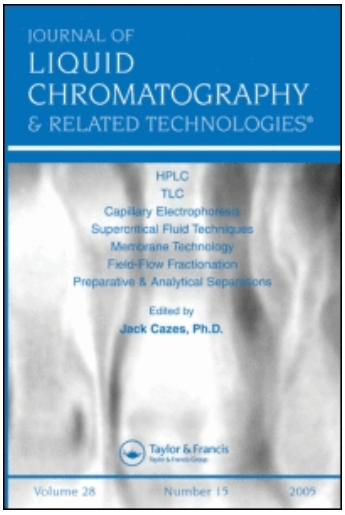


Journal of Liquid Chromatography & Related Technologies

Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713597273

Prediction of the Chromatographic Retention (Lipophilicity) of Some New Methyl-Thiazole-Oxadiazoline Derivatives by Multivariate Regression Methods

B. Tiperciuc[a]; C. Sârbu[b]

[a] Faculty of Pharmacy, Iuliu Hatieganu University of Medicine and Pharmacy, Cluj-Napoca, Romania [b] Faculty of Chemistry and Chemical Engineering, Babeş-Bolyai University, Cluj-Napoca, Romania

To cite this Article Tiperciuc, B. and Sârbu, C.(2006) 'Prediction of the Chromatographic Retention (Lipophilicity) of Some New Methyl-Thiazole-Oxadiazoline Derivatives by Multivariate Regression Methods', Journal of Liquid Chromatography & Related Technologies, 29: 15, 2257 — 2270

To link to this Article: DOI: 10.1080/10826070600832939

URL: http://dx.doi.org/10.1080/10826070600832939

Journal of Liquid Chromatography & Related Technologies®, 29: 2257–2270, 2006

ISSN 1082-6076 print/1520-572X online
DOI: 10.1080/10826070600832939

Prediction of the Chromatographic Retention (Lipophilicity) of Some New Methyl-Thiazole-Oxadiazoline Derivatives by Multivariate Regression Methods

B. Tiperciuc
Iuliu Hatieganu University of Medicine and Pharmacy,
Faculty of Pharmacy, Cluj-Napoca, Romania

C. Sârbu
Babeş-Bolyai University, Faculty of Chemistry and Chemical
Engineering, Cluj-Napoca, Romania

Abstract: Retention indices for a new series of methyl-thiazole-oxadiazolines were determined by reversed phase high performance thin layer chromatography on C_{18} plates with methanol-water in different volume proportions as mobile phase. Comparisons based on the multiple regression methods including multiple linear regression (MLR), principal component regression (PCR) and partial least squares (PLS) have been applied to the modelling of chromatographic lipophilicity (R_{Mo} and b values) by means of 16 different descriptors obtained by using Alchemy package software. The results achieved concerning the prediction of retention indices are highly significant and are in very good agreement with the molecular structure of the compounds investigated. The largest positive charge over the atoms in a molecule (Q_+), the sum of absolute values of the charges on each atom of the molecule (SQ), and the sum of absolute values of the charges on the nitrogens and oxygens in the molecule (SQ_{NO}) seem to be dominant in the retention mechanism.

Keywords: Methyl-thiazole-oxadiazolines, Lipophilicity, TLC, QSRR, MLR, PCA, PCR, PLS

Address correspondence to C. Sârbu, Babeş-Bolyai University, Faculty of Chemistry and Chemical Engineering, Arany Janos 11, 400028 Cluj-Napoca, Romania. E-mail: csarbu@chem.ubbcluj.ro

INTRODUCTION

Quantitative structure-activity relations (QSAR) describe how the molecular structure, in terms of descriptors – lipophilic, electronic, and steric – affects the biological activity of a compound.[1–4] Similarly, quantitative structure – retention relations (QSRR) relate these descriptors to chromatographic retention. Finally, the quantitative retention – activity relations (QRAR) imply that conclusions concerning biological activity can be based on chromatographic experiments.[5–11] In this regard of QRAR, it is considered that the same basic intermolecular actions determine the behaviour of chemical compounds in both biological and chromatographic environments. As a consequence, the chromatographic approach has been quite successful in duplicating Log P data derived by the traditional "shake-flask" technique or other procedures. The relations themselves are usually based on correlation analysis.

Another form of computational analysis used for the correlation of chemical or biological activity and chromatographic retention with different molecular descriptors are Multiple Linear Regression (MLR),[11–13] Principal Component Analysis (PCA),[14–16] Partial Least Squares (PLS),[17–19] or Artificial Neural Networks (ANN).[20–22] In the case of PCA and PLS, for example, starting from a multidimensional space described by different variables, a quantitative model is derived that transforms the axes of the hypersystem. The first principal component (PC1) defines as much of the variation in the data as possible. The second principal component (PC2) describes the maximum amount of residual variation after the first PC has been taken into consideration, and so on. By using only a limited number of PCs, the dimensionality of the data space is reduced, thereby simplifying further analysis.

In this paper we discuss and apply three multivariate regression methods to develop comparative studies, and to provide a QSAR-QSRR model for the characterization and classification of some new methyl-thiazole-oxadiazoline derivatives with a potential antibacterial, antimycotic, and anti-inflammatory activity.

Principal Component Analysis

Principal components analysis (PCA) is also known as eigenvector analysis, eigenvector decomposition, or Karhunen-Loéve expansion. Many problems from chemistry and other scientific fields are strongly related to PCA.[23] The main purpose of PCA is to represent, in an economic way, the location of the samples in a reduced coordinate system where, instead of m-axes (corresponding to m characteristics), only p ($p < m$) can usually be used to describe the data set with maximum possible information.

Principal component analysis practically transforms the original data matrix (X_{nxm}) into a product of two matrices, one of which contains the

information about the objects (S_{nxm}) and the other about the variables (V_{mxm}). The S matrix contains the scores of the *n* objects on *m* principal components (the scores are the projection of the objects on principal components). The V matrix is a square matrix and contains the loadings of the original variables on the principal components (the loadings are the weights of the original variables in each principal component).

Moreover, it may well turn out that usually two or three principal components provide a good summary of all the original variables. Loadings and scores plots, respectively, are very useful as a display tool for examining the relationships between characteristics as well as between compounds, looking for trends, groupings, or outliers.

Multiple Linear Regression

Multiple linear regression (MLR) is an extension of simple linear regression consisting of two or more independent variables (e.g., chemical descriptors) and a numeric dependent variable (e.g., chromatographic retention index). MLR attempts to model the relationship between the independent variables and a response variable by fitting a linear equation to observed data in the following equation:

$$R = a_o + \sum_{i=1}^{k} a_i x_i \quad (1)$$

where a_o, a_i are the estimated regression parameters.

Principal Component Regression

Principal component regression (PCR) is a two-step multivariate calibration method: in the first step, a principal component analysis of the data matrix X is performed. The measured or calculated variables (e.g., descriptors) are converted into new ones (scores on latent variables). This is followed by a multiple linear regression step, MLR, between the scores obtained in the PCA step and the characteristic R to be modelled.

Partial Least Squares Regression

Partial least squares (PLS) is a statistical multivariate regression procedure that is widely applied in chemistry and many other scientific fields. Being a multivariate procedure, PLS provides the ability to predict multiple components of interest simultaneously. The PLS procedure simultaneously estimates underlying factors (loading factors or eigenvectors) that represent

the variation patterns (trends) in both the descriptor data X (in our study) and the retention values R (R_{Mo} and b in our study). These loading factors are used to define a subspace in X that better models R. This is accomplished by using the columns of R matrix to estimate the loading factors of X matrix. At the same time, the columns of X matrix are used to estimate the loading factors for R matrix. The resulting models are shown in Equation (2) and (3).

$$X = TP + E \tag{2}$$

$$R = TQ + F \tag{3}$$

PLS resolves matrices X and R into products of smaller matrices, namely P (the X loading matrix, which contains the directions of the principal components or axes) and T (the scores matrix, which includes the coordinates for the new axes) and Q (the R loading matrix). The factors for X and R are associated through the following relationship:

$$q = bt + \varepsilon \tag{4}$$

where q and t equal the column vectors of Q and T matrices, respectively, and ε equals errors associated with the u-t relationship.

Using both descriptor and retention information to determine the loading factors is the main difference between the PLS and other statistical multivariate procedures, such as multiple linear regression and principal component regression. This feature makes the PLS prediction models more robust for complex data. The number of loading factors used in the final regression model was selected through a cross-validation procedure.[27]

EXPERIMENTAL

The chromatographic behavior of the compounds was studied on the C_{18} silica gel bonded plates. Glass HPTLC plates (20 × 20 cm) were obtained as a gift

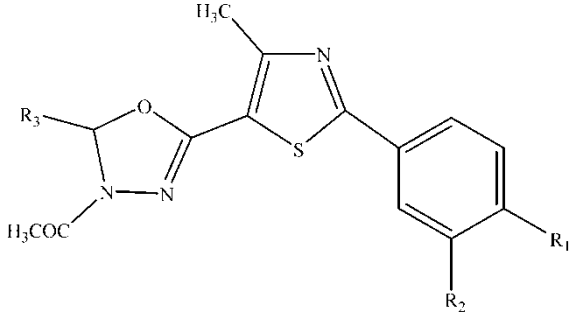
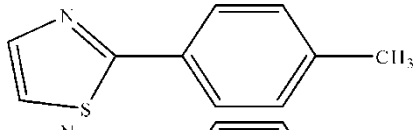
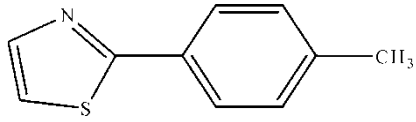
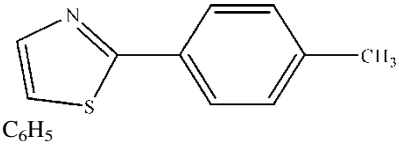


Figure 1. Chemical structure of methyl-thiazole-oxadiazolines.

from Macherey-Nagel (Düren, Germany). Methanol for chromatography was supplied from Reactivul (Bucharest, Romania). Methanolic solutions of each compound, in Table 1, were prepared at a concentration of 1 mg mL^{-1}. Chromatograms were developed by an ascending technique at room temperature; the developing distance being 10 cm. The mobile phase was a mixture of methanol–water with various content from 45 to 70% (v/v) in 5% steps, as the studied compounds differed considerably in their retention. After being developed, the dried plates were examined under a UV lamp ($\lambda = 254$ nm). The R_M values of each compound were obtained by using the well known following equation:

$$R_M = \log (1/R_f - 1) \tag{5}$$

Table 1. The structures of the substituents in Figure 1

Compound	R_1	R_2	R_3
1	CH_3	H	C_6H_4Cl(o)
2	H	H	C_6H_4Cl(m)
3	H	CF_3	C_6H_4Cl(p)
4	H	H	$C_6H_4OCH_3$(p)
5	H	CF_3	
6	CH_3	H	
7	CH_3	H	$C_6H_4OCH_3$(o)
8	H	CF_3	$C_6H_4OCH_3$(m)
9	H	CF_3	$C_6H_4OCOCH_3$(p)
10	H	H	$C_6H_4OCOCH_3$(o)
11	CH_3	H	$C_6H_4OCOCH_3$(m)
12	H	H	C_6H_5
13	CH_3	H	C_6H_4Cl(p)
14	H	H	
15	H	CF_3	C_6H_5
16	H	CF_3	$C_6H_4OCOCH_3$(o)
17	H	H	C_6H_4Cl(o)
18	H	CF_3	C_6H_4Cl(o)
19	H	CF_3	$C_6H_4OCH_3$(o)
20	CH_3	H	$C_6H_4OCOCH_3$(p)

Linear correlation between R_M values and the concentration of organic modifier in the mobile phase were calculated separately for each compound according to the equation:

$$R_M = R_{Mo} + bC, \quad (6)$$

where C is the concentration of methanol in the mobile phase. The R_{Mo} and b values (related to the molecular lipophilicity) are listed in Table 2.

Description of Methyl-Thiazole-Oxadiazolines

The investigated methyl-thiazole-oxadiazoline derivatives were synthesized by the procedure described earlier.[24,25] The molecular structure of methyl-thiazole oxadiazoline series studied in this paper is depicted in Table 1.

In order to define the character of the compound structure, the following descriptors available in the ALCHEMY 2000 programs[26] were taken into consideration and used as independent variables. The partition coefficient (*Log P*), the first-order ($^1\chi$) and the third-order ($^3\chi$) connectivity index, the zero-order ($^0\chi^v$) and the first-order ($^1\chi^v$) valence order connectivity index, the third-order shape index for molecule ($^3K_\alpha$), the Wiener (*W*) index based on the graph of the molecule, volume (*V*), molar mass (*M*), dipole moment (*DM*), molecular polarizability (*MP*), specific molar polarizability (*SP*), the largest positive charge over the atoms in a molecule, in electrons (Q_+), the largest negative charge over the atoms in a molecule, in electrons (Q_-), the sum of absolute values of the charges on each atom of the molecule, in electrons (*SQ*), and the sum of absolute values of the charges on the nitrogens and oxygens in the molecule, in electrons (SQ_{NO}). The values obtained are presented in Table 2.

RESULTS AND DISCUSSION

By reducing the number of features from 18 original descriptors (including retention indices) to three principal components (latent variables), the information preserved is enough to permit a primary examination of the similarities and differences between descriptors and methyl-thiazole-oxadiazoline derivatives. The contribution of the first component represents 55.71% of the total variance and 75.11% of the total variance. The first three components reproduce approximately 88% of the total variance and the first six even 99.31%, and the eigenvalues become negligible after seventh component.

All the statements above are well supported by the 2D- and 3D-representations of the loadings. The projection of the 3D-representation (Figure 2) gives a more complete pattern: it is clear, for example, that the majority of the descriptors considered in this study form two close clusters: the first one includes $^1\chi$, $^3\chi$, *W*, R_{Mo}, *M*, and the second one encompasses *MP*, $^0\chi^v$, $^1\chi^v$,

Table 2. The descriptors computed for the methyl-thiazole-oxadiazoline derivatives investigated in this paper

	Log P	$^1\chi$	$^3\chi$	$^0\chi^v$	$^1\chi^v$	$^3K_\alpha$	W	V	M	DM	MP	SP	Q_+	Q_-	SQ	SQ_{NO}	R_{Mo}	b
1	4.111	13.44	10.60	17.21	10.0	3.72	2059	348.84	411.91	3.80	43.86	0.126	0.215	−0.377	3.760	1.253	2.74	−4.35
2	3.941	13.02	10.04	16.29	9.6	3.61	1856	332.36	397.87	4.67	42.03	0.126	0.215	−0.376	3.683	1.256	2.63	−4.22
3	4.207	14.63	11.41	17.85	10.3	4.53	2782	361.07	465.88	4.28	43.59	0.121	0.424	−0.376	4.502	1.254	3.38	−5.33
4	3.335	13.56	10.53	16.56	9.6	3.73	2103	343.01	393.47	5.72	42.57	0.124	0.215	−0.382	4.099	1.635	2.54	−4.07
5	4.248	17.10	13.59	20.96	12.5	5.07	4327	424.07	528.58	5.45	51.97	0.122	0.424	−0.376	5.124	1.549	3.55	−5.32
6	4.637	15.88	12.72	20.33	12.2	4.39	3405	411.67	474.61	3.10	52.25	0.127	0.219	−0.376	4.338	1.549	3.06	−4.76
7	3.594	13.97	10.91	17.49	10.0	3.82	2260	359.52	407.49	3.13	44.41	0.123	0.215	−0.381	4.200	1.633	2.30	−3.88
8	3.563	15.17	11.75	18.12	10.3	4.64	3024	372.14	461.47	3.02	44.14	0.119	0.424	−0.382	4.939	1.637	2.71	−4.47
9	2.718	16.02	12.05	19.03	10.8	5.28	3718	390.11	489.48	8.03	46.06	0.118	0.424	−0.376	5.428	1.906	2.87	−4.66
10	2.558	14.43	10.75	17.47	10.1	4.20	2455	361.03	421.48	5.49	44.49	0.123	0.216	−0.375	4.657	1.900	1.78	−3.36
11	3.315	14.81	11.10	18.40	10.5	4.57	2806	377.26	435.51	3.74	46.33	0.123	0.215	−0.376	4.670	1.911	2.16	−3.77
12	3.302	12.63	9.71	15.23	9.1	3.30	1676	318.38	363.44	3.47	40.10	0.126	0.215	−0.376	3.672	1.256	2.04	−3.48
13	4.180	13.42	10.53	17.21	10.0	3.85	2101	349.16	411.91	5.26	43.86	0.126	0.215	−0.376	3.717	1.254	2.81	−4.50
14	4.606	15.49	12.30	19.41	11.8	4.17	3098	394.78	460.58	2.98	50.41	0.128	0.219	−0.376	4.296	1.549	2.70	−4.30
15	4.020	14.24	11.00	16.79	9.8	4.20	2523	346.99	431.44	5.04	41.66	0.120	0.424	−0.376	4.499	1.256	2.64	−4.33
16	3.171	16.04	12.04	19.03	10.8	5.13	3526	390.42	489.48	3.65	46.06	0.118	0.424	−0.376	5.487	1.900	2.29	−4.00
17	3.733	13.04	10.19	16.29	9.6	3.49	1836	331.85	397.89	3.66	42.03	0.127	0.215	−0.376	3.719	1.253	2.84	−4.55
18	4.188	14.65	11.48	17.85	10.3	4.39	2734	361.29	465.88	4.40	43.59	0.121	0.424	−0.376	4.546	1.253	3.10	−4.89
19	3.929	15.18	11.79	18.12	10.3	4.50	2976	372.31	461.47	5.65	44.14	0.118	0.424	−0.381	4.985	1.633	2.61	−4.33
20	2.675	14.81	11.17	18.40	10.5	4.57	2890	377.03	435.51	6.07	46.33	0.123	0.215	−0.376	4.642	1.906	2.20	−3.77

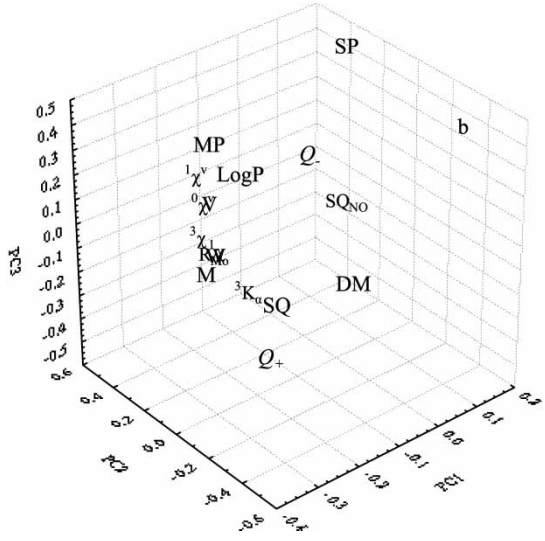


Figure 2. PC1, PC2, and PC3 loading plot of the autoscaled all data in Table 2.

Log P, *V*; $^3K_\alpha$, *SQ*, Q_+, *DM*, SQ_{NO}, Q_-, *SP*, and *b* appear more or less as outliers and are the most discriminating.

The scatter plot of scores in the space described by PC1, PC2, and PC3 (see Figure 3) shows interesting results. Three clusters appear to be

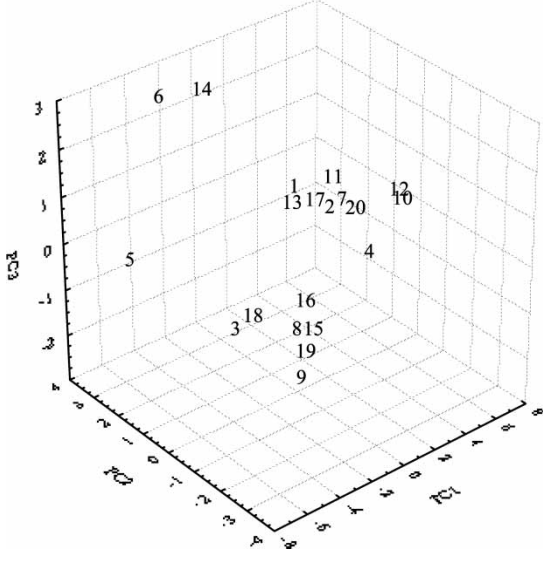


Figure 3. PC1, PC2, and PC3 score plot of the autoscaled all data in Table 2.

well defined and in good agreement to the structure of compounds: one of them corresponds to the compounds 5, 6, 14 (the largest molecules in the series) in the above left part of the graph, the second including the group of fluorine derivatives (3, 8, 9, 15, 16, 18, 19), with the exception of compounds 5, 6, and 16, is located in the middle bottom of the graph and the third group (1, 2, 4, 7, 10, 11, 12, 13, 17, 20), respectively (right).

In order to describe the relationship between the chromatographic retention indices of tested compounds (R_{Mo} and b values) and the calculated structural descriptors, a multivariate regression analysis was performed. By forward stepwise multiple regression analysis, the following high quality regression equations were obtained:

$$R_{Mo} = 3.365 - 0.203\text{LogP} + 0.021\text{M} + 3.268\ SQ_{NO} - 4.094\text{SQ} + 10.937Q_{+} + +0.0003\text{W} \quad (7)$$

$$(R^2 = 0.9438, n = 20, F = 36,\ p < 0.0000, s = 0.1260)$$

$$b = -2.477 - 0.203\text{LogP} + 0.029\text{M} - 3.625SQ_{NO} + 4.393\text{SQ} - 11.999Q_{+} - 0.026\text{DM} - 2.094\text{SP} \quad (8)$$

$$(R^2 = 0.9402, n = 20, F = 27,\ p < 0.0000, s = 0.1635)$$

where n is the number of compounds, R^2 the determination coefficient, F the F-test value, p is the significance level of all the equation, and s is standard error of estimates.

The F and p values of Equation (7) and (8) show that the equations are very significant, having a high determination coefficient and relatively small s values.

The results also suggest that the largest positive charge over the atoms in a molecule (Q_{+}), the sum of absolute values of the charges on each atom of the molecule (SQ) and the sum of absolute values of the charges on the nitrogens and oxygens in the molecule (SQ_{NO}), seem to be dominant in the retention mechanism and, as a consequence, control the lipophilicity.

By eliminating the descriptors having statistical non significant coefficients (according to t-test), the following simpler models may be obtained (Equation (9) and (10)):

$$R_{Mo} = 1.132 - 0.021\text{M} + 3.112SQ_{NO} - 3.487\text{SQ} + 9.593Q_{+} \quad (9)$$

$$(R^2 = 0.9191, n = 20, F = 43,\ p < 0.0000, s = 0.1408)$$

$$b = -2.369 - 0.025\text{M} - 3.777SQ_{NO} + 4.125\text{SQ} - 11.900Q_{+} \quad (10)$$

$$(R^2 = 0.9196, n = 20, F = 43,\ p < 0.0000, s = 0.1696)$$

For PCR and PLS methods, the original 16 descriptors were used for the selection of the optimum number of factors by using the cross-validation procedure.

The prediction error was calculated as a root mean square (RMS) error according to ref. [27]. The values of RMS are graphed in Figure 4. By using the cross-validation procedure, it is easy to observe that its numerical values were minimized in the case of the first eight factors for PCR and seven factors for PLS, respectively.

In order to compare the MLR, PCR, and PLS predictive models, the linear regression was applied considering observed and predicted values of retention indices. Good correlations between observed and predicted retention indices of the three methods were found, as it is easy to observe in Table 3. In all cases, the intercept a_o is statistically zero (t-test) and the slope a_1 is not statistically different from 1 (t-test). As a direct consequence, one can conclude that all three straight lines are parallel and pass through the origin (Figure 5 and 6). In addition, high F values and very small s values were obtained in all cases with a special remark in the case of MLR, which appears to be the most effective for the estimation of retention indices.

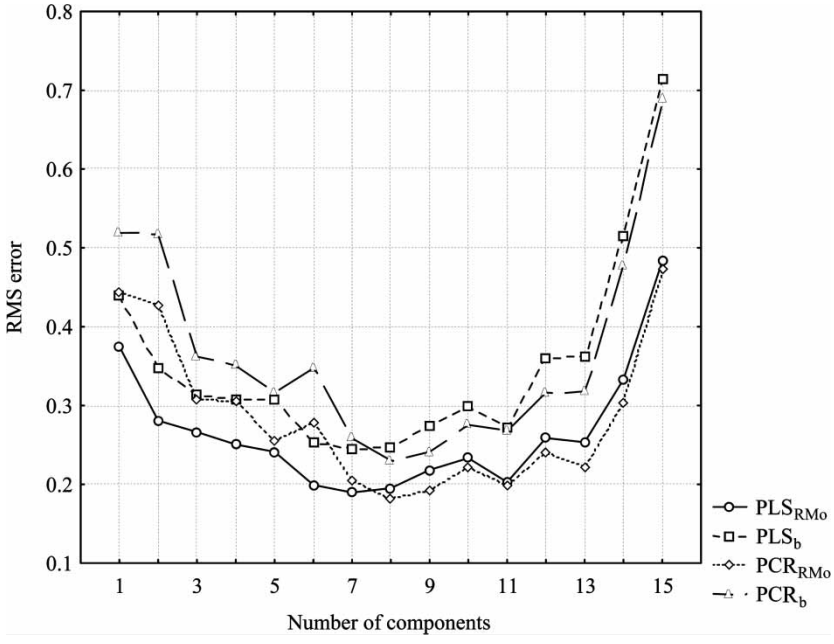


Figure 4. Root mean square errors of estimation of R_{Mo} and b using PCR and PLS.

Table 3. Statistical parameters to evaluate linear relation between observed and predicted retention indices considering all three multiple regression methods (MLR, PCR and PLS)

	MLR		PCR		PLS	
Statistical	R_{Mo}	b	R_{Mo}	b	R_{Mo}	b
R^2	0.9438	0.9402	0.8254	0.8099	0.8164	0.7926
F	302	283	85	77	80	69
p	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
s	0.1042	0.1294	0.1845	0.2311	0.1944	0.2493
a_o	0.1486	−0.2581	0.3145	−0.5708	0.2636	−0.4828
a_1	0.9438	0.9402	0.8875	0.8737	0.9073	0.8930

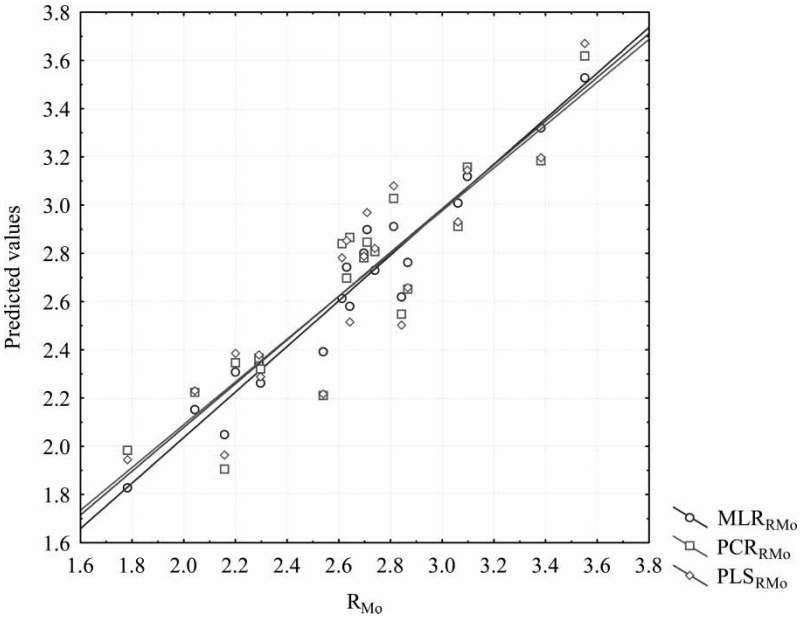


Figure 5. Correlation between R_{Mo} determined by HPTLC and predicted values by MLR, PCR, and PLS.

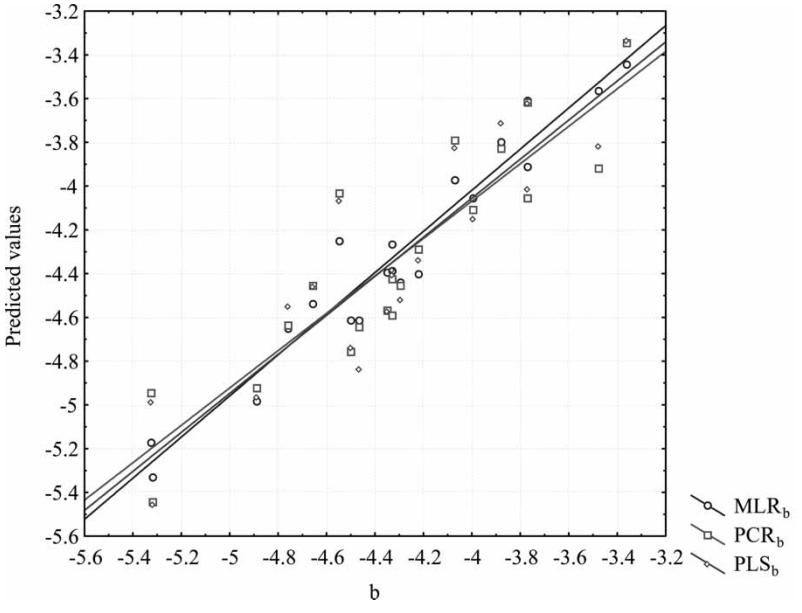


Figure 6. Correlation between b determined by HPTLC and predicted values by MLR, PCR, and PLS.

CONCLUSIONS

Correlation obtained between chromatographic retention indices and structure descriptors for substituted methyl-thiazole-oxadiazolines are highly significant and might be used to predict the retention behavior, and as a consequence, the lipophilicity of other members of the series. By comparing the multivariate regression methods used in this study, the forward stepwise MLR appeared to be the most effective in predicting retention indices for compounds investigated. The largest positive charge over the atoms in a molecule (Q_+), the sum of absolute values of the charges on each atom of the molecule (SQ), and the sum of absolute values of the charges on the nitrogens and oxygens in the molecule (SQ_{NO}) seem to be dominant in the retention mechanism and, hence, these descriptors control the lipophilicity.

REFERENCES

1. Leo, A.; Hansch, C.; Elkins, D. Partition coefficients and their uses. Chem. Rev. **1971**, *71*, 525–616.
2. Rekker, R.F.; Mannhold, R. *Calculation of Drug Lipophilicity*; VCH: Weinheim, 1992.
3. Karelson, M. *Molecular Descriptors in QSAR/QSPR*; Wiley & Sons: New York, 2000.
4. Chiriac, A.; Ciubotariu, D.; Simon, Z. *QSAR-Quantitative Relationships*; Mirton Publishing House: Timişoara, 1995.
5. Kaliszan, R. Quantitative structure retention relationships. Anal. Chem. **1992**, *64*, 619A–631A.
6. Pyka, A.; Miszczyk, M. Chromatographic evaluation of the lypophilic properties of selected pesticides. Chromatographia **2005**, *61*, 37–42.
7. Kaliszan, R. Chromatography and capillary electrophoresis in modelling the basic processes of drug action. Trends Anal. Chem. **1999**, *18*, 401–409.
8. Wang, Q.S.; Zhang, L. Review of research on quantitative structure-retention relationships in thin-layer chromatography. J. Liq. Chromatogr. & Rel. Technol. **1999**, *22*, 1–14.
9. Garcia Alvarez-Coque, M.C.; Torres Lapasio, J.R. Quantitation of hydrophobicity in micellar liquid chromatography. Trends Anal. Chem. **1999**, *18*, 533–543.
10. Djaković-Sekulić, T.L.J.; Sârbu, C.; Perišić-Janjić, N.U. A Comparative study of the lipophilicity of benzimidazole and benztriazole derivatives by RPTLC. J. Planar Chromatogr. **2005**, *18*, 432–436.
11. Dai, J.; Jin, L.; Yao, S.; Wang, L. Prediction of partition coefficient and toxicity for benzaldehyde compounds by their capacity factors and various molecular descriptors. Chemosphere **2001**, *42*, 899–907.
12. Detroyer, A.; Heyden, Y.V.; Carda-Broch, S.; Garcia-Alvarez-Coque, M.C.; Massart, D.L. Quantitative structure-retention and retention-activity relationships of β-blocking agents by micellar liquid chromatography. J. Chromatogr. A **2001**, *912*, 211–221.
13. Sârbu, C.; Casoni, D.; Dărăbanţu, M.; Maiereanu, C. Quantitative structure-retention and retention-activity relationships of some 1,3-oxazolidine systems by RP-HPTLC and PCA. J. Pharm. Biomed. Anal. **2004**, *35*, 213–219.

14. Sârbu, C.; Todor, S. Determination of lipophilicity of some non-steroidal anti-inflammatory agents and their relationships by using principal component analysis based on thin-layer chromatographic retention data. J. Chromatogr. A **1998**, *822*, 263–269.
15. Sârbu, C.; Kuhajda, K.; Kevresan, S. Evaluation of lipophilicity of bile acids and their derivatives by Thin Layer Chromatography and Principal Component Analysis. J. Chromatogr. A **2001**, *917*, 361–366.
16. Detroyer, A.; Schoonjans, V.F.; Questier, F.; Heyden, Y.V.; Borosy, A.P.; Guo, Q.; Massart, D.L. Exploratory chemometric analysis of the classification of pharmaceutical substances based on chromatographic data. J. Chromatogr. A **2000**, *897*, 23–36.
17. Martens, H.; Naes, T. *Multivariate Calibration*; Wiley: Chichester, 1991.
18. Zhou, Y.; Xu, L.; Wu, Y.; Liu, B. A QSAR study of the antiallergic activities of some benzamides and their structures. Chemometr. Intell. Lab. Syst. **1999**, *45*, 95–100.
19. Li, T.; Mei, H.; Cong, P. Combining nonlinear PLS with the numeric genetic algorithm for QSAR. Chemometr. Intell. Lab. Syst. **1999**, *45*, 177–184.
20. Zhang, R.; Yan, A.; Liu, M.; Liu, H.; Hu, Z. Application of artificial networks for prediction of the retention indices of alkylbenzenes. Chemometr. Intell. Lab. Syst. **1999**, *45*, 113–120.
21. Zhao, R.H.; Yue, B.F.; Ni, J.Y.; Zhou, H.F.; Zhang, Y.K. Application of an artificial neural network in chromatography-retention behaviour prediction and pattern recognition Chemometr. Intell. Lab. Syst. **1999**, *45*, 163–170.
22. Chen, Y.; Chen, D.; He, C.; Hu, S. Quantitative structure-activity relationships study of herbicides using neural networks and different statistical methods. Chemometr. Intell. Lab. Syst. **1999**, *45*, 267–276.
23. Sârbu, C.; Pop, H.F. Fuzzy Soft-Computing Methods and Their Applications in Chemistry. In *Reviews in Computational Chemistry*; Lipkowitz, K.B., Boyd, D.B., Cundari, T.R., Eds.; Wiley-VCH, 2004; 249–332.
24. Tiperciuc, B.; Oniga, O.; Palage, M.; Chiran, D.; Chirtoc, I. Heterocicli 81. Obţinerea şi caracterizarea unor 3-N-acetil-2-R-5-[2′-fenil-4′–metil-tiazol-5′-il]-Δ_2-1,3,4-oxadiazoline. Clujul Medical **1999**, *LXXII*, 99–103.
25. Tiperciuc, B.; Parvu, A.; Palage, M.; Oniga, O.; Chirtan, D. Heterocycles 82. The synthesis and the study of the anti-inflammatory activity of some 3-N-acetyl-2-R-5-[2′-aryl-4′-methyl-thiazole-5′-yl]-Δ_2-1,3,4-oxadiazoline. Farmacia **1999**, *XLVII*, 77–83.
26. http://www.tripos.com/software/alchemy.html.
27. Brereton, R.G. *Chemometrics; Data Analysis for the Laboratory and Chemical Plant*; Wiley & Sons: Chichester, 2003.

Received February 22, 2006
Accepted March 24, 2006
Manuscript 6848