TOPS-MODE Versus DRAGON Descriptors in QSAR. 1. Skin Permeation.

Maykel Pérez González [1, 2, *], Aliuska Morales Helguera [3, 2].

[1] Unit of Services, Experimental Sugar Cane Station "Villa Clara-Cienfuegos", Ranchuelo, Cuba.

[2] Department of Drug Design, Chemical Bioactives Center, Central University of Las Villas, Santa Clara, Villa Clara. Cuba.

[3] Department of Chemistry, Faculty of Chemistry and Pharmacy, Central University of Las Villas, Santa Clara, Villa Clara, Cuba.

Abstract

The TOPological Sub-Structural MOlecular DEsign (TOPS-MODE) approach has been applied to the study of the permeability coefficient of various compounds through human skin. A model able to describe more than 93% of the variance in the experimental permeability of 37 organic compounds was developed with the use of the mentioned approach. In contrast, none of nine different approaches, including the use of constitutional, topological, BCUT, 2D autocorrelations, geometrical, RDF, 3D Morse, GETAWAY and WHIM descriptors was able to explain more than 90% of the variance in the mentioned property with the same number of descriptors. In addition the TOPS MODE allows a simple interpretation of the model in comparison with others methodologies.

Keywords: **Molecular descriptors, Permeability coefficients, QSPR, TOPS-MODE.**

***Corresponding Author:** mpgonzalez76@yahoo.es
Tel: 53-42-281473; Fax: 53-42-281130

1 INTRODUCTION

The barrier function of human skin is important both to the transdermal administration of drugs and to the uptake of toxic chemicals following dermal exposure. As a result, several models to predict molecular transport through human skin have been developed [1-3].

Various synthetic membranes have been employed in drug release studies. The most commonly used artificial membranes are polydimethylsiloxane (PDMS) and cellulose acetate [4- 10].

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 [11].

Cellulose acetate membranes have similarly found use in such experiments and also in the characterization of iontophoretic delivery [12-16]. However, these membranes have often been shown to overestimate significantly the flux across skin and their use is significantly limited. Further, Cronin et al. [17], 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 [18, 19]. Patel et al. [20] demonstrated in an excellent paper as

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 [21-25].

The successful applications of this theoretical approach to the modeling of physical and physical-chemical properties [26, 27] have inspired us to perform a more exhaustive study in order to test and/or validate the TOPS MODE applicability in this area.

Therefore, the aim of this study was to investigate the role that TOPS-MODE and other molecular descriptors calculated from the molecular structure plays on the explanation of such property using a data set of 37 organic compounds.

2 MATERIALS AND METHODS

2.1 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 [21 - 24]. The TOPS-MODE approach has been recently reviewed in the literature [28], and both the methodology and its software implementation have been described [29].

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 [30],
4. To find QSPR by using a suitable linear or non-linear multivariate statistical technique, such as multi-linear regression analysis (MRA), etc. to obtain an equation of the form:

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

   where P is the property measurement, μ_k is the *k*th 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.

2.2 Data Sets and Computational Strategies.

A data set of 37 compounds for which the permeability coefficients were reported in the literature was selected [31]. The parameter studied is log(p) where p is the permeability

coefficient through human epidermis. The names of the compounds, as well as the calculated and experimental values of log(p) are shown in Table 1.

Table 1 comes about here

TOPS-MODE [29] and DRAGON [32] computer softwares were employed to calculate the molecular descriptors. In the case of TOPS-MODE software, the polar surface, the dipole moment, the Gasteiger-Marsili charges and hidrophobicity were used to weigh the bond adjacency matrix. The selection of only these four types of descriptors from the whole pool of ten types included in 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 skin layers. The total number of descriptors used for obtaining this model was 64 spectral 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 [33] included in MOPAC 6.0 [34]. Nine other models were developed using the computer software Dragon [32], and calculating the Constitutional, Topological, BCUT, 2D autocorrelations, Geometrical, RDF, 3D-MORSE, GETAWAY and WHIM descriptors [35]. 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.

3 RESULTS AND DISCUSSION

3.1 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.

$$\log(p) = -5.87 - 1.12 \cdot 10^{-9} \cdot \mu_{15}{}^{D} + 1.11 \cdot 10^{-7} \cdot \mu_{15}{}^{GM} + 0.76 \cdot \mu_{1}{}^{H} - 0.02 \cdot \mu_{1}{}^{PS} \quad \textbf{(Eq. 2)}$$

$N = 37 \quad S = 0.24 \quad R^2 = 0.938 \quad F = 121.06 \quad p = 0.000 \quad q^2 = 0.907 \quad S_{cv} = 0.351$

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., D for dipole moment, PS for polar surface and H for hifrophobicity.

From the statistical point of view this model is a robust one as can be seen from the statistical parameters of the cross-validation.

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, 9 other models were developed using the same data set and the same number of variables that was included in the TOPS-MODE QSPR model. The results obtained with the use of constitutional,

topological, BCUT, 2D autocorrelations, geometrical, RDF, 3D Morse, GETAWAY and WHIM descriptors are given in Table 2.

Table 2 comes about here.

3.2 Comparison with other Approaches

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 93% of permeability the rest of the models are unable to explain beyond 90% of such variance.

The TOPS-MODE model is superior to the other nine 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 ($q^2 = 0.907$) was the highest of all.

3.3 Interpretation of the Model

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.

According to the equation 2, the permeation coefficient decreases as the polar surface increases in the molecule and an increase of the hidrophobicity increase the permeability. 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 human skin. This behavior was reported by Moss et al. in an excellent review [37] where the main role of hidrophobicity in accounting for this property was explained.

However, the contributions of the heteroatom are also dependent on its volume [37]. 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 thence the result of this effect is a delay in the permeation process [38].

Finally, this decrease of the permeability when increase the polarity of the molecule also should be due to the oxygen and nitrogen present the possibility to form a hydrogen bond with the polar compounds in the human skin. Potts and Guy [31] and Patel et al. [20] 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. [39] where the hydrogen bond acceptors sites could potentially hamper skin permeation.

This study demonstrated that the passage of chemicals throughout a human skin in this set of compounds is controlled by the hydrophobicity and the polarity. In addition we demonstrated that the TOPS MODE is a excellent tools for the prediction of the permeability is this type of matrix.

4 CONCLUDING REMARKS

We have shown that the TOPS-MODE approach is able to describe the permeability of different compounds through human skin. In fact, we have developed a model for predicting the permeability coefficient of a data set of 37 permeants, which is both statistically and chemically sound. This model explains more than 93 % 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 other nine different methodologies used to predict this property. Therefore, the spectral moments show a better performance than other kind of descriptors, which suggests that they can be used in new QSPR applications.

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.

Acknowledgments

Maykel Pérez thanks the collaboration of the owner of the software Modeslab 1.0 for the donation of this valuable tool and his valuable contributions to the completion of this work.

5 REFERENCES

1. G. B. Kasting, R. L. Smith, E. R. Cooper. Effect of lipid solubility and molecular size on percutaneous absorption. In Skin Pharmacokinetics, B. Shroot and H. Schaefer (eds.), Karger, Basel, **1987,** 138-153.
2. B.D. Anderson, W. I. Higuchi, P. V. Raykar. Heterogeneity effects on permeability-partition coefficient relationships in human stratum corneum. Pharm. Res. **1988**, 5 566-573.
3. B.D. Anderson, W. I. Higuchi, P. V. Raykar. Solute structure-permeability relationships in human stratum corneum. J. Invest. Dermatol. **1989,** 93, 280-286.
4. A.D. Woolfson, D.F. McCafferty. Percutaneous local anaesthesia: drug release characteristics of the amethocaine phase-change system. Int. J. Pharm. **1993,** 94, 75–80.
5. Y. Kurosaki, N. Nagahara, W.A. Taniza, H. Nishimura, T. Nakayama, T. Kimura. Use of lipid disperse systems in transdermal drug delivery - comparative study of flufenamic acid permeation among rat abdominal skin, silicon rubber membrane and stratum corneum sheet isolated from hamster cheek pouch. International Journal of Pharmaceutics **1991,** 67, 1–9.
6. Megrab, N.A., Williams, A.C., Barry, B.W. Estradiol permeation through human skin and Silastic membrane — effects of propylene glycol and supersaturation. Journal of Controlled Release **1995,** 36, 277–294.

7. N.A. Megrab, A.C. Williams, B.W. Barry. Estradiol permeation through human skin, Silastic and snake skin membranes the effects of ethanol-water co-solvent systems. International Journal of Pharmaceutics **1995,** 116, 101–112.
8. P.W. Stott, A.C. Williams, B.W. Barry. Characterisation of complex coacervates of some tricyclic antidepressants and evaluation of their potential for enhancing transdermal flux. Journal of Controlled Release **1996**, 41, 215–227.
9. D. van Hal, A. van Rensen, T. Vringer, H. Junginger, J. Bouwstra. Diffusion of estradiol from non-ionic surfactant vesicles through human stratum corneum in vitro. STP Pharma Sciences **1996,** 6, 72–78.
10. P. Minghetti, A. Casiraghi, F. Cilurzo, L. Montanari, M. Marazzi, L. Falcone, V. Donati. Comparison of different membranes with cultures of keratinocytes from man for percutaneous absorption of nitroglycerine. Journal of Pharmacy and Pharmacology **1999,** 51, 673–678.
11. M.T.D. Cronin, J.C. Dearden, G.P. Moss, G. Murray-Dickson. Investigation of the mechanism of flux across human skin in vitro by quantitative structure-permeability relationships. European Journal of Pharmaceutical Sciences **1999,** 7, 325–330.
12. R. Cichon, S. Janicki. Effect of polyoxyethylene glycols (PEG) on properties of the matrix model of transdermal therapeutic system (TTS) with testosterone. Pharmazie **1991,** 46, 719–723.
13. S.H. Yuk, S.J. Lee, T. Okano, B. Berner, S.W. Kim. One-way membrane for transdermal drug delivery systems 1. Membrane preparation and characterisation. International Journal of Pharmaceutics **1991,** 77, 221–229.

14. D. Foley, J. Corish, O.I. Corrigan. Iontophoretic delivery of drugs through membranes including human stratum corneum. Solid State Ionics **1992,** 53, 184–196.

15. A.M.R. Bayon, J. Corish, O.I. Corrigan. In vitro passive and iontophoretically assisted transport of salbutamol sulfate across synthetic membranes. Drug Development and Industrial Pharmacy **1993,** 19, 1169–1181.

16. J. Ramis, L. Conte, X. Segado, J. Forn, J. Lauroba, A. Calpena, E. Escribano, J. Domenech. Influence of formulation on the in vitro transdermal penetration of flutrimazole. Arzneimittel-Forschung/ Drug Research **1997**, 47, 1139–1144.

17. M.T.D. Cronin, J.C. Dearden, R. Gupta, G.P. Moss. An investigation of the mechanism of flux across polydimethylsiloxane membranes by use of quantitative structure-permeability relationships. Journal of Pharmacy and Pharmacology **1998,** 50, 143–152.

18. R.J. Scheuplein, I.H. Blank. Permeability of the skin. Physiological Reviews **1971,** 51, 702–747.

19. M.S. Roberts, R.A. Anderson, J. Swarbrick,. Permeability of human epidermis to phenolic compounds. Journal of Pharmacy and Pharmacology **1977,** 29, 677–683.

20. H. Patel, W. ten Berge, M.T.D. Cronin. Quantitative structure-activity relationships (QSARs) for the prediction of skin permeation of exogenous chemicals. Chemosphere. **2002,** 48, 603-613.

21. E. Estrada. Spectral moments of the edge adjacency matrix in molecular graphs. 1. Definition and applications to the prediction of physical properties of alkanes. *J. Chem. Inf. Comput. Sci*., **1996**. 36, 844-849.

22. E. Estrada. Spectral moments of the edge adjacency matrix in molecular graphs. 2. Molecules containing heteroatoms and QSAR applications. *J. Chem. Inf. Comput. Sci.*, **1997**. 37, 320-328.
23. E. Estrada. Spectral moments of the edge adjacency matrix in molecular graphs. 3. Molecules containing cycles. *J. Chem. Inf. Comput. Sci*., **1998**. 38, 23-27.
24. E. Estrada. Modeling the diamagnetic susceptibilities of organic compounds by a substructural graph theoretical approach. *J. Chem. Soc. Faraday Trans*., **1998.** 94, 1407-1411.
25. E. Estrada. Generalized Spectral Moments of the interacted line graph sequence. A novel approach to QSPR studies. *J. Chem. Inf. Comput. Sci*., **1998**. 39, 90-95.
26. E. Estrada. and Y. Gutiérrez. Modeling chromatographic parameters by a novel graph theoretical sub-structural approach. *J. Chromatogr. A*, **1999,** 858, 187-199.
27. E. Estrada, Y. Gutiérrez and H. González. Modeling diamagnetic and magnetooptic properties of organic compounds with the TOSS-MODE approach. *J. Chem. Inf. Comput. Sci*., **2000**, 40, 1386-1399.
28. Y. Gutierrez and E. Estrada. TOSS-MODE (1997) (Topological Sub-Structural Molecular Design) for Windows Version 4.0, Universidad de Santiago de Compostela, Spain.
29. Y. Gutierrez and E. Estrada. Modes Lab ®, **2002**, version 1.0 b.
30. E. Estrada, Edge Adjacency relationship and a novel topological index related to molecular volume, *Inf. Comput. Sci*., **1995**, 35, 31-33.
31. R. O. Potts and R. H. Guy. A predictive algorithm for skin permeability : The effects of molecular size and hydrogen bond activity. Pharmaceutical Research. **1995**, 12, 1628-1633.
32. R. Todeschini, V. Consonni and M. Pavan. 2002 Dragon. Software version 2.1.
33. J.S. Michael, E. Dewar, G. Zoebisch, F. Eamonn, and J.P. Stewart. J. Am. Chem. Soc., **1985,** 107, 3902.
34. J. J. P. Stewar. **1990** MOPAC manual, 6th ed., p 189, Frank J. Seiler Research Laboratory, U. S. Air Force academy, Colorado Springs, CO.

35. R. Todeschini and V. Consonni. Handbook of molecular descriptors. Wiley-VCH, Weinheim, Germany, **2000**.

36. R B. Kowalski and S. Wold. Pattern recognition in chemistry, In: Handbook of statistics. Eds: Krishnaiah, P. R. and Kanal, L. N. North Holland Publishing Company, Amsterdam. **1982,** 673-697.

37. G.P. Moss, J.C. Dearden, H. Patel, M.T.D. Cronin Quantitative structure–permeability relationships (QSPRs) for percutaneous absorption. Toxicology in Vitro. **2002**, 16, 299-317.

38. M. P. Gonzalez and A. M. Helguera. TOPS-MODE versus DRAGON descriptors to predict permeability coefficients through Low-density polyethylene. J. Comp. Aided Mol. Design. (*in press*)

39. C.A. Lipinski, F. Lombardo, B.W. Dominy, P.J. Feeney. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. Adv. Drug Deliv. Rev. **1997,** 23, 3–25.

Table 1. Observed, predicted, and residual values of permeability coefficients (cm/s) through human epidermis for the 37 compounds used to derive the QSPR [31].

Number	Compounds	Observed	Predicted	Deleted Residuals
1	Water	-6.130	-5.871	-0.356
2	Methanol	-6.680	-7.165	0.603
3	methanoic acid	-7.080	-7.402	0.405
4	Ethanol	-6.660	-6.635	-0.029
5	ethanoic acid	-7.010	-6.946	-0.073
6	n-propanol	-6.410	-6.362	-0.052
7	n-propanoic acid	-7.010	-6.633	-0.442
8	butane-2-one	-5.900	-5.590	-0.402
9	Benzene	-4.510	-4.336	-0.200
10	diethyl ether	-5.350	-5.003	-0.381
11	n-butanol	-6.160	-6.090	-0.075
12	n-butanoic acid	-6.360	-6.340	-0.024
13	Phenol	-5.640	-5.518	-0.128
14	Toluene	-3.560	-4.118	0.651
15	Styrene	-3.750	-3.860	0.130
16	n-pentanol	-5.780	-5.817	0.040
17	phenylmethanol	-5.780	-5.776	-0.005
18	n-pentanoic acid	-6.010	-6.064	0.063
19	2-chlorophenol	-5.040	-4.719	-0.338
20	4-chlorophenol	-5.000	-4.954	-0.049
21	m-cresol	-5.380	-5.245	-0.141
22	o-cresol	-5.360	-5.068	-0.304
23	p-cresol	-5.290	-5.271	-0.020
24	4-bromophenol	-5.000	-4.823	-0.190
25	4-nitrophenol	-5.810	-5.728	-0.098
26	3-nitrophenol	-5.810	-5.624	-0.228
27	2-nitrophenol	-4.560	-4.796	0.580
28	Ethylbenzene	-3.480	-3.886	0.481
29	n-hexanol	-5.450	-5.545	0.101
30	n-hexanoic acid	-5.440	-5.791	0.413
31	b-naphthol	-3.700	-3.714	0.017
32	n-heptanol	-5.050	-5.273	0.239
33	n-heptanoic acid	-5.280	-5.518	0.286
34	n-octanol	-4.840	-5.000	0.176
35	n-octanoic acid	-5.210	-5.246	0.045
36	n-nonanol	-4.770	-4.728	-0.048
37	n-decanol	-4.660	-4.456	-0.242

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

Descriptors	Variables[a]	S	R^2	F	p	q^2
Spectral moments	$\mu_{15}D$, $\mu_{15}GM$, μ_1H, μ_1PS	0.240	0.938	121.06	0.000	0.907
Constitutional	nC, nN, nO, nX	0.378	0.843	43.22	0.000	0.741
Topological	SPI, Jhete, PW4, SEigv	0.370	0.851	45.75	0.000	0.801
BCUT	BELe3, BELe4, BELp6, BELp5	0.381	0.841	42.415	0.000	0.791
2D autocorrelations	ATS1e, ATS4e, ATS4p, GATS1p	0.440	0.789	29.908	0.000	0.701
Geometrical	MAXDP, G2, SPAM, G(N..O)	0.340	0.873	55.291	0.000	0.824
RDF	RDF010u, RDF020e, RDF010p, RDF020p	0.415	0.812	34.577	0.000	0.742
3D-MORSE	Mor28v, Mor26u, Mor32m, Mor31u	0.293	0.906	77.628	0.000	0.861
GETAWAY	H2u, H1e, R1e, R1p	0.349	0.867	52.147	0.000	0.817
WHIM	G2p, Ts, Au, Ap	0.527	0.702	18.873	0.000	0.632

[a] The definition of the terms appears largely explained in reference 35.

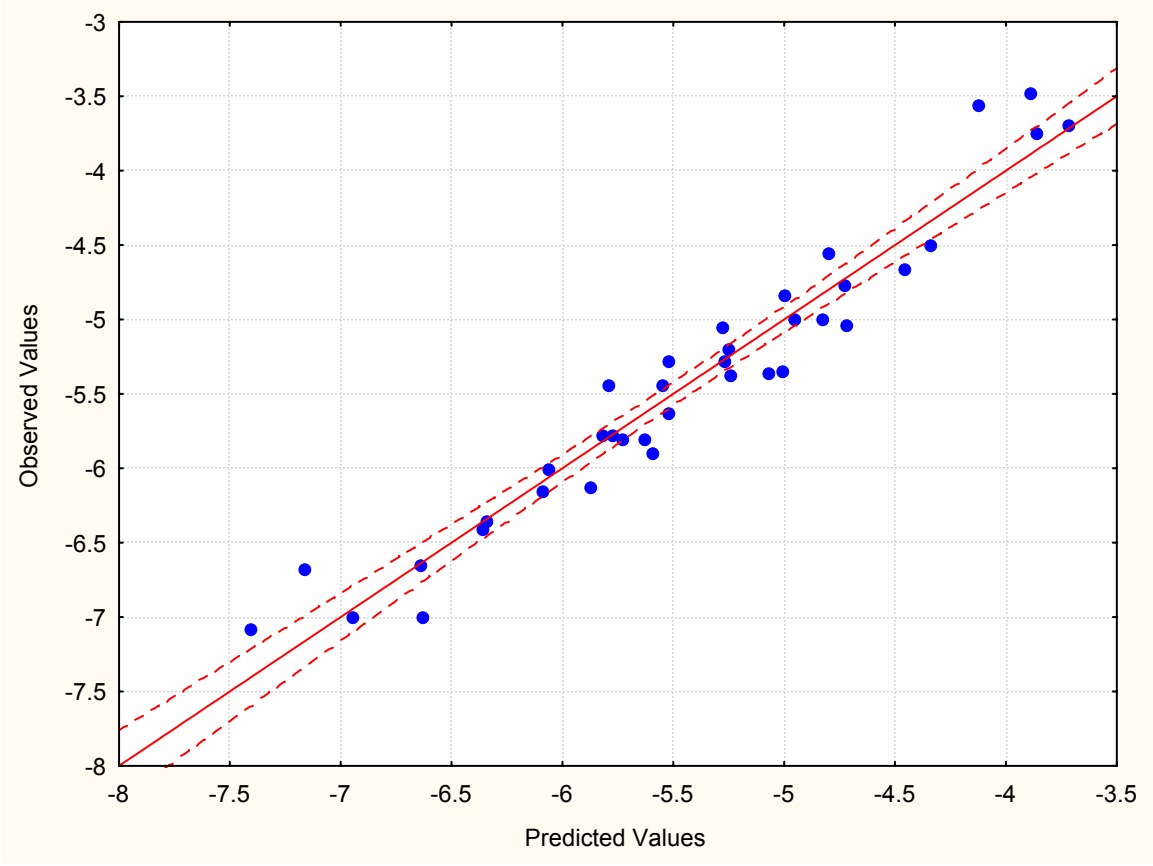

Figure 1. The linear relation between observed and predicted permeability for the compounds of the training set.