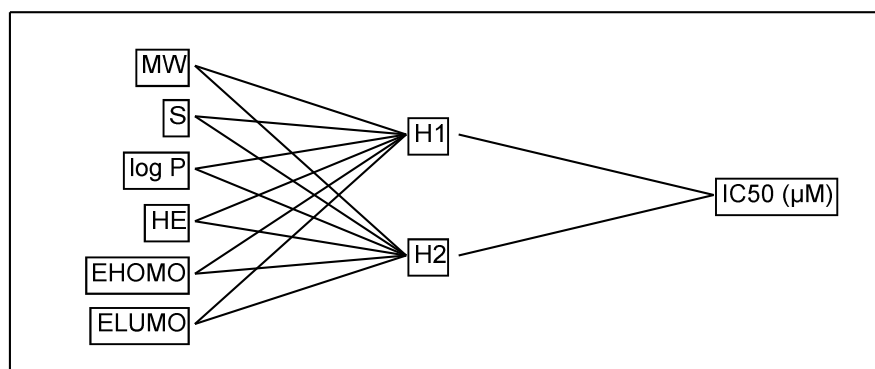


Figure 4: Structure of ANN.

In this work, ANN contained six inputs corresponding to the six descriptors selected from the correlation matrix, two hidden neurons, and one output neuron which is IC_{50} (Figure 4). The number of artificial neurons in the hidden layer was adjusted experimentally [37], two neurons in the hidden layer permitted to attain the best correlation between experimental and predicted data. Then, the ANN was trained using Gauss Newton method. A good correlation of experimental and predicted IC_{50} by ANN is found. This is shown in Figure 5, and illustrated by R^2 and ARE values of $R^2=0.99$ and ARE= 0.99.

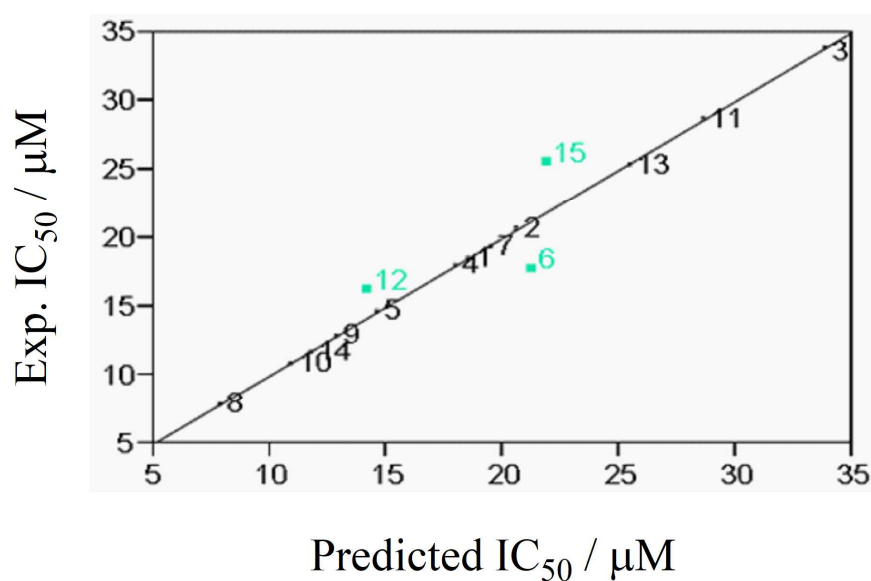


Figure 5: Correlation of experimental and predicted IC_{50} as calculated by ANN.

From both results of training and test sets (Figure 5), we can conclude that the ANN model with (6-2-1) architecture is able to establish a satisfactory relationship between the six descriptors and the antioxidant activity. For instance, all test molecules (6, 12 and 15) are in a good agreement with the two models.

4. Drug-likeness screening

Drug-likeness is a qualitative concept used in drug design [38]. It is described to encode the balance among the molecular properties of a compound that influences its pharmacodynamics, pharmacokinetics and ADME (Absorption, Distribution, Metabolism and Excretion) in human body [39]. Several rules have been proposed to evaluate drug-likeness. The most commonly used are Lipinski's rule of five [40] (see NOTE 4), Veber rules [41] (see NOTE 5), and lipophilicity indices [42, 43], which are viewed to be quite effective and efficient [44]. These represent guidelines and not absolute values to attest that a compound is a drug-like or not. Indeed, pharmacokinetic parameters are strongly influenced in vivo by the physicochemical properties of a drug [45]. Moreover, successful drug discovery not only requires the optimization of physicochemical parameters, but also more complex parameters related to toxicity and bioavailability defined as lipophilic efficiency indices [46]: ligand lipophilic efficiency LipE (see NOTE 4), ligand efficiency (LE see NOTE 6) and ligand-efficiency dependent lipophilicity LELP (see NOTE 7) [42]. At present, we evaluated the oral bioavailability of the fifteen benzofuran derivatives under study. Our parameters can be used to increase the overall efficiency of drugs at different stages of discovery [45].

Table 5: Drug-likeness parameters and lipophilicity indices of benzofuran derivatives. See text and NOTES 2-6 [40- 43] for the definition of the parameters given here.

Compound	Lipinski rules				Veber rules		Lipophilicity indices	
	MW (amu)	logP	HBD	HBA	PSA (Å ²)	NRB	LipE	LELP
	<500	<5	<5	<10	<140	<10	>5	-10 < LELP < 10
1	242.23	-2.44	3	3	73.82	1	7.17	-6.63
2	321.13	-2.39	3	3	73.82	1	7.08	-6.92
3	305.13	-1.36	2	2	53.6	1	5.83	-3.91
4	321.13	-2.39	3	3	73.82	1	7.13	-6.84
5	337.13	-3.41	4	4	94.05	1	8.24	-10.07
6	321.13	-2.39	3	3	73.82	1	7.12	-6.85
7	416.02	-3.36	4	4	94.05	1	8.07	-10.70
8	400.02	-2.33	3	3	73.82	1	7.44	-6.52
9	258.23	-3.46	4	4	94.05	1	8.35	-9.60
10	242.23	-2.44	3	3	73.82	1	7.41	-6.32
11	321.13	-2.39	3	3	73.82	1	6.93	-7.14
12	340.33	-2.27	3	4	100.13	5	7.04	-8.49
13	340.33	-2.27	3	4	100.13	5	6.86	-8.82
14	356.33	-3.29	4	5	120.36	5	8.22	-12.39
15	340.33	-2.27	3	4	100.13	5	6.85	-8.85

Table 5 presents the drug-likeness parameters and lipophilicity indices of the benzofuran derivatives under investigation. We give also the criteria to satisfy the lipophilicity indices and the Lipinski and Veber rules. As can be seen there, all compounds satisfy the rules, except compounds 5, 7 and 14 which show a slight out of range LELP (<-10) [47]. This table shows also that these compounds have molecular weights of less than 500 Da. Moreover, Table 5 gives the values of log P, which is a parameter used to estimate molecular hydrophobicity. For all studied compounds, log P is less than zero, so these molecules possess high water solubility and, in principle, poor membrane permeation capacity, which may be partly remedied by the small molecular weight. Hydrogen bonds increase solubility in water and must be broken for the compound to permeate into and through the lipid bilayer membrane. Thus, an increasing number of hydrogen bonds reduces partitioning from the aqueous phase into the lipid bilayer membrane for permeation by passive diffusion [44]. We see that all the screened molecules have a number of hydrogen bonds in the appropriate interval of Lipinski. In addition, all benzofuran derivatives show PSA values in the range of 53 - 120 Å², which permit them to correlate very well with the human intestinal absorption, Caco-2 monolayer's permeability, and blood-brain barrier penetration [48]. We can note also that compounds 12, 13, 14 and 15 are likely to be the most flexible with a number of rotatable bonds of 5.

In the studied set, LipE takes values between 5.83 and 8.35. Thus, this set would have a high quality as lead compounds, provided that LELP is situated in the suggested range $-10 < \text{LELP} < 10$ to guarantee that molecules reach both size and lipophilicity criteria, which is the case for all the studied compounds except three molecules (5, 7 and 14) having LELP inferior than the minimum limit ($-10.07, -10.70, -12.39$).

In general, according to the results listed in Table 5, we can estimate that all benzofuran derivatives satisfy the Lipinski and Veber rules, so they would not have problems with oral bioavailability. 80 % of compounds reach an LELP in the suggested range $-10 < \text{LELP} < 10$.

5. Summary and conclusion

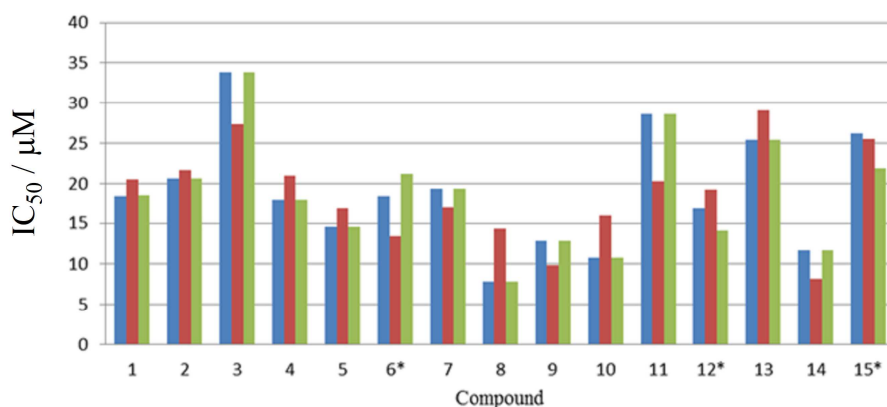


Figure 6: Comparative bar chart between experimental (in blue) and predicted IC₅₀ (in μM) using MLR (in red) and ANN (in green) methods.

A quantitative analysis of the structure- antioxidant activity relationship of fifteen antioxidant benzofuran derivatives is performed. QSAR models were obtained using multiple linear regression (MLR) and artificial neural network (ANN) techniques. Our work shows that this series of compounds obey the Lipinski's rule of five. These molecules satisfy also Veber's rules with NRB <10 and PSA <140. In addition, the calculated LELP values, these derivatives, expect compounds 5, 7 and 14, could have promising results in the clinic tests.

From a mythological point of view, we used first principles approaches to treat the furan and benzofuran subunits. The comparison of the spectroscopic parameters of these entities with experiment showed that the B3LYP/6-31 G(d,p) approach is sufficiently reliable. In addition, its low computational cost allows for investigating larger molecular systems. Afterwards, we showed that both ANN and MLR methods provide similar QSAR model accuracy. For instance, Figure 6 reveals that MLR and ANN predict almost coinciding IC_{50} s, whereas some systematic deviations can be observed with the measured IC_{50} s. In sum, to improve the QSAR models, research needs to be conducted with more activity data of similar compounds and molecular parameters [49].

Acknowledgements

The authors would like to extend their sincere appreciation to the Deanship of Scientific Research at King Saud University for funding the research through the Research Group Project No. RGP-333.

References

- [1] M. Percival, *Clin. Nut. Insights*. 1, 1 (1998).
- [2] B. Poljsak, U. Glavan, R. Dahmane, *Int. J. Cancer Res. Prev.* 4, 193 (2011).
- [3] T. Bahorun, M. A. Soobrattee, V. L. Ramma, O. I. Aruoma, *Internet J. Med. Update*. 1, 25 (2006).
- [4] Z.Chen, C. Zhong, *Neurosci. Bull.* 30, 271 (2014).
- [5] Md. N. Alam, N. J. Bristi, M. Rafiquzzaman, *Saudi Pharm. J.* 21, 144 (2013).
- [6] J.F. Hsieh, W.J. Lin , K.F. Huang , J. H. Liao , M. J. Don, C. C. Shen , Y. J. Shiao, W. T. Li, *Eur. J. Med. Chem.* 93, 446 (2015).
- [7] H. Khanam, Shamsuzzaman. Bioactive Benzofuran derivatives: A review. *Eur. J. Med. Chem.* 97, 483-504 (2015).
- [8] Database: <http://www.molinspiration.com>.
- [9] R. Darnag, B. Minaoui, M. Fakir, *Arab. J. Chem.* 10, 600-608 (2017).
- [10] K. Dermeche, N. Tchouar, S. Belaidi, T. Salah, *J. Bionosci.* 9, 4 (2015).
- [11] M. Alloui, S. Belaidi, H. Othmani, N. E. Jaidane, M. Hochlaf, *Chem. Phys. Lett.* 696, 70-78 (2018).
- [12] P. Xuan, Y. Zhang, T. J. Tzeng, X. F. Wan, F. Luo, *Glycobiology*. 22, 554 (2012).
- [13] S. Kothiwale, C. Borza, A. Pozzi, J. Meiler, *Molecules* 22, 6 (2017).
- [14] Z. Hajimahdi, A. Ranjbar, A. A. Suratgar, A. Zarghi, *Iran. J. Pharm. Res.* 14, 73 (2014).
- [15] J. M. Díaz-González, F. J. Aguirre-Crespo, X. García-Mera, F. J. Prado-Prado, *MOL2NET*. 1, 2 (2015).
- [16] M. Lazar, A. Mouzdahir, M. Badia, M. Zahouily, *J. Pharm. Res.* 8, 119 (2014).
- [17] E. Pournasheer, S. Vahdani, D. Malekzadeh, R. Aalizadeh, A. Ebadi, *Iran. J. Pharm. Res.* 16, 970 (2017).
- [18] C. Lee, W. Yang, R. G. Parr, *Phys. Rev. B* 37, 785-789 (1998).
- [19] B. Bak, D. Christensen, W.B. Dixon, L. Hansen-Nygaard, J. Rastrup-Andersen, M. Schottlander, *J. Mol. Spectrosc.* 9, 124 (1962).
- [20] A. Hartford, A. R. Muirhead, J. R. Lombardi, *J. Mol. Spectros.* 35, 199–213 (1970).
- [21] M.J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, & H. Nakatsuji, *Gaussian 09*, Revision E. 01, 2009, Gaussian. Inc., Wallingford CT.
- [22] H.-J. Werner, P. J. Knowles, et al., *MOLPRO*, ab initio programs package, 2015, <http://www.molpro.net>.
- [23] T. D. Klots, W. B. Collier. *Spectrochim. Acta A Mol. Biomol. Spectrosc.* 51, 1291-1316 (1995).
- [24] T. D. Klots, R. D. Chirico, W. V. Steele, *Spectrochim. Acta A Mol. Biomol. Spectrosc.* 50, 765-795 (1994).
- [25] A. Mellouki, J. Liévin, and M. Herman, *Chem. Phys.* 271, 239 (2001).
- [26] W.B. Collier, T.D. Klots, *Spectrochim. Acta Part A* 51, 1255 (1995).
- [27] F. Soualmia, S.Belaidi, H. Belaidi, N. Tchouar, Z. Almi, *J. Bionosci.* 11, 584 (2017).
- [28] B. Jhanwarb, V. Sharmaa, R. K. Singla, B. Shrivastava, *Pharmacologyonline* 1, 306 (2011).
- [29] Y. Zhang, H. Wang, Z. Yang, J. Li, *PLoS One* 9, 4 (2014).
- [30] JMP 8.0.2, SAS Institute Inc., (2009).
- [31] P. Liu, W. Long, *Int. J. Mol. Sci.* 10, 1979 (2009).

- [32] K. Roy, S. Kar, R. N. Das, A Primer on QSAR/QSPR Modeling, Springer Briefs in Molecular Science. USA, (2015).
- [33] Hyper Chem (Molecular Modeling System) Hypercube, Inc., 1115 NW, 4th Street, Gainesville, FL 32601, USA (2008).
- [34] MarvinSketch17.1.2, Chemaxon (2017), (<http://www.chemaxon.com>).
- [35] L. Terfloth, J. Gasteiger. DDT. 6, S102 (2001).
- [36] F. Cheng, V. Sutariya, Clin. Exp. Pharmacol. 2, 1 (2012).
- [37] M. Larif, A. Adad, R. Hmammouchi, A. I. Taghki, A. Soulaymani, A. Elmidaoui, M. Bouachrine, T. Lakhlifi, Arab. J. Chem.10, 948 (2017).
- [38] M. Joydeep, C. Raja, S. Saikat, V. Sanjay, D. Biplab, T. K. Ravi, Der. Pharma. Chemica. 1, 188-198 (2009).
- [39] G. Vistoli, A. Pedretti, B. Testa, Drug. Discov. Today. 13, 285-294 (2008).
- [40] C. A. Lipinski, F. Lombardo, B. W. Dominy, P. J. Feeney, Adv. Drug Deliv. Rev. 23, 3 (1997).
- [41] D. F. Veber, S. R. Johnson, H. Y. Cheng, B. R. Smith, K. W. Ward, K. D. Kopple, J. Med. Chem. 45 2615-2623 (2002).
- [42] A. L. Hopkins, G. M. Keserü, P. D. Leeson, D. C. Rees, C. H. Reynolds, Nat. Rev. Drug Disc. 13, 105-121 (2014).
- [43] P.D. Leeson, B. Springthorpe, Nat. Rev. Drug Discov. 6, 881–890 (2007).
- [44] E. H. Kerns, L. Di, Drug-like Properties: Concepts, Structure Design and Methods: from ADME to Toxicity Optimization, Academic Press, Elsevier, USA, 2008.
- [45] A. Zerroug, S. Belaidi , I. Ben Brahim , L. Sinha, S. Chtita, J. King Saud University – Science (2018), in press, <https://doi.org/10.1016/j.jksus.2018.03.024>
- [46] J. A. Arnott, R. Kumar, S. L. Planey, J. Appl. Biopharm. Pharmacokinet. 1, 31-36 (2013).
- [47] G. M. Keserü, G. M. Makara, Nat. Rev. Drug Discov. 8, 203-212 (2009).
- [48] P. Ertl, B. Rohde, P. Selzer, J. Med. Chem. 43, 3714–3717 (2000).
- [49] J. Luo, J. Hu, L. Fu, C. Liu, X. Jin, Procedia Eng. 15, 5162-5163 (2011).

Notes

NOTE 1: The experimental values of bond length and angles are estimated from data concerning the 5-membered rings taken from microwave studies of thiophene and furan, and data for the 6-membered ring taken from a crystal study of 3-indolylacetic acid [20].

NOTE 2: The MLR and ANN models were generated using the software JMP 8.0.2 [34]. The accuracy of models is mainly evaluated by the correlation coefficient R^2 and the average relative error ARE. The relations

of R^2 [10] and ARE [35] are given by the following equations: $R^2 = 1 - \frac{\sum_{m=1}^N (\hat{y}_m - y_m)^2}{\sum_{m=1}^N (y_m - \bar{y})^2}$ and $ARE = 1 - \frac{\sum_{m=1}^N \frac{|y_m - \hat{y}_m|}{|y_m|}}{N}$, where: y_m is the desired output, \hat{y}_m is the predicted value by model, \bar{y} is the mean of dependent variable, and N is the number of the molecules in the data set.

NOTE 3: According to Lipinski's rule [23], a good absorption is achieved if: (i) molecular weight is under 500 Da; (ii) log P is under 5; (iii) there are less than 5 H-bond donors; (iv) there are less than 10 H-bond acceptors.

NOTE 4: Veber [24] found that reduced molecular flexibility, as measured by the number of rotatable bonds (NRB), and low polar surface area (PSA) or total hydrogen bond count (sum of donors and acceptors) are important predictors of good oral bioavailability, independent of molecular weight. He suggested that a high probability of good oral bioavailability is obtained if: (i) the polar surface area is equal to or less than 140 Å²; (ii) there are 10 or fewer rotatable bonds.

NOTE 5: LipE = pIC₅₀ - log P. It combines both potency and lipophilicity. It gives an estimation on the binding of a ligand to a given target. It is suggested to target a LipE in a range of 5-7 or even higher [26]. There is a greater likelihood of achieving good in vivo performance when potency can be increased without increasing logP [42].

NOTE 6: LE is defined as LE=1.4 x pIC₅₀ / N_H, where N_H is the number of heavy atoms. It gives a measure of the properties of molecules, such as size and lipophilicity, needed to increase their binding affinity to a drug target. [25]

NOTE 7: LELP = log P/ LE. The optimal LELP scores are -10 < LELP < 10 [43].