Turkish Journal of Chemistry

Volume 33 | Number 1 Article 15

1-1-2009

Quantitative Structure--Retention Study of Some 2,4-dioksotetrahydro- 1,3--thiazole Derivatives Using the Partial Least Squares Method

TATJANA DJAKOVIC-SEKULIC

COSTEL SARBU

NADA PERISIC-JANJIC

ZAGORKA LOZANOV-CRVENKOVIC

Turk J Chem
33 (2009) , 149 – 157.

doi:10.3906/kim-0705-9

Quantitative Structure–Retention Study of Some 2,4-dioksotetrahydro-1,3–thiazole Derivatives Using the Partial Least Squares Method

Tatjana DJAKOVIĆ SEKULIĆ[1,*], Costel SÂRBU[2], Nada PERIŠIĆ JANJIĆ[1], Zagorka LOZANOV CRVENKOVIĆ[3]

[1] *Department of Chemistry, Faculty of Sciences, Trg Dositeja Obradovića 3, 21000 Novi Sad, SERBIA*
e-mail: tanja@ih.ns.ac.yu

[2] *Babeş-Bolyai University, Faculty of Chemistry and Chemical Engineering, Arany Janos 11, 400028 Cluj-Napoca, ROMANIA*

[3] *Department of Mathematics and Informatics, Trg D. Obradovića 3, 21000 Novi Sad, SERBIA*

Received 17.05.2007

Thin-layer chromatography on rice starch support and aqueous ammonia–organic modifier (methanol, dioxane, and acetone) mobile phases was used to study the effect on retention of the chromatographic system and the physicochemical properties of twelve 2,4-dioxotetrahydro-1,3-thiazoles. A multivariate approach to the retention behaviour of the investigated compounds with 3 quite different organic solvents was used to explain the interactions between the 1,3-thiazoles and the mobile phases. Partial least-squares (PLS) regression was used to quantify differences between the observed data and recognize the molecular properties with the greatest effect on retention for each modifier. Good correlation was obtained between experimental and calculated retention data.

Key Words: 2,4-dioksotetrahydro-1,3-thiazoles, Quantitative Structure-Retention Relationships (QSRR), Molecular descriptors, Partial Least Squares Regression (PLS).

Introduction

The biological activities and commercial applications of many thiazoles, especially 2,4–dioksotetrahydro–1,3–thiazoles derivatives, have received much attention. Besides well known fungicidal or anticorrosive activity

*Corresponding author

of thiazoles, recent investigations reveal a spectrum of additional activities, such as antioxidant, antiviral, antimicrobial, and antiproliferative activities.[1–6]

Lipophilicity is one of the important physico-chemical parameters determining the activity. Hence, the estimation of the lipophilic character of new, potentially biologically active compounds is regarded as one of the first parameters to be determined at the earliest possible opportunity. It has long been recognized that the retention of a compound in reversed-phase liquid chromatography (RP-LC) is governed by its lipophilicity and thus chromatography may be successfully used for the determination of lipophilicity.[7–9] Although thin-layer chromatography (TLC) is a relatively old LC technique, it offers several practical advantages compared to the traditional shake-flask method, including a rapid and simple way of lipophilicity determination, reproducibility, broader dynamic range, insensitivity to impurities or degradation products, and a wide choice of adsorbents and solvents.

The partitioning (chromatographic) process is a result of characteristic properties of the molecule, which may be quantified using a variety of molecular descriptors. Such descriptors can be created for any compound, and provide a wide range of information about molecular structure and properties to be encoded. Through mathematical relationships, it is possible to describe how a molecular structure, represented by the descriptors, affects the biological activity of a compound.[10–13] These relationships are known as quantitative structure–activity relationships (QSAR). In the same way, quantitative structure–retention relations (QSRR) relate descriptors to chromatographic retention data.[14–17] The QSRR can be applied: (1) for prediction of retention; (2) for identification of the most revealing structural descriptors (regarding properties); (3) to gain insight into the retention mechanism of separation; (4) for evaluation of complex physicochemical properties of analytes, other than chromatographic properties, e.g., lipophilicity; and (5) for prediction of relative biological activities.

In the case of TLC, the QSRR studies are usually based on R_M value, which was defined by Bate-Smith and Westall[18] equation:

$$R_M = \log\left(\frac{1}{R_F} - 1\right) \tag{1}$$

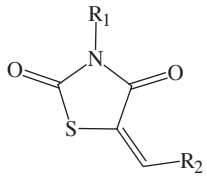
where R_F is the retardation factor, defined as the ratio of the distance travelled by the centre of the spot to the distance simultaneously travelled by the mobile phase. The R_M values are related to the lipophilicity through the following equation:

$$R_M = R_M^0 - b\varphi \tag{2}$$

where φ stands for the volume fraction of the organic component in the mobile phase, and b is the slope. In Eq. (2) the R_M^0 correspond to the retention extrapolated to pure water and represent a chromatographic lipophilicity parameter, which is commonly used as quantitative TLC retention descriptors in QSRR.

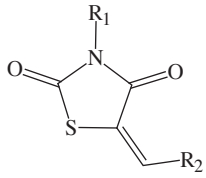
In this paper, particular attention was given to a newly synthesized 3-(4'-bromobenzilidene)-2,4-dioxotetrahydro-1,3-thiazole derivatives. The chromatographic lipophilicity parameters of the investigated compounds were determined using thin-layer chromatography on rice starch stationary phase and 3 different mobile phases. Obtained lipophilicity parameters were compared. To enable better understanding of retention, the partial least-squares regression was used to quantify differences between the observed data and molecular descriptors generated for each of the compounds. This showed the effect of different interactions between the analyte and the mobile phases used and enabled assessment of the suitability of PLS for this particular QSRR study.

Table 1. The investigated compounds.



Compound	R_1	R_2	ACD/logP
1A	H	C_6H_3-2'-F,6'-Cl	2.54
2A	H	1'-$C_{10}H_7$	3.57
3A	H	C_6H_4-4'-OCH_3	2.28
4A	H	C_6H_4-2'-OH	1.60
5A	H	2'-C_4H_3S	2.02
6A	H	C_6H_4-4'-$N(CH_3)_2$	2.45
7A	H	C_6H_4-4'-Br	3.11
8A	H	C_6H_3-3',4'-OCH_3	2.17
9A	H	C_6H_4-4'-$CH(CH_3)_2$	3.68
10A	H	C_6H_4-4'-OC_2H_5	2.82
11A	H	C_6H_5	2.34
12A	H	-	-0.54
1B	$COOC_2H_5$	C_6H_3-2'-F,6'-Cl	2.91
2B	$COOC_2H_5$	1'-$C_{10}H_7$	3.91
3B	$COOC_2H_5$	C_6H_4-4'-OCH_3	2.65
4B	$COOC_2H_5$	C_6H_4-2'-OH	1.96
5B	$COOC_2H_5$	2'-C_4H_3S	2.38
6B	$COOC_2H_5$	C_6H_4-4'-$N(CH_3)_2$	2.81
7B	$COOC_2H_5$	2'-C_8H_5-N-$COOC_2H_5$	2.14
8B	$COOC_2H_5$	C_6H_3-3',4'-$(OCH_3)_2$	2.53
9B	$COOC_2H_5$	C_6H_4-4'-$CH(CH_3)_2$	4.04
10B	$COOC_2H_5$	C_6H_5	2.70
1C	CO-2'-C_4H_3O	C_6H_3-2'-F,6'-Cl	2.84
2C	CO-2'-C_4H_3O	1'-$C_{10}H_7$	3.86
3C	CO-2'-C_4H_3O	-	-0.25
4C	CO-2'-C_4H_3O	C_6H_4-2'-OH	1.89
5C	CO-2'-C_4H_3O	2'-C_4H_3S	2.31
6C	CO-2'-C_4H_3O	C_6H_4-4'-$N(CH_3)_2$	2.74
7C	CO-2'-C_4H_3O	2'-C_4H_2O-5'-CH_3	2.25
8C	CO-2'-C_4H_3O	C_6H_3-3',4'-$(OCH_3)_2$	2.46
9C	CO-2'-C_4H_3O	C_6H_4-4'-$CH(CH_3)_2$	3.97
10C	CO-2'-C_4H_3O	C_6H_3-3',4'-OCH_2O	2.49
11C	CO-2'-C_4H_3O	C_6H_5	2.63

Table 1. Continued.



Compound	R_1	R_2	ACD/logP
1E	COO-CH_2-C_6H_5	-	1.07
2E	COO-CH_2-C_6H_5	C_6H_4-4'-$N(CH_3)_2$	4.05
3E	COO-CH_2-C_6H_5	1'-$C_{10}H_7$	5.18
4E	COO-CH_2-C_6H_5	C_6H_4-2'-OH	3.21
5E	COO-CH_2-C_6H_5	C_6H_3-3',4'-OCH_2O	3.81
6E	COO-CH_2-C_6H_5	C_6H_5	3.95
7E	COO-CH_2-C_6H_5	3'-C_4H_3S	3.62
8E	COO-CH_2-C_6H_5	C_6H_3-3',4'-$(OCH_3)_2$	3.77
9E	COO-CH_2-C_6H_5	C_6H_4-4'-Br	4.72
10E	COO-CH_2-C_6H_5	2'-C_4H_2O-5'-CH_3	3.57
11E	COO-CH_2-C_6H_5	C_6H_4-4'-$OCH_2C_6H_5$	5.55
12E	COO-CH_2-C_6H_5	C_6H_4-4'-OC_2H_5	4.42

Experimental

The compounds investigated (listed in Table 1) were dissolved in acetone at a concentration of 10 mg•mL^{-1}.

TLC was performed on 20×20 cm glass-backed plates of rice starch prepared in our laboratory, with the addition of 0.2% fluorescence indicator F_{254} (Merck). Layers were prepared as described elsewhere.[19] Samples were spotted on the plates by means of a micropipette.

The mobile phases used were mixtures of aqueous ammonia - organic modifier (methanol, acetone, or dioxane). The volume fraction of organic modifiers range between 0% and 24% (v/v) in steps of 4%. Chromatograms were developed by the ascending technique, at room temperature, without previous saturation of the chamber with the solvent. After being developed, the dried plates were examined under UV illumination (λ = 254 nm). The R_M^0 and b values of Eq. (2) are listed in Table 2.

Molecular descriptors and partial least square analysis

Molecular descriptors were calculated using QSAR and SciLogP option of the following molecular modelling computer programs: ALCHEMY 2000[20], Interactive Analysis LogP and LogW predictor website,[21] and SC ChemDraw software.

PLS was performed using Statistica 7.1 software[22] using non-iterative partial least squares (NIPLAS) algorithm, optimized by the cross-validation procedure. The models were interpreted with the help of VIPs (variable importance in the projection).

Table 2. The values of R_M^0 and b of linear equation $R_M = R_M^0 - b\varphi$ calculated for 3 different mobile phase modifiers.

	methanol		dioxane		acetone	
Comp.	R_M^0	b	R_M^0	b	R_M^0	b
1	0.341	1.481	0.222	3.651	0.259	3.347
2	0.649	1.969	0.537	4.660	0.433	3.875
3	0.512	1.794	0.449	4.013	0.480	4.265
4	0.090	0.588	-0.230	0.669	-0.321	1.396
5	0.547	1.312	0.391	3.738	0.348	3.020
6	0.612	1.461	0.482	4.768	0.360	3.279
7	0.978	2.524	0.740	5.165	0.631	3.882
8	0.275	1.435	0.108	3.475	0.017	2.884
9	0.871	2.426	0.772	4.998	0.808	4.997
10	0.669	2.056	0.501	5.083	0.428	4.047
11	0.293	0.774	0.313	3.068	0.241	2.767
12	-	-	-	-	-0.301	0.359
13	0.470	1.935	0.332	3.482	0.343	4.812
14	0.726	2.391	0.558	4.435	0.555	4.906
15	0.464	1.195	0.238	2.242	0.238	3.449
16	0.089	0.524	-0.171	0.657	-0.198	2.109
17	0.650	2.186	0.463	3.961	0.439	4.802
18	0.656	1.699	0.527	4.108	0.508	3.993
19	0.586	2.137	0.453	3.044	0.471	4.198
20	0.348	2.150	0.171	3.809	0.020	2.247
21	0.979	2.802	0.856	5.072	0.763	4.449
22	0.469	1.998	0.375	3.712	0.341	3.697
23	0.300	1.160	0.301	3.006	0.335	4.230
24	0.564	1.834	0.548	4.234	0.625	5.098
25	-0.164	0.521	-0.103	2.329	-0.120	1.128
26	-0.016	0.400	-	-	-0.166	0.849
27	0.212	0.778	0.172	1.568	0.425	3.703
28	0.415	1.390	0.532	3.913	0.528	3.585
29	0.470	1.538	0.455	3.241	0.182	2.749
30	0.119	0.814	0.124	2.867	0.177	2.917
31	0.833	2.250	0.821	4.896	0.889	4.965
32	0.673	1.749	0.488	3.700	0.535	3.562
33	0.347	1.010	0.317	2.762	0.349	3.008
34	-0.023	0.704	-0.353	1.660	-0.533	0.270
35	0.742	2.117	0.634	6.666	0.472	3.733
36	0.724	2.751	0.606	6.185	0.553	4.615
37	0.122	0.795	-0.096	4.450	-0.360	0.439
38	0.777	2.200	0.684	5.743	0.526	4.004
39	0.401	1.509	0.382	4.862	0.241	2.687
40	0.431	1.524	0.349	3.584	0.334	2.906
41	0.319	1.974	0.158	4.380	0.050	2.899
42	0.292	1.421	0.863	7.394	0.644	3.911
43	0.288	0.526	0.290	4.414	0.134	1.729
44	0.835	3.078	0.837	9.264	1.089	5.438
45	0.708	2.271	0.618	7.018	0.497	4.451

Results and Discussion

Liquid chromatography, especially TLC, offers a variety of stationary and mobile phases that can be used for lipophilicity determination. Unfortunately, due to the variability of experimental conditions, a universal lipophilicity scale does not exist. Table 2 lists the chromatographic lipophilicity parameters, R_M^0, of the investigated compounds (Eq. (2)). It can be seen that they significantly differ between chromatographic systems. More precisely, the retention changes affected by mobile phase composition are less sensitive to organic modifier volume fraction variation when this is methanol rather than dioxane or acetone; the mean slope of Eq. (2), b values are 1.622, 4.092, and 3.326 for methanol, dioxane, and acetone, respectively. Additionally, a different R_M^0 values were obtained for methanol (average value 0.470) compared to dioxane (average value 0.389) or acetone (average value 0.317). These differences in the mean slope and intercept values reflect the chemical nature of organic modifiers used. Lower slope values in the case of methanol as a modifier may be due to a stronger preferential adsorption of methanol on stationary phase and a higher affinity of analyte to the methanol-solvated stationary phase than to the acetone or dioxane-solvated stationary phase.

In typical partition RP systems, R_M^0 and b values are usually in linear correlation. We found that such correlation stands and for unconventional rice starch TLC support, Eqs. (3-5):

$$R_M^0(\text{acetone}) = -0.463(\pm 0.056) + 0.235(\pm 0.016)b \tag{3}$$

$$(r = 0.9159; SD = 0.1384; n = 45)$$

$$R_M^0(\text{dioxane}) = -0.168(\pm 0.076) + 0.136(\pm 0.017)b \tag{4}$$

$$(r = 0.7760; SD = 0.1873; n = 43)$$

$$R_M^0(\text{methanol}) = -0.102(\pm 0.050) + 0.353(\pm 0.030)b \tag{5}$$

$$(r = 0.88737; SD = 0.12829; n = 44)$$

Equations (3-5) are the confirmation of the partition retention mechanism that is one of the main circumstances for chromatographic determination of lipophilicity.

Since in chromatography linear addition is usually expected, methods such as PLS may provide insight into the complex interactions that occur between a solute and the mobile and stationary phases. Since the retention (and hence lipophilicity) of each solute depends on structural features of the molecule (beside the chromatographic system used), it is convenient to apply various molecular descriptors to express different molecular properties. For that purpose a set of 37 descriptors were calculated. Since some of them were clustered in 2D and 3D loading plots (which means that they describe similar information), for the subsequent calculations, number of descriptors was reduced by choosing a representative one for each cluster of variables.

PLS analysis of retention data

The reduced data set of molecular descriptors pointing to the physicochemical properties of the compounds were used as the input data for PLS analysis. The PLS modelling yielded 2 significant PLS component models for acetone and dioxane as modifiers. Statistical data for PLS model are presented in Table 3. Note that the

data for methanol as mobile phase modifier are not shown in Table 3 because the PLS model does not give statistically significant results for methanol.

Table 3. Partial least squares analysis summary for acetone and dioxane modifiers.

Component	R^2X(Cumul.)	Eigenvalues	R^2Y(Cumul.)	Significance	Iterations
	Acetone				
1	0.375	2.309	0.447	S	1
2	0.167	1.257	0.694	S	1
	Dioxane				
1	0.2937	1.6380	0.5158	S	1
2	0.4903	1.2811	0.6840	S	1

Calculated PLS model gives a good correlation between the experimental and calculated retention data (Figure 1) with R^2 values of 0.9070 and 0.8982 for acetone and dioxane, respectively. Both regression lines have intercept almost equal to zero.

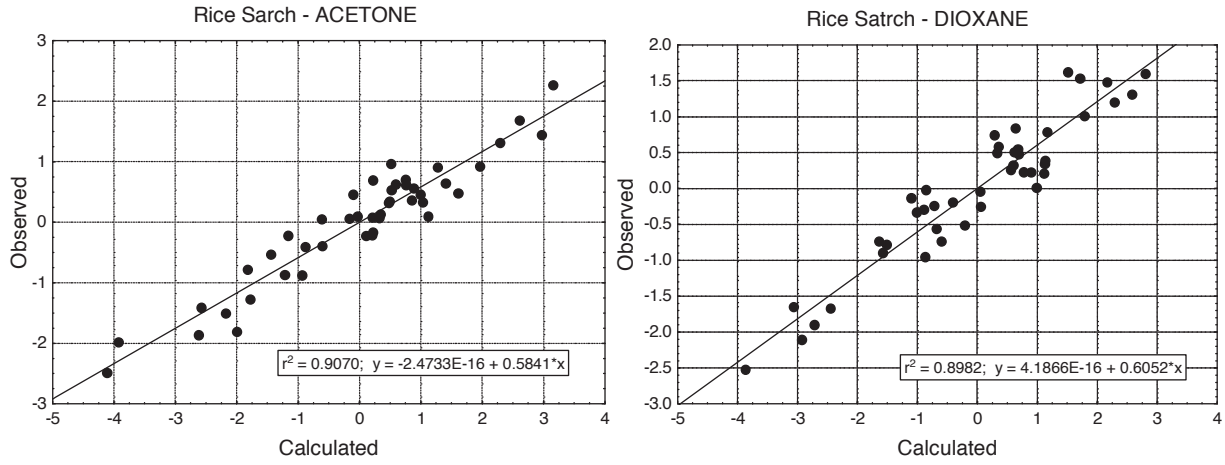


Figure 1. Correlation between the retention determined by TLC and predicted values by PLS method for 2 different mobile phase modifiers.

By means of PLS, we were able to identify the most significant molecular characteristic (descriptor) that is important for the retention. Descriptors were described by its VIP, which is normalized so that they can be compared (Table 4). Descriptors with higher VIPs are more significant for the retention of studied compounds in a particular mobile phase.

Table 4 lists the VIPs of the descriptors. As a result of similar molecular interactions between thiazoles from one side and acetone and dioxane mobile phase modifiers from the other side, similar descriptors were identified as most important for the retention. However, certain differences exist between modifiers. For example, with acetone as modifier, descriptor with higher VIP is molecular polarizability. Molecular polarizability (Polar) represents a measure of the overall electronic charge distribution and also depends on the molecular

volume. However, with dioxane as modifier, the polarizability is not selected as important for retention. On the other hand, with both modifiers, the largest negative charge over the atoms of the molecule (*MaxQneg*) and dipole moment of a molecule are of importance for retention.

In general, physico-chemical properties that highly influence the retention are lipophilicity and electronic characteristics of the molecule. This is not surprising because the thiazoles are specific solutes that are capable of polar interactions.[23]

Table 4. Variable importance (VIP) of selected descriptors for acetone and dioxane modifiers.

	Acetone		Dioxane	
Descriptor*	VIP	importance	VIP	importance
Polar	0.840	1	-	-
Dipole	0.750	2	0.833	1
logP	0.694	3	0.729	3
logW	0.623	4	0.536	4
MaxQneg	0.588	5	0.744	2
Gibbs	0.398	7	0.303	5
ABSQon	0.287	6	0.139	7
HOMO	0.158	8	0.147	6

*Polar (polarizability of the molecule), Dipol (dipole moment of the molecule), logP (the octanol–water partition coefficient), log W (logarithm of water solubility), MaxQneg (the largest negative charge over the atoms of the molecule), Gibbs (standard Gibbs-free energy), ABSQon (the sum of absolute values of charges on the nitrogens and oxygens in the molecule), and HOMO (highest occupied molecular orbital energy).

Conclusions

The chromatographic lipophilicity parameters of 45 investigated thiazole derivatives on TLC rice starch depends on the mobile phase modifier used. Intercept and slope values of the linear relationship between retention and the volume fraction of the modifier in mobile phase were in mutual correlation. This represents an important indication of partitioning and the main requirement for chromatographic determination of lipophilicity, and leads to 3 chromatographic lipophilicity scales. Using partial least-squares (PLS) regression, the comparison of those lipophilicity scales was possible. Molecular properties that are important for the retention (and hence for the lipophilicity) were quantified by various molecular descriptors. Each lipophilicity scale was influenced by a certain molecular property which is dominant for dioxane (charge dependent (logP, Dipole)) and for acetone (3D structure dependent descriptors (i.e. molecular polarizability)). The proposed model might also be useful for the prediction of the lipophilicity of the investigated compounds.

This proves that the chromatographic lipophilicity scales obtained with different mobile phase modifiers are not the same. However, the question about which lipophilicity parameters are most reliable for QSAR study still remains unanswered. The answer may be given by future regression models employed in a QSAR study.

Acknowledgment

This work was performed within the framework of the research project: Investigation of the Synthesis, Structure, and Activity of Natural and Synthetic Organic Compounds (Grant number 1420632), supported by the Ministry of Science of Serbia."

References

1. A. Andreani, M. Granaiola, M. Guardigli, A. Leoni, A. Locatelli, R. Morigi, M. Rambaldi and A. Roda, **Eur. J. Med. Chem**. **40**, 1331-34 (2005).
2. H-L. Liu1, Z. Li1 and T. Anthonsen, **Molecules 5**, 1055-61 (2000).
3. Gy. Vastag, E. Szöcs, A. Shaban and E. Kálmán, **Pure Appl. Chem 73**, 1861-9 (2001).
4. J. J. Harnett, V. Roubert, C. Dolo, C. Charnet, B. Spinnewyn, S. Cornet, A. Rolland, J.-G. Marin, D. Bigg and P. Chabrier, **Bioorg. Med. Chem. Lett. 14** 157-60 (2004).
5. P. Vicini, A. Geronikaki, M. Incerti, B. Busonera, G. Poni, C. A. Cabras and P. La Colla, **Bioorg. Med. Chem. 11**, 4785-9 (2003).
6. K.-Y. Jung, S.-K. Kim, Z.-G. Gao, A. S. Gross, N. Melman, K. A. Jacobson and Y.-C. Kim, **Bioorg. Med. Chem. 12**, 613-23 (2004).
7. R. Kaliszan, **Chem. Rev. 107**, 3212–3246 (2007).
8. A. Nasal, D. Siluk, R. Kaliszan, **Curr. Med. Chem. 10**, 381-426 (2003).
9. S. K. Poole, C. F. Poole, **J. Chromatogr. B 797**, 3–19 (2003).
10. A. Leo and C. Hansch, D. Elkins, **Chem. Rev. 71,** 525-616 (1971).
11. W. J. Lyman, W. F. Reehl and D.H. Rosenblatt, "**Handbook of Chemical Property Estimation Method**", McGraw-Hill, New York, 1982.
12. R. F. Rekker and R. Mannhold, "**Calculation of Drug Lipophilicity**", VCH, Weinheim, 1992.
13. A. Chiriac, D. Ciubotariu and Z. Simon, "**QSAR-Quantitative Relationships**", Mirton Publishing House, Timişoara, 1995.
14. R. Kaliszan, **Anal. Chem. 64**, 619-31 (1992).
15. L.-G. Danielsson and Y.-H. Zhang, **Trends Anal. Chem. 15**, 188-96 (1996).
16. R. Kaliszan, **Trends Anal. Chem. 18**, 401-16 (1999).
17. M. C. Garcia Alvarez-Coque and J. R. Torres Lapasio, **Trends Anal. Chem. 18**, 533-43 (1999).
18. E. C. Bate-Smith, R. G. Westall, **Biochim. Biophys. Acta 4**, 427–440 (1950).
19. N. U. Perišić-Janjić, T. Lj. Djaković and S. M. Petrović, **Chromatographia 40**, 96-8 (1995).
20. http://www.tripos.com/software/alchemy.html
21. http://www.logp.com
22. http://www.statsoft.com
23. E. Brzezinska, G. Koska and K. Walczynski, **J. Chromatogr. A 1007**, 145-55 (2003).