A comparative Study of Molecular Lipophilicity Indices of Some Formyl- and Acetylpyridine-3-Thiosemicarbazone Derivatives and Various Computed Descriptors

CRISTINA ONIŞOR[1], DIMITRA KOVALA-DEMERTZI[2], MAVROUDIS A. DEMERTZIS[2], COSTEL SARBU [1]*

[1]Babes-Bolyai University, Faculty of Chemistry and Chemical Engineering, 11 Arany Janos, 400028 Cluj-Napoca, Romania;

[2]University of Ioannina, Department of Chemistry, Sector of Inorganic and Analytical Chemistry, 45110 Ioannina, Greece

Retention indices for formyl- and acetylpyridine-3-thiosemicarbazone derivatives were determined by RP-HPTLC experiments and were performed with methanol–water mixtures in different volume proportions as mobile phase on two stationary phases of different polarity: RP-18F$_{254s}$ and RP-18WF$_{254s}$ plates. The correlation coefficients obtained on the RP-18W stationary phase are somewhat higher than those obtained on the RP-18 plates and it can be concluded that the methanol concentration in the mobile phase have a larger influence on the interactions of the compounds with the RP-18 stationary phase than with the RP-18W stationary phase. In order to describe the relationship between the chromatographic retention indices of investigated compounds (RMo, b, ϕ_o and scores corresponding to the first principal component) and the computed physico-chemical and topological descriptors, a multivariate regression analysis was performed. Excellent structure–retention index models show the suitability of these descriptors in the structure–retention index correlation. In all cases, the path/walk 5 Randic shape index, the highest occupied and the lowest unoccupied molecular orbital play a dominant role in determining retention indices on both stationary phases with different polarities, and RP-18W phase denotes the contribution of the specific polarizability.

*Key words**:** lipophilicity**,** molecular descriptors, multiple linear regression analysis, PCA*

The partition coefficient between n-octanol and water, *P*, is extensively used in environmental and biomedical sciences as a descriptor of the lipophilicity properties of different compounds [1, 2]. The lipophilicity of a substance affects its biological activity. As such, relationships between biological activities and log *P* can be demonstrated [3-5]. Reversed-phase high performance thin-layer chromatography (RP-HPTLC) is an alternative method for estimation of lipophilicity [6].

Correlation between thin layer chromatographic retention indices and molecular structure (QSRR) can provide more profound insights into interactions between eluents and stationary phases from a theoretical viewpoint, and also give very important information about the effect of chemical structures on the retention behaviour and possible mechanism of absorption and elution. A large number of structure–retention index correlations have been developed for different compounds using various known physico-chemical parameters and molecular descriptors such as geometric, electronic or electrostatic, polar, steric, and topological ones. Among these molecular descriptors, topological indices are particularly of interest because they can be readily calculated directly from the molecular structures [7, 8-14].

In the last decades, considerable attention has been given to thiosemicarbazones and on first row of transition metal complexes with such ligands due to their interesting biological activities: anticarcinogenic, antibacterial, anti-HIV, anticancer, fungicides, antiviral, antifungal, antitumour, anti-inflammatory, antiparasitic, antituberculosis and antileukemic properties [15-18]. In this field, the 2-formylpyridinethiosemicarbazones (HPATS) probably has gotten the most attention due to its marked antitumour properties [19].

The objective of this work was to analyze and discuss the correlations found between the different chromatographic retention indices of tested compounds (RMo, *b*, ϕ_o and the scores corresponding to the first principal component) and the calculated descriptors obtained through different software's of the formyl- and acetylpyridine-3-thiosemicarbazone derivatives. In addition, the results obtained in this study may be used in interpreting the molecular mechanism of interactions between eluents and stationary phases with different polarities and to explain the chromatographic behaviour of compounds.

Experimental part

Chemicals

The formyl- and acetylpyridine-3-thiosemicarbazone derivatives and their palladium and zinc complexes (table 1) were synthesized in the Department of Chemistry, Section of Inorganic and Analytical Chemistry, University of Ioannina, (Greece). Their chemical structure has been investigated by classical methods including elemental analysis and NMR [20-22]. These compounds can be classified in two major groups: the first one having the amino- and ethyl-amino N-terminal of the thiosemicarbazone derivatives, with OH and CH_3 groups attached on the pyridinic nucleus; the second one having azepane, pyridine-piperazine and cyclohexane N-terminal of the thiosemicarbazone derivatives, with OH and CH_3 groups attached on the pyridinic nucleus.

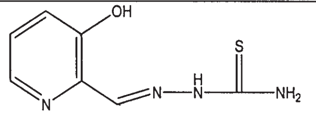
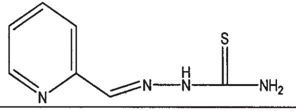
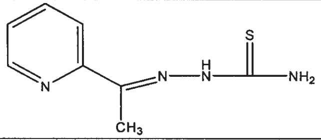
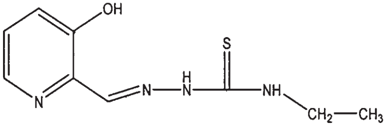
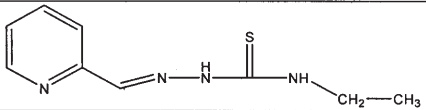
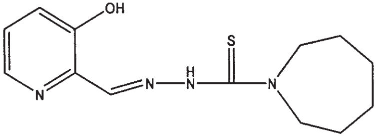
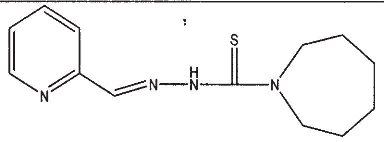
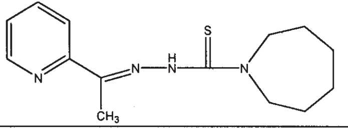
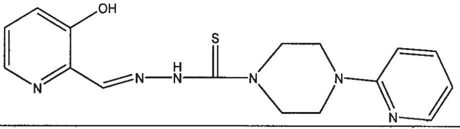
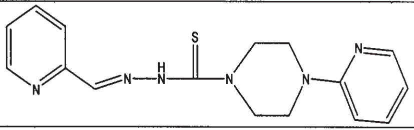
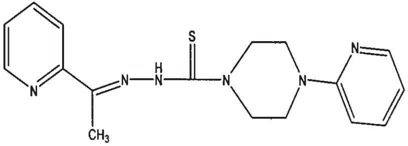
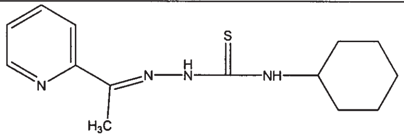
Chromatography

The chromatographic behaviour of the formyl- and acetylpyridine-3-thiosemicarbazone derivatives was studied on two stationary phases: RP-18F$_{254S}$ (20x20 cm)

---

* *email: costelsrb@yahoo.co.uk*

Table 1
CHEMICAL STRUCTURE OF FORMYL- AND ACETYLPYRIDINE-3-THIOSEMICARBAZONE DERIVATIVES

Cpds	Chemical structure
1	
2	
3	
4	
5	
6	
7	
8	
9	
10	
11	
12	

and RP-18WF$_{254S}$ (10x20 cm) silica gel bounded plates from Merck (Darmstadt, Germany). RP-18 stationary phases are still the most popular and most widely used phases for RP-HPTLC, due to their ability to separate a wide range of solutes with good resolution, selectivity and efficiency. RP-18W plate differs from the RP-18 plate by the inclusion of a polar end-capping agent. The polar end-capping agent increases the polarity near the surface of the silica, and therefore should allow better penetration of highly aqueous mobile phases into the stationary phase [23].

The compounds were dissolved in methanol, reaching a final concentration of 3mg/mL in the corresponding solvent. A 2μL aliquot of each solution was applied manually to the origin of the plates by means of a 10μL Hamilton (Switzerland) microliter syringe. Chromatography was performed in a normal developing chamber at room temperature (~20 °C), the developing distance being 8.5 cm in all cases. A standard mobile phase (methanol-water) was used, methanol in the concentration range 25-45% (v/v) for RP-18F$_{254S}$ plate and 15-35% for RP-18WF$_{254S}$ plate, in steps of 5% in each case, as the studied compounds differed considerably in their retention. Each measurement was performed in duplicate.

After development, the plates were dried in air and colored zones appeared on a colorless background and fluorescent blue-orange under UV lamp (λ = 254 nm).

Structural and topological descriptors

All of the molecules were drawn into Hyperchem [24] and optimized using the MM+ molecular mechanics force field. The optimized geometries were loaded into the ALCHEMY 2000 program, and Chem3D Ultra 8.0 and DRAGON Plus version 5.4 software's and in order to define the character of the compounds structure, various descriptors were taken into consideration and used as independent variables.

We derived a set of 112 descriptors from which 86 descriptors were given by the software Dragon 5.4 [25] that include 74 topological descriptors and 12 molecular descriptors, and the other 26 additional descriptors that take into account the number of functional groups and their position in the molecule, together with physical or physico-chemical and quantum-chemical descriptors such as molecular dipole moment, total charges over atoms, and HOMO-LUMO energies (provided by Alchemy 2000 [26] and ChemDraw Ultra 8.0) [27].

Methods

RP-HPTLC

RP-HPTLC provides a variety of descriptors that can be used as lipophilicity estimators. The most popular lipophilicity indices measured by RP-HPTLC are derived from the retardation factor, R_F according to the following formula [28, 29]:

$$R_M = R_{Mo} + bC \quad (1)$$

where:

$$R_M = \log(1/R_F - 1) \quad (2)$$

Another retention-related parameter has been introduced recently, the isocratic chromatographic hydrophobicity index, ϕ_o. The isocratic chromatographic hydrophobicity index, ϕ_o, represents the volume fraction of organic solvent in the mobile phase for which the amount of solute in the mobile phase is equal to that in the stationary phase, and in RP-HPTLC parameter ϕ_o is calculated as the ratio of the intercept and slope [30-31]:

$$\varphi_0 = \frac{R_{M_0}}{b} \quad (3)$$

In addition, we used also the lipophilicity scale obtained by applying Principal Component Analysis (PCA) directly to the obtained R_F values resulted for all compounds and combinations of methanol-water and the scores and loadings of the covariance matrix were obtained. Thus, a significant one component PC1 model explaining 93.66% and 95.53% of the total variance for RP-18F$_{254S}$, respectively

Table 2

R_F VALUES OF INVESTIGATED COMPOUNDS FOR THE CORRESPONDING VOLUME FRACTION OF METHANOL ON RP-18F_{254S} AND RP-18WF_{254S} SILICA GEL BOUNDED PLATES

Cpd	R_F values on RP-18F_{254s}					R_F values on RP-18WF_{254s}				
	25%	30%	35%	40%	45%	35%	30%	25%	20%	15%
1	0.376	0.447	0.588	0.665	0.671	0.765	0.612	0.528	0.188	0.253
2	0.394	0.471	0.602	0.885	0.753	0.747	0.606	0.539	0.353	0.194
3	0.282	0.424	0.574	0.853	0.718	0.706	0.559	0.518	0.306	0.141
4	0.171	0.282	0.456	0.764	0.694	0.659	0.456	0.382	0.082	0.088
5	0.129	0.271	0.435	0.733	0.659	0.606	0.441	0.382	0.176	0.059
6	0.029	0.106	0.212	0.506	0.512	0.253	0.112	0.071	0.047	0.012
7	0.053	0.053	0.129	0.412	0.488	0.318	0.124	0.065	0.071	0.021
8	0.029	0.101	0.194	0.500	0.294	0.447	0.253	0.194	0.153	0.024
9	0.021	0.067	0.151	0.371	0.376	0.200	0.120	0.081	0.076	0.024
10	0.036	0.084	0.175	0.424	0.406	0.359	0.164	0.118	0.035	0.012
11	0.072	0.065	0.153	0.399	0.485	0.340	0.053	0.035	0.024	0.024
12	0.018	0.053	0.112	0.371	0.441	0.353	0.218	0.112	0.024	0.011

Table 3

MOLECULAR LIPOPHILICITY INDICES OF THE LINEAR REGRESSION EQUATION 1 AND THE REGRESSION COEFFICIENT

Cpd	RP-18F_{254s}					RP-18WF_{254s}				
	R_{Mo}	*b*	φ_0	$PC1_{RF}$	r	R_{Mo}	*b*	φ_0	$PC1_{RF}$	r
1	0.92	-0.03	-31.89	-0.434	**0.966**	1.47	-0.06	-26.25	-0.523	**0.9304**
2	1.33	-0.05	-29.27	-0.600	0.8352	1.34	-0.05	-25.62	-0.547	**0.9878**
3	1.61	-0.05	-32.02	-0.495	0.8736	1.52	-0.06	-27.27	-0.464	**0.9738**
4	2.16	-0.06	-36.06	-0.277	**0.9442**	2.20	-0.07	-30.78	-0.241	**0.9519**
5	2.30	-0.06	-37.11	-0.217	**0.9449**	2.07	-0.07	-30.99	-0.222	**0.9712**
6	3.41	-0.08	-42.49	0.180	**0.9712**	2.80	-0.07	-42.31	0.363	**0.9776**
7	3.20	-0.07	-44.87	0.295	**0.9567**	2.49	-0.06	-42.53	0.317	**0.9534**
8	2.95	-0.06	-45.75	0.276	0.8842	2.37	-0.07	-35.70	0.084	**0.9317**
9	3.47	-0.08	-45.59	0.352	**0.9725**	2.17	-0.04	-48.30	0.372	**0.9516**
10	3.08	-0.07	-45.10	0.288	**0.9698**	3.08	-0.08	-37.78	0.255	**0.9885**
11	2.85	-0.06	-45.24	0.278	**0.9534**	2.76	-0.06	-45.54	0.372	0.8623
12	3.86	-0.09	-44.86	0.355	**0.9859**	3.30	-0.09	-36.87	0.234	**0.9875**

R_{Mo}-molecular lipophilicity index; *b*-specific hydrophobic surface area; C-concentration of the organic component of the mobile phase; φ_0-isocratic chromatographic hydrophobicity index; $PC1_{RF}$-scores corresponding to the first component of the R_F values

RP-18WF_{254S} stationary phase. The scores corresponding to the first principal component appeared to be the best solution for the lipophilicity scale resulted from retention data [32, 33].

Multiple Linear Regression

An important stage of the multilinear regression in QSAR/QSPR/QSRR methodology is the search for the best multilinear equation among a given pool of descriptors. The strategy used to develop physically meaningful multilinear QSAR/QSPR/QSRR equations from the very large pool of descriptors is a combination of the multilinear regression and forward stepwise selection procedures (descriptors with insignificant variance were rejected according to the statistical criteria). This helps to speed up the descriptor selection and reduces the probability of including irrelevant descriptors by chance. The best multilinear equation selects the best two-parameter regression equation, the best three-parameter regression equation, etc., based on the statistical criteria in the stepwise regression procedure [34]. Statistical analysis of the results was performed using the Statistica 7.0 program [35].

Results and discussion

To analyze the relationships between molecular lipophilicity and the structure of the molecule and various molecular properties of formyl- and acetylpyridine-3-thiosemicarbazone derivatives, we have tested the topological indices along with the classical physico-chemical properties.

The R_F values of each compound are listed in table 2 and table 3 presents the chromatographic indices estimated from retention data and shows that the correlation coefficients concerning chromatographic behaviour obtained on the RP-18WF_{254S} stationary phase are somewhat higher than those obtained on the RP-18F_{254S} plates (see the r values). This means that the compounds enter into stronger interactions with the more nonpolar RP-18F_{254S} stationary phase. Also, from the larger variation in R_{Mo} on the RP-18F_{254S} stationary phase, it can be concluded that the structure of the formyl- and acetylpyridine-3-thiosemicarbazone derivatives as well as the methanol

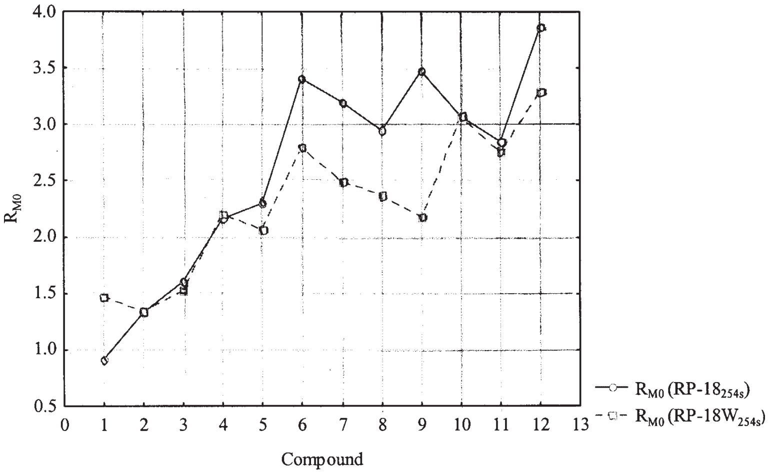


Fig 1. Profiles of R_{M0} values of investigated compounds on RP-18F_{254s} and RP-18WF_{254s} stationary phase

Table 4
TABLE OF CORRELATION CONSIDERING PHYSICOCHEMICAL DESCRIPTORS COMPUTED WITH ALCHEMY 2000 AND CHEMDRAW ULTRA PLUS SOFTWARE'S

Variable	RP-18F_{254s}				RP-18WF_{254s}			
	R_{Mo}	*b*	ϕ_o	$PC1_{RF}$	R_{Mo}	*b*	ϕ_o	$PC1_{RF}$
X1	**0.79**	**-0.66**	**-0.90**	**0.90**	**0.74**	-0.17	**-0.91**	**0.91**
VX0	**0.84**	**-0.72**	**-0.94**	**0.94**	**0.79**	-0.24	**-0.91**	**0.93**
VX1	**0.88**	**-0.76**	**-0.96**	**0.96**	**0.82**	-0.27	**-0.91**	**0.96**
LogPAlch	**0.75**	**-0.67**	**-0.79**	**0.79**	**0.86**	**-0.60**	-0.56	**0.75**
Volume	**0.85**	**-0.73**	**-0.95**	**0.94**	**0.80**	-0.24	**-0.91**	**0.94**
WienI	**0.69**	**-0.56**	**-0.81**	**0.81**	**0.64**	-0.08	**-0.86**	**0.83**
X3	**0.76**	**-0.63**	**-0.88**	**0.88**	**0.71**	-0.13	**-0.90**	**0.89**
MW	**0.79**	**-0.67**	**-0.90**	**0.90**	**0.74**	-0.17	**-0.91**	**0.91**
Sp.Pol	**-0.77**	**0.73**	**0.74**	**-0.74**	**-0.71**	0.46	0.57	**-0.69**
DM	-0.35	0.49	0.12	-0.12	-0.35	0.37	0.06	-0.19
Polar	**0.82**	**-0.70**	**-0.93**	**0.92**	**0.78**	-0.22	**-0.90**	**0.92**
Q+	-0.29	0.17	0.50	-0.46	-0.17	-0.35	**0.61**	-0.54
Ka3	**0.89**	**-0.87**	**-0.84**	**0.83**	**0.83**	-0.49	**-0.71**	**0.82**
ABSQ	0.53	-0.41	**-0.64**	**0.65**	0.48	0.05	**-0.78**	**0.70**
$ABSQ_{ON}$	0.29	-0.19	-0.40	0.41	0.25	0.20	**-0.62**	0.49
Q-	0.06	-0.07	-0.07	0.05	0.11	-0.25	0.15	-0.04
LogPDraw	**0.92**	**-0.83**	**-0.96**	**0.95**	**0.86**	-0.36	**-0.86**	**0.95**
CAA	**0.84**	**-0.72**	**-0.93**	**0.92**	**0.80**	-0.25	**-0.90**	**0.93**
CMA	**0.84**	**-0.73**	**-0.94**	**0.93**	**0.80**	-0.25	**-0.90**	**0.93**
Ovality	**0.69**	**-0.60**	**-0.81**	**0.78**	**0.67**	-0.16	**-0.83**	**0.82**
PartCoeff	**0.60**	**-0.62**	-0.43	0.48	0.52	-0.44	-0.27	0.43
HOMO	**0.83**	**-0.88**	**-0.64**	**0.66**	**0.69**	-0.49	-0.50	**0.62**
LUMO	**0.79**	**-0.70**	**-0.85**	**0.85**	**0.77**	-0.40	**-0.66**	**0.79**
R_{Mo}^{a}	1.00	**-0.97**	**-0.92**	**0.93**	**0.88**	-0.43	**-0.83**	**0.93**
b^{a}		1.00	**0.81**	**-0.82**	**-0.83**	0.46	**0.74**	**-0.85**
ϕ_o^{a}			1.00	**-0.99**	**-0.85**	0.34	**0.88**	**-0.96**
$PC1_{RF}^{a}$				1.00	**0.86**	-0.35	**-0.89**	**0.96**
R_{Mo}^{b}					1.00	**-0.73**	**-0.67**	**0.85**
b^{b}						1.00	-0.02	-0.26
ϕ_o^{b}							1.00	**-0.95**
$PC1_{RF}^{b}$								1.00

[a] lipophilicity indices corresponding to RP-18F_{254s} stationary phase
[b] lipophilicity indices corresponding to RP-18WF_{254s} stationary phase

concentration in the mobile phase have a larger influence on the interactions of the compounds with the RP-18F_{254S} stationary phase than with the RP-18WF_{254S} stationary phase. In the case of RP-18F_{254S} stationary phase the amino- and ethylamino- N-terminal thiosemicarbazone derivatives, OH and CH_3 groups bring a negative, respectively a positive contribution to retention indices, while on RP-18WF_{254S} stationary phase there is a positive contribution of both groups. As for the aromatic N-terminal thiosemicarbazone derivatives, OH and CH_3 groups bring a positive, respectively a negative contribution on both stationary phases (fig. 1).

A visual comparison of the studied compounds demonstrates that the introduction of an additional OH group decreases R_{Mo} values for amino- and ethyl-amino N-terminal thiosemicarbazone derivatives by the formation of an intramolecular hydrogen bond and increases molecular lipophilicity indices for azepane, pyridine-piperazine and cyclohexane thiosemicarbazone derivatives due to the presence of heterocycles. On RP-18 and RP-18W, the R_{Mo} values obtained indicate that compound **9** shows the highest lipophilicity (3.86 and 3.30), whereas compound **1** the lowest lipophilicity (0.92).

Their retention results from the combination of the hydrophobic contribution to the retention which strongly depends upon the nature and size of radicals attached to the N-terminal. The difference in lipophilicity may be seen between compounds **1** and **4**, respectively **2** and **5**. These compounds have the same active thiosemicarbazone units, but while compounds **4** and **5** have the long lipophilic chain, compounds **1** and **3** have no hydrocarbon chain that greatly decreases the lipophilic potential. These differences may be attributed to the highly polar free amino group,

Table 5

TABLE OF CORRELATION CONSIDERING TOPOLOGICAL DESCRIPTORS COMPUTED WITH DRAGON SOFTWARE PROGRAM

Variable	RP-18F$_{254s}$				RP-18WF$_{254s}$			
	R_{Mo}	b	ϕ_o	$PC1_{RF}$	R_{Mo}	b	ϕ_o	$PC1_{RF}$
ZM1	**0.76**	**-0.63**	**-0.88**	**0.88**	**0.71**	-0.14	**-0.90**	**0.89**
ZM1V	**0.56**	-0.43	**-0.69**	**0.69**	0.52	0.03	**-0.81**	**0.74**
ZM2	**0.75**	**-0.62**	**-0.87**	**0.87**	**0.70**	-0.13	**-0.89**	**0.88**
ZM2V	**0.60**	-0.47	**-0.73**	**0.73**	0.56	-0.01	**-0.82**	**0.77**
Qindex	**0.70**	-0.56	**-0.83**	**0.83**	**0.65**	-0.08	**-0.87**	**0.85**
SNar	**0.79**	**-0.66**	**-0.90**	**0.90**	**0.74**	-0.16	**-0.90**	**0.91**
HNar	**0.83**	**-0.72**	**-0.91**	**0.91**	**0.77**	-0.23	**-0.86**	**0.92**
GNar	**0.83**	**-0.71**	**-0.92**	**0.93**	**0.77**	-0.21	**-0.89**	**0.93**
Xt	**-0.86**	**0.74**	**0.95**	**-0.95**	**-0.81**	0.24	**0.92**	**-0.96**
Dz	**0.76**	**-0.64**	**-0.87**	**0.87**	**0.71**	-0.14	**-0.90**	**0.89**
Ram	**0.60**	-0.47	**-0.73**	**0.74**	0.56	-0.02	**-0.80**	**0.75**
Pol	**0.78**	**-0.64**	**-0.90**	**0.90**	**0.71**	-0.11	**-0.92**	**0.92**
LPRS	**0.77**	**-0.65**	**-0.88**	**0.88**	**0.72**	-0.15	**-0.9**	**0.90**
VDA	**0.74**	**-0.62**	**-0.85**	**0.85**	**0.70**	-0.13	**-0.88**	**0.87**
MSD	**-0.73**	**0.59**	**0.85**	**-0.85**	**-0.67**	0.15	**0.80**	**-0.83**
SMTI	**0.69**	-0.56	**-0.81**	**0.81**	**0.64**	-0.07	**-0.86**	**0.83**
SMTIV	**0.64**	-0.51	**-0.76**	**0.76**	**0.58**	-0.02	**-0.84**	**0.79**
GMTI	**0.68**	-0.55	**-0.80**	**0.80**	**0.63**	-0.07	**-0.86**	**0.83**
GMTIV	**0.58**	-0.46	**-0.70**	**0.70**	0.52	0.03	**-0.81**	**0.74**
Xu	**0.80**	**-0.67**	**-0.90**	**0.90**	**0.75**	-0.19	**-0.91**	**0.91**
SPI	0.20	-0.17	-0.24	0.23	0.13	0.12	-0.38	0.26
W	**0.69**	-0.56	**-0.81**	**0.81**	**0.64**	-0.08	**-0.86**	**0.83**
WA	**0.79**	**-0.67**	**-0.88**	**0.88**	**0.75**	-0.20	**-0.89**	**0.90**
Har	**0.77**	**-0.64**	**-0.89**	**0.88**	**0.72**	-0.15	**-0.90**	**0.90**
Har2	**0.76**	**-0.63**	**-0.88**	**0.88**	**0.71**	-0.13	**-0.90**	**0.89**
QW	**0.72**	**-0.59**	**-0.83**	**0.83**	**0.67**	-0.10	**-0.88**	**0.85**
TI1	**0.64**	-0.51	**-0.77**	**0.77**	**0.60**	-0.03	**-0.83**	**0.80**
TI2	**0.89**	**-0.80**	**-0.93**	**0.93**	**0.83**	-0.30	**-0.91**	**0.95**
STN	**0.75**	**-0.61**	**-0.87**	**0.87**	**0.69**	-0.11	**-0.89**	**0.89**
HyDp	**0.64**	-0.52	**-0.77**	**0.77**	**0.60**	-0.04	**-0.84**	**0.79**
RHyDp	**0.77**	**-0.64**	**-0.88**	**0.88**	**0.72**	-0.14	**-0.90**	**0.90**
W	**0.71**	**-0.58**	**-0.83**	**0.83**	**0.65**	-0.06	**-0.88**	**0.85**
Ww	**0.69**	-0.56	**-0.80**	**0.81**	**0.62**	-0.03	**-0.88**	**0.84**
Rww	**0.72**	**-0.63**	**-0.79**	**0.78**	**0.72**	-0.30	**-0.77**	**0.77**
D/D	**0.72**	**-0.60**	**-0.84**	**0.83**	**0.69**	-0.14	**-0.86**	**0.85**
Wap	**0.59**	-0.46	**-0.72**	**0.72**	0.54	0.02	**-0.82**	**0.75**
WhetZ	**0.70**	**-0.58**	**-0.82**	**0.82**	**0.66**	-0.09	**-0.87**	**0.84**
Whetm	**0.70**	**-0.58**	**-0.82**	**0.82**	**0.66**	-0.09	**-0.87**	**0.84**
Whetv	**0.69**	-0.57	**-0.81**	**0.81**	**0.64**	-0.08	**-0.87**	**0.84**
Whete	**0.70**	**-0.58**	**-0.82**	**0.82**	**0.66**	-0.09	**-0.87**	**0.84**
Whetp	**0.69**	-0.57	**-0.81**	**0.81**	**0.64**	-0.07	**-0.87**	**0.84**
J	**-0.80**	**0.68**	**0.88**	**-0.90**	**-0.74**	0.17	**0.89**	**-0.91**
JhetZ	**-0.86**	**0.74**	**0.93**	**-0.94**	**-0.79**	0.23	**0.91**	**-0.95**
Jhetm	**-0.86**	**0.74**	**0.93**	**-0.94**	**-0.80**	0.23	**0.91**	**-0.95**
Jhetv	**-0.86**	**0.74**	**0.93**	**-0.94**	**-0.80**	0.23	**0.91**	**-0.95**
Jhete	**-0.86**	**0.74**	**0.93**	**-0.94**	**-0.79**	0.23	**0.91**	**-0.94**
Jhetp	**-0.86**	**0.75**	**0.93**	**-0.94**	**-0.80**	0.23	**0.91**	**-0.95**
MAXDN	-0.51	0.41	**0.65**	**-0.61**	**-0.61**	0.48	0.35	-0.50
MAXDP	0.11	-0.08	-0.11	0.14	0.03	0.20	-0.31	0.20
DELS	0.36	-0.28	-0.40	0.43	0.27	0.16	**-0.61**	0.50
TIE	**0.81**	**-0.70**	**-0.86**	**0.87**	**0.73**	-0.18	**-0.89**	**0.89**
S0K	**0.74**	**-0.61**	**-0.85**	**0.85**	**0.70**	-0.14	**-0.88**	**0.87**
S1K	**0.83**	**-0.72**	**-0.93**	**0.92**	**0.78**	-0.22	**-0.92**	**0.93**
S2K	**0.91**	**-0.81**	**-0.97**	**0.96**	**0.85**	-0.33	**-0.91**	**0.96**
S3K	**0.89**	**-0.87**	**-0.84**	**0.83**	**0.83**	-0.49	**-0.71**	**0.82**
PHI	**0.92**	**-0.85**	**-0.96**	**0.94**	**0.87**	-0.40	**-0.86**	**0.93**
BLI	**0.66**	**-0.68**	-0.56	0.55	**0.59**	-0.51	-0.31	0.49
PW2	0.41	-0.26	-0.54	0.57	0.38	0.06	**-0.61**	0.56
PW3	0.79	**-0.67**	**-0.9**	**0.89**	**0.74**	-0.19	**-0.89**	**0.90**
PW4	**0.88**	**-0.77**	**-0.96**	**0.96**	**0.83**	-0.29	**-0.92**	**0.96**
PW5	**0.71**	**-0.59**	**-0.8**	**0.81**	**0.66**	-0.17	**-0.75**	**0.79**
PJI2	**-0.63**	0.50	**0.77**	**-0.75**	-0.49	-0.09	**0.78**	**-0.79**
CSI	**0.73**	**-0.6**	**-0.84**	**0.84**	**0.69**	-0.13	**-0.87**	**0.86**
ECC	**0.73**	**-0.61**	**-0.84**	**0.84**	**0.69**	-0.14	**-0.87**	**0.86**
AECC	**0.77**	**-0.66**	**-0.86**	**0.86**	**0.74**	-0.21	**-0.87**	**0.87**
DECC	**0.61**	-0.50	**-0.72**	**0.72**	**0.59**	-0.09	**-0.78**	**0.74**
MDDD	**0.65**	-0.53	**-0.78**	**0.77**	**0.62**	-0.09	**-0.83**	**0.80**
UNIP	**0.75**	**-0.63**	**-0.86**	**0.86**	**0.70**	-0.13	**-0.89**	**0.88**
CENT	**0.67**	-0.54	**-0.80**	**0.80**	**0.63**	-0.08	**-0.85**	**0.82**
VAR	**0.68**	-0.56	**-0.81**	**0.80**	**0.67**	-0.16	**-0.83**	**0.82**
BAC	**-0.78**	**0.67**	**0.82**	**-0.84**	**-0.70**	0.20	**0.76**	**-0.84**
Lop	**-0.82**	**0.68**	**0.89**	**-0.91**	**-0.74**	0.20	**0.84**	**-0.90**
ICR	**0.77**	**-0.67**	**-0.85**	**0.84**	**0.76**	-0.28	**-0.83**	**0.85**
D/Dr06	0.54	-0.42	**-0.65**	**0.65**	0.54	-0.09	**-0.71**	**0.67**
R_{Mo}^{a}	1	**-0.97**	**-0.92**	**0.93**	**0.88**	-0.43	**-0.83**	**0.93**
b^{a}		1	**0.81**	**-0.82**	**-0.83**	0.46	**0.74**	**-0.85**
ϕ_o^{a}			1	**-0.99**	**-0.85**	0.34	**0.88**	**-0.96**
$PC1_{RF}^{a}$				1	**0.86**	-0.35	**-0.89**	**0.96**
R_{Mo}^{b}					1	**-0.73**	**-0.67**	**0.85**
b^{b}						1	-0.02	-0.26
ϕ_o^{b}							1	**-0.95**
$PC1_{RF}^{b}$								1

[a] lipophilicity indices corresponding to RP-18F$_{254s}$ stationary phase
[b] lipophilicity indices corresponding to RP-18WF$_{254s}$ stationary phase

when replaced with the ethyl group there is total gain of activity. Compounds **6-12** have similar central unit with different substituent on the N-terminal, and the difference in lipophilicity may lie in the steric and electronic properties of this substituent.

The correlation matrix shows that much higher correlation coefficients were obtained between the physico-chemical and topological descriptors and the lipophilicity indices (RMo, b, ϕ_o and scores corresponding to first principal component) of studied compounds for RP-18F$_{254S}$ stationary phase as compared to RP-18WF$_{254S}$. These values range between 0.75 and 0.96 in the majority of cases, so the physico-chemical and topological parameters have a statistically significant correlation with lipophilicity. Also, we have to mention that in this case the isocratic chromatographic hydrophobicity index, ϕ_o, appeared to correlate better with the computed descriptors, compared to RMo in both cases of stationary phases used. Best results and higher statistically significant values are given by the combination of lipophilicity indices (RMo, ϕ_o, and scores corresponding to the first component) with physico-chemical descriptors (table 3) compared to the topological descriptors calculated with Dragon software (table 4, 5), and this combination is adequate to explain the relationship between the structure of the molecule and the molecular lipophilicity. Lower correlation coefficients within the range of 0.06 to 0.54 for RP-18F$_{254S}$ and RP-18WF$_{254S}$ plates were found, this implying a minimal contribution of charge descriptors (DM-dipole moment, ABSQ-sum of absolute charges on atoms, ABSQON-sum of absolute charges on oxygen and nitrogen, $Q+$ and $Q-$, largest positive and negative charges over each atom) and some molecular descriptors (MAXDP- maximal electrotopological positive variation, DELS-molecular electrotopological variation, TPSA(NO) and TPSA(Tot)-topological polar surface area using N, O, respectively N, O, S and P polar contributions, Ui- unsaturation index) over RMo.

In order to describe the relationship between the chromatographic retention indices of tested compounds (RMo, ϕ_o and scores corresponding to the first component) and the calculated descriptors, a multivariate regression analysis was performed and leave-one-out cross-validation procedure was applied in order to verify the reliability of our results.

By forward stepwise multiple regression analysis, the following high-quality regression equations were obtained for RP-18 stationary phase (eq. 3-5).

Regression ecuation	Statistic parameters				
	r	Q^2	F	s_e	Eq
R_{Mo} = -50.777 + 0.075*DM* – 3.313*HOMO* + 2.056*MAXDN* + 3.592*S3K* + 164.359*PW5* – -0.054*AMR*	0.9977	0.9594	128	0.102	(3)
φ_o = -25.338 – 0.412*TIE* – 1.210*ALOGP2*	0.9644	0.8758	60	1.848	(4)
PC1 = 0.625 – 0.087*MAXDN* + 0.007*DELS* – -0.009*MLOGP2* + 0.010*ALOGP2*	0.9908	0.9551	95	0.007	(5)

The results also suggest that the highest occupied molecular orbital (HOMO), dipole moment (DM), molecular electrotopological variation (DELS), maximal electrotopological negative variation (MAXDN), path/walk 5 Randic shape index (PW5), 3-path Kier alpha-modified shape index (S3K) and E-state topological parameter (TIE) seem to be dominant in the retention mechanism. As a consequence the topological descriptors seem to control the lipophilicity in the case of RP-18F$_{254S}$ stationary phase.

Lower correlation coefficients were obtained for RP-18WF$_{254S}$ phase excepting the case when the R_{Mo} was estimated in dependence of the variables, and the regression equations are the following (eq. 6-8):

Regression ecuations	Statistic parameters				
	r	Q^2	F	s_e	Eq
R_{Mo} = -27.397 – 0.444*ALOGP* + 0.182*Rww* + +0.841*LogP* + 2.485*LUMO* + 0.775*MAXDN*+ +158.057*PW5* – 1.632*HOMO* – 139.202*SP* + +63.374*MSD* – 0.534*Ram*	1.0000	0.8363	2171 x10^4	0.0001	(6)
φ_o = -94.984 + 2.389*LogP* + 184.591*Xt*	0.9366	0.8091	32	2.9990	(7)
PC1= 0.372 + 0.011*PartCoeff* + 0.034*S2K*	0.9846	0.9177	143	0.0096	(8)

As for RP-18WF$_{254S}$ stationary phase, the highest correlations were given by the following descriptors: highest occupied and the lowest unoccupied molecular orbital (HOMO and LUMO), specific polarizability (SP), octanol-water partition coefficient (LogP), and partition-coefficient (PartCoeff), path/walk 5 Randic shape index and (PW5), 2-path Kier alpha-modified shape index (S2K) total structure connectivity index (Xt) ramification index (Ram), reciprocal hyper-detour index (Rww), mean square distance index (MSD) and Ghose-Crippen octanol-water partition coefficient (ALOGP) seem to be dominant in the retention mechanism. Lower correlation coefficients were obtained when the isocratic coefficients were estimated in dependence of the descriptors.

The statistical details of the QSAR models given above speak for their good statistical quality which indicates that these equations represent a satisfactory model for some of the calculated descriptors and have definite physical meaning corresponding to different intermolecular interactions. All the r values are above 0.9, which suggest that a good percentage of the total variance in lipophilicity is accounted by the model.

The goodness of the correlation is tested by the regression coefficient (r), the cross-validated (leave-one-out) $Q^2_{(LOO)}$, the F-test (F) and the standard error of estimate (s_e). The correlation coefficient, r, is a measure of the goodness of fit of the regression model. The F-test reflects the ratio of the variance explained by the model and the variance due to the error in the model. The standard error measures the model error.

The models predict the lipophilicity indices (RMo and scores along PC1) with correlation coefficients of 0.9846 and 1, and lower values were obtained when the hydrophobicity index, ϕ_o, was estimated.

A highly positive contribution upon the retention mechanism on both kind of phases was given by path/walk 5 Randic shape index (PW5), and high negative contribution of specific polarizability was observed on RP-18WF$_{254S}$ which states that specific polarizability accounts the favorable effects of dipole-dipole interactions between the solutes and the bulk phases. The negative energies of HOMO and LUMO approximate the electron affinity of molecules and mutagenicity. The dipole moment indicates the intrinsic polarity of the molecule and its magnitude is also a good indicator of lipophilicity; the larger the magnitude, the higher is its hydrophilicity.

The Kier symmetry index S3K measures the disorder in the system. It is known that lower values of S3K lead to greater topological symmetries, and thus to a lower tendency for changing the molecular freedom. Moriguchi octanol–water partition coefficient (MLOGP) is a significant descriptor for efflux inhibitory activity.

Based on a 2D representation of loadings representing both lipophilicity indices and descriptors shows an interesting clustering, the lipophilicity parameters form a cluster, which indicates the predominant role of lipophilicity in the reversed-phase retention mechanism; almost all descriptors form a large cluster comprising the majority of them according to their similarity, with the exception of a smaller compact cluster which includes: SP, MAXDN, MAXDP, DM, SPI, DELS, ABSQon, TPSA(NO), TPSA(Tot); also, Q+, MSD, BLI, BAC, Lop, PJI2 appear to form an outlier group and are the most discriminating.

It was also shown that score plots can be used to search for structural similarities within groups of compounds, since similar structures are grouped. The scatterplot of scores shows interesting results (fig. 2). Three clusters appear to be well defined and in a good agreement to the structure

Table 6
TABLE OF CORRELATION CONSIDERING MOLECULAR DESCRIPTORS COMPUTED WITH DRAGON SOFTWARE PROGRAM

Variable	RP-18F_{254s}				RP-18WF_{254s}			
	R_{Mo}	b	ϕo	$PC1_{RF}$	R_{Mo}	b^{18W}	ϕo	$PC1_{RF}$
Ui	0.35	-0.23	-0.50	0.50	0.35	0.09	**-0.63**	0.54
Hy	**-0.84**	**0.76**	**0.90**	**-0.87**	**-0.77**	0.32	**0.78**	**-0.87**
AMR	**0.82**	**-0.69**	**-0.92**	**0.92**	**0.77**	-0.20	**-0.91**	**0.92**
TPSA(NO)	-0.49	0.49	0.49	-0.46	-0.49	0.44	0.22	-0.39
TPSA(Tot)	-0.49	0.49	0.49	-0.46	-0.49	0.44	0.22	-0.39
MLOGP	**0.88**	**-0.79**	**-0.94**	**0.93**	**0.84**	-0.41	**-0.77**	**0.89**
MLOGP2	**0.74**	**-0.60**	**-0.87**	**0.87**	**0.79**	-0.43	**-0.66**	**0.78**
ALOGP	**0.94**	**-0.86**	**-0.94**	**0.95**	**0.88**	-0.47	**-0.77**	**0.90**
ALOGP2	**0.90**	**-0.82**	**-0.90**	**0.92**	**0.86**	-0.49	**-0.71**	**0.86**
BLTF96	**-0.88**	**0.79**	**0.94**	**-0.93**	**-0.84**	0.41	**0.77**	**-0.89**
BLTD48	**-0.88**	**0.79**	**0.94**	**-0.93**	**-0.84**	0.41	**0.77**	**-0.89**
BLTA96	**-0.88**	**0.79**	**0.94**	**-0.93**	**-0.84**	0.41	**0.77**	**-0.88**
R_{Mo}^{a}	1.00	**-0.97**	**-0.92**	**0.93**	**0.88**	-0.43	**-0.83**	**0.93**
b^{a}		1.00	**0.81**	**-0.82**	**-0.83**	0.46	**0.74**	**-0.85**
ϕ_o^{a}			1.00	**-0.99**	**-0.85**	0.34	**0.88**	**-0.96**
$PC1_{RF}^{a}$				1.00	**0.86**	-0.35	**-0.89**	**0.96**
R_{Mo}^{b}					1.00	**-0.73**	**-0.67**	**0.85**
b^{b}						1.00	-0.02	-0.26
ϕ_o^{b}							1.00	**-0.95**
$PC1_{RF}^{b}$								1.00

[a]lipophilicity indices corresponding to RP-18F_{254s} stationary phase
[b]lipophilicity indices corresponding to RP-18WF_{254s} stationary phase

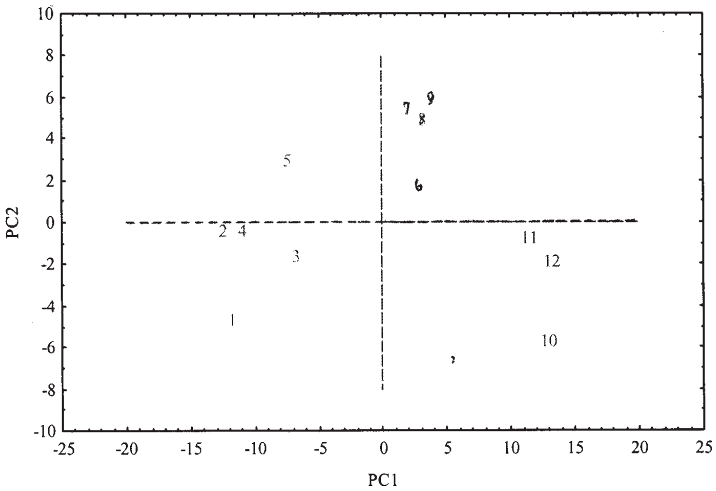


Fig. 2. Scatterplot of scores

of compounds: one of them corresponds to the compounds 1, 2, 3, 4 and 5 (formyl and acetyl derivatives), the second includes the hexamethylenimine derivatives (6, 7 and 8) and cyclohexyl derivative (9), which has a distinct structure, and the third group corresponds to the piperazinyl derivatives (10, 11 and 12).

Conclusions

In the present study, the relationship between chromatographic retention indices (RMo, *b*, ϕ_o and scores corresponding to the first principal component), estimated by RP-HPTLC, and the various descriptors computed through different software of the formyl- and acetylpyridine-3- thiosemicarbazone derivatives has been investigated.

The lipophilicity of most of the compounds tested probably resides in the action of the N-terminal at some reactive site, especially the heterocyclic ones which may be more active because of the enhanced ability of the compound to penetrate the silica gel stationary phase.

Excellent structure–retention index models show the efficiency of these indices in the structure–retention index correlations. In all cases, the path/walk 5 Randic shape index and (PW5), highest occupied and the lowest unoccupied molecular orbital (HOMO) and (LUMO) and specific polarizability (SP) play a dominant role in chromatographic behaviour on both stationary phases with different polarities. Much more, principal component analysis applied to the data matrix showed an interesting clustering of descriptors and also of compounds in a good agreement to their chemical structure and chromatographic behaviour.

The results of these investigations also indicate that some topological and charge descriptors are a most useful basis for study of QSAR/QSPR/QSRR for formyl- and acetylpyridine-3-thiosemicarbazone derivatives, which assures broad evaluations in these domains.

Acknowledgement: This research has been supported by the Romanian National Council for Scientific Research in Higher Education (CNCSIS) through the PNII-IDEI research grant ID_550/2007.

References

1. MAZZOBRE, M.F., ROMÁN, M.V., MOURELLE, A.F., CORTI, H.R., Carbohydr. Res., **340**, nr.6, 2005, p.1207
2. BENFENATI, E., GINI, G., PICLIN, N., RONCAGLIONI, A., VARI, M.R., Chemosphere, **53**,nr.9, 2003, p.1155
3. HANSCH, C., LEO, A., Exploring QSAR. Fundamentals and Applications in Chemistry and Biology, ACS: Washington, DC, 1995
4. KARELSON, M., Molecular Descriptors in QSAR/QSPR, John Wiley & Sons, Inc.: New York, 2000
5. WANG, R., FU, Y., LAI, L., J. Chem. Inf. Comput. Sci., **37** 1997, p.615
6. KALISZAN, R., Anal. Chem., **64**, nr.11, 1992, p.619A.
7. MOHAJERI, A., DINPAJOOH, M.H., J. Mol. Struct. THEOCHEM, http://dx.doi.org/10.1016/j.theochem.2007.12.037, in print, 2007
8. REN, B., Chemom. Intell. Lab. Syst., **66**, 2003, p.29
9. DUVENBECK, C., ZINN, P., J. Chem. Inf. Comput. Sci., **33**, 1993, p.211
10. KATRITZKY, A.R., IGNATCHENKO, E.S., BARKOCK, R.A., LOBANOV, V.S., KARELSON, M., Anal. Chem., **66**, 1994, p.1799.
11. LIU, S.S., YIN, C.S., CAI, S.X., LI, Z.L., Chemometr. Intell. Lab. Syst., **61,** nr.1, 2002, p.3
12. TRINAJSTIC, N., Chemical Graph Theory, 2nd ed.; CRC Press: Boca Raton, 1992
13. ESTRADA, E., Chem. Phys. Lett., **336,** nr.3/4, 2001, p.248
14. PUT, R., VANDER HEYDEN, Y., Anal. Chim. Acta, **602**, nr.2, 2007, p.164
15. KHAN, S.A., KUMAR, P., JOSHI, R., IQBAL, P.F., SALEEM, K., Eur. J. Med. Chem., http://dx.doi.org/10.1016/j.ejmech.2007.12.004, in press, 2008
16. DE AQUINO, T.M., LIESEN, A.P., DA SILVA, R.E.A., LIMA, V.T., CARVALHO, C.S., DE FARIA, A.R., DE ARAÚJO, J.M., DE LIMA, J.G., ALVES, A.J., DE MELO GÓES, E.J.T., GÓES, A.J.S., Bioorg. Med. Chem., **16**, nr.1, 2008, p.446
17. CHANDRA, S., KUMAR, A., Spectrochim. Acta, Part A, **68**, 2007, p.1410
18. PRABHAKARAN, R., RENUKADEVI, S.V., KARVEMBU, R., HUANG, R., MAUTZ, J., HUTTNER, G., SUBACHKUMAR, R., NATARAJAN, K., Eur. J. Med. Chem., **43**, nr.2, 2008, p.268
19. ALI, M.A., IBRAHIM, N.E.H., BUTCHER, R.J., JASINSKI, J.P., JASINSKI, J.M., BRYAN, J.C., Polyhedron, **17**, nr.11/12, 1998, p.1803
20. KOSTAS, I.D., HEROPOULOS, G.A., KOVALA-DEMERTZI, D., YADAV, P.N., JASINSKI, J. P. DEMERTZIS, M. A. ANDREADAKI, F. J., VO-THANH, PETIT, G. A., LOUPY, A., Tetrahedron Lett., **47,** nr. 26, 2006, p.4403
21. KOSTAS, I.D., ANDREADAKI, F.J. KOVALA-DEMERTZI, D., PRENTJAS, C., DEMERTZIS, M.A., Tetrahedron Lett., **64**, nr.12, 2005, p.1967
22. KOVALA-DEMERTZI, D., VIJDELUC, C., DEMERTZIS, M.A., SIAPI, E., MAVROMOUSTAKOS, T., Thermochim. Acta, **424**, nr.1/2, 2004, p.53
23. KAYILLO, S., DENNIS, G.R., SHALLIKER, R.A., J. Chromatogr. A, **1126,** nr.1/2, 2006, p.283.
24. *** HyperChem, release 7.5 for Windows, Molecular Modeling System; Hypercube, Inc. and Autodesk, Inc.
25. *** http://www.talete.mi.it.
26. *** http://www.tripos.com.
27. *** http://www.cambridgesoft.com
28. KOSSOY, A.D., RISLEY, D.S., KLEYLE, R.M., NUROK, D., Anal. Chem., **64**, 1992, p. 1345
29. DILLI, S., HADDAD, P.R., HTOON, A.K., J. Chromatogr. A, **500**, 1990, p. 313
30. SÂRBU, C., DJAKOVIC SEKULIC, T., PERIŠIC-JANJIC, N., J. Pharm. Biomed. Anal., **30**, nr.3, 2002, p.739
31. DJAKOVIC-SEKULIC, T., ACANSKI, M.M., PERISIC-JANIC, N., J. Chromatogr. B, **766**, 2002, p.67
32. JÓZWIAK, K., SZUMILO, H., SOCSEWIŃSKI, E., Wiad. Chem., **55**, nr.11/12, 2001, p.1047
33. SÂRBU, C., TODOR, S., J. Planar. Chromatogr-Mod. TLC, **11**, 1998, p.123
34. KATRITZKY, A.R., STOYANOVA-SLAVOVA, I.B., DOBCHEV, D.A., KARELSON, M., J. Mol. Graphics Modell., **26**, nr.2, 2007, p.529
35. *** http://www.statistica.com.

Manuscript received: 3.10.2008