

Definition of a novel atomic index for QSAR: the refractotopological state.

R. Carrasco

Pharmaceutical Chemistry Center, Atabey, Playa Apto, La Habana, Cuba

J. A. Padrón

Faculty of Chemistry, Havana University, Cuba

J. Gálvez

Chemical-Physics Department, Valencia University, Spain

Received 23 January 2003, Revised 25 November 2003, Accepted 9 January 2004

ABSTRACT. Purpose. In this work, a novel descriptor of atoms in molecules is introduced. The Refractotopological State Index for atoms (R-state, \mathfrak{R}), rectifies the atomic refractivity values reported by Crippen et al with the atomic refractivity values of the topological environment of each skeletal atom in the molecule. **Method.** The R-state (\mathfrak{R}_i), for atom i is an hybrid index that is defined as the intrinsic refractivity value of the atom i (AR_i) plus a perturbation term ΔAR_i in the non-hydrogen depleted graph. **Results.** The variations of the \mathfrak{R} values in different molecules are showed. QSAR examples previously reported by other authors are given for benzimidazole inhibition of Lee strain flu virus and receptor binding affinity of β -carboline. **Conclusions.** The index does not only describe the representation of the atomic dispersive forces related to the molar refractivity but also the influence of bounded and unbounded atoms as a measure of the distance-effect of the other groups in the molecule. The R-state index has proved a good performance, either alone or combined with the electro topological (E)-state index. This implies that in those cases this representation of dispersive forces between the molecule and the active site is a valid approach to the biological problem.

INTRODUCTION

The final objective of any quantitative chemical structure-biological activity relationship (QSAR) is to portray the receptor by the structural, physicochemical and biological properties of the ligand. It is common knowledge that the relationship between the binding

energy of the ligand (related to the molecular forces involved in binding) with different physicochemical properties for different parts of the ligand can be establish. Therefore, the identification of the appropriate physicochemical property for each molecular force is an important task.

On the other hand, a central concept in chemistry states that the properties of a molecule are intimately related to its molecular structure. The molecular skeleton is frequently used to obtain information about the number and kinds of atoms and the connections between them. The pattern of these connections within the skeleton is what actually determines the three-dimensional architecture of any molecule. To establish a QSAR model it is necessary not only to completely define the molecular structure, but also the influences governing measured values of the physical properties or biological activities. The electronic features of the molecules are portrayed by molecular orbital calculations and the descriptors include the net charge on the atoms, the energy of frontier orbitals, maps of electrostatic potentials, etc. The second feature of the chemical structure's influences is broadly described by a great number of numerical indices associated to the skeleton itself of a compound as a whole, or to a fragment of it. Most of these indices are based on chemical graph theory (1).

An index that includes both of these features of structural attributions intrinsic to a molecule is the electrotopological state index for atoms in a molecule (2-4). This index recognizes that every atom in a molecule is unique, and this uniqueness arises from differences in the electronic and topological environment of each atom (4). This descriptor is formulated as an intrinsic value I_i plus a perturbation given by the electronic

Corresponding Author: Ramon Carrasco Velar, Pharmaceutical Chemistry Center, Ave. 25 y 158 Rpto. Atabey, Playa Apto, 16042, La Habana, Cuba. carrasco@cqf.co.cu

influence of the topological environment of the molecule.

One of the most important chemico-physical properties used in QSAR studies is the molar refractivity (MR). It has been shown to be related to lipophilicity, molar volume and steric bulk (5). The importance of splitting the MR into its atomic components for QSAR studies oriented to three-dimensional molecules was demonstrated by Crippen et al (6-7). A method for the estimation of molar refractivity, based on the assignment of 22 atomic contributions obtained by classification of each atomic fragment according to the number and nature of the connected atoms to it, was developed by those authors (8-9). The reported correlation coefficient between observed and predicted values of the property using this method is greater than 0.999, with standard deviations values of 0.774(8). This method has been widely used for MR estimation and QSAR studies. However, the atomic contributions to MR are not suitable as one-atom descriptors for QSAR studies since they do not change when the chemical environment is changed, for example if a substitution occurs.

Recently, a molecular index based on a combination of these same values of atomic refractivity was defined. (10) The capability of this index to afford good results in QSAR studies was demonstrated. In this paper, we propose a Refractotopological State (R-state, \mathfrak{R}) descriptor of atoms in molecules. This index corrects the atomic refractivity values reported by Crippen et al using the atomic refractivity values of the topological environment of each skeletal atom in the molecule as in the Electrotological State index defined by Kier and Hall (2-4). The applicability of this index in QSAR studies is also demonstrated.

MOLAR REFRACTIVITY AND POLARIZABILITY

Molar refractivity

The molar refractivity is a constitutive-additive property that is calculated by the Lorenz-Lorentz formula (eq. 1):

$$MR = \frac{n^2 - 1}{n^2 + 2} \cdot \frac{M}{\rho}$$

where M is the molecular weight, n the refraction index and ρ the density, which, at a given tempera-

ture, depends only on the wavelength of the light used to measure the refraction index. For a radiation of infinite wavelength, the molar refractivity represents the real volume of the molecules contained in one mole of the substance.

Polarizability of the molecules.

The electrons and nuclei of molecule are mobile to a limited degree. For that reason, when a polar or non-polar molecule is placed in an electric field a small displacement of the charge will take place. As a result, a dipole would be introduced in the molecule, in addition to the permanent one that may exist.

The molar polarization is given by:

$$P = \frac{D - 1}{D + 2} \cdot \frac{M}{\rho}$$

where P is the molar polarization, induced polarization or polarization of distortion and D is the dielectric constant of the environment. As D is an abstracted number, P represents a molar volume. This equation, called Mosotti-Clausius equation is only an approximation. In 1831, Maxwell related the dielectric constant with the refraction index measured at long wavelength:

$$D = n_{\infty}^2$$

and equation 2 becomes:

$$P = \frac{n_{\infty}^2 - 1}{n_{\infty}^2 + 2} \cdot \frac{M}{\rho}$$

where the right term of the equation [4] is the molar refractivity measured at long wavelengths and the magnitude is the average polarizability of a non-polar molecule in an electric field.

Therefore, the value of n_{∞} cannot be obtained by extrapolation to infinite wavelengths starting from the refraction indexes measured with visible light, because this radiation can only displace electrons, while the positively loaded nuclei remain without changes. In fact, if n is the refraction index measure using visible light, then:

$$P_E = \frac{n^2 - 1}{n^2 + 2} \cdot \frac{M}{\rho}$$

where P_E is the electronic polarization. This is similar to the molar refractivity measure using visible light

and represents the part of the total induced polarization P caused by the distortion or deformation of the molecule electronic cloud. This deformation originates from the presence of instantaneous dipoles of an approaching ligand. That is why the molar refractivity is related not only to the volume of the molecules, but also to the London dispersive forces that, for instance, play a key role in the drug-receptor interaction.

The atomic contribution to molar refractivity calculated by the Ghose and Crippen method.

In order to model the dispersive interactions, Ghose and Crippen defined 110 atom types, (6-8) representing most commonly occurring atomic states of carbon, hydrogen, oxygen, nitrogen, halogens, and sulphur in organic molecules. They stated that this classification partially differentiates the polarizing effects of heteroatoms and the effect of overlapping with non-hydrogen atoms, although they accepted that this classification might be weak in differentiating the conjugation effects. In the explanation of their method of calculation, the authors said that the classification may not completely cover all organic molecules, and that the addition of atom types is possible. Further on, the 110 atom types were reduced to 22. The evaluation of the individual atomic value is based on the idea that the sum of the atomic values (α_i) being related to the molecular value of the molar refractivity:

$$MR_{calc.} = \sum n_i \alpha_i$$

The Electropotological State Index (E-State)

The electropotological state index (E-state) is developed from chemical graph theory and uses the chemical graph (hydrogen-suppressed skeleton) for generation of atom-level structure indices. The index is based on the electronic effect of each atom on the other atoms in the molecule as modified by molecular topology (2-4). Each atom has an assigned intrinsic state value I_i calculated as follows:

$$I_i = [(2/N)^2 \delta^v + 1] / \delta$$

where N is the principal quantum number of the atom i , δ^v the number or valence electrons in the skeleton ($Z^v - h$), and δ the number of σ electrons in the skeleton (σh). For a skeleton atom, Z^v is the number of valence electrons, σ the number of electrons in σ orbitals, and h the number of bounded hydrogen atoms. The E-state $S(A_i)$ for the atom is the modified

intrinsic value:

$$S(A_i) = I_i + \Delta I_i$$

where ΔI_i quantifies the perturbing effect on the intrinsic atom value. This perturbation is assumed to be a function of the difference in the intrinsic values I_j and I_i :

$$\Delta I_i = \sum_{j=1}^N (I_i - I_j) / r_{ij}^2$$

where r_{ij} is the number of atoms in the shortest path between atoms i and j including both i and j . The difference in intrinsic values, ΔI_{ij} , for a pair of skeletal atoms encodes both electronic and topological attributes that arise from electronegativity differences and skeletal connectivity. Derived from this electronegativity difference, the E-state value for an atom is related but not limited to the concept of atomic partial charge.

The Refractotopological State Index (R-State).

The refractotopological state index R-state is developed from the chemical graph theory and the partition of the molar refractivity defined by Ghose and Crippen. The index is based on the influence of dispersive forces of each atom on the other atoms in the molecule, modified by molecular topology.

\mathfrak{R}_i Definition

The R-state (\mathfrak{R}_i), for atom i is defined as follows,

$$\mathfrak{R}_i = AR_i + \Delta AR_i$$

where AR_i is the intrinsic refractivity value of the atom i , and ΔAR_i is a perturbation term defined by:

$$\Delta AR_i = \sum_{j=1}^N (AR_i - AR_j) / r_{ij}^2$$

where the sum is all over j adjacent vertices in the graph, AR_i and AR_j are the intrinsic refractivity values of the atoms i and j respectively, and r_{ij}^2 is the number of atoms in the shortest path between atoms i and j including both i and j that's to say, the distance between atoms counted as the graph distance (d_{ij}) plus one (2-4). As in the E-state, this quadratic topological distance, indicates that there must be a decrease in the interaction effect with the separation distance between atoms. Atomic refractivity values (7) used in the analysis is shown in table 1. The intrinsic values of the

refractivity for each one of the heavy atoms include the atomic refractivity value of the hydrogen atoms bonded to them.

Table 1: Atomic refractivity values used in the analysis.

Atom Type	Atomic Refractivity	Atom Type	Atomic Refractivity
C _{sp₃}	2.8158	N (A _r)	2.7662
C _{sp₂}	3.8278	NO ₂	3.5054
C _{sp}	3.8974	Ar-N=X	3.8095
C (A _r)	3.5090	F	1.0632
C=X	3.0887	Cl	5.6105
H	0.9155	Br	8.6782
-O-	1.6351	I	13.8741
=O	1.7956	S _{sp₂}	7.3190
O=N	2.1407	S _{sp₃}	9.1680
N _{sp₃}	3.0100	R-SO-R	6.0762
N _{sp₂} , N _{sp}	3.2009	R-SO ₂ -R	5.3321

Method of calculation.

In all cases \mathcal{R}_i index calculation involves three steps: (a) molecular structure generation, (b) classification of the atom types according to the 22 reported types (table 1) and (c) \mathcal{R}_i index calculation. The RIGA program (10) and software developed in the author's laboratories were used to perform the (a), (b) and (c) steps, respectively. The STATISTICA v.4.0 software was used in all the statistical analysis of the data (11).

To illustrate the calculation steps of the index the procedure is shown in table 2 for 2-chloro-2 butanol.

Note that the sum of the R-state index of all the atoms (A_i) in the molecule is equal to the sum of the intrinsic atomic contribution values (ΣAR_i), but the value itself is rather dissimilar. Using this approach the methyl groups 1 and 4 are marked as non equal as expected since their environments are different and the dispersive forces in the methyl group 4 should be higher because of the number and nature of the groups present in that part of the molecule. It is also relevant the decrease of the value on the atom number 3, which is buried in the skeleton and hence unable to have a direct interaction with a hypothetical active site or molecule from a dispersive-sterical point of view.

Table 2: Refractotopological state calculation for 2-chloro-2 butanol.

Intrinsic Values

AR(1)=AR(C _{sp3})+3 AR(H)=2.8158+2.7465=5.562	A _i (1)=6.2811
AR(2)=4.6468	A _i (2)=4.8999
AR(3)=2.8158	A _i (3)=0.7340
AR(4)=5.562	A _i (4)=6.6795
AR(5)=2.5506	A _i (5)=1.3886
AR(6)=5.6105	A _i (6)=6.7647

$(AR_i - AR_j)/r_{ij}$ matrix									
i	1	2	3	4	5	6	$\Sigma \Delta AR_i$ values	ΣAR_i Values (Crippen)	A_i values ($AR_i + \Delta AR_i$)
1	0.0000	0.2288	0.3051	0.0000	0.1882	-0.0030	0.7191	5.562	6.2811
2	-0.2288	0.0000	0.4578	-0.1017	0.2329	-0.1071	0.2531	4.6468	4.8999
3	-0.3051	-0.4578	0.0000	-0.6866	0.0663	-0.6987	-2.0818	2.8158	0.7340
4	0.0000	0.1017	0.6866	0.0000	0.3346	-0.0054	1.1175	5.562	6.6795
5	-0.1882	-0.2329	-0.0663	-0.3346	0.0000	-0.3400	-1.1620	2.5506	1.3886
6	0.0030	0.1071	0.6987	0.0054	0.3400	0.0000	1.1542	5.6105	6.7647
						Total	0.0000	26.7477	26.7477

EXAMPLES AND DISCUSSION

When compared to other topological methods, which do not consider the influence of the hydrogen bonded to the atoms, this index includes this influence on the intrinsic values of the heavy atoms. According to the nature of the studied phenomenon, this inclusion reflects the total interaction volume of the group with the protein ligand and it is important not only to the intrinsic value of the direct bonded atom but also to other groups present on the molecule.

Effect of the position. Buried atoms.

The index presents a large decrease on its value as the atom changes from primary to quaternary as represented in figures 1 and 2, where a series of pentanes and amines are presented along with the R-state values for each skeletal atom.

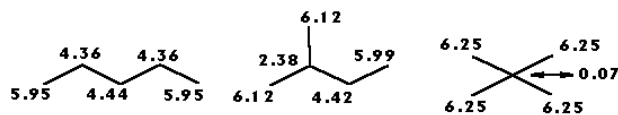


Figure 1: Changes on the R-state values in pentane isomers.

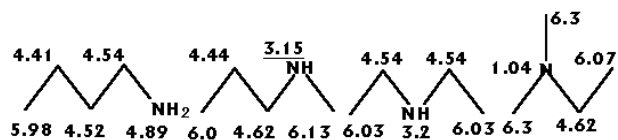


Figure 2: Changes on the R-state values in amines.

Due to the nature of the dispersive forces, it is logical that an atom will interact less with the protein ligand as it becomes buried in the skeleton and hence a smaller total interaction volume. Note that the amine group R-state values reflect changes following small variations on its position.

Figure 3 shows R-state values for alcohols. As the OH group has an R-state intrinsic value lower than the methyl group, its value is influenced accordingly.

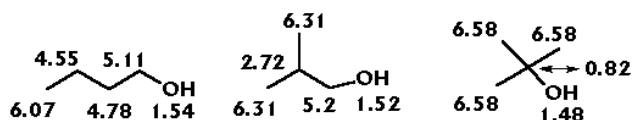


Figure 3: Changes on the R-state values in alcohols.

Effect of the atom nature in aromatic compounds.

In figure 4 a series of halo-substituted toluenes is shown. The effect of changing the halogen substituents is reflected by the R-state values: the higher values correspond to the less electronegative groups.

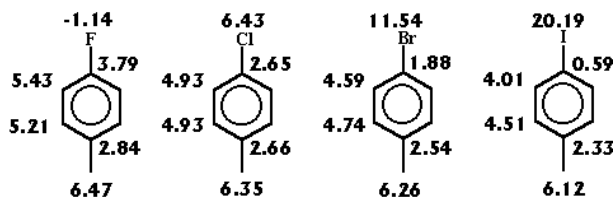


Figure 4: Changes on the R-state values in halogen toluene derivatives.

Effect of the position isomers in aromatic compounds.

The R-state index is also able to portray the changes on the relative position, as is shown on figure 5. These changes affect not only the halogen atom but also the different atoms in the molecule. For example, the more buried C₁, the less corresponding value of the index.

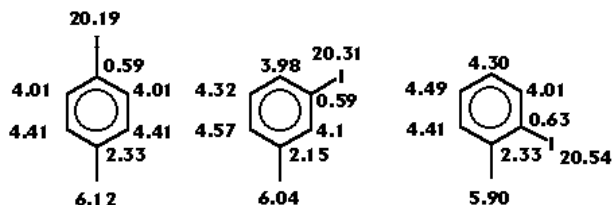


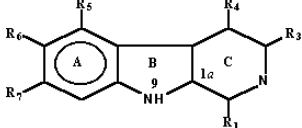
Figure 5: Changes on the R-state values in iodotoluene isomers.

APPLICATION TO QSAR STUDIES

Receptor binding affinity of β -carbolines.

To test the capability of the index to describe interactions in active sites, the receptor binding of a series of β -carbolines was investigated (12). The affinities were determined by the drug displacement of 50% of 3H-flunitrazepam from synaptosomal rat membranes (pIC_{50}) and are given in Table 3.

Table 3: Reported data of the binding of β -carbolines to the benzodiazepine receptor*.



R ₁	R ₃	R ₄	R ₅	R ₆	R ₇	pIC ₅₀ ^a	Calc ^b	Res ^c
CH ₃	H	H	OH	H	H	-0.450	-0.685	0.235
CH ₃	H	H	H	H	OCH ₃	-1.980	-1.789	-0.191
CH ₃	H	H	H	H	OH	-1.900	-1.747	-0.153
H	CH ₂ OH	H	H	H	H	1.591	2.341	-0.750
H	COOC ₂ H ₅	H	H	H	H	2.097	2.135	-0.038
H	COOC ₂ H ₅	H	H	H	H	2.155	2.090	0.065
H	COOC ₃ H ₇	H	H	H	H	1.921	2.045	-0.124
H	COOC ₂ H ₅	C ₂ H ₅	H	OCH ₃	OCH ₃	2.400	1.941	0.459
H	COOC ₂ H ₅	CH ₂ OCH ₃	OCH ₂ C ₆ H ₅	H	H	2.959	2.816	0.143
H	COOC ₂ H ₅	CH ₃	O=C ₂ H ₇	H	H	2.960	2.547	0.413
H	COONHCH ₃	H	H	H	H	1.780	2.071	-0.291
CH ₃	H ₂	H ₂	H	H	OCH ₃	-2.813	-2.906	0.093
CH ₃	H ₂	H ₂	H	H	OH	-2.748	-2.841	0.093
H	COOC ₂ H ₅	CH ₂ OCH ₃	H	OCH ₂ C ₆ H ₅	H	2.959	2.815	0.144
H	H	H	H	H	H	-1.000	-1.432	0.432
H	H	OOCH	H	H	H	-1.491	-0.961	-0.530

*The ring C is non-aromatic, only for compounds 12 and 13. ^a The negative log of the IC₅₀ value for binding. ^b Activity Value computed from eq. 15. ^c Res = pIC₅₀-Calc.

The 13 common ring atoms in all structures were calculated and the corresponding R-state values obtained were correlated with the biological activity using forward stepwise regression analysis. The obtained model was then compared with the previously reported by Kier and Hall (3).

The best one-variable model obtained by using the R-state value includes the atom 11 as independent variable ($r=0.86$, $s=1.16$, $F=40$, $n=16$). This result is worse than the reported with the use of E-state ($r=0.893$, $s=1.03$, $F=55$, $n=16$). Nevertheless, the inclusion of the R-state value of the atom 8 results in a good model ($r=0.977$, $s=0.50$, $F=138$, $n=16$), better than the two variable equation reported using E-state ($r=0.92$, $s=0.93$, $F=36$, $n=16$). The three-variable

model (Eq. 12) is also better than the equivalent previously reported (Eq. 13).

$$pIC_{50} = -15.328 (\pm 2.895) - 1.746 (\pm 0.129) \mathfrak{R}_1 - 6.269 (\pm 0.914) \mathfrak{R}_{1,4} - 0.525 (\pm 0.158) \mathfrak{R}_7$$

$$r = 0.989 \quad F = 167 \quad s = 0.375 \quad n = 16$$

$$r^2_{press} = 0.947 \quad S_{press} = 0.522$$

$$pIC_{50} = -58.070 (\pm 11.484) + 0.934 (\pm 0.075) S_{CO} + 13.248 (\pm 2.740) S_{NN} + 3.426 (\pm 0.888) S_{14}$$

$$r = 0.965 \quad F = 54 \quad s = 0.644 \quad n = 16$$

where S_{CO} , S_{NN} and S_{C14} are linear combinations of several other variables (3). In Figure 1, it can be observed the performance of the relative errors.

Our results indicate that the best distribution is achieved through Eq. 12. Only in three cases were the errors slightly higher with the use of \mathfrak{R} than with E-STATE.

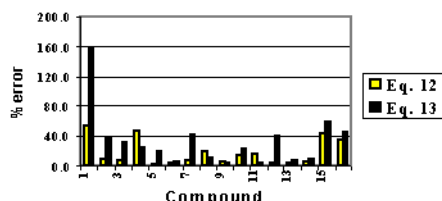


Figure 6: Relative errors for equations 12 and 13.

The results show that the R-state index allows for a better prediction of the binding affinities than the E-state index, therefore the dispersive forces play a predominant role in the specific interaction site between the drugs and the benzodiazepine receptor, rather than the electronic factor.

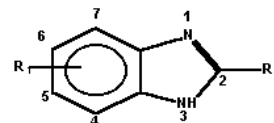
After the study, the atoms flagged, as the most relevant are those numbered 3, 1 α , and 7. Those positions reflect the changes in activity along with the different substitution patterns since they cover the β -carboline three-ring system.

According to the results, to enhance the activity it is necessary to reduce the atom 3 R-state value and increase atoms 1 α and 7 R-state values with adequate substitutions that follows this objectives.

Inhibition of Lee strain flu virus by benzimidazoles.

Another example of the use of the R-state in QSAR studies is the inhibition of Lee strain flu virus by benzimidazoles (13). This was studied after the computation of the R-state values for the seven common skeletal atoms in the series (Table 4).

Table 4: Inhibition of Lee strain flu virus by benzimidazoles.



Substituent	Equation 14		Equation 15		Equation 16		
	PK _i	Calc ^b Res ^c	Calc ^d Res ^e	Calc ^f Res ^g	Error (%)	Error (%)	
H	2.14	2.34 -0.20	9.35	2.29 -0.15	7.01	2.28 -0.14	6.54
2-methyl-	2.51	2.60 -0.09	3.59	2.55 -0.04	1.59	2.55 -0.04	1.59
5-methyl-	2.72	2.57 0.15	5.51	2.57 0.15	5.51	2.56 0.16	5.88
5,6-dimethyl-	2.72	2.80 -0.08	2.94	2.85 -0.13	4.78	2.83 -0.11	4.04
4,6-dimethyl-	2.82	2.93 -0.11	3.90	2.89 -0.07	2.48	2.90 -0.08	2.84
2,5-dimethyl-	2.89	2.82 0.07	2.42	2.83 0.06	2.08	2.82 0.07	2.42
4,5-dimethyl-	2.96	2.93 0.03	1.01	2.81 0.15	5.07	2.84 0.12	4.05
2,5,6-trimethyl-	3.05	3.04 0.01	0.33	3.19 -0.14	4.59	3.16 -0.11	3.61
2,4,5-trimethyl-	3.20	3.17 0.03	0.94	3.15 0.05	1.56	3.16 0.04	1.25
5,6-diethyl-	3.39	3.27 0.12	3.54	3.37 0.02	0.59	3.32 0.07	2.06
2-propyl-5-methyl-	3.60	3.73 -0.13	3.61	3.51 0.09	2.50	3.57 0.03	0.83
2,4,5,6,7-pentamethyl-	3.66	3.74 -0.08	2.19	3.88 -0.22	6.01	3.87 -0.21	5.74
2-ethyl-5-methyl-	3.74	3.40 0.34	9.09	3.72 0.02	0.53	3.64 0.10	2.67
2-butyl-5-methyl-	3.77	3.95 -0.18	4.77	3.79 -0.02	0.53	3.84 -0.07	1.86
2-isopropyl-5-methyl-	3.77	3.75 0.02	0.53	3.55 0.22	5.84	3.61 0.16	4.24
Mean			3.58		3.38		3.31

^a Experimental value for the inhibition constant. ^b Values computed for inhibition from eq. 14 ^cres=pKi-calc. ^d Values computed for inhibition from eq. 15 ^eres=pKi-calc. ^f Values computed for inhibition from eq. 16 ^gres=pKi_{exp}- pKi_{calc}

The regression equation and statistical parameters for the correlation biological property-R-state values are shown (Eq. 14).

$$pK = 34.085 (\pm 2.683) - 0.100 (\pm 0.035) \mathfrak{R}_{C2} - 8.042 (\pm 0.721) \mathfrak{R}_{N1,3}$$

$$r = 0.969 \quad F = 92 \quad s = 0.137 \quad n = 15$$

$$r^2_{press} = 0.910 \quad S_{press} = 0.152$$

where r, s and F represent the correlation coefficient, standard deviation of regression, and Fisher ratio, respectively. The high value of the cross-validation regression coefficient (r^2_{press}), and the similarity of "s" and S_{press} , tells us about the high stability of the model. The variable \mathfrak{R}_{C2} is the R-state of the atom numbered as 2, and $\mathfrak{R}_{N1,3}$ is the average of the R-state values of the nitrogen atoms numbered 1 and 3. The best model previously reported is the one obtained by Hall (4), which includes the E-state values of the atom 2, and the

average E-state values of the nitrogens 1 and 3 (eq. 15).

$$pK = -15.995 (\pm 2.044) + 4.267 (\pm 0.436) S_{N1,3} + 0.786 (\pm 0.193) S_{C2}$$

$$r = 0.956 \quad F = 63 \quad s = 0.162 \quad n = 15$$

$$r^2_{press} = 0.850 \quad S_{press} = 0.190$$

In this case, the R-state values afford a better model than the E-state values with the same variables. That is explained by the best values of each statistical parameter in the eq. 14 as compared with eq. 15. This suggests that the biological response of benzimidazoles against flu virus involve not only electrostatic interactions explained by E-state, but also local dispersive forces represented by the R-state. On the other hand, we have found that the combination of both descriptors results in an even better model (eq. 16), as can be seen in the higher values of r , F , and r^2_{press} and in the lesser values of s and S_{press} respectively.

$$pK = 20.903 (\pm 5.090) - 4.267 (\pm 1.006) \mathfrak{R}_{N1,3} + 1.216 (\pm 0.350) S_{N1,3}$$

$$r = 0.974 \quad F = 110 \quad s = 0.125 \quad n = 15$$

$$r^2_{press} = 0.923 \quad S_{press} = 0.144$$

Therefore, it is possible to express the biological activity of these drugs as a combination of both electronic and dispersive forces in the interaction. In this combination there must be a balance between the dispersive value of the substituents to decrease the nitrogens 1 and 3 R-state values (negative coefficient) and the electronegativity of the substituents to increase the E-state of the same atoms (positive coefficient), in order to optimize the activity of the drugs under study. In a box and whisker plot (Figure 2), is seen that with the combined use of \mathfrak{R} and the E-state, it is possible to obtain a better quality for the relative error.

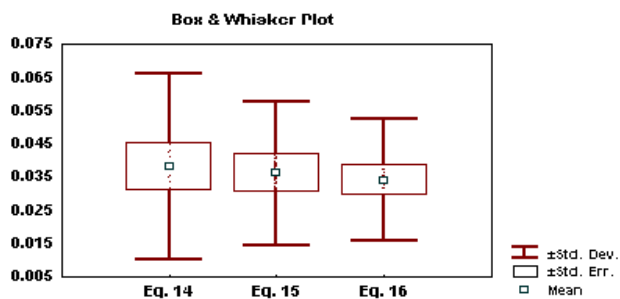


Figure 7: Box and whisker plot of the standard deviation, the standard error and the mean of the errors in equations 14, 15 and 16.

CONCLUSIONS

An index based on the correction of the atomic refractivity values, defined by Crippen et al with the atomic refractivity values of the topological environment of each skeletal atom in molecule, is proposed. The new atomic descriptor R-state promises to be useful in quantitative structure-activity relationships studies, especially in those involving dispersive atomic interactions with active sites. The index has a direct structural interpretation since it expresses the atomic contribution of an important physico-chemical property (molar refractivity) related to dispersive forces in biologically active sites and also contains topological information about the chemical structure of the atomic environment. In addition, the R-state index can be successfully used in QSAR studies with the Electrotopological State index that it is very simple to calculate since no quantum chemical calculations are involved.

ACKNOWLEDGEMENTS

This work was supported by grants from The Ministry of Science, Technology and Environment of the Republic of Cuba. R.C. thanks Dr. D. José Elguero Bertollini the helpful discussions for the conclusion of the manuscript. Thanks to M. Encinosa for his kind help and incisive critics.

REFERENCES AND NOTES

- [1] For goods reviews on this topic see: a) Kier, L.B.; Hall, L.H. *Molecular Connectivity in Chemistry and Drug Research*; Academic Press, New York, 1976. b) Kier, L.B.; Hall, L.H. *Molecular Connectivity in Structure-Activity Analysis*; John Wiley, London, 1986. c) Trinajstić, N.; *Chemical Graph Theory*; CRC, Boca Raton, FL, 1983; Vols. I and II.
- [2] Hall, L. H., Mohnhey, B.; Kier, L.B., *The Electrotopological State: Structure Information at the Atomic Level for Molecular Graphs*, J. Chem. Inf. Comput. Sci., 1991, 31, 1, 76-82.
- [3] Hall, L. H., Mohnhey, B.; Kier, L.B., *The Electrotopological State: An Atom Index for QSAR*, Quant. Struct.-Act. Relat., 1991 10, 43-51.
- [4] Kier, L.B. and Hall, L. H., *An Electrotopological State Index for Atoms in Molecules*, Pharmac. Res., 1990, 7, 8, 801-807.
- [5] Kubinyi, H., QSAR: Hansch Analysis and Related Approaches, p. 40; *Methods and Principles in Medicinal Chemistry*, R. Mannhold, P. Krogsgaard-Larsen and H. Timmerman, eds. VCH 1993.

- [6] Ghose A.K.; Crippen, G.M. *Atomic Physicochemical Parameters for Three-Dimensional Structure-Directed Quantitative Structure-Activity Relationships. I. Partition Coefficients as a Measure of Hydrophobicity.* J. Comput. Chem., 1986, 7, 4, 565-577.
- [7] Ghose A.K.; Crippen, G.M. *Atomic Physicochemical Parameters for Three-Dimensional Structure-Directed Quantitative Structure-Activity Relationships. 2. Modeling Dispersive and Hydrophobic Interactions.* J. Chem. Inf. Comput. Sci., 1987, 27, 1, 21-35
- [8] Viswanadhan, V.N.; Ghose A.K. ; Revankar, G.R. and Robins, R.K. J. Chem. Inf. Comput. Sci, 1989, 29, 163.
- [9] Padrón, J.A.; Carrasco, R and Pellón, R.F., J. Pharm. Pharmaceut. Sci. *Molecular descriptor based on a molar refractivity partition using Randic-type, graph theoretical invariant.* (www.ualberta.ca/~csps) 5(3)267-274, 2002.
- [10] RIGA Editor, Institute of Organic Synthesis, Latvian Academy of Science, Riga, Latvia, 1991
- [11] STATISTICA for WINDOWS, version 4.0, Statsoft, Inc.(1993)
- [12] Borea, P. A., Pietogrande, M.C. Biagi, G.L., Guerra, M.C., Barbaro, A.M. and Recanatini, M.; *QSAR: Quantitative Structure-Activity Relationships in Drug Design*, Alan R. Liss, Inc. New York, pp. 361-364 (1989).
- [13] Tamm, I; Folkers, K; Shunk, C.H.; Horofall, F.L.; J. Exp. Med., 1953, 98, 245.