TOPS-MODE *versus* DRAGON descriptors to predict permeability coefficients through low-density polyethylene

Maykel Pérez González[a,b,]* & Aliuska Morales Helguera[c]
[a]*Unit of Services, Experimental Sugar Cane Station 'Villa Clara-Cienfuegos', Ranchuelo, Cuba;* [b]*Department of Drug Design, Chemical Bioactives Center, Central University of Las Villas, Santa Clara, Villa Clara, Cuba;* [c]*Department of Chemistry, Faculty of Chemistry and Pharmacy, Central University of Las Villas, Santa Clara, Villa Clara, Cuba*

Summary

The TOPological Sub-Structural MOlecular DEsign (TOPS-MODE) approach has been applied to the study of the permeability coefficient of various compounds through low-density polyethylene at 0 °C. A model able to describe more than 92% of the variance in the experimental permeability of 38 organic compounds was developed with the use of the mentioned approach. In contrast, none of eight different approaches, including the use of constitutional, topological, BCUT, 2D autocorrelations, geometrical, RDF, 3D Morse, and GETAWAY descriptors was able to explain more than 75% of the variance in the mentioned property with the same number of descriptors. In addition, the TOPS-MODE approach permitted to find the contribution of different fragments to the permeability coefficients, giving to the model a straightforward structural interpretability.

Introduction

The penetration of molecules through polymer films is called 'permeability'. In the literature, many dimensions and units can be found for the general expression 'permeability' [1].

The permeability coefficient, in a strict sense, is not only a function of the chemical structure of the polymer. It varies also with the morphology of the polymer and depends on many physical properties such as density, crystallinity and orientation of the polymer chains. However, the chemical structure of a polymer and the permeant can be considered to be the predominant factor that controls the magnitude of the permeability coefficient [2].

If a permeant does not interact with the polymer under investigation, then the permeability coefficient is characteristic for the permeant–polymer system. This is the case with the permeation of many gases, such as H_2, He, N_2, O_2 and CO_2, through many polymers. On the other hand, if a permeant interacts with the polymer, the permeation coefficient is no longer a constant, and this process depends on the special conditions of the measurement and on the history of the polymer film. In such cases, a single value of the permeability coefficient does not represent the characteristic permeability of the polymer and it is necessary to know the dependency of the permeability coefficient of all possible variables in order to obtain the complete profile of the permeability of the polymer [3].

Various synthetic membranes have been employed in drug release studies [4–6]. The most commonly used artificial membranes are polydimethylsiloxane (PDMS) and cellulose acetate [7–12].

PDMS (for example, Silastic) is an isotropic polymer widely employed as an alternative model barrier for *in vitro* percutaneous penetration. It behaves according to Fick's first law of diffusion and possesses lipid-like properties, making it a good model for the stratum corneum [4, 13, 14].

---

*To whom correspondence should be addressed. E-mail: mpgonzalez76@yahoo.es

Cellulose acetate membranes have similarly found use in such experiments and also in the characterization of iontophoretic delivery [15–19]. However, these membranes have often been shown to significantly overestimate the flux across skin and their use is rather limited. Further, Cronin et al. [20], in a mechanistic study of penetration across a PDMS membrane, indicated that penetration is related primarily to the ability of the penetrants to form hydrogen bonds and not to their lipophilicity, as suggested by similar studies on skin *ex vivo*.

Early quantitative structure–activity relationship (QSAR) studies to predict skin permeation of chemicals revealed that hydrophobicity was correlated linearly with increasing permeability [21, 22]. Patel et al. [23] demonstrated in an excellent paper how the hydrophobicity, molecular size and the hydrogen bonding capability of a molecule affect its ability to permeate skin.

In the context of *in silico* methods for modeling physicochemical and biological properties of chemicals the topological sub-structural molecular design (TOPS-MODE) approach has been introduced [24–28].

The successful applications of this theoretical approach to the modeling of physical and physical–chemical properties [27, 29] have inspired us to perform a more exhaustive study in order to test and/or validate the TOPS-MODE applicability in this area.

Since low density polyethylene is not an alternative to other polymers such as cellulose acetate and PDMS to study skin permeation, no global conclusions can be reached, by extrapolation, about the behavior of the permeants in skin. Nevertheless, we thought it might prove interesting to demonstrate the capability of TOPS-MODE descriptors to model this property in this type of membrane.

The permeability through low-density polyethylene was previously studied using a set of physicochemical properties as molecular descriptors for a limited data set of compounds [3].

Therefore, the aim of this study was to investigate the role that TOPS-MODE and other molecular descriptors calculated from the molecular structure play in the explanation of such a property using an extended data set of 38 organic compounds.

The TOPS-MODE approach

TOPS-MODE is based on the computation of the spectral moments of the bond matrix, the mathematical basis of which has been described previously [24–26]. The TOPS-MODE approach has been recently reviewed in the literature [30], and both the methodology and its software implementation have been described [31].

According to the authors, the application of the TOPS-MODE approach to the study of quantitative structure–permeability relationships (QSPR) can be summarized in the following steps:

(1) To draw the hydrogen-depleted molecular graphs for each molecule of the data set,

(2) To use appropriate bond weights in order to differentiate the molecular bonds, e.g., bond distance, bond dipoles, bond polarizabilities, etc.,

(3) To compute the spectral moments of the bond matrix with the appropriate weights for each molecule in the data set, generating a table in which rows correspond to the compounds and columns correspond to the spectral moments of the bond matrix. Spectral moments are defined as the trace of the different powers of the bond matrix [32],

(4) To find QSPR by using a suitable linear or nonlinear multivariate statistical technique, such as multilinear regression analysis (MRA), etc. to obtain an equation of the form:

$$P = a_0\mu_0 + a_1\mu_1 + a_2\mu_2 + a_3\mu_3 \ldots a_k\mu_k + b, \quad (1)$$

where P is the property measurement, μ_k is the kth spectral moment, and a_k's are the coefficients obtained by the MRA,

(5) To test the predictive capability of the QSPR model by using cross-validation techniques.

(6) To compute the contributions of different groups of interest in order to determine their quantitative contribution to the permeability of the molecules under study.

The computation of fragment contributions to the permeability property being studied is probably the most important advance of the TOPS-MODE approach when compared to other traditional QSAR and QSPR methods. The procedure consists of calculating the spectral moment for all the fragments contained in a given substructure, and by subtraction of these spectral moments to obtain the contribution of the substructure. The general algorithm for this computational approach is as follows:

First, we select the substructure whose contribution

Table 1. Observed, predicted, and residual values of permeability coefficients (cm/h) through low-density polyethylene for the 38 compounds used to derive the QSPR [2].

Number	Compound	Observed	Predicted	Deleted residuals
1	Acetaldehyde	0.127	0.178	−0.051
2	Acetic acid	−0.853	−0.926	0.072
3	Acetic anhydride	−1.292	−0.587	−0.704
4	Acetone	−0.259	0.062	−0.322
5	Allyl alcohol	−1.200	−1.237	0.037
6	Amyl acetate	−0.657	−0.418	−0.239
7	Aniline	−1.008	−0.641	−0.367
8	Benzaldehyde	−0.823	−0.201	−0.622
9	Benzene	1.247	1.288	−0.040
10	n-Butyl alcohol	−1.397	−1.386	−0.010
11	s-Butyl alcohol	−1.301	−1.386	0.085
12	t-Butyl alcohol	−1.698	−1.385	−0.313
13	Butyraldehyde	−0.455	−0.053	−0.402
14	n-Butyric acid	−1.124	−1.098	−0.026
15	Carbon tetrachloride	1.301	1.494	−0.193
16	Chlorobenzene	1.356	1.238	0.117
17	p-Chloro toluene	1.071	1.109	−0.037
18	Cyclohexane	1.093	1.056	0.036
19	Decane	0.568	0.536	0.031
20	Dibutyl ether	0.643	0.562	0.080
21	Diethyl ether	1.276	1.049	0.226
22	Ethyl acetate	−0.124	−0.090	−0.034
23	Ethyl mercaptan	1.008	1.131	−0.122
24	Formic acid	−1.000	−0.841	−0.158
25	n-Heptane	1.281	0.922	0.358
26	n-Hexane	1.276	1.050	0.225
27	Methyl alcohol	−1.000	−1.091	0.091
28	Methyl ethyl ketone	0.161	−0.053	0.214
29	Nitrobenzene	−0.823	−1.358	0.534
30	Nitroethane	−0.420	−1.112	0.692
31	Octyl alcohol	−1.408	−1.782	0.373
32	i-Pentane	1.276	1.179	0.097
33	Phenol	−1.397	−1.435	0.037
34	n-Propyl alcohol	−1.522	−1.288	−0.234
35	i-Propylamine	−0.045	−0.492	0.447
36	Tetradecane	−0.173	0.022	−0.196
37	Toluene	1.356	1.159	0.196
38	o-Xylene	1.152	1.030	0.121

to the moments we would like to determine. Then, we generate all the fragments, which are contained in the corresponding substructure, and calculate the spectral moments for both, the substructure and all their fragments. The contribution of the substructure to the spectral moments is finally obtained as the difference between the spectral moments of the substructure and all those from their fragments. Once the contribu-

tions of the different structural fragments are obtained, we only need to substitute these contributions into the quantitative model developed to describe the property studied.

Table 2. Statistical parameters of the lineal regressions models obtained for the nine kinds of descriptors.

Descriptor	Variables[a]	S	R^2	F	P	q^2
Spectral moments	μ_2SP, $\mu_1 \cdot \mu_2$SP, $\mu_1 \cdot \mu_5$SP, μ_1VDW	0.306	0.926	102.504	0.000	0.897
Constitutional	Mv, Ms, nN, nO	0.800	0.673	9.077	0.000	0.641
Topological	BAC, IVDE, IC1, SEigp	0.686	0.649	15.300	0.000	0.614
BCUT	BEHm1, BELm5, BEHv7, BELe4	0.891	0.408	5.701	0.000	0.379
2D autocorrelations	ATS1p, MATS3e, MATS4e, GATS2e	0.716	0.618	13.360	0.000	0.582
Geometrical	W3D, J3D, MAXDP, FDI	0.661	0.670	17.106	0.000	0.664
RDF	RDF035m, RDF045m, RDF100m, RDF030p	0.991	0.268	3.023	0.031	0.222
3D-MORSE	Mor10u, Mor15m, Mor25p, Mor11e	0.604	0.729	22.148	0.000	0.701
GETAWAY	ISH, H4m, HATS1e, HATS1p	0.717	0.617	13.284	0.000	0.567

[a]The definition of the terms appears largely explained in reference 39.

Data sets and computational strategies

A data set of 38 compounds for which the permeability coefficients were reported in the literature was selected [2]. The parameter studied is log(p), where p is the permeability coefficient through low-density polyethylene. The names of the compounds, as well as the calculated and experimental values of log(p) are shown in Table 1.

TOPS-MODE [31] and DRAGON [33] computer software were employed to calculate the molecular descriptors. In the case of TOPS-MODE software, only the polar surface and van der Waals radii were used to weigh the bond adjacency matrix. The selection of only these two types of descriptors from the whole pool of six types included in the TOPS-MODE methodology was carried out for the sake of simplicity and on the belief that steric and polarity parameters influence the permeability of compounds through polymer layers. Besides the traditionally used linear TOPS-MODE descriptors we also use here their multiplication as independent variables [34–36]. In this case we only multiplied μ_1 for the first five spectral moments, obtaining six new variables. The total number of descriptors used for obtaining this model was 42 (30 spectral moments + 12 multiplications of moments). On the other hand, we carry out geometry optimization calculations for each compound used in this study using the quantum chemical semiempirical method AM1 [37] included in MOPAC 6.0 [38]. Eight other models were developed using the computer software Dragon [33], and calculating the Constitutional, Topological, BCUT, 2D autocorrelations, Geometrical, RDF, 3D-MORSE and GETAWAY descriptors [39]. The statistical processing to obtain the QSAR models was carried out by using the forward stepwise regression methods.

The statistical significance of the models was determined by examining the regression coefficient, the standard deviation, the number of variables, the cross-validation leave-one-out statistics and the proportion between the cases and variables in the equation.

Quantitative structure – permeation relations

The best QSPR model obtained with the TOPS-MODE descriptors is given below together with the statistical parameters of the regression.

$$\log(p) = 1.828 - 0.006 \cdot \mu_2^{PS} + 6.13 \cdot 10^{-5} \cdot$$

$$\mu_1\mu_2^{PS} \quad + 8.90 \cdot 10^{-1} \cdot \mu_1\mu_5^{PS} - 0.03 \cdot \mu_1^{VDW}$$

$n = 38$, S $= 0.306$, $R^2 = 0.926$, F $= 102.504$, $P = 0.000$, $q^2 = 0.897$, $S_{cv} = 0.341$, where n is the number of compounds included in the model, R^2 is the correlation coefficient, S the standard deviation of the regression, F the Fisher ratio, q^2 the correlation coefficient of the cross-validation, P is the significance of the variables in the model and S_{cv} is the standard deviation of the cross-validation.

The variables included in the model are designated as follows: the sub-index represents the order of the spectral moment and the super-index the type of bond weight used, i.e., PS for polar surface and VDW for van der Waals areas.

The structural significance of this model will be more evident later when we analyze the contribution of the different structural fragments to the permeability through the polyethylene layer. From the statistical point of view this model is a robust one, as can be seen from the statistical parameters of the cross-validation.
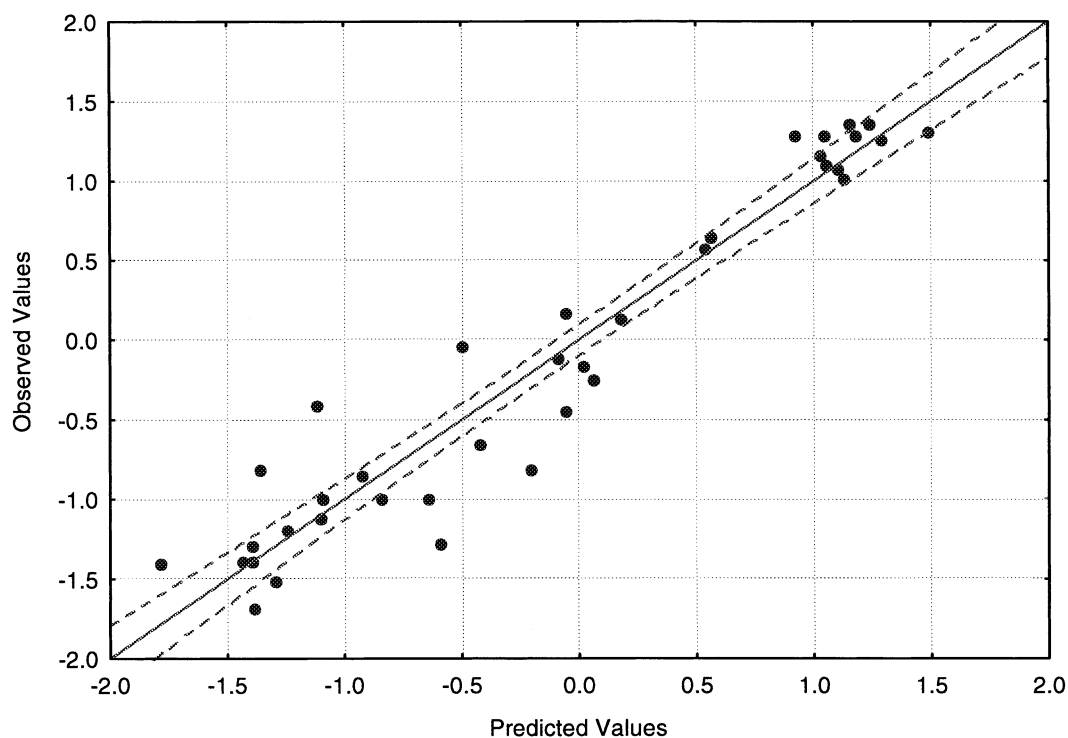
Figure 1. The linear relation between observed and predicted permeability for the compounds of the training set.

As we previously mentioned, one of the objectives of the current work was to compare the reliability of the TOPS-MODE approach to describe the property under study as compared to other different descriptors and methods. Consequently, eight other models were developed using the same data set and the same number of variables that were included in the TOPS-MODE QSPR model. The results obtained with the use of constitutional, topological, BCUT, 2D autocorrelations, geometrical, RDF, 3D Morse, and GETAWAY descriptors are given in Table 2.

As can be seen, there are remarkable differences concerning the explanation of the experimental variance given by these models compared to the TOPS-MODE one. While the TOPS-MODE QSPR model explains more than 92% of permeability, the rest of the models are unable to explain beyond 75% of such variance.

The TOPS-MODE model is superior to the other eight models not only in the statistical parameters of the regression but also, and more importantly, in its stability upon inclusion/exclusion of compounds as measured by the correlation coefficient and standard deviation of the cross-validation. Because of the structural variability of the compounds in the data set these

statistics of the leave-one-out cross-validation might be considered as a good measurement of the predictive ability of the models. As can be seen in Table 2, the value of the determination coefficient of leave-one-out cross-validation for the model obtained with the spectral moments ($q^2 = 0.897$) was the highest of all.

Fragment contributions

One of the most important advantages that TOPS-MODE brings to the study of QSPR and QSAR is concerned with the structural interpretability of the models. This interpretability comes from the fact that the spectral moments can be expressed as linear combinations of structural fragments. In such a way, we can learn what fragments are making a positive or negative contribution to the property under study, which can be interpreted in terms of the physicochemical or biological processes influencing it. In Table 3 and Figure 2 we show the fragments and their contributions to the permeation coefficient across the low-density polyethylene as calculated from Equation 2.

Here we have studied only some small fragments that are present in the structures of the compounds in

Table 3. Contributions of different groups to the permeability coefficients through low-density polyethylene.

Studied fragments	Group contribution	Studied fragments	Group contribution
F1	0.74	F11	0.03
F2	0.55	F12	0.001
F3	0.41	F13	0.002
F4	0.30	F14	0.49
F5	0.21	F15	0.43
F6	0.41	F16	0.66
F7	0.43	F17	0.59
F8	0.44	F18	0.53
F9	0.01	F19	0.023
F10	0.002	F20	0.017
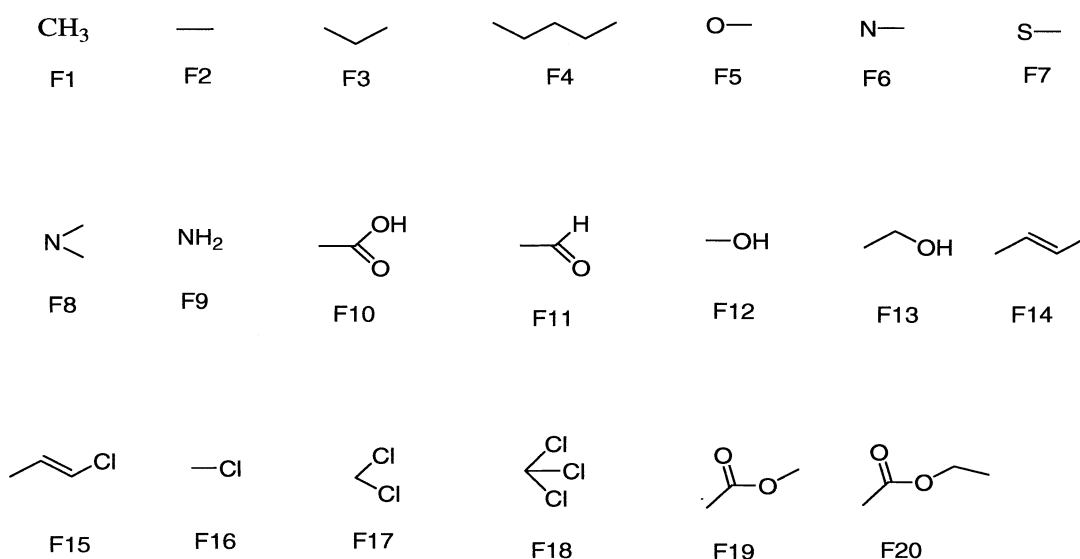
Figure 2. Structures of selected fragments and their contributions to the permeability coefficient.

the data set. However, extension of this study to other fragments in such molecules or even to fragments in molecules not contained in this data set should be straightforward, as shown for other particular cases elsewhere [8, 10, 11, 19].

According to the contributions of fragments F1 to F4 (see Figure 2), the permeation coefficient decreases as the number of carbons increases in a chain. This is a logical size effect involving the length of the chains in compounds as a factor limiting the permeability through polyethylene layers. A similar trend is observed when we analyze fragments F16 to F18. In this case, when the number of halogens in the fragment increases the permeation coefficient decreases. Therefore, the result of this study indicates that an increase

of the molecular size decreases the permeation across this polymer. Similar results were found by Potts et al. [40], Barratt [41] and Cronin et al. [14] where the percutaneous absorption across excised human skin *in vitro* was shown to be governed by molecular size.

On the other hand, it is observed from the contributions of fragments F5 to F7 that an increase of the polarity of the heteroatom produces a decrease in permeation. The polarity of the atoms produces a higher interaction of the permeant with the polymer and therefore an increase of the hydrophobicity leads to a higher flux across the polymer. This behavior was reported by Moss et al. in an excellent review [42] where the main role of hydrophobicity in accounting for this property was explained.

However, the contributions of the heteroatom are also dependent on its volume as can be observed from the smaller contribution difference between fragments F6 and F7. The atomic volume of sulfur is larger than that of nitrogen, but the polarity of the latter atom is higher than that of the former and thus the result of this effect is a delay in the permeation process.

Finally, when fragments F10 and F19 are compared, it can be observed that F10 delays approximately ten times more the process of permeation than F19 does, despite the fact that the former is smaller than the latter. This should be because the methyl group decreases the possibility of the oxygen atom in F10 to form a hydrogen bond with the polymeric matrix. Abraham et al. [43] and Patel et al. [23] pointed out that the hydrogen bonding capability of a molecule affects its ability to permeate the skin. Similar results were obtained by Lipinski et al. [44] where the hydrogen bond acceptors sites could potentially hamper skin permeation.

This study demonstrated that the passage of chemicals throughout a low-density polyethylene membrane is controlled by the molecular volume and the polarity. The non-correspondence among this polymer, the PDMS and the skin, justify the different mechanisms that control the passage of chemicals through these matrixes.

Concluding remarks

We have shown that the TOPS-MODE approach is able to describe the permeability of different compounds through low-density polyethylene. In fact, we have developed a model for predicting the permeability coefficient of a data set of 38 permeants, which is both statistically and chemically sound. This model explains more than 92% of the variance in the experimental permeability coefficients and shows good predictive ability in cross-validation. These features are significantly better than those obtained for eight other different methodologies used to predict this property. Therefore, the spectral moments show a better performance than other kinds of descriptors, which suggests that they can be used in new QSPR applications.

On the other hand, the main advantage of using a TOPS-MODE approach in QSPR/QSAR has been confirmed again in this work. This approach is able to derive group contributions and simultaneously provides the means of interpreting them, thus contributing to our understanding of the physicochemical or biological processes involved.

Finally, the present results were compared to others obtained in previous works and evidence was obtained on the similarity of the properties that explain the phenomenon. Carrying out a complete comparison is not feasible, however, due to the fact that membrane and polymer permeability data are different in each work. Nevertheless, we hope to develop additional models of these data in the future with a view to clarifying this phenomenon.

Acknowledgements

References

1. Minoura, N., Tani, S. and Nakagawa, T., J. Polym. Sci., 22 (1978) 833.
2. Brandrup, J. and Immergut, E.H., Polymer Handbook (3rd. ed.). Wiley & Sons, New York, 1989, p. V122.
3. Patel, H., Tokarski J. and Hopfinger A., Pharm. Res., 10 (1997) 1349.
4. Garret, E.H. and Chemburkar, P.B., J. Pharm. Sci., 57 (1968) 949.
5. Friend, D.R., J. Contr. Rel., 18 (1992) 235.
6. Woolfson, A.D. and McCafferty, D.F., Int. J. Pharm., 94 (1993) 75.
7. Kurosaki, Y., Nagahara, N., Taniza, W.A., Nishimura, H., Nakayama, T. and Kimura, T., Int. J. Pharm., 67 (1991) 1.
8. Megrab, N.A., Williams, A.C. and Barry, B.W., J. Contr. Rel., 36 (1995) 277.
9. Megrab, N.A., Williams, A.C. and Barry, B.W., Int. J. Pharm., 116 (1995) 101.
10. Stott, P.W., Williams, A.C. and Barry, B.W., J. Contr. Rel., 41 (1996) 215.
11. van Hal, D., van Rensen, A., de Vringer, T., Junginger, H. and Bouwstra, J., STP Pharma Sci., 6 (1996) 72.
12. Minghetti, P., Casiraghi, A., Cilurzo, F., Montanari, L., Marazzi, M., Falcone, L. and Donati, V., J. Pharm. Pharmacol., 51 (1999) 673.
13. McCafferty, D.F., Woolfson, A.D., McClelland, K.H. and Boston, V., Br. J. Anaesth., 60 (1988) 64.
14. Cronin, M.T.D., Dearden, J.C., Moss, G.P. and Murray-Dickson, G., Eur. J. Pharm. Sci., 7 (1999) 325.
15. Cichon, R. and Janicki, S., Pharmazie, 46 (1991) 719.
16. Yuk, S.H., Lee, S.J., Okano, T., Berner, B. and Kim, S.W., Int. J. Pharm., 77 (1991) 221.
17. Foley, D., Corish, J., and Corrigan, O.I., Solid State Ionics, 53 (1992) 184.

672

18. Bayon, A.M.R., Corish, J. and Corrigan, O.I., Drug Dev. Industr. Pharmacy, 19 (1993) 1169.

19. Ramis, J., Conte, L., Segado, X., Forn, J., Lauroba, J., Calpena, A., Escribano, E. and Domenech, J., Arzneimittel-Forschung/Drug Res., 47 (1997) 1139.

20. Cronin, M.T.D., Dearden, J.C., Gupta, R. and Moss, G.P., J. Pharm. Pharmacol., 50 (1998) 143.

21. Scheuplein, R.J. and Blank, I.H., Phys. Rev., 51 (1971) 702.

22. Roberts, M.S., Anderson, R.A. and Swarbrick, J., J. Pharm. Pharmacol., 29 (1977) 677.

23. Patel, H., ten Berge, W. and Cronin, M.T.D., Chemosphere, 48 (2002) 603.

24. Estrada, E., J. Chem. Inf. Comput. Sci., 36 (1996) 844.

25. Estrada, E., J. Chem. Inf. Comput. Sci., 37 (1997) 320.

26. Estrada, E., J. Chem. Inf. Comput. Sci., 38 (1998) 23.

27. Estrada, E., J. Chem. Soc. Faraday Trans., 94 (1998) 1407.

28. Estrada, E., Gutiérrez, Y. and González, H., J. Chem. Inf. Comput. Sci., 40 (2000) 1386.

29. Estrada, E. and Gonzalez, H., J. Chem. Inf. Comput. Sci., 42 (2003) 75.

30. Estrada, E., SAR & QSAR Environ. Res., 11 (2000) 55.

31. Gutierrez, Y. and Estrada, E., TOSS-MODE (Topological Sub-Structural Molecular Design) for Windows Version 4.0. Universidad de Santiago de Compostela, Spain, 1997.

32. Estrada, E., J. Chem. Inf. Comput. Sci., 35 (1995) 31.

33. Todeschini, R., Consonni, V. and Pavan, M., Dragon Software version 2.1, 2002.

34. González M.P., Gonzalez H., Molina R., Cabrera M.A., and Ramos R., J. Chem. Inf. Comput. Sci., 43 (2003) 1192.

35. Estrada, E. and Peña, A., Bioorg. Med. Chem., 8 (2000) 2755.

36. Estrada, E. and Uriarte E., SAR and QSAR Environ. Res., 12 (2001) 309.

37. Michael, J.S., Dewar, E., Zoebisch, G., Eamonn, F. and Stewart, J.P., J. Am. Chem. Soc., 107 (1985) 3902.

38. Stewar, J.J.P., MOPAC manual, 6th ed. Frank J. Seiler Research Laboratory, U.S. Air Force Academy, Colorado Springs, CO, 1990, p. 1890.

39. Todeschini R. and Consonni V., Handbook of Molecular Descriptors. Wiley-VCH, Weinheim, Germany, 2000.

40. Potts, R.O. and Guy, R.H., Pharm. Res., 9 (1992) 663.

41. Barratt, M.D., Toxicol. In Vitro, 9 (1995) 27.

42. Moss, G.P., Dearden, J.C., Patel, H. and Cronin, M.T.D., Toxicol. In Vitro, 16 (2002) 299.

43. Abraham, M.H., Chadha, H.S. and Mitchell, R.C., J. Pharm. Pharmacol., 47 (1995) 8.

44. Lipinski, C.A., Lombardo, F., Dominy, B.W. and Feeney, P.J., Adv. Drug Deliv. Rev., 23 (1997) 3.

45. Kowalski R.B. and Wold S., Pattern recognition in chemistry In: Handbook of Statistics (Eds: Krishnaiah, P.R. and Kanal, L.N.) North Holland Publishing Company, Amsterdam, 1982, pp. 673–697.