ХИМИЯ

LINEAR REGRESSIONS BASED ON QSRR MODELS FOR THE PREDICTION OF CHROMATOGRAPHIC RETENTION OF SOME FIVE-MEMBERED NITROGEN-CONTAINING HETEROCYCLES

© 2010   A.V. Bulanova,[1]   Yu.L. Polyakova[2]

This work deals with twenty substances composed of five-membered nitrogen-containing heterocycles. The relationships between the chromatographic retention factor (k) and those physicochemical and topological properties which are relevant in quantitative structure-retention relationship (QSRR) studies, such as the polarizability (α), molar refractivity (MR), lipophilicity $(log\,P)$, dipole moment (μ), molecular surface area (S_M), molecular connectivity indexes $(^0\chi \sim {}^5\chi)$, Wiener index (W), Kier flexibility index (ϕ), and Harary index (H) were investigated. The accuracy of the simple linear regressions between the chromatographic retention (CR) and some descriptors for all of the compounds was satisfactory. The QSRR models of these compounds could be predicted with a multiple linear regression (MLR) equations having the statistical index, $r^2 = 1.0$. This work demonstrated the successful application of the MLR approaches through the development of accurate predictive equations for k in liquid chromatography.

Key words: QSRR, five-membered nitrogen-containing heterocycles, retention factor, correlation.

Introduction

An a priori prediction of properties of chemical substances or their structural formulas is a fundamental aim, however still quite open problem of the chemical science. Initial conditions for deriving quantitative structure-property relationships, allowing for reliable property predictions, are determined

---

[1]Bulanova Angela Vladimirovna (`bulanova@ssu.samara.ru`), Dep. Of Chem., Samara State University, Samara, 443011, Russia.

[2]Polyakova Yulia Lvovna (`yulia.polyakova@gmail.com`), Center for Advanced Bioseparation Technology and Dept. of Chem. Eng., Inha University, 402-751, South Korea.

by the correct, reproducible parameters, and the exact, identically defined structural performance of the chemical objects under consideration encoding peculiar information on their individual property appearances, on the other hand. The pioneer research applying quantitative structure–property/activity relationships (QSPR/QSAR) was due to H. Wiener [1] Modeling and prediction of the chromatographic retention (CR) as a function of the molecular property, has been the subject of study for the last decades. As a result, quantitative structure-retention relationships (QSRR) have, since their introduction in the late 1970s, been considered a model approach to establish strategy of property predictions, to test the performance of various methods as well as property predictor potency of theoretically limitless number of molecular descriptors obtainable by computational methods [2]. Some reviews distinctly verified a good performance, especially in quantitative comparison of several chromatographic properties, of the general QSRRs models employing the descriptors of analyte [3, 4]. QSRR are also an important issue in organic chemistry, physical chemistry, molecular design and medicinal chemistry, as well as in drug discovery [5]. Over the past few years, different methods have been used to correlate CR behavior with various molecular structure and physicochemical parameters, such as the Van der Waals and molecular volumes [6], the molecular connectivity indexes [7, 8], the number of carbon atoms [9], the n-octanol/water partition coefficient (log P) [10, 11], the hydrophobic substituent constant (log S) [12], the solubility parameters [13], and some others [14, 15].

The incorporation of the five-membered heterocyclic nuclei, imidazole and triazole, into various compounds therapeutic agents is the main synthetic strategy in drugs discovery [16]. Imidazole and triazole drugs have a broad scope in remedying various disorders in clinical medicine [17].

In the present paper, a QSRR analysis was conducted to investigate the quantitative effect of the physicochemical and structural properties of five-membered nitrogen-containing molecules on their chromatographic retention. It was governed that the retention factor (k) of investigated substances was assumed by their descriptors such as polarizability (a), molar refractivity (MR), lipophilicity (log P), dipole moment (μ), molecular surface area (S_M), molecular connectivity indexes ($^0\chi \sim {}^5\chi$), Wiener index (W), Kier flexibility index (ϕ), and Harary index (H). The examination of the linear and multiple linear regression (MLR) equations between descriptors and k was the purpose of this work.
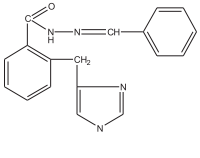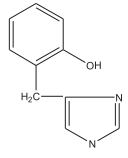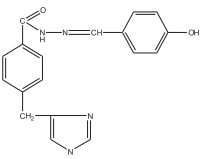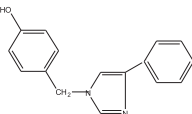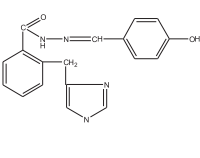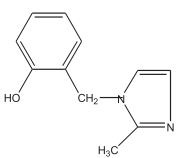
Experimental Section

Reagents. Twenty substances were kindly provided by Professor P. Purygin (Russia, Samara State University). Chemical structures and name of compounds are specified in Table 1. HPLC grade acetonitrile

Table 1

Chemical structures and names of five-membered nitrogen-containing heterocycles

№	Structure	Name	№	Structure	Name
1		methyl 4-((1H-imidazol-4-yl)methyl)benzoate	11		2-((1H-imidazol-4-yl)methyl)-N′-(2-methoxyben-zylidene)benzo-hydrazide
2		methyl 4-((3H-benzoi-midazol-5-yl)methyl)benzoate	13		2-((1H-imidazol-4-yl)methyl)-N′-(4-hydroxy-3-methoxybenzyli-dene)benzohydra-zide
3		methyl 4-((3H-benzo[1,2,3]triazol-5-yl)methyl)ben-zoate	14		2-((1H-imidazol-4-yl)methyl)-N′-(4-methoxyben-zylidene)benzo-hydrazide
4		4-((1H-imidazol-4-yl)methyl)ben-zohydrazide	15		4-((1H-imidazol-4-yl)methyl)-N′-(4-hydroxy-3-methoxybenzyli-dene)benzohydra-zide
5		2-((1H-imidazol-4-yl)methyl)ben-zohydrazide	16		4-((1H-imidazol-4-yl)methyl)-N′-(3-hydroxy-4-methoxybenzyli-dene)benzohydra-zide
7		4-((1H-imida-zol-4-yl)me-thyl)-N′-benzylidene-benzohydrazide	17		2-((3H-benzoi-midazol-5-yl)methyl)phenol

8		2-((1H-imida-zol-4-yl)me-thyl)-N′-ben-zylidenebenzo-hydrazide	18		2-((1H-imida-zol-4-yl)me-thyl)phenol
9		4-((1H-imida-zol-4-yl)me-thyl)-N′-(4-hy-droxybenzyli-dene)benzohy-drazide	19		4-((4-phenyl-1H-imidazol-1-yl)methyl) phenol
10		2-((1H-imida-zol-4-yl)me-thyl)-N′-(4-hy-droxybenzyli-dene)benzohy-drazide	20		2-((2-methyl-1H-imidazol-1-yl)methyl) phenol

was obtained from Panreac Corp. (Spain). Trifluoroacetic acid (TFA) and dimethylsulfoxide (DMSO) were purchased from Sigma-Aldrich Corp. (St. Louis, MO, USA). All water used in the experiments was filtered by a Millipore ultra pure water system (Millipore, Belford, MA, USA).

Sample Preparation. 5 mg of each of the twenty-four analytes was dissolved in 1 ml of water or the 1ml of mixture (water/DMSO). A constant injection volume (5 μl) was used for the mixture solution throughout. The solutions were stored at 277 K and the working standards were re-prepared every 2 days to avoid the potential errors from decomposition of the targets.

Apparatus and Methods. The research was performed using an Agilent 1100 series HPLC–MS system (Agilent technologies, USA). The HPLC system included an Agilent (Wilmington, Delaware, USA) 1100 binary pump with an online degasifier, and an autosampler. The HPLC system was used with UV-Vision diode array detection (DAD), evaporative light scattering detection (ELSD) and mass spectrometric (MS) detection. The chromatographic separations were performed using a Betasil 10 μm octadecyl silica (C_{18}) column (4,6 × 150 mm) (Phenomenex, USA). The column temperature was 293 K permanently. The eluent consisted of a mixture of 0,1 % TFA in water (A) with 0,1 % TFA in acetonitrile (B) and was delivered at a flow-rate of 2,5 ml/min. The gradient program consisted of increasing the concentration of B from 5 % to 90 % after 0,5 min over a period of 5,5 min and then holding it at this concentration for another 1,5 min. The run time, from injection to injection, was 9,5 min. The single quadrupole MS was equipped

with an electrospray source and controllable in selected ion-monitoring (SIM) mode. The electrospray was operated in positive ion mode to produce mainly protonated molecular mass ions $[M-H]^{+1}$. The scanning mass to charge values ranged from 90 to 900 m/z with a scanning speed of 13000 m/z per second. The MS parameters were set to facilitate the ionization process and achieve the optimal sensitivity. These conditions allowed for minimal fragmentation of the target compound, optimum yield of its quasi-molecular ion and minimum interference from the matrix. An Agilent 1100 LC–MS system managed by Windows NT based HP ChemStation software (Rev 8,04) was used. The chromatograms were integrated using the HP ChemStation software package. Retention factors, k, were calculated using the formula $k = (t_R - t_0)/t_0$ where t_R is the retention time of the analyte and t_0 is the t_R of the non-retained peak (taken as the first deviation of the baseline following the injection of 5 μl KNO2). Three replicate injections were made to determine the t_R, and the average values were used to calculate the ks. Evaluation of the results of the experiment was carried out by mathematical statistics techniques. The relative error of a single measurement did not exceed 5 %. All experimental procedures were performed at 293 K.

Theoretical Background

The polarizability (a), molar refractivity (MR), lipophilicity (log P), dipole moment (μ), and molecular surface area (S_M,) were drawn using HyperChem software (Hypercube Inc. FL, USA, version 7) on a PC, using semi empirical AM1 calculations for the determination of the molecular dynamics (293 K) of the molecules. Extending the empirical approach can produce methods with good accuracy and excellent computational performance. In this extended method, the calculation is generally sufficiently reliable and the results in good agreement with those obtained experimentally [18]. Connectivity indexes ($^0\chi \times \times {}^5\chi$), Winner index ($W$), Kier flexibility index (ϕ), and Harary index (H) were obtained by the preADME program software. All values of descriptors and k were listed in Table 2. The statistical evaluation of the data was performed using the Origin program package. To test the quality of the regression equation, the squire of correlation coefficient (r^2) was utilized as the statistical parameter.

Results and Discussion

In this paper, the k were estimated using the mathematical relationships expressed as a linear relationship in terms of their descriptors. We collect the t_R obtained from ELSD, DAD, and MS detectors for each substance. The display one of them (DAD, 254 nm) was taken for calculation of k, because instrumentation indications from ELSD and MS were very close with DAD's

data. Typical chromatograms of one of the solute are shown in Figure 1. The absence of strict homology in a sequence (set) of investigated substances poses a serious problem in the creation of a QSRR. In such cases, the linear correlation equation is not always applicable to the description of the CR within the limit of a full nonhomologous set. From this point of view, the best results can be obtained by the formal division of all of the investigated substances into groups of substances having some common features. In the present work, the principle of structural analogy was used for the division of the substances into groups. Molecules having similar structural fragments were classified into the same group. In this way, all of the investigated molecules were classified into 5 groups depending upon their chemical structures. The data in Table 2 served as the basis for statistical analysis.

Table 2

The values of the retention factors (k) and descriptors of five-membered nitrogen-containing heterocycles

№	k	$\log P$	$MR,$ A^3	$\alpha\ A^3$	$S_M,$ A^3	$\mu,$ D
Group I						
1	2,66	-0,60	65,32	23,49	429	4,61
2	3,81	-0,29	84,64	29,67	495	4,19
3	5,68	0,75	86,22	28,96	485	1,79
Group II						
4	0,81	-1,51	66,86	23,72	432	4,47
5	1,38	-1,51	66,86	23,72	411	5,18
6	1,79	-1,20	86,18	29,90	495	3,24
Group III						
7	3,38	-0,16	94,65	34,79	583	3,68
8	3,75	-0,16	94,65	34,79	550	5,54
9	2,73	-1,19	96,25	35,43	609	3,84
10	2,79	-1,13	100,47	35,43	581	6,27
Group IV						
11	3,68	-1,10	105,24	37,27	613	4,57
12	3,15	-2,13	106,84	37,90	556	4,42
13	3,12	-2,13	106,84	37,90	622	2,60
14	3,86	-1,10	105,24	37,27	614	3,84
15	2,90	-2,13	106,84	37,90	621	1,85
16	3,01	-2,13	106,84	37,90	634	2,87
Group V						
17	3,47	-0,74	75,48	25,91	428	1,98
18	2,26	-1,05	56,16	19,37	362	3,88
19	3,62	-0,04	85,06	29,39	489	3,35
20	2,62	-0,53	59,34	21,56	388	4,44

№	$^0\chi$	$^1\chi$	$^2\chi$	$^3\chi$	$^4\chi$	$^5\chi$	W	ϕ	H
Group I									
1	11,3	7,7	6,5	5,4	4,0	2,6	491	3,2	45,5
2	12,9	9,7	8,5	7,4	5,8	4,3	909	3,4	64,4
3	13,9	9,7	8,5	7,4	5,8	4,3	909	3,3	64,4
Group II									
4	11,3	7,7	6,5	5,4	4,0	2,6	491	3,2	45,5
5	11,3	7,8	6,4	5,4	4,2	2,8	443	3,2	46,7
6	13,9	9,7	8,5	7,4	5,8	4,3	909	3,3	64,5
Group III									
7	15,9	11,3	9,5	7,8	6,1	4,3	1487	4,8	75,2
8	16,9	11,3	9,5	7,8	6,3	4,5	1355	4,8	77,1
9	16,7	11,7	10,2	8,2	6,2	4,6	1689	5,0	79,8
10	16,7	11,7	10,1	8,2	6,5	4,8	1545	5,0	81,8
Group IV									
11	17,4	12,2	10,2	8,6	7,0	4,9	1691	5,5	87,3
12	18,3	12,6	10,8	9,0	7,3	5,0	1939	5,7	91,9
13	18,3	12,6	10,8	9,0	7,2	5,1	1922	5,7	92,0
14	17,4	12,2	10,2	8,6	6,8	4,9	1759	5,5	86,3
15	18,3	12,6	10,9	9,1	7,0	5,0	2090	5,7	89,9
16	18,3	12,6	10,9	9,0	7,1	4,8	2107	5,7	89,7
Group V									
17	11,6	8,3	7,3	6,3	5,1	3,9	533	2,4	52,1
18	9,0	6,3	5,4	4,4	3,4	2,2	253	2,2	34,2
19	13,0	9,3	8,1	6,7	5,3	4,2	787	3,2	59,6
20	9,6	6,7	5,9	4,9	3,7	2,5	306	2,4	38,5

To obtain the optimal QSRR model using the subset of the molecular structural descriptors the next followed procedures were realized in this work. Firstly, a linear mathematical analysis with all molecular descriptors as predictor variables was performed (Table 2). These data served as a basis for further statistical analysis. The linear regression equations were obtained for each of the groups and a summary of these models is given in Tables 3 and 4. The theoretical descriptors calculated in this study include a wide spectrum of molecular properties, which all have high values, because each of them influences the CR to some extent. Some of the results had a high value of r^2, which indicate that these equations represent a satisfactory model for some of the calculated descriptors. It could be seen that log P, MR, and a were well correlated with the k and the linear models were adequate $(0,8 \leqslant r^2 \leqslant 1,0)$. Most of the regression equations showed that the size, steric factors, and polar effects primarily influenced the retention. However, some of the simple linear correlations shown in Table 3 gave a low (inadequate) r^2, of less than 0,80.

A comparison between $\log P$ and the k of the azoles using linear equations is shown in Figure 2. In all of the groups, the k increased with increasing $\log P$. Thus, it is possible to assert that the investigated azoles are eluted in increasing order of their hydrophobicity.
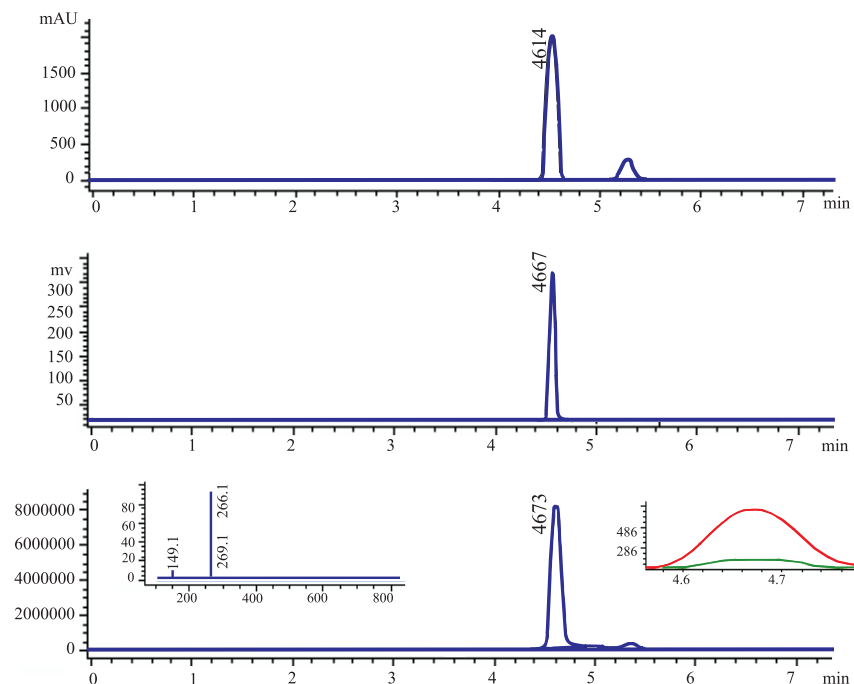

Fig. 1. Chromatograms of methyl 4-((3H-benzo[1,2,3]triazol-5-yl)methyl)benzoate in gradient mode with DAD 254 nm (a), ELSD (b) , and MS (c) detectors

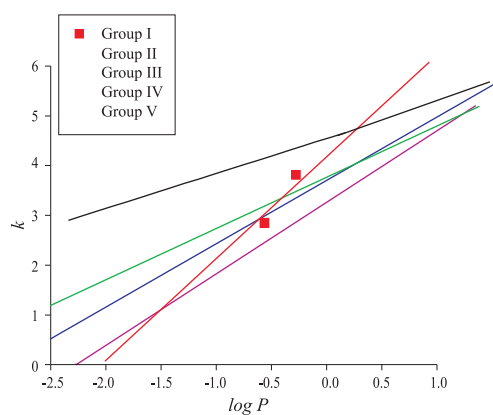Fig. 2. Simple linear regressions between experimental k and log P for the five groups (k was calculated from DAD data)

The simple linear regressions ($k = a+bD$, D indicates topological descriptor) of each group are listed in Table 4. The regression coefficient (r^2) of the group I (0,788 × 0,803), group III (0,531 × 0,963, except 0,289), and group IV (0,796 ×

$\times$ 0,872, except 0,347) are high, but the group II and group V are showed the low r^2. It indicated that these linear equations represents satisfactory model for some group compound, but not all of the groups. It could be seen that some order of the connectivity indexes ($^0\chi$ and $^1\chi$), Wiener index (W), and Harary index (H) had good correlation with the k.

Table 3

Linear mathematical analysis ($k = a + bD$) of the physicochemical descriptors (D) for the five groups

Group	D	a	b	r^2
I	$\log P$	3,879	1,911	0,891
	MR	-1,876	0,078	0,755
	α	-2,346	0,239	0,674
	S_M	-6,813	0,024	0,668
	μ	6,694	-0,830	0,848
II	$\log P$	3,232	1,375	0,962
	MR	-3,226	0,065	0,819
	α	-3,597	0,200	0,738
	S_M	-4,879	0,015	0,635
	μ	4,280	-0,649	-0,772
III	$\log P$	3,698	0,803	0,952
	MR	15,297	-0,126	-0,710
	α	47,166	-1,254	-0,951
	S_M	13,044	-0,017	-0,843
	μ	3,110	0,012	0,032
IV	$\log P$	4,544	0,702	0,963
	MR	51,327	-0,452	-0,963
	α	46,543	-1,148	-0,963
	S_M	3,586	-0,001	-0,035
	μ	2,466	0,244	0,691
V	$\log P$	3,896	1,712	0,842
	MR	-1,412	0,063	0,932
	α	-1,658	0,189	0,936
	S_M	-3,545	0,016	0,904
	μ	3,320	-0,113	-0,112

A common practice employed in the development of QSRR methodology has been to identify relationships employing self-chosen single or more descriptors. However, regardless of which type of variable is chosen, it is usually difficult to predict in advance the descriptor variables of value in describing the properties variation. Understandably, it is very difficult, or even impossible, to handle a large number of descriptors while aiming to produce an unambiguous and understandable selection. The latter fact is one of the disadvantages of QSRR.

Table 4

Linear equations ($k = a + bD$) of the structural descriptors (D) with groups

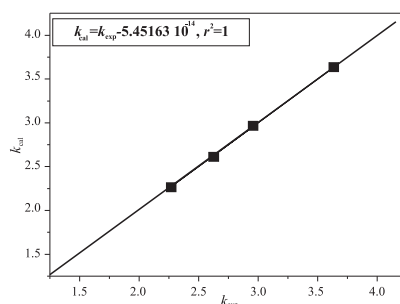| Group	D	a	b	r^2	D	a	b	r^2
I	$^0\chi$	-9,712	0,999	0,798	W	-1,169	0,006	0,804
	$^1\chi$	-8,485	1,297	0,801	ϕ	-42,108	13,726	0,720
	$^2\chi$	-6,599	1,258	0,803	H	-4,650	0,137	0,797
	$^3\chi$	-5,612	1,329	0,797	$^4\chi$	-4,443	1,480	0,793
	$^5\chi$	-2,408	1,472	0,788				
II	$^0\chi$	11,681	-0,510	0,469	W	8,312	-0,003	0,764
	$^1\chi$	14,742	-0,990	0,421	ϕ	14,674	-2,299	0,469
	$^2\chi$	13,346	-1,016	0,631	H	5,826	-0,032	0,190
	$^3\chi$	8,899	-0,692	0,331	$^4\chi$	-1,411	0,736	0,294
	$^5\chi$	7,567	-0,926	0,316				
III	$^0\chi$	17,819	-0,803	0,955	W	6,977	-0,002	0,893
	$^1\chi$	25,534	-1,774	0,952	ϕ	24,965	-3,799	0,958
	$^2\chi$	15,283	-1,123	0,950	H	15,311	-0,134	0,800
	$^3\chi$	17,040	-1,537	0,963	$^4\chi$	11,395	-1,146	0,531
	$^5\chi$	8,063	-0,958	0,289				
IV	$^0\chi$	-6,756	0,935	0,815	w	-0,167	0,008	0,837
	$^1\chi$	-5,610	1,183	0,850	ϕ	-7,088	4,223	0,347
	$^2\chi$	-4,598	1,187	0,819	H	-2,610	0,131	0,849
	$^3\chi$	-3,610	1,240	0,851	$^4\chi$	-2,362	1,281	0,872
	$^5\chi$	-1,078	1,334	0,843				
IV	$^0\chi$	20,567	-0,948	0,633	w	14,818	-0,006	0,685
	$^1\chi$	-18,541	1,735	0,468	ϕ	4,538	-0,200	0,164
	$^2\chi$	1,218	0,200	0,077	H	-25,753	0,321	0,673
	$^3\chi$	2,690	0,085	0,039	$^4\chi$	-0,673	0,575	0,434
	$^5\chi$	-1,348	0,930	0,490				

Fig. 3. Comparison between the calculated retention factors (k) from DAD data and the corresponding experimental values in group V

Table 5

The MLR models $k = f(D) = a_0 + a_1 \ log \, P + a_2 MR + a_3 \alpha + a_4 S_M + a_6 \mu$
containing physicochemical descriptors

Group	a_0	a_1	a_2	a_3	a_4	a_6	r
I	8,525	-6,006	0,084	-0,292	0,003	-3,461	1,000
II	-2,654	3,293	-0,043	0,130	-0,001	0,924	1,000
III	-1,837	7,338	-0,099	0,862	0,003	0,143	1,000
IV	-287,47	3,041	-3,768	0,997	0,022	0,560	1,000
V	-6,607	-1,500	-0,066	-0,375	-0,004	0,147	1,000

Table 6

The MLR models
$$k = f(D) = a_0 + a_1{}^0\chi + a_2{}^1\chi + a_3{}^2\chi + a_4{}^3\chi + a_5{}^4\chi + a_6{}^5\chi + a_7 W + a_8 \phi + a_9 H$$
containing structural descriptors

Group	I	II	III	IV	V
a_0	46408	-43,627	-4,712	-28,636	-4,985
a_1	2008	-0,235	-5,019	4,416	0,103
a_2	-1806	7,003	5,356	3,072	0,580
a_3	-8305	6,946	-13,186	12,976	-0,203
a_4	-1723	-6,338	5,272	12,407	-1,601
a_5	6938	1,126	-5,870	-6,017	-0,276
a_6	5877	-18,453	-4,795	-0,999	0,674
a_7	54,3	0,012	0,040	-0,034	-0,001
a_8	7,429	5,634	-7,108	-34,073	-0,563
a_9	-554	0,662	0,860	-0,718	-0,045
r^2	1,000	1,000	1,000	1,000	1,000

As a result, in this study, we attempted to consider the entire spectrum of the calculated descriptors. Therefore, the multiple linear statistical analysis technique was used, in order to obtain a reliable expression, which considers all of the descriptors calculated in this work. The MLR was obtained by regression techniques using the multilinear forms:

$$P = f(D) = a_0 + a_1 D_1 + a_2 D_2 + \ldots a_n D_n \qquad (1)$$

where P is the property (in our case, P is the k), D_1, D_2 and D_n are the descriptors and n is the number of descriptors. The intercept (a_0) and regression coefficients of the descriptors $(a_1, a_2, \ldots, a_n)$ were determined using the least squares method. Eq. 1 was examined and the numeric coefficients were determined by linear regression. The MLRs of the k with the physicochemical properties of the azoles and their r^2 are listed in Table 5. The result of CR was estimated with calculated values of regression coefficients, and the result showed clearly that each obtained descriptor with different molecular properties

has different effect. It is not difficult to see that coefficients a_2 (except group II) and a_4 are small and close to zero. This testifies, that the physicochemical parameters such as it MR and S_M weakly influence the retention. However, even such low-informative descriptors should be considered also, since their use unambiguously increases the predictive ability of the equations. The obtained MLRs showed that they possess the best predictive ability, in comparison with the linear equations. The calculated t_R times are compared with the corresponding experimental values in Figure 3. The MLRs obtained for the observed and calculated k are of excellent quality. The MLRs of the k with topological properties of compounds and regression coefficients were listed in Table 6. The small errors of prediction make it possibility for these models to be applied for the peak identification on a chromatogram, particularly for various t_R. The advantages of approach, used in this paper, are that they are not restricted to closely related compounds, they can be easily obtained, and describe clearly defined molecular properties. Because of their advantages, they are widely used in the QSRR studies. It is undoubtedly possible to further improve the ability of this technique to identify unknown azole's compounds. Unfortunately, in the present study, the small data sets limited our ability to predict the retention data, due to the inevitable statistical irregularities. Another approach to identifying unknown azoles peaks in a chromatogram is to use a similar QSPR approach to correlate the structural properties of various azoles compounds. The use of the t_R and structural correlations should offer the possibility of making further significant progress in this regard.

Conclusion

The t_R of 20 azoles were modeled using simple-linear and multi-linear models based on the calculated topological and physicochemical descriptors. The obtained models showed perfect correlation and predictive ability. The resulting MLRs for the t_R of the azoles can be used to predict the t_R and indices of unknown but structurally similar compounds with a considerable degree of confidence. Also, this study clearly demonstrates that QSPR models can be used to predict the t_R and indices without the need for chemical standards.

References

[1] Wiener H. Structural Determination of Paraffin Boiling Points // J. Am. Chem. Soc. 1947. № 69. P. 17–20.
[2] Kaliszan R. Structure and Retention in Chromatography–A Chemometric Approach. Harwood; Amsterdam, 1997. 211 p.

[3] Kaliszan R. J. Quantitative structure-retention relationships applied to reversed-phase high-performance liquid chromatography // J. Chromatogr. A. 1993. V. 656. P. 417–435.

[4] Polyakova Y., Row K. Quantitative Structure-Retention Relationships Applied to Reversed-Phase High-Performance Liquid Chromatography // Med. Chem. Res. 2006. № 14(8/9). P. 488–522.

[5] Kaliszan R. Quantitative Structure–Chromatographic Retention Relationships. New York: Wiley, 1987. 437 p.

[6] Prediction of high-performance liquid chromatography retention of peptides with the use of quantitative structure-retention relationships / R. Kaliszan [et al.]. Proteomics. 2005. № 5(2). P. 409-415.

[7] Lee S., Polyakova Y., Row K. Evaluation of Predictive Retention Factors for Phenolic Compounds with QSPR Equations // J. Liq. Chromatogr. Related Technol. 2004. № 4(27). P. 629–639.

[8] Retention of substituted indole compounds on RP-HPLC: correlation with molecular connectivity indices and quantum chemical / J. Dai [et al.] // Chromatogr. Related Technol. 1999. № 22(15). P. 2271–2282.

[9] Hinze W.L., Weber S.G. Why the Relationship between the Logarithm of k and Homologue Number in Micellar Chromatography Is Not Linear // Anal. Chem. 1991. № 63. P. 1808–1811.

[10] Yamagami C., Oguda T., Takao N. Relationship between capacity factors and octanol-water partition coefficients for monosubstituted pyrazines and the related pyridines // J. Chromatogr. A. 1990. № 514. P. 123–136.

[11] A comprehensive method for determining hydrophobicity constants by reversed-phase high-performance liquid chromatography / D.J. Minick [et al.] // J. Med. Chem. 1988. № 31. P. 1923–1933.

[12] Braumann T. Determination of hydrophobic parameters by reversed-phase liquid chromatography: Theory, experimental techniques, and application in studies on quantitative structure-activity relationships // J. Chromatogr. A. 1986. № 373. P. 191–225.

[13] Yamagami C., Fujita T. Hydrophobicity parameter of heteroaromatic compounds derived from various partitioning systems. In Classical and Three-dimensional QSAR in Agro-chemistry; Hansch, C., Fujita, T., Eds.: ACS Symposium Series 606. Washington: American Chemical Society. 1995. P. 36–47.

[14] Zheng J., Polyakova Y., Row K. Prediction of Chromatographic Retention of Nucleic Acids Based on QSPR Model // Chromatographia. 2006. № 64. P. 129–137.

[15] Relationships "biological activity-physicochemical property" for imidazolide and triazolide sulfuric acids, Samara State University Bulletin, Special Edition (Natural Science) / K. Egorova [et al.] // Вестник Самарского государственного университета. Специальный выпуск. 2002. P. 124–131.

[16] Townsend L.B. Imidazole nucleosides and nucleotides // Chem. Rev. 1967. № 67. P. 533–563.

[17] Pharmaceutical Substances / A. Kleeman [et al.]. 3rd ed. New York; Stuttgart, 1999.

[18] AM1: A new general purpose. Quantum mechanical model / M.J.S. Dewar [et al.] // J. Am. Chem. Soc. 1985. № 107. P. 3902–3909.

ЛИНЕЙНЫЕ РЕГРЕССИИ, ОСНОВАННЫЕ НА ПРИМЕНЕНИИ ХРОМАТОГРАФИЧЕСКИХ ХАРАКТЕРИСТИК УДЕРЖИВАНИЯ ДЛЯ ПРЕДСКАЗАНИЯ ХРОМАТОГРАФИЧЕСКОГО УДЕРЖИВАНИЯ НЕКОТОРЫХ ПЯТИЧЛЕННЫХ АЗОТСОДЕРЖАЩИХ ГЕТЕРОЦИКЛИЧЕСКИХ СОЕДИНЕНИЙ

© 2010   А.В. Буланова,[3]   Ю.Л. Полякова[4]

Работа посвящена установлению количественных соотношений "структура-хроматографическое удерживание" пятичленных азотсодержащих гетероциклов. Найдены зависимости между хроматографическим фактором удерживания, физико-химическими и структурными характеристиками 20 веществ, такими как: поляризуемость, молярная рефракция, липофильность, дипольный момент, индексы молекулярной связанности, индекс Винера, индекс Харари и др. Показано успешное применение многопараметрического линейного регрессионного анализа для предсказания хроматографического поведения азолов.

Ключевые слова: хроматографические характеристики удерживания, пятичленные азотсодержащие гетероциклические соединения, фактор удерживания, взаимосвязь "структура – свойство".

[3]Буланова Анджела Владимировна (bulanova@ssu.samara.ru), кафедра химии Самарского государственного университета, 443011, Россия, г. Самара, ул. Акад. Павлова, 1.

[4]Полякова Юлия Львовна (yulia.polyakova@gmail.com), Центр прогрессивных технологий биоразделения, химический факультет Инха университета, 402–751, Ю. Корея.