ORIGINAL PAPER

Berenice da Silva Junkes · Anna Celia Silva Arruda
Rosendo Augusto Yunes · Ledilege Cucco Porto
Vilma Edite Fonseca Heinzen

Semi-empirical topological index: a tool for QSPR/QSAR studies

Abstract The semi-empirical topological index (I_{ET}), developed to predict the chromatographic retention of a series of organic compounds, is extended to predict other properties and biological activities of aliphatic alcohols. This topological index takes into account the contribution of each individual atom type to the property considered and is able to encode information about structural features of the molecules. The efficiency of this index is verified by high quality Structure - Property and structure - Activity Relationships (QSPR/QSAR) models obtained for several representative physicochemical properties, biological activities and toxicities of aliphatic alcohols. Most of the properties investigated are well modeled ($r > 0.98$) employing the I_{ET}. Cross-validation using the more general leave-one-out method demonstrates that these models are highly statistically reliable. The proposed I_{ET} index promises to be a useful descriptor in the QSPR/QSAR studies.

Keywords Topological index · QSPR · QSAR · Semi-empirical topological index · Aliphatic alcohols

B. S. Junkes · R. A. Yunes
L. C. Porto · V. E. F. Heinzen (✉)
Departamento de Química,
Universidade Federal de Santa Catarina,
Campus Universitário, Trindade,
88040-900 Florianópolis, SC, Brazil
E-mail: heinzen@qmc.ufsc.br

A. C. S. Arruda
Departamento de Química,
Universidade de Caxias do Sul, Cidade Universitária,
Cinqüentenário,
95070-560 Caxias do Sul, RS, Brazil

Introduction

Structure-Property and Structure-Activity Relationships (QSPR/QSAR) remain the focus of many studies aimed at the modeling and prediction of physicochemical and biological properties of molecules [1–5]. This kind of study is based on the main paradigm of chemistry where the physicochemical and biological properties of organic compounds depend on their molecular structures.

The main contribution to the widespread use of QSPR/QSAR models comes from the development of novel structural descriptors and statistical equations relating various physical, chemical, and biological properties to the chemical structures. The success of the QSPR/QSAR approach can be explained by the insight offered into the structural determination of chemical compounds, and the possibility to estimate the properties without the need to synthesize and test them [6].

Several topological, geometric, electronic, and quantum chemical descriptors have been used in QSPR/QSAR research [7–9]. The topological descriptors have shown their usefulness in the prediction of diverse physicochemical and biological properties of various types of compounds [10–23]. In general, these indices are numbers containing relevant information about the steric structure of molecules. Most of the measured physicochemical properties are steric properties, and consequently they may be reasonably well described by topological indices. However, in some cases, these indices also contain structural information related to the electronic and/or polar features of molecules [24, 25].

The molecular size, shape, polarity, and the ability of the molecule to participate in hydrogen bonding are among the different factors that can contribute to the physicochemical properties or biological activities of a molecule. It is well known that these factors are related to intermolecular interactions such as van der Waals forces and hydrogen-bonding interactions.

In recent years, our research group introduced a new topological descriptor, called the semi-empirical topological index (I_{ET}), to predict the chromatographic retention on low polarity stationary phases for different classes of organic compounds [26–33]. This topological descriptor takes into account the contributions of each individual atom type or group to the property considered. The I_{ET} is able to encode information about structural features such as the presence and position of the heteroatom and the size and branching of the molecules.

In this context, alcohols, which show interesting biological properties, represent an attractive class of organic compounds for QSPR/QSAR studies considering the influence of the hydrogen-bonding interaction. In a recent study [31] the I_{ET} was extended to estimate the chromatographic retention of aliphatic alcohols, yielding high-quality structure-property relationships.

The aim of this study is to demonstrate the applicability of I_{ET} to predict physicochemical properties and especially biological activities that depend on the strength of intermolecular forces.

Methods

Calculation of the semi-empirical topological index

The development of I_{ET} for aliphatic alcohols has been described in a previous paper [31].

In the present approach, the molecules are represented by hydrogen-suppressed molecular graphs based on chemical graph theory, where the carbon atoms and the C–OH group are considered as vertices of the molecular graph of these compounds. The contribution coming from carbon atoms and from the C–OH group to the property considered is represented by a single symbol, C_i, as may be observed from Eq. 1. Thus, the I_{ET} is expressed as:

$$I_{ET} = \sum_i (C_i + \delta_i), \tag{1}$$

where C_i is the value attributed to C–OH fragments and/or to each carbon atom i in the molecule; and δ_i is the sum of the logarithm of the values for each adjacent carbon atom (C_1, C_2, C_3, and C_4) and/or the logarithm of the value of the adjacent C–OH group [31].

The values attributed to the carbon atoms and to the functional groups (C_i) for alcohols were calculated by numerical approximation, based on the experimental data and supported by theoretical considerations. These values (C_i) are listed in Ref. [31].

The molecule 2-methyl-3-hexanol is taken as an example of the calculation of the semi-empirical topological index, using Eq. 1:

Structure 1
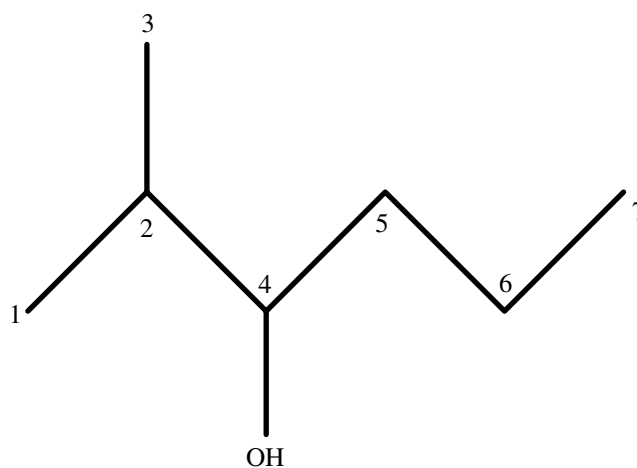
Structure 1

Annexe "a"

$C_1 = C_3 = -CH_3 = 2(1.0 + \log 0.75) = 1.7501$
$C_2 = \, >CH- = 0.75 + \log 1.0 + \log 1.0 + \log 1.78 = 1.0004$
$C_4 = \, >CH-OH = 1.78 + \log 0.75 + \log 0.9 = 1.6093$
$C_5 = -CH_2- = 0.9 + \log 1.78 + \log 0.9 = 1.1047$
$C_6 = -CH_2- = 0.9 + \log 0.9 + \log 1.0 = 0.8542$
$C_7 = -CH_3- = 1.0 + \log 0.9 = 0.9542$
$I_{ET} = 1.7501 + 1.0004 + 1.6093 + 1.1047 + 0.8542 + 0.9542 = 7.2730$

Data set

The physicochemical properties and biological activities of aliphatic alcohols selected in this study were: boiling points, BP (°C); molar volumes, MV (cm^3/mol^1); molar refractions, MR (cm^3/mol^1); molecular total surface area, TSA ($\hat{A}^2$); water solubility, log (1/S), where S is the solubility in mol/L; octanol/water partition coefficient, log P; narcosis activities on barnacle larvae, toxicities towards tomatoes and spiders, pC and odor threshold, log T.

Physicochemical properties

The experimental boiling points of 130 alcohols, used to develop the structure-boiling point model were taken from Ref. [15]. Molar Volumes, molar refractions, and molecular total surface area were obtained from Ref. [16]. The water solubility and the octanol/water partition coefficient, the most frequently used measure of hydrophobicity (or lipophilicity) of organic compounds, were taken from the literature [15]. The interest in these properties is because of the toxic action of these com-

pounds is mainly dependent on their solubility in water. The compounds and the experimental data values for all these physicochemical properties are listed in Table S1 (Supplementary material).

Biological activities and toxicities

In this study, the narcosis activity of alcohols in relation to barnacle larvae and toxicity of alcohol steams towards tomatoes and spiders were selected in order to verify the applicability of I_{ET} to estimate biological activity and toxicity. The data set for the narcosis activity of 14 alcohols are recorded in terms of pC values, where $pC = \log(1/C)$ and C is the molar concentration that elicits a constant biological response [15]. The toxicity of organic compounds is one of the biological activities of particular interest to the scientific community due to its impact on environmental and human health. The toxicities of 14 alcohols towards tomatoes and spiders are taken directly from the literature [15], where the toxicity (pC) is 50% inhibitory growth impairment concentration ($-\log LC_{50}$). The odor-threshold values for a set of 49 alcohols, used in the present investigation, were reported [34] with high and low values in units of μmol/L. The average odor threshold value was selected as the dependent variable. The range of average thresholds spanned five orders of magnitude (0.047–980). Therefore, it was necessary to take the log of the average threshold to avoid leverage problems and artificially good model statistics [34]. The biological activity and toxicity values for each compound are shown in Table S2 (Supplementary material).

Regression analysis

The *Origin* and Bilin [35] program packages were used in the regression analysis. To test the quality of the regression equation, the coefficient of determination (r^2), the coefficient of correlation (r) and the standard deviation (SD) were used as statistical parameters. To verify the validity and stability of the model obtained, a cross-validation test (r^2_{cv}), using the "leave-one-out" method [36, 37] was performed, and a further examination of external stability of the model was carried out by means of a procedure in which the entire data group is systematically divided into three subgroups and each subgroup is predicted by using the other two as the training group [38].

Results and discussion

The good results obtained employing the I_{ET} to predict the chromatographic retention of aliphatic alcohols can be considered as an initial step towards forthcoming QSPR/QSAR studies.

Analyzing the influence of structural features on the chromatographic behavior of the organic compounds, it is possible to verify that the retention mainly depends on the number of carbon atoms, the degree of branching, the presence of heteroatoms and the position of functional groups in the carbon chain. Our topological index was able to reflect these structural factors, and it is therefore expected that physicochemical properties and biological activities, which are related to these factors, will also show good relationships using the I_{ET}.

To illustrate the potential of this index in QSPR/QSAR studies, two series of examples of applications were analyzed. First, several representative properties such as the normal boiling points (BP), water solubility, octanol/water partitioning alcohols with a wide range of nonhydrogen atoms were selected for this case. The other series of examples was related to biological activities and toxicities of alcohols.

Correlations between the properties studied

The correlations between the properties and activities examined are shown in Table 1. As can be seen, most of properties are highly correlated with one another, with the exception of odor threshold of alcohols and toxicities of alcohols towards tomatoes and spiders. The remaining seven properties/activities: BP, log (1/S), log P, TSA, MR, MV and narcosis activity of alcohols in relation to barnacle larvae, (pC) have correlation coefficients greater than 0.95. It can be noted that the high colinearity between most of the properties/activities studied suggests that similar interactions play an important role in the properties/activities examined in this work.

Correlations between I_{ET} and physicochemical properties/biological activities

Physicochemical properties

We studied six structure-property models that had been reported previously in the literature [15, 16]. Boiling points and retention indices are typical "surface"-dependent properties, while molar volumes and molar refractions are "molecular volume"-dependent properties.

The boiling point of a compound is predetermined by the intermolecular interactions in the liquid and by the difference in the molecular internal partition function in the gas phase and in the liquid at the boiling temperature. Therefore, it is expected to be related directly to the chemical structure of the molecule, and indeed numerous methods have been developed over the years for estimating the normal boiling point of a compound from its structure [1]. The molar refraction is related to the bulk and polarizability of a molecule and is also a useful physical parameter in the field of chemical, biological and pharmaceutical sciences. The molecular total

surface area defined as the cavity dimension of the solute when placed in a water media, is a practically valuable property in the estimation of the aqueous solubility of organic compounds. Aqueous solubility and octanol/water partition coefficient are particularly important properties in medicinal chemistry, toxicology and pharmaceutical or environmental science. They are also valuable in understanding drug transport and environmental impact. The partition coefficient for octanol–water has become the preferred measurement for lipophilicity in the development of biologically active molecules, in which transport across biological membranes is often critical. Lipophilicity is a measure of the degree to which a given molecule prefers hydrophobic nonpolar environments to water [39].

The simple linear regressions obtained using the I_{ET} for the six selected properties are of good quality, as can be seen in Table 2. The validity and the stability of the QSPR models were tested in a cross-validation-like procedure, with the computation of r^2_{cv} (Table 2).

To prove the external stability of the QSPR models obtained for the properties (BP, log P, and log (1/S)) we systematically selected from the entire data set of alcohols studied, three different subgroups. Each subgroup was predicted by using the other two subgroups as the training group. The results obtained are listed in Table 3 with an average training quality of r^2 for the three tested properties (BP, log P, and log (1/S)). The cross-validation correlation coefficient (r^2_{cv}) in comparison with the coefficient of determination (r^2) indicates the stability of the QSPR model obtained.

The correlation coefficient (r) and the standard deviation (SD), in general, measure the quality of the QSPR models. Mihalic and Trinajstic [40] suggested that a good QSPR model must have $r > 0.99$ and SD $< 5.0°C$ for BP. In the present study the QSPR model

Table 1 Correlations between the physicochemical properties and biological activities of aliphatic alcohols (correlation coefficients and number of compounds)

	BP	log (1/S)	log P	TSA	MV	MR	pC_a	pC_b	pC_c	log T
BP[a]	1.0000									
log (1/S)[b]	0.9794 (61)	1.0000								
log P[c]	0.9885 (64)	0.9890 (54)	1.0000							
TSA[d]	0.9821 (51)	0.9830 (48)	0.9892 (47)	1.0000						
MV[e]	0.9597 (41)	0.9643 (33)	0.9733 (37)	0.9709 (29)	1.0000					
MR[f]	0.9632 (40)	0.9674 (33)	0.9764 (36)	0.9725 (29)	0.9996 (41)	1.0000				
pC_a[g]	0.9586 (14)	0.9736 (11)	0.9773 (14)	0.8249 (9)	0.8941 (13)	0.9122 (13)	1.0000			
pC_b[h]	0.9278 (14)	0.9609 (11)	0.9627 (14)	0.7908 (9)	0.8907 (13)	0.9008 (13)	0.9699 (14)	1.0000		
pC_c[i]	0.9772 (14)	0.9927 (11)	0.9902 (14)	0.9831 (9)	0.9713 (13)	0.9757 (13)	0.9759 (11)	0.9773 (11)	1.0000	
log T[j]	−0.7841 (52)	−0.7765 (38)	−0.8284 (41)	−0.7798 (30)	−0.8767 (30)	−0.8801 (29)	−0.7590 (11)	−0.8038 (11)	−0.9244 (12)	1.0000

[a] Normal boiling point
[b] Water solubility
[c] Octanol/water partition coefficient
[d] Molecular total surface area
[e] Molar volume
[f] Molar refraction
[g] Toxicities of alcohols towards spiders
[h] Toxicities of alcohols towards tomatoes
[i] Narcosis activities of alcohols on Barnacle Larvae
[j] Average odor threshold

Table 2 Simple linear regressions for the properties/activities using I_{ET} and statistical parameters

Property	Regression equation	Statistical parameters			
		N[k]	r[l]	SD[m]	r^2_{CV}[n]
log P[a]	−2.5941 + 0.6443 I_{ET}	64	0.9937	0.1830	0.987
BP[b]	−7.4137 + 21.6421 I_{ET}	146	0.9908	5.6442	0.981
TSA[c]	67.2002 + 37.565 I_{ET}	52	0.9876	11.827	0.974
log (1/S)[d]	−3.9608 + 0.7158 I_{ET}	63	0.9868	0.171	0.971
MR[e]	−5.5001 + 5.4583 I_{ET}	41	0.9751	2.457	0.944
MV[f]	−6.1876 + 19.4297 I_{ET}	42	0.9712	9.381	0.936
log T[g]	4.7172–0.5456 I_{ET}	49	0.8748	0.490	0.747
pC_a[h]	1.6838 + 0.3696 I_{ET}	14	0.9713	0.0968	0.897
pC_b[i]	1.5331 + 0.3763 I_{ET}	14	0.9626	0.112	0.892
pC_c[j]	−1.8482 + 0.5964 I_{ET}	14	0.9860	0.173	0.961

[a] Octanolwater partition coefficient
[b] Normal boiling point
[c] Molecular total surface area
[d] Water solubility
[e] Molar refraction
[f] Molar volume
[g] Average odor threshold
[h] Toxicities of alcohols towards spiders
[i] Toxicities of alcohols towards tomatoes
[j] Narcosis activities of alcohols on Barnacle Larvaes
[k] Number of compounds
[l] Coefficient of determination
[m] Standard deviation
[n] Cross-validation correlation coefficient

Table 3 Verification of statistical validity of the QSPR/QSAR models

Property/Activity	Training groups	r^2	r^2_{cv}	Predicted groups	r^2
BP[a]	1 and 2	0.981	0.980	3	0.970
	1 and 3	0.977	0.976	2	0.979
	2 and 3	0.974	0.973	1	0.984
	Average	0.977	0.976		0.978
log P[b]	1 and 2	0.987	0.985	3	0.989
	1 and 3	0.986	0.984	2	0.991
	2 and 3	0.990	0.988	1	0.983
	Average	0.988	0.986		0.988
log (1/S)[c]	1 and 2	0.975	0.973	3	0.975
	1 and 3	0.982	0.980	2	0.947
	2 and 3	0.957	0.952	1	0.987
	Average	0.971	0.968		0.970
log T[d]	1 and 2	0.749	0.722	3	0.743
	1 and 3	0.760	0.729	2	0.689
	2 and 3	0.705	0.641	1	0.784
	Average	0.738	0.697		0.739

[a]Normal boiling point
[b]Octanol/water partition coefficient
[c]Water solubility
[d]Average odor threshold

obtained through simple linear regression with BP, log P, and log (1/S) can explain 98.2, 98.7, and 97.4% of the variances in the experimental values (and the predicted variances are 98.1, 98.8, and 97.1%), respectively (Table 2). These results represent satisfactory QSPR models that can be used to predict these properties.

The I_{ET} index was developed to take into account the contributions of each individual atom type or group to the property considered. Thus, the good results indicate the importance of the separate contributions to the physical properties of a molecule.

In our previous paper [31] it was observed that dispersive forces play a more important role in the process of chromatographic retention of aliphatic alcohols than hydrogen-bonding interactions. The same behavior can be attributed to the physicochemical properties studied in this work, such as BP, log P, and log (1/S), where the size and the branching of the molecules are the dominant factors and the hydrogen-bonding interaction formed with the –OH group may be considered as a secondary factor. This fact can be observed by the good results obtained for these properties (Table 2). The semiempirical topological index encodes information on the hydrogen-bonding interaction, which is important to explain some processes, such as aqueous solubility, octanol/water partitioning and other biological processes.

The plots of calculated BP, log P, and log (1/S) versus experimental data for aliphatic alcohols are shown in Fig. 1.

The QSPR model obtained through simple linear regression with MR, MV, and TSA can explain 95.1, 94.3, and 97.5% of the variances in the experimental values (and the predicted variances are 94.4, 93.6, and 97.4%), respectively, as can be observed in Table 2.

Biological activities and toxicities

In this section, we will provide other examples of the application of the new topological index, I_{ET} with the aim to further verify its applicability to predicting biological activities and toxicities.

The models and statistical parameters for predicting toxicities of 14 alcohols towards tomatoes and spiders and for narcosis activities of alcohols in relation to barnacle larvae are on listed in Table 2. In general, the values for correlation coefficients are very good, the models explain 92.7, 94.3, and 97.2% of variances in the experimental values (the predicted variances are 89.2, 89.7, and 96.1%) respectively. The plot of calculated versus observed data for these biological activities are shown in Fig. 2. It is interesting to note that the experimental values for biological activity generally have more uncertainty than those for the physicochemical properties. As is well known, the interaction of –OH groups present in small molecules with biological macromolecules is carried out through the hydrogen bonding formed between them. The strength of hydrogen bonding interactions is mainly controlled by the position and steric hindrance of the –OH group in a molecule. The good results obtained show that the I_{ET} was able to differentiate and characterize the –OH group in terms of C_i values attributed to this fragment [31].
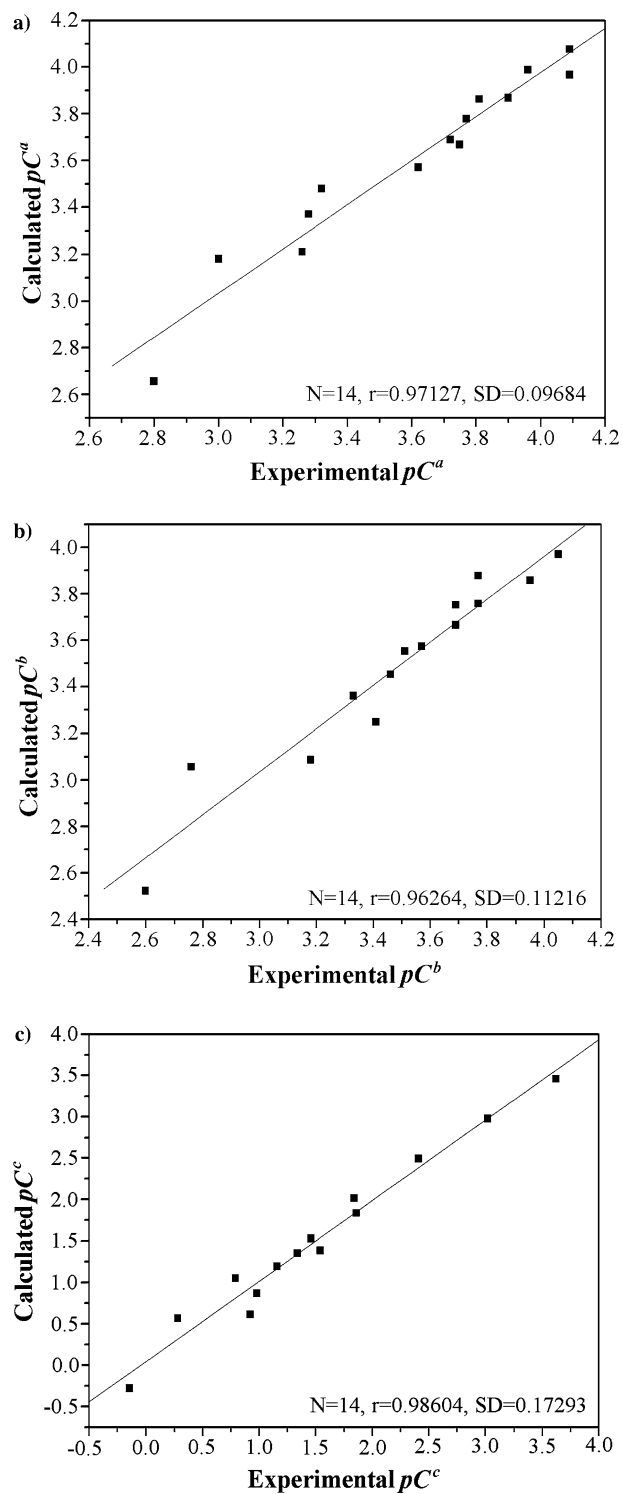
The present study further verifies the high potential of this index for application to predicting biological activities or toxicities within the structural class studied in this work.

The correlation between the I_{ET} and the log of the average odor threshold values for alcohols is shown in Table 2. The model explains more than 76.5% of the variance in the experimental values ($r = 0.8748$, $n = 49$, the predicted variance is 74.7%). The values for the statistical parameters obtained with this simple model are of a similar order and quality to those obtained by Anker et al. [34] using four descriptors in a multiple linear regression ($r = 0.929$, SD $= 0.372$ and $n = 49$). Four compounds were considered as outliers (22, 35, 54, and 55) as can be seen in Table S2. This anomalous behavior would be explained by the conformational or steric effects observed for these compounds that our index was not able to encode.
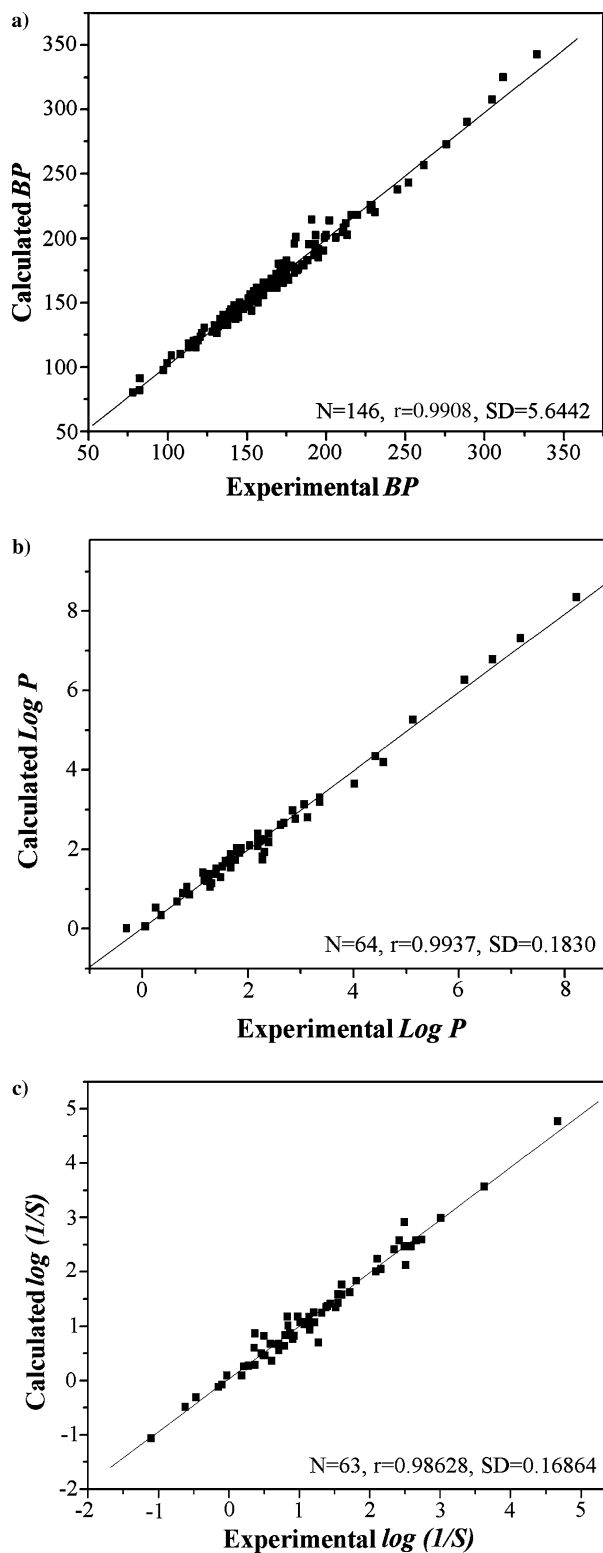
Fig. 1 The plots of the calculated versus experimental properties for aliphatic alcohols: **a** Normal boiling point (BP); **b** Octanol/water partition (log P); **c** Water solubility (log ($1/S$))

Fig. 2 The plots of the calculated versus experimental biological activities for aliphatic alcohols: **a** Toxicities of alcohols on spiders (pC_a); **b** Toxicities of alcohols on tomatoes (pC_b); **c** Narcosis activities of alcohols on Barnacle Larvaes (pC_c)

The good predictive ability of the models obtained is indicated by the cross-validated correlation coefficients, r^2_{cv} and the stability of the QSPR/QSAR models in comparison with determination coefficient, r^2 (Table 2).

For the properties/activities studied, the QSPR/QSAR models using the I_{ET} through a simple linear

regression show statistical qualities similar to those obtained by recent studies employing multiple linear regressions [15, 16] These results indicate again that the I_{ET} is very suitable for modeling the properties/activities of aliphatic alcohols.

Conclusion

This study demonstrates the successful application of the semi-empirical topological index to predict selected physicochemical properties and biological activities of a large group of aliphatic alcohols.

Statistical analysis shows that the application of QSPR/QSAR models has high internal stability, as established by cross-validation (r^2_{cv}). The results indicate that the physicochemical properties of alcohols are dominated by the molecular size showing that the structural features are very important in determining these properties.

This study indicates that the theoretical predictive models for physicochemical and biological properties of alcohols based on I_{ET}, have powerful predictive capability, general applicability and the advantage of the simplicity and the use of only one parameter.

This method may be extended to QSPR/QSAR for other compounds. Further researches along these lines are in progress and will continue.

References

 1. Katritzky AR, Maran U, Lobanov VS, Karelson M (2000) J Chem Inf Comput Sci 40:1–18
 2. Katritzky AR, Petrukhin R, Tatham D (2001) J Chem Inf Comput Sci 41:679–685
 3. Smeeks FC, Jurs PC (1990) Anal Chim Acta 233:111–119
 4. Chenzhong C, Zhiliang L (1998) J Chem Inf Comput Sci 38:1–7
 5. Romanelli GP, Cafferata LFR, Castro EA (2000) J Mol Struc (Theochem) 504:261–265
 6. Ivanciuc O, Ivanciuc T, Balaban AT (2002) Internet Electron J Mol Des 1:252–268 http://www.biochempress.com
 7. Katritzky AR, Gordeeva EV (1993) J Chem Inf Comput Sci 33:835–857
 8. Karelson M, Lobanov VS, Katritzky AR (1996) Chem Rev 96:1027–1044
 9. Kier LB, Hall LH (1990) Pharm Res 7:801–807
10. Sabljic A, Trinajstic N (1981) Acta Pharm Jugosl 31:189–214
11. Randic M (2001) J Mol Graph Modell 20:19–35
12. Estrada E (2001) Curr Med Chem 8:1573–1588
13. Randic M (1993) Croat Chem Acta (CCACAA) 66:289–312
14. Estrada E (2001) Chem Phys Lett 336:248–252
15. Ren B (2002) J Chem Inf Comput Sci 42:858–868
16. Ren B (2002) Comput Chem 26:223–235
17. Ren B (2002) J Mol Struc (Theochem) 586:137–148
18. Heinzen VEF, Yunes RA (1993) J Chromatogr A 654:83–89
19. Arruda AC, Heinzen VEF, Yunes RA (1993) J Chromatogr A 630:251–256
20. Heinzen VEF, Yunes RA (1996) J Chromatogr A 719:462–467
21. Heinzen VEF, Cechinel Filho V, Yunes RA (1999) IL Farmaco 54:125–129
22. Amboni DM, Junkes BS, Yunes RA, Heinzen VEF (2000) J Agric Food Chem 48:3517–3521
23. Kier LB, Hall LH (1986) Molecular connectivity in structure-activity studies. Research Studies Press, Letchworth
24. Hall LH, Mohney B, Kier LB (1991) J Chem Inf Comput Sci 31:76–82
25. Galvez J, Garcia R, Salabert MT, Soler R (1994) J Chem Inf Comp Sci 34:520–525
26. Heinzen VEF, Soares MF, Yunes RA (1999) J Chromatogr A 849:495–506
27. Junkes BS, Amboni RDM, Heinzen VEF, Yunes RA (2002) Chromatographia 55:75–81
28. Junkes BS, Amboni RDM, Yunes RA, Heinzen VEF (2002) Chromatographia 55:707–714
29. Amboni RDM, Junkes BS, Yunes RA, Heinzen VEF (2002) J Mol Struc (Theochem) 579:53–62
30. Amboni RDM, Junkes BS, Yunes RA, Heinzen VEF (2002) J Mol Struc (Theochem) 586:71–80
31. Junkes BS, Amboni RDM, Yunes RA, Heinzen VEF (2003) Anal Chim Acta 477:29–39
32. Junkes BS, Amboni RDM, Yunes RA, Heinzen VEF (2003) Internet Electron J Mol Des 2:33–49 http://www.biochem-press.com
33. Junkes BS, Amboni RDM, Yunes RA, Heinzen VEF (2004) J Braz Chem Soc 15:183–189
34. Anker LS, Jurs PC, Edwards PA (1990) Anal Chem 62:2676–2684
35. Kubiny H (1993) QSAR: hansch analysis and related approaches. In: Mannhold R, Gaard-Larsen PK, Timmerman H (eds) VCH, Weinheim
36. Muresan S, Bologa C, Mracec M, Chiriac B, Jastorff B, Simon Z, Náray-Szabó G (1995) J Mol Struc (Theochem) 342:161–171
37. van de Waterbeend H (1995) Chemometric methods in molecular design. VCH, Weinheim, Chapter 5
38. Katritzky AR, Chen U, Maran K, Carlson DA (2000) Anal Chem 72:101–109
39. Peruzzo PJ, Marino DJG, Castro EA, Toropov AA (2003) Internet Electron J Mol Des 2:334–347 http://www.biochem-press.com
40. Mihalic Z, Trinajstic NA (1992) J Chem Educ 69:701–712