Journal of Chromatography A, 1218 (2011) 5120–5130

Contents lists available at ScienceDirect

Journal of Chromatography A

journal homepage: www.elsevier.com/locate/chroma

Thermodynamic vs. extrathermodynamic modeling of chromatographic retention

Roman Kaliszan*,[1], Paweł Wiczling, Michał J. Markuszewski, Mehdi A. Al-Haj

Department of Biopharmaceutics and Pharmacodynamics, Medical University of Gdańsk, Gen. J. Hallera 107, 80-416 Gdańsk, Poland

ARTICLE INFO

Article history:
Received 28 March 2011
Received in revised form 17 May 2011
Accepted 21 May 2011
Available online 27 May 2011

Keywords:
Extrathermodynamics
LFER
Liquid Chromatography
Molecular descriptors
Phase equilibrium
QSRR
Retention modeling
Reversed-phase HPLC
Solvophobic theory
Thermodynamics

ABSTRACT

To predict a given physicochemical or biological property, and hence, to design rationally requested chemical entity, the relationships must be identified between the chemical structure and the desired property. Unfortunately, classical thermodynamics never predicts any property by itself, even so simple one like chromatographic retention. Therefore progress in understanding and describing molecular equilibrium between phases requires a combination of experimental measurements and correlations by means of empirical equations and approximate theories. In this work the retention prediction performance was tested of the well thermodynamically founded solvophobic theory of Horváth and co-workers of reversed-phase HPLC. The retention parameters of four series of analytes were modeled with regard to their chemical structure by: (1) observing the rules of classical thermodynamics; (2) applying an extrathermodynamically derived correction to the model based on the thermodynamic hermeneutics; (3) using extrathermodynamic, chemical intuition-based Quantitative Structure–Retention Relationships (QSRR). The combined thermodynamic/extrathermodynamic model with empirical correction accounting for the number of polar atoms provided an improvement of the agreement between the observed and the predicted retention parameters. However, a purely extrathermodynamic QSRR model, employing analyte descriptors from calculation chemistry, produced similar retention predictions. Both thermodynamic and QSRR models accounted well for abilities of analyte to participate in nonspecific, dispersive intermolecular interactions. Less reliable appeared descriptors of analyte polarity. The approach presented here can be further developed to search for proper polarity parameters, necessary to correctly predict complex physicochemical and biological properties of chemical compounds.

1. Introduction

In 1979 Prausnitz [1] formulated the often cited phrases "Classical thermodynamics is revered, honored and admired, but in practice it is inadequate"; "Gibbsian thermodynamics, because of its breadth and elegance, is one of the highlights of classical science, but at the same time it possesses severe limitations for practical work"; "Unfortunately, classical thermodynamics never predicts any property by itself but only relates one property to another".

There is an approach to chemical problems which lacks the rigor of thermodynamics but can provide otherwise inaccessible information. This *extrathermodynamic* approach combines detailed models of physicochemical processes with certain concepts of thermodynamics [2]. Extrathermodynamic in character are linear free-energy relationships (LFER) acknowledged by chemists [3].

Assuming LFER, one can treat a physicochemical property measuring system, for example, chromatographic system, as a "a free-energy transducer" translating differences in chemical potentials of analytes, arising from differences in structure, into quantitative differences in the property, for example, chromatographic retention [4]. The retention parameters are assumed to depend on the free-energy change associated with the analyte distribution between the stationary and the mobile phase of the separation system.

The separative equilibria in chromatography involve adjustments in concentrations of the analyte i in two partitioning systems α and β to satisfy the criterion of equal chemical potentials, $\mu_i^s - \mu_i^m$, where superscript denotes stationary, s, and mobile phase, m, phases. There is a well known contribution to the actual chemical potential μ_i, of the standard-state chemical potential, μ_i^m, and of dilution, c_i, namely: $\mu_i = \mu_i^0 + RT \ln c_i$, where R is gas constant and T is absolute temperature. Certainly, μ_i^0 is a function of chemical structure of interacting entities. Unfortunately, thermodynamics is helpless as regards its calculating [5].

* Corresponding author. Tel.: +48 58 349 3260; fax: +48 58 349 3262.
E-mail address: roman.kaliszan@gumed.edu.pl (R. Kaliszan).
[1] Present address: College of Pharmacy, University of Aden, Yemen. Tel.: +967 7773735/5.

0021-9673/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.chroma.2011.05.077

Change of μ_i^0 with partitioning between two phases, $\Delta\mu_i^0 = \mu_i^{0\alpha} - \mu_i^{0\beta}$, cannot be related rigorously by any means to molecular parameters. Therefore, as concludes Giddings [5]: "Because pure theory is impractical, progress in understanding and describing molecular equilibrium between phases requires a combination of careful experimental measurements and correlations by means of empirical equations and approximate theories". Such an approach to the analysis of chromatographic data will be presented here and discussed in terms of predictive capabilities of the thermodynamics as compared to extrathermodynamics, which is at the basis of the so-called Quantitative Structure–Retention Relationships (QSRR) [6].

In QSRR analysis one relates a set of quantitatively comparable retention parameters for a sufficiently large series of analytes to a set of their structural parameters (descriptors). Through the use of computerized chemometrics techniques, retention parameters are characterized in terms of various analyte descriptors.

The number of structural descriptors which can be assigned to an individual analyte is practically unlimited [7–9]. The problem often arises with physical meaning of the individual descriptors, especially if descriptors of obscure meaning provide QSRR of a well proven statistically retention prediction potency [2,8]. In our opinion statistical criteria alone are insufficient and the descriptors applied in QSRR should be in agreement with thermodynamics hermeneutics: formally or at least according to chemical intuition. Such descriptors are used in the QSRR analysis considered here.

2. Results and discussion

The first comprehensive theoretical frameworks that account for the governing factors in reversed-phase liquid chromatography were given in years 1976–1987 [10–15]. The approach of Horváth et al. [11], further developed [16,17] and discussed [18], got a wide recognition under the name of solvophobic theory.

According to the theory the following simplified expression describes retention coefficient, k, for a given analyte when the eluent composition is changed at constant temperature and flow rate within a fixed column:

$$\ln k = A + B\delta + C\gamma + D(k_e - 1)V^{2/3}\gamma + E + \ln\frac{RT}{P_0 V} \quad (1)$$

where A and C may be regarded as constants and can be determined experimentally; B, D, E and the last term are also constant and can be estimated by the relations given in the original work [11]. P_0 – the atmospheric presume; V – the mole volume of the eluent; γ – its surface tension; κ^e is a parameter related to the internal energy change associated with the vaporization of the solvent.

When the same eluent and column are used, the retention coefficient of different analytes may be obtained at a fixed temperature and flow rate from:

$$\log k = A' + B'\frac{1-\lambda}{2\lambda}\frac{{\mu_s}^2}{V_A}\frac{1}{1-\alpha_s/V_A} + C'\Delta A \quad (2)$$

where A', B' and C' are constants; μ_A is the analyte static dipole moment; V_A – the molecular volume of the analyte; α_A – its polarizability; ΔA – the contact surface area of associated species A (analyte) and L (hydrocarbon ligand of stationary phase), i.e. $\Delta A = \Delta AL - \Delta L$; V_A – molecular volume of an analyte; and λ is the proportionality factor reflecting the volume change occurring upon binding of the analyte to the ligand of the stationary phase.

With some simplifying assumptions and approximations Horváth et al. [11] arrived at the following expressions for the coefficients A', B' and C' of Eq. (2):

$$A' = \frac{(\Phi - (\Delta F_{vdv,assoc}/RT) + (\Delta F_{vdv,S}/RT) + (4836N^{1/3}(k_e - 1)V^{2/3}\gamma/RT) + \ln(RT/P_0V))}{2.303} \quad (3)$$

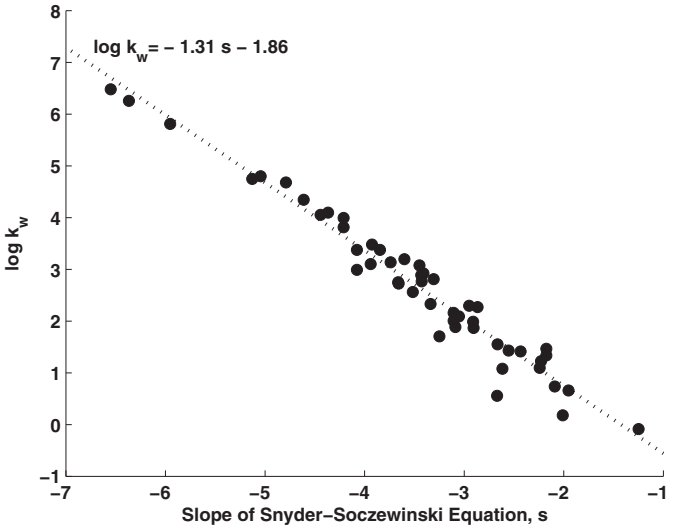

Fig. 1. Relationship between intercepts, log k_w, and slopes, s, of the relationships between logarithms of retention coefficients, k, and percent of methanol, φ, in methanol–buffer eluent for 49 nondissociated analytes (pKa < 8 for acids and pKa > 6 for bases) chromatographed [19] on an Inertsil column.

$$B' = \frac{N\tilde{D}}{4\pi\varepsilon_0 RT 2.303} \quad \tilde{D} = \frac{2(\varepsilon - 1)}{2\varepsilon + 1} \approx 1 \quad (4)$$

$$C' = \frac{N\gamma}{RT 2.303} \quad (5)$$

where $\Delta F_{vdw,assoc}$ is the free-energy change of association of analyte A and ligand L in hypothetical gas phase without any influence of eluent; $\Delta F_{vdw,S}$ is the van der Waals component of the interaction of the analyte with the eluent; N is Avogadro number; ε and ε_0 are electric permitivites of the eluent and the vacuum, respectively; Φ is a characteristic constant: logarithm of the so-called phase ratio. In case of water solvent ($\log k = \log k_w$), the theoretical values for B' and C' are: $C' = 0.0816\,\text{Å}^{-2}$ and $B' = 10.549\,\text{Å}^3/\text{D}^2$.

Let us now analyze a first set of experimental reversed-phase HPLC data previously reported from our laboratory [19] and collected in Table 1. Consistency of the data considered well illustrates Fig. 1. To calculate $\log k$ corresponding to 100% v/v of buffer as the mobile phase, i.e. $\log k_w$, let us first apply Eq. (2) with two variable terms:

$$\beta = \frac{1-\lambda}{2\lambda}\frac{\mu_s^2}{V_A}\frac{1}{1-\alpha_s/V_A} \quad \text{and} \quad \alpha = \Delta A \quad (6)$$

For ΔA, as a first approximation, can be substituted the water accessible surface of analyte molecule, ΔA_{WAS}. Parameter β can be calculated from Eq. (6) if value of λ is represented by $\lambda = (695 + V_A)/V_A$ where 695 approximates volume of octadecyl ligand, $C_{18}H_{37}$, of stationary phase (in Å^3) and V_A is molecular volume of individual analytes. Hence, the following regression equation of the type of Eq. (2) results:

$$\log k_w = -1.96(\pm 0.74) + 164(\pm 31)\beta + 0.0161(\pm 0.002)\Delta A_{WAS}$$
$$r^2 = 0.719 \quad SD = 0.810 \quad F = 58.8 \quad p < 0.0001 \quad n = 49 \quad (7)$$

where r denotes correlation coefficient, SD – standard deviation error of regression, F – value of the F-test of significance; p – significance level, n – number of regressions and in parenthesis are given the standard deviations of regression coefficients.

The coefficients at the dependent variables in Eqs. (7) and (2) can be mutually compared: the theoretically calculated $C' = 0.0816$ is of the same order of magnitude like the statistically derived value

Table 1
Intercepts, log k_w, and slopes, s, of the linear relationship between logarithms of retention coefficients determined on an octadecylsilica column (Inertsil) and percent of methanol in aqueous eluent for a series of test analytes [19], along with their structural parameters from molecular modeling: ΔA_{AWS} – water accessible surface area; ΔA_{HSA} – hydrophobic surface area; N_H – number of hydrogen-bonding atoms; μ_A – total dipole moment; δ_{min} – electron excess charge of the most negatively charged atom; V_A – molecular volume of analyte, β is analyte parameter calculated from Eq. (6).

No.	Name	log k_w	s	V_A	ΔA_{WAS}	ΔA_{HSA}	N_H	μ	δ_{min}	β
1	Acridone	2.09	−3.05	589.4	384.0	324.0	2	4.220	−0.324	−0.0085
2	Aniline	1.10	−2.24	367.5	265.0	195.5	1	1.583	−0.412	−0.0023
3	Anisole	2.30	−2.95	407.9	288.2	269.5	1	1.249	−0.212	−0.0012
4	Anthracene	4.75	−5.13	585.4	381.6	365.3	0	0.000	−0.127	0.0000
5	Anthraquinone	3.20	−3.60	602.1	389.1	322.5	2	0.000	−0.286	0.0000
6	Benzamide	1.07	−2.62	418.2	292.7	194.4	2	3.583	−0.433	−0.0099
7	Benzene	2.27	−2.86	331.8	245.0	240.0	0	0.000	−0.130	0.0000
8	Benzonitrile	1.86	−2.91	389.7	277.6	216.7	1	3.336	−0.135	−0.0094
9	Benzyl chloride	2.77	−3.42	425.2	295.6	287.2	0	1.494	−0.128	−0.0017
10	Biphenyl	4.05	−4.44	538.3	358.4	345.7	0	0.000	−0.131	0.0000
11	4,4′-Bipyridine	1.70	−3.25	511.7	340.6	285.6	2	0.000	−0.168	0.0000
12	1-Bromonaphthalene	4.09	−4.37	515.1	342.0	329.6	0	1.414	−0.154	−0.0012
13	Caffeine	0.74	−2.09	569.4	369.6	305.2	6	3.709	−0.362	−0.0069
14	Carbazole	3.37	−4.08	546.1	360.5	346.3	1	1.206	−0.245	−0.0008
15	2-Chloroaniline	1.99	−2.91	405.6	285.4	263.5	1	1.676	−0.401	−0.0023
16	1-Chloroanthraquinone	3.08	−3.45	639.4	408.0	392.1	2	0.899	−0.286	−0.0003
17	Chlorobenzene	2.89	−3.43	375.3	269.6	262.8	0	1.306	−0.129	−0.0015
18	4-Chlorophenol	2.33	−3.34	396.2	280.8	273.5	1	1.477	−0.248	−0.0018
19	2-Chloropyridine	1.41	−2.43	362.4	262.4	239.6	1	2.823	−0.182	−0.0075
20	4-Cyanophenol	1.55	−2.67	409.9	289.2	201.8	1	3.311	−0.244	−0.0087
21	3-Cyanopyridine	0.66	−1.95	375.0	269.2	185.8	1	2.892	−0.186	−0.0075
22	Cyclohexanone	1.23	−2.23	384.0	269.5	228.1	1	2.972	−0.294	−0.0076
23	Dibenzothiophene	3.99	−4.21	556.0	364.5	351.2	0	0.522	−0.271	−0.0001
24	3,5-Dichlorophenol	3.14	−3.74	440.1	306.0	273.9	1	1.407	−0.243	−0.0014
25	2,2-Dinaphthyl ether	6.26	−6.36	826.6	496.2	493.0	1	1.070	−0.170	−0.0003
26	2,2′-Dipyridil	1.47	−2.18	521.1	349.7	316.3	2	2.978	−0.175	−0.0050
27	Hexachlorobutadiene	4.80	−5.04	516.8	342.8	338.4	0	0.000	−0.073	0.0000
28	*n*-Hexylbenzene	5.81	−5.95	646.6	423.2	409.2	0	0.349	−0.211	−0.0001
29	Hydroquinone	0.56	−2.67	374.1	269.7	209.1	2	0.000	−0.253	0.0000
30	Indazole	1.89	−3.09	405.0	285.5	241.3	2	1.547	−0.203	−0.0019
31	Indole	2.16	−3.11	420.1	292.6	233.1	1	1.883	−0.219	−0.0027
32	4-Iodophenol	2.75	−3.66	434.3	300.4	147.9	1	1.585	−0.302	−0.0018
33	Isopropylbenzene	3.81	−4.21	479.3	321.9	314.8	0	0.247	−0.206	0.0000
34	1-Methyl-2-pyrrolidinone	−0.09	−1.25	381.5	271.7	216.3	2	3.593	−0.353	−0.0112
35	Naphthalene	3.38	−3.84	459.0	313.3	302.6	0	0.000	−0.128	0.0000
36	2-Naphthol	2.56	−3.52	408.4	325.3	264.5	1	1.460	−0.252	−0.0017
37	1,4-Naphthaquinone	2.01	−3.11	480.2	325.2	247.9	2	1.332	−0.270	−0.0011
38	1-Naphthylacetonitrile	2.73	−3.66	556.4	365.8	302.0	1	3.031	−0.138	−0.0048
39	Nicotinamide	0.18	−2.01	404.7	283.9	180.7	3	4.906	−0.432	−0.0194
40	Phenanthrene	4.34	−4.61	577.6	376.3	357.7	0	0.020	−0.128	0.0000
41	Phenol	1.43	−2.56	353.1	256.2	201.3	1	1.233	−0.253	−0.0015
42	Phenylhydrazine	1.34	−2.18	404.3	282.2	184.3	2	1.557	−0.233	−0.0020
43	4-Phenylphenol	3.10	−3.94	559.5	370.0	306.1	1	1.183	−0.249	−0.0007
44	Pyrene	4.68	−4.79	618.6	393.8	375.4	0	0.000	−0.127	0.0000
45	Toluene	2.81	−3.31	384.7	274.4	265.6	0	0.264	−0.179	−0.0001
46	2,4,6-Trichloroaniline	3.48	−3.93	487.1	331.3	274.4	1	2.059	−0.385	−0.0027
47	3-Trifluoromethylphenol	2.99	−4.08	432.4	300.9	149.3	1	2.092	−0.245	−0.0032
48	1,3,5-Triisopropylbenzene	6.47	−6.55	777.1	477.8	477.5	0	0.004	−0.206	0.0000
49	Xanthene	2.92	−3.41	581.4	375.0	365.1	1	1.146	−0.152	−0.0006

0.0161 at ΔA_{WAS} in Eq. (7). The thermodynamically based value of $B' = 10.549$ Å^3/D^2 is significantly smaller than the corresponding coefficient in Eq. (7), which is equal to 164, however.

An improving trend can be observed if the parameter ΔA_{WAS} in Eq. (7) is replaced by ΔA_{HCA} – the hydrophobic contact area, i.e. the water accessible molecular surface area calculated after neglecting polar atoms N, O and F. The following regression equation based on data from Table 1 results:

$$\log k_w = -1.39(\pm 0.52) + 94.8(\pm 29.4)\beta + 0.0152(\pm 0.0016)\Delta A_{HCA}$$
$$r^2 = 0.791 \quad SD = 0.699 \quad F = 86.9 \quad p < 0.0001 \quad n = 49 \tag{8}$$

A further improvement of retention prediction we managed to obtain after applying an approach based on the modified solvophobic theory [17]. To elaborate it Vailaya and Horvàth [17] started from defining retention coefficient as $k = K\Phi$, where K is the equilibrium constant described by the known relationship: $\ln K = -\Delta G_R^0/R/T$, where ΔG_R^0 denotes the standard free energy change associated with analyte transfer from the mobile to the stationary phase. Now, the basic modification of the original solvophobic theory consists in the following assumption: since the change in surface area upon binding of analyte to the hydrocarbonaceous stationary phase ligand cannot be measured, the linear relationship between ΔA and ΔA_{np} must be considered.

$$\Delta A = \Delta A_{np}\alpha' + \beta' \tag{9}$$

where ΔA_{np} is the nonpolar water accessible surface area of the analyte and α' and β' are coefficients which are constant for a set of homologues.

With the assumption expressed by Eq. (9), the fundamental equation of the modified solvophobic theory is as follows:

$$\Delta G_R^0 = [(\kappa_H^g\gamma_H^g - \kappa_{HE}^g\gamma_{HE}^g)\alpha' - 0.25a + a]\Delta A_{np} - RT\ln\frac{RT}{V} + b' \tag{10}$$

where γ_H and γ_{HE} are hydrocarbon part of analyte surface tension and hydrocarbon/eluent interfacial tension, respectively, κ_H^g and κ_{HE}^g convert the respective surface or interfacial tension to the

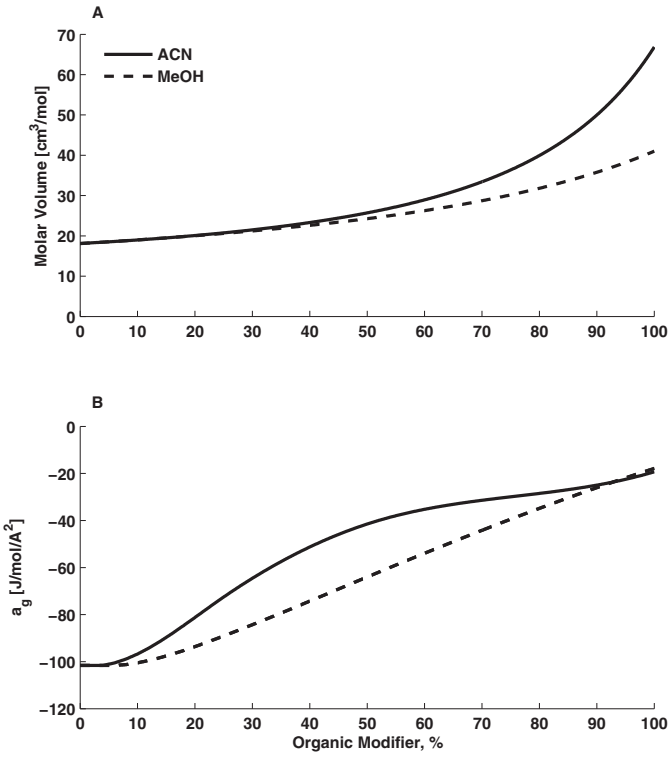

Fig. 2. (A) Molar volume of aqueous eluent in function of concentration (% v/v) of methanol and acetonitrile as organic modifiers. (B) Contributions due to the mobile and stationary phase effects, a_g, to the free energy of retention as a function of concentration (% v/v) of methanol and acetonitrile as organic modifiers of aqueous eluents. The numerical values of a_g were digitalized from Fig. 8 in the original Ref. [8].

microthermodynamic value applicable to molecular dimensions, b' is the constant arising from replacement of ΔA by ΔA_{np} and it is negligible for nonpolar and weakly polar analytes. V – the molar volume of the eluent, depends on volume fraction of organic modifier linearly: in a wider concentration range in the case of methanol than for acetonitrile (Fig. 2).

The constant a in Eq. (10) originates from the following relationship:

$$\Delta G_R^0 = a\Delta A_{np} + b \tag{11}$$

which accounts for the change in free energy of association of analytes in gas phase.

The term a_g at ΔA_{np}, $a_G = [(\kappa_H^g\gamma_H^g - \kappa_{HE}^g\gamma_{HE}^g)\alpha' - 0.25a + a]$ represents the contribution to retention due to the mobile and the stationary phase effects to the free energy change. To get numerical values of a_g we digitalized relationships presented in Fig. 8 in the original Ref. [8] using program GetData Graph Digitizer v. 2.4. The resulting data were plotted in Fig. 2. It is evident that in case of methanol a_g changes linearly with the modifier concentration above 10% v/v.

Having in mind that and $k = K\Phi$, Eq. (10) can be transformed into: $\ln K = -\Delta G_R^0/R/T$

$$\ln k = \ln \Phi - a_g(\varphi)\frac{\Delta A_{np}}{RT} + \ln\left(\frac{RT}{V(\varphi)}\right) - \frac{b'}{RT} \tag{12}$$

The $\log k$ eqals $\log k_w$ if $\varphi = \varphi_0$. Therefore:

$$\log k_w = \log \Phi - [a_g(\varphi) - a_g(\varphi_0)]\frac{\Delta A_{np}}{2.303RT} + \log\left(\frac{RT}{V(\varphi_0)}\right) - \frac{b'}{2.303RT} \tag{13}$$

Combining the last two equations one obtains a dependence of $\log k$ on φ in relation to $\log k_w$:

$$\log k = \log k_w - (a_g(\varphi) - a_g(\varphi_0))\frac{\Delta A_{np}}{2.303RT} + \log\left(\frac{V(\varphi_0)}{V(\varphi)}\right) \tag{14}$$

From Eq. (14) one gets a relationship describing the slope, s, of the Snyder–Soczewiński relationship, $\log k = \log k_w + s\varphi$, by derivation:

$$s = \frac{d}{d\varphi}\log k(\varphi) = -\frac{\Delta A_{np}}{2.303RT}\frac{d}{d\varphi}a_g(\varphi) + \frac{d}{d\varphi}\log\left(\frac{V(\varphi_0)}{V(\varphi)}\right) \tag{15}$$

As evident from Fig. 2, the $a_g(\varphi)$ term can be well approximated by a straight line in the range of methanol concentrations from 10 to 100% v/v. The appropriate linear equation is:

$$a_g(\varphi) \approx -112 + 96.21\varphi \tag{16}$$

Thus, the first derivative equals to 96.21:

$$\frac{\partial}{\partial\varphi}a_g(\varphi) \approx 96.21 \tag{17}$$

Assuming that, as a first approximation, the relationship in Fig. 2 between molar volume of eluent and the content of methanol is linear, we get for the last term in Eq. (14) the following linear regression:

$$\log\left(\frac{V(\varphi_0)}{V(\varphi)}\right) \approx -0.360\,\varphi + 0.0394 \tag{18}$$

Its first derivative equals:

$$\frac{\partial}{\partial\varphi}\log\left(\frac{V(\varphi_0)}{V(\varphi)}\right) \approx -0.360 \tag{19}$$

By combining Eqs. (17) and (19) with Eq. (15) one obtains:

$$s = -\frac{\Delta A_{np}}{2.303RT}96.21 - 0.360 = -0.0169\Delta A_{np} - 0.360 \tag{20}$$

Coming back to Eq. (13) and putting in the numerical values for all the known physicochemical parameters, with setting the term $(b'/2.303RT) = 0$ one gets the relationship for $\log k_w$ as:

$$\log k_w = \log\phi + 0.0196\Delta A_{np} + 3.13 \tag{21}$$

Now, combining Eqs. (20) and (21) one gets for the data presented in Table 1:

$$\log k_w = \log\phi + 2.71 - 1.160s \tag{22}$$

The slope −1.16 in Eq. (22) is in a good agreement with the slope −1.31 for experimental data presented in Fig. 1.

The problem with the use of Eqs. (20) and (21) is the lack of precise computational methods for ΔA_{np} determinations for individual analytes. For the nonpolar analytes, the water accessible surface area, ΔA_{WAS}, can be treated as actually ΔA_{np}. Of 49 analytes in Table 1, 15 are nonpolar. For them $\log k_w$ well correlates with ΔA_{WAS} (Fig. 3). The slope of this relationship is 0.0184, which is in excellent agreement with the value 0.0196 derived for Eq. (21).

To overcome the difficulty with calculating ΔA_{np} for polar analytes we undertook an attempt to introduce a correction factor, α'', for the number of hydrogen acceptor/donor atoms, N_H, in analyte molecule. Hence, the following modification of Eqs. (20) and (21) was proposed and tested for whole set of analytes in Table 1:

$$\log k_w = \log\phi + 0.0196(\Delta A_{WAS} + \alpha'' N_H) + 3.13 \tag{23}$$

$$s = -0.0169(\Delta A_{WAS} + \alpha'' N_H) - 0.360 \tag{24}$$

The above equations can be converted into the corresponding regular QSRR regression equations:

$$\log k_w = k_0 + k_1 A_{WAS} + k_2 N_H \quad \text{where} \quad k_0 = \log\phi + 3.13, \quad k_1 = 0.0196, \quad k_2 = 0.0196\alpha' \tag{25}$$

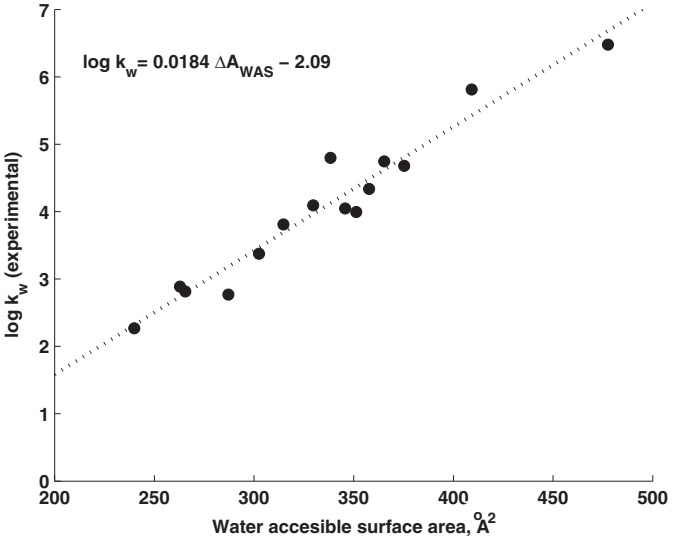

Fig. 3. The correlation between experimental log k_w for nonpolar analytes from Table 1 and their water accessible surface area, ΔA_{WAS}.

$$s = k_{S0} + k_{S1}A_{WAS} + k_{S2}N_H \quad \text{where} \quad k_{S0} = -0.360, \quad k_{S1} = -0.0169, \quad k_{S2} = -0.0169\alpha' \tag{26}$$

The regression coefficients determined for the set of 49 analytes of Table 1 are:

$$\log k_w = -2.59(\pm 0.43) + 0.0193(\pm 0.0013)A_{AWS} - 0.820(\pm 0.069)N_H \quad r^2 = 0.889 \quad SD = 0.508 \quad F = 186 \quad p < 0.0001 \quad n = 49 \tag{27}$$

$$s = 0.594(\pm 0.388) - 0.0147(\pm 0.0012)A_{AWS} + 0.555(\pm 0.062)N_H \quad r^2 = 0.840 \quad SD = 0.457 \quad F = 121 \quad p < 0.0001 \quad n = 49 \tag{28}$$

The values of log k_w and s predicted from Eqs. (27) and (28), respectively, vs. corresponding experimental data from Table 1, are presented in Fig. 4.

There is a striking agreement of the coefficients at molecular surface descriptors in Eq. (23) (0.0196) and Eq. (27) (0.0193) describing log k_w as well as in the pair of equations accounting for s: Eq. (24) (−0.0169) and Eq. (28) (−0.0147). That proves that a thermodynamically guided approach to statistical modeling of retention of analytes in relation to their structure helps not only to get insight into physicochemical mechanism of reversed-phase HPLC but also allows for improvement of retention predictions and hence, for rational optimization of separations.

Now, it appeared interesting to compare the above presented thermodynamically-based retention models with QSRR models, i.e. those derived statistically with a trial-and-error approach employing various molecular structural descriptors of analytes provided by computation chemistry. The simple and most robust extrathermodynamic QSRR model applicable to reversed-phase HPLC was built [19,20] using the following analyte descriptors: μ – total dipole moment, $\delta_{\min}$ – electron excess charge of the most negatively atom, A_{WAS} – water accessible molecular surface area:

$$\log k_w = k'_1 + k'_2\mu^2 + k'_3\delta_{\min} + k'_4A_{WAS} \tag{29}$$

where $k'_1 - k'_4$ are regression coefficients. The descriptors used were calculated by HyperChem Chem Plus software (HyperCube, Waterloo, Canada) after geometry optimization by the molecular mechanics MM+ force field method, followed by quantum chemical calculations by the semiempirical AM1 method.

For the analytes listed in Table 1 the following QSRR equation was derived:

$$\log k_w = -1.66(\pm 0.72) - 0.100(\pm 0.023)\mu^2 + 3.15(1.24)\delta_{\min} + 0.0167(\pm 0.00190)\Delta A_{WAS} \quad r^2 = 0.779 \quad SD = 0.727 \quad F = 52.8 \quad p < 0.0001 \quad n = 49 \tag{30}$$

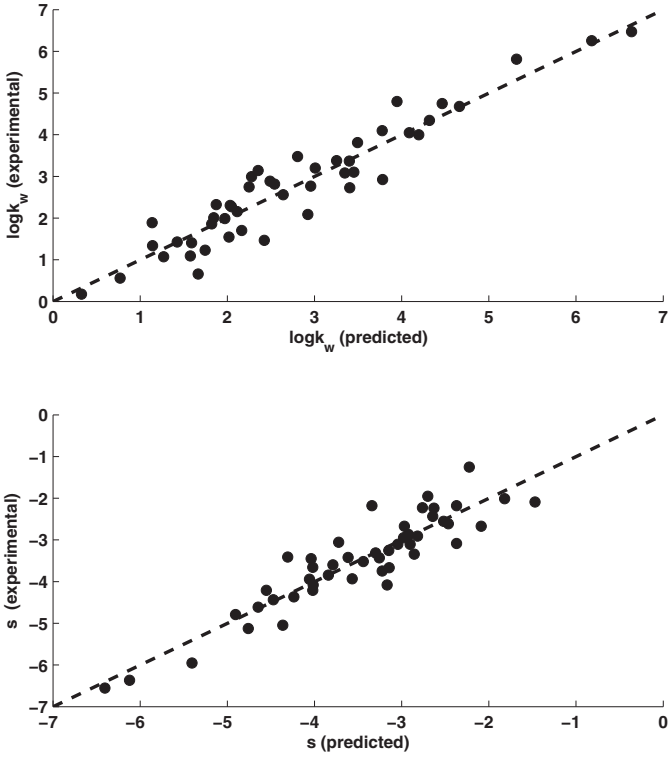

Fig. 4. The parameters log k_w and s, experimental from Table 1 vs. predicted by Eqs. (27) and (28), respectively.

In case of slope, s, the highly correlated to log k_w parameter (see Fig. 1), the corresponding QSRR equation is:

$$s = 0.008846(\pm 0.575) + 0.0700(\pm 0.0181)\mu^2 - 1.92(1.07)\delta_{\min} - 0.0129(\pm 0.0015)\Delta A_{WAS} \quad r^2 = 0.750 \quad SD = 0.578 \quad F = 44.9 \quad p < 0.0001 \quad n = 49 \tag{31}$$

Actually, the free term in Eq. (31) is insignificant but we leave it for the sake of consistency. Plots of experimental vs. predicted from Eqs. (30) and (31) data are given in Fig. 5.

Eqs. (30) and (31) make a good physical sense, although they were proposed based on chemical intuition rather than on thermodynamic hermeneutics. According to Eq. (30), log k_w increases with size ("bulkiness") of analytes, reflected by the magnitude of ΔA_{WAS} parameter and decreases with polarity of analytes, reflected by μ. The reverse is true in case of s (Eq. (31)). All that is in agreement with the observations and can be explained in terms of intermolecular interactions determining chromatographic separations.

If A_{WAS} is to account for nonpolar interactions of analytes (mostly dispersive London–Hall interactions), then one agrees that such interactions (which are ever attractive) will be stronger between the analytes and the bulky octadecyl moieties of the stationary phase than between the same analytes and the small molecules of aqueous eluent. Hence, one can rationalize increase of log k_w with carbon number within a series of homologues. As regards dipolar interactions (dipole–dipole Debay interactions and dipole-induced dipole Keesom interactions), these will be stronger between the analytes and the polar eluent molecules than the nonpolar hydrocarbonaceus stationary phase. Hence, log k_w, for corresponding analytes of homologous series will be lower for those possessing a more polar functionality, as it has been clearly demonstrated, e.g. in Refs. [11,12], and in particular in a review on solvophobic theory by Vailaya and Horváth [21].

Similar argumentation can be applied to Eq. (31) to explain the reverse trend in changes of slope, s, with the analyte size and polarity.

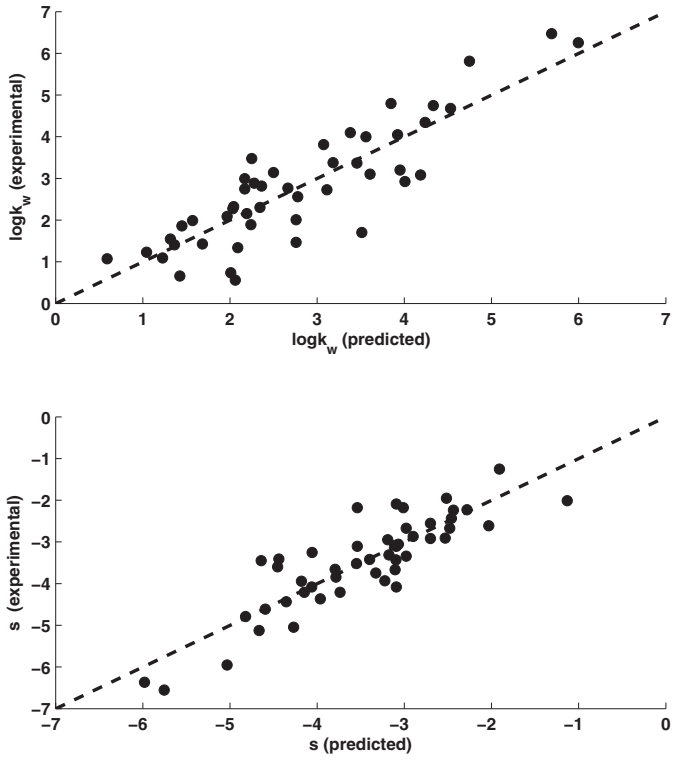

Fig. 5. The parameters log k_w and s, experimental from Table 1 vs. predicted by QSRR Eqs. (30) and (31), respectively.

Mechanistic interpretation of Eqs. (30) and (31) fully applies to Eqs. (27) and (28) if N_H is considered a polarity descriptor of analytes.

Statistical quality and hence the retention modeling reliability of QSRR (Eqs. (30) and (31)) is certainly limited but it is even better than that of the corresponding equations based on thermodynamic hermeneutics, i.e. Eqs. (7) and (8). However, combining of thermodynamic modeling with an extrathermodynamically derived correcting factor results in the models (Fig. 4) of a quality acceptable from the point of view of elucidation of physicochemical mechanism of reversed-phase HPLC separation and prediction of the log k_w and s parameters. Of course, having these parameters calculated from the structure of individual analytes one can evaluate their retention at specified eluent composition and thus optimize the conditions of their separation.

The approach to the reversed-phase HPLC theory here developed has been tested on another sets of retention data reported by various laboratories. In Table 2 are given log k_w and s parameters determined on a column home-packed with a Merck RP-18 material in a methanol-gradient mode by Schoenmakers et al. [22] We excluded from present study the nitro derivatives (five), and anthracene and naphathalene, which high log k_w values (above 12) have been questioned by the original authors themselves. In Table 2 are also given structural parameters of the analytes calculated by us, analogously as in Table 1.

The relationship between log k_w and s is illustrated in Fig. 6. There is a nearly perfect agreement of the slope of that relationship (−1.30) with analogous parameter for data presented in Fig. 1 (−1.31).

The theoretical values for parameters B' and C' of Horváth equation (Eq. (2)) are: $B' = 10.549\ Å^3/D^2$ and $C' = 0.0816\ Å^{-2}$. Agreement of these theoretical B' and C' values with the corresponding coefficients of the equations analogous to Eq. (7) (involving the calculated analyte dependent variables β and ΔA_{WAS}) and Eq. (8) (with variables β and ΔA_{HCA}) was similarly limited as previously discussed.

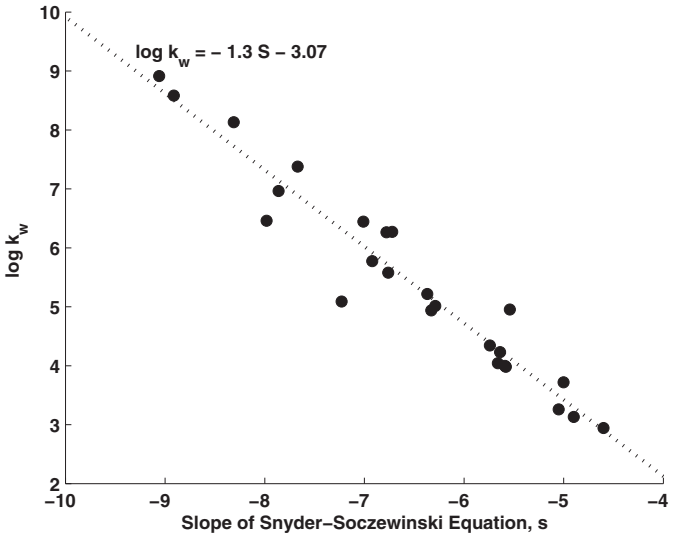

Fig. 6. The relationship between intercepts, log k_w, and slopes, s, of the relationship between logarithms of retention coefficients, k, and percent of methanol, φ, in methanol–water eluent for 25 test analytes chromatographed [22] on a C18 Merck stationary phase in methanol-gradient mode.

A combined thermodynamic/extrathermodynamic model with the correction of analyte hydrophobic contact area for hydrogen-bonding atoms, provided the following relationships for the data listed in Table 2:

$$\log k_w = -1.48(\pm 0.91) + 0.0280(\pm 0.0034)\Delta A_{WAS} - 0.702(\pm 2.289)N_H \quad (32)$$
$$r^2 = 0.754 \quad SD = 0.869 \quad F = 33.8 \quad p < 0.0001 \quad n = 25$$

$$s = -1.13(\pm 0.689) - 0.0207(\pm 0.0026)\Delta A_{WAS} + 0.239(\pm 0.143)N_H \quad (33)$$
$$r^2 = 0.745 \quad SD = 0.0.658 \quad F = 32.23 \quad p < 0.0001 \quad n = 25$$

The retention modeling potency of Eqs. (32) and (33) is illustrated in Fig. 7. It is comparable to that of Eqs. (27) and (28)

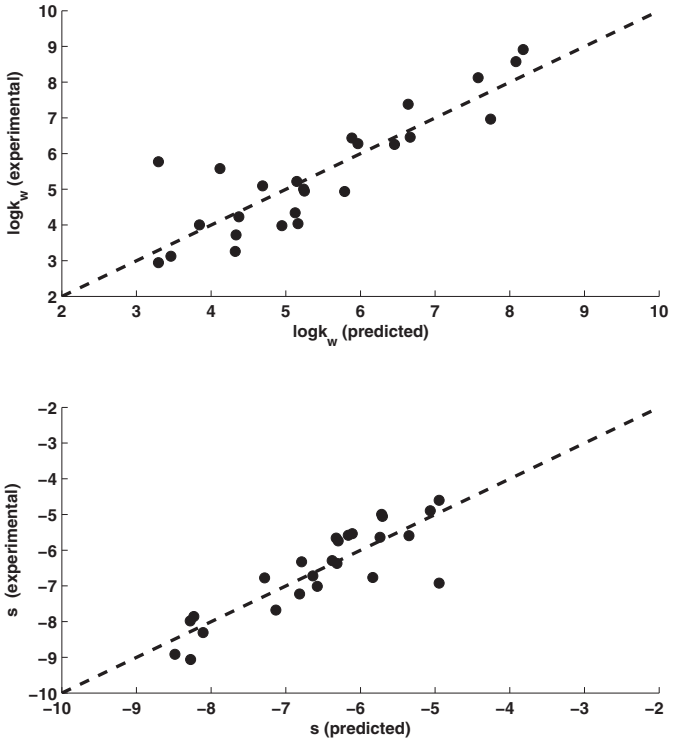

Fig. 7. The parameters log k_w and s, experimental from Table 2 vs. predicted by Eqs. (32) and (33), respectively.

Table 2
Intercepts, log k_w, and slopes, s, of the linear relationship between logarithms of retention coefficients determined on an octadecylsilica column (home-packed with RP-18 Merck material) and percent of methanol in aqueous eluent for a series of test analytes [22], along with their structural parameters from molecular modeling: ΔA_{WAS} – water accessible surface area; ΔA_{HSA} – hydrophobic surface area; N_H – number of hydrogen-bonding atoms; μ_A – total dipole moment; δ_{min} – electron excess charge of the most negatively charged atom; V_A – molecular volume of analyte, β is analyte parameter calculated from Eq. (6).

No.	Name	log k_w	s	V_A	ΔA_{WAS}	ΔA_{HSA}	N_H	μ	δ_{min}	β
1	Acetophenone	4.34	−5.74	435.1	292.9	260.7	1	2.843	−0.306	−0.0059
2	Anethole	8.13	−8.31	549.4	359.5	348.1	1	1.240	−0.210	−0.0008
3	Aniline	2.94	−4.60	367.5	258.1	195.5	1	1.584	−0.412	−0.0023
4	Anisole	5.77	−6.92	519.7	348.3	195.5	1	1.281	−0.215	−0.0009
6	Benzene	3.72	−5.00	386.7	268.6	232.6	1	2.930	−0.289	−0.0074
7	Benzaldehyde	4.95	−5.54	332.0	240.0	240.0	0	0.000	−0.130	0.0000
8	Benzonitrile	4.00	−5.59	388.7	270.0	215.0	1	3.336	−0.135	−0.0095
9	Benzophenone	6.96	−7.86	591.1	375.0	354.0	1	2.579	−0.318	−0.0032
10	Benzyl alcohol	3.26	−5.05	406.7	277.4	232.1	1	1.457	−0.326	−0.0017
11	Biphenyl	8.91	−9.06	538.0	344.6	344.6	0	0.000	−0.131	0.0000
12	Chlorobenzene	6.44	−7.01	375.0	262.6	262.6	0	1.306	−0.130	−0.0015
13	o-cresol	4.23	−5.64	403.3	275.2	233.8	1	0.955	−0.254	−0.0007
14	Diethyl o-phthalate	6.46	−7.98	696.2	436.8	390.9	4	5.891	−0.303	−0.0129
15	N,N-Dimethylaniline	6.26	−6.78	464.5	308.2	308.2	1	1.507	−0.310	−0.0015
16	2,4-Dimethylaniline	5.22	−6.37	455.2	302.9	261.5	1	1.027	−0.254	−0.0007
17	Dimethyl o-phthalate	5.09	−7.23	583.2	375.0	320.3	4	6.026	−0.300	−0.0175
19	Diphenyl ether	8.58	−8.91	572.8	369.5	366.2	1	0.761	−0.189	−0.0003
20	Ethylbenzene	7.38	−7.67	432.0	289.6	289.6	0	0.321	−0.211	−0.0001
21	N-methylaniline	5.01	−6.29	420.4	288.3	264.7	1	1.646	−0.361	−0.0021
27	Phenol	3.13	−4.90	353.2	250.3	201.3	1	1.233	−0.253	−0.0015
28	1-Phenylethanol	4.04	−5.66	451.7	302.3	262.0	1	1.626	−0.325	−0.0018
29	2-Phenylethanol	3.98	−5.58	456.1	306.7	254.3	1	1.345	−0.330	−0.0012
30	3-Phenylpropanol	4.94	−6.33	490.3	326.2	284.2	1	2.599	−0.288	−0.0042
31	Quinolone	5.58	−6.76	463.5	308.9	249.9	2	5.278	−0.330	−0.0188
32	Toluen	6.27	−6.72	384.7	265.6	265.6	0	0.264	−0.135	−0.0001

illustrated in Fig. 4, although significance of the N_H term is much lower in the former pair of equations.

The QSRR modeling of data from Table 2 resulted in the following equations:

$$\log k_w = -0.212(\pm 1.57) - 0.051(\pm 0.0229)\mu^2 + 10.21 \rightleftarrows (2.53)\delta_{min} + 0.0279(\pm 0.0048)\Delta A_{WAS} \quad (34)$$
$$r^2 = 0.710 \quad SD = 0.966 \quad F = 17.2 \quad p < 0.0001 \quad n = 25$$

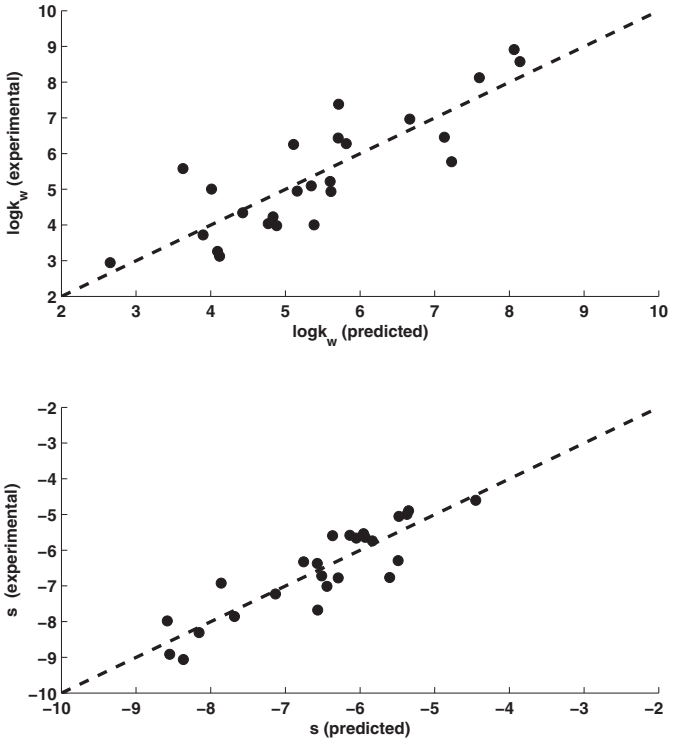

Fig. 8. The parameters log k_w and s, experimental from Table 2 vs. predicted by QSRR Eqs. (34) and (35), respectively.

$$s = -1.263(\pm 0.990) + 0.0225(\pm 0.0144)\mu^2 - 6.60(1.59)\delta_{min} - 0.0231(\pm 0.0030)\Delta A_{WAS} \quad (35)$$
$$r^2 = 0.791 \quad SD = 0.610 \quad F = 26.4 \quad p < 0.0001 \quad n = 25$$

The retention predicting capacity of Eqs. (34) and (35) is illustrated in Fig. 8.

The forms of Eqs. (32)–(35) and their mechanistic interpretation are analogous to those of the corresponding Eqs. (27)–(29) and (31). The main quantitative difference is in a lesser significance of the polarity terms N_H, μ^2 and δ_{min} for description of the retention data determined on the older type RP-18 Merck column in comparison to the highly reliable modern Inertsil material.

A slightly worse fitting of the data from Table 2 to the solvophobic theory-based model is also reflected by a larger difference of the coefficients at the contact area parameters in the thermodynamics-based equations (0.0196 in Eq. (23) describing log k_w and −0.0169

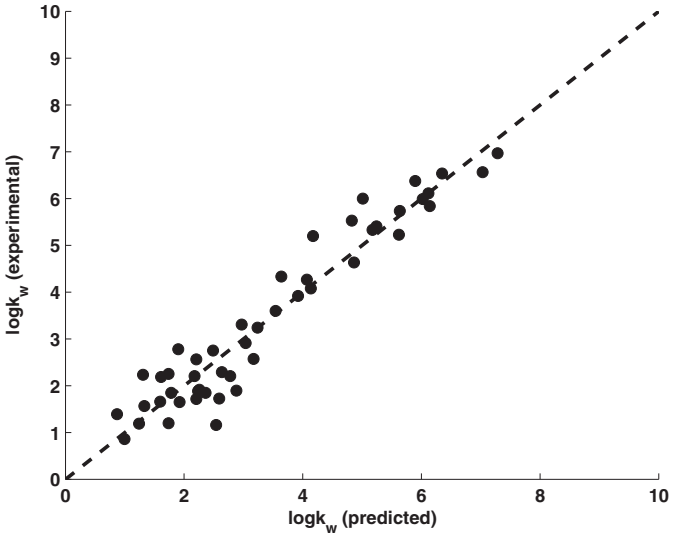

Fig. 9. The parameter log k_w, experimental from Table 3 vs. predicted by Eq. (36).

Table 3
Intercepts, log k_w, of the linear relationship between logarithms of retention coefficients determined on an octadecylsilica kolumn (Acquity BEHShield) and percent of methanol in aqueous eluent for a series of test analytes [23], along with their structural parameters from molecular modeling: ΔA_{WAS} – water accessible surface area; ΔA_{HSA} – hydrophobic surface area; N_H – number of hydrogen-bonding atoms; μ_A – total dipole moment; δ_{min} – electron excess charge of the most negatively charged atom; V_A – molecular volume of analyte, β is analyte parameter calculated from Eq. (6).

No.	Name	log k_w	V_A	ΔA_{WAS}	ΔA_{HSA}	N_H	M	δ_{min}	β
1	Acetophenone	1.71	435.0	292.9	260.7	1	2.843	−0.306	−0.0059
2	2-Aminobiphenyl	2.91	564.2	357.0	312.3	1	1.422	−0.406	−0.0010
3	2-Aminonaphthalene	2.2	494.1	322.9	259.3	1	1.868	−0.410	−0.0021
4	Anisole	1.85	408.4	281.7	270.4	1	1.249	−0.212	−0.0012
5	Benz[a]anthracene	5.33	704.5	422.2	422.2	0	0.045	−0.128	0.0000
6	Benzidine	1.85	608.9	382.6	256.9	2	0.000	−0.410	0.0000
7	Benzyl alcohol	1.2	406.4	277.4	232.1	1	1.457	−0.326	−0.0017
8	Benzyl benzoate	4.08	682.2	429.2	402.3	2	2.176	−0.357	−0.0018
9	Benzyl cyanide	1.65	437.2	294.2	243.6	1	2.875	−0.137	−0.0060
10	Bibenzyl	4.63	645.4	403.1	403.1	0	0.000	−0.135	0.0000
11	Biphenyl	3.92	538.2	344.6	344.6	0	0.000	−0.131	0.0000
12	Butyl acetate	1.88	469.9	320.9	285.0	2	1.889	−0.355	−0.0023
13	2-Chloroaniline	1.66	405.6	278.6	223.4	1	1.676	−0.401	−0.0023
14	4-Chlorobenzyl alcohol	1.91	449.8	301.9	263.8	1	2.508	−0.326	−0.0044
16	3-Chlorophenol	2.18	396.2	273.5	224.4	1	1.477	−0.248	−0.0018
17	3-Chlorophenylacetic acid	2.25	501.7	331.0	254.2	2	0.856	−0.360	−0.0004
18	2,4′-DDD	5.84	798.4	482.2	482.2	0	0.571	−0.127	−0.0001
19	4,4′-DDE	6.37	767.1	467.0	467.0	0	0.179	−0.139	0.0000
20	4,4′-DDT	6.53	825.7	494.5	494.5	0	1.048	−0.128	−0.0003
21	Dibenz[a,h]anthracene	6.11	824.1	480.7	480.7	0	0.000	−0.127	0.0000
22	*m*-Dichlorobenzene	3.31	417.9	286.2	286.2	0	1.233	−0.125	−0.0012
23	N,N-Diethylacetamide	1.16	452.8	303.3	281.4	1	3.699	−0.372	−0.0094
24	Ethyl acetate	0.86	355.6	258.0	208.2	2	4.462	−0.307	−0.0190
25	N-Ethylaniline	1.89	476.0	319.9	302.6	1	1.867	−0.360	−0.0022
26	Ethyl benzoate	2.75	513.7	340.5	300.4	2	4.639	−0.316	−0.0124
27	Hexabromobenzene	5.53	658.8	400.9	400.9	0	0.000	−0.128	0.0000
28	Hexachlorobenzene	5.2	564.6	360.4	360.4	0	0.000	−0.052	0.0000
29	Mesitylene	3.6	489.2	321.2	321.2	0	0.056	−0.179	0.0000
30	3-Methylcholanthrene	5.99	813.0	474.7	474.7	0	0.786	−0.177	−0.0002
31	Naphthalene	3.24	458.9	302.6	302.6	0	0.000	−0.128	0.0000
32	1-Naphthoic acid	2.78	524.6	335.7	264.3	2	2.497	−0.377	−0.0035
36	(Pentabromoethyl)benzene	6	688.4	412.4	412.4	0	1.577	−0.214	−0.0009
37	Pentamethylbenzene	4.27	564.2	353.8	353.8	0	0.121	−0.180	0.0000
38	Perthan	6.56	907.9	536.7	536.7	0	2.015	−0.211	−0.0010
39	Perylene	5.4	730.8	426.2	426.2	0	0.000	−0.128	0.0000
40	Phenol	1.19	353.0	250.3	201.3	1	1.233	−0.253	−0.0015
41	1-Phenylpropan-2-one	1.72	485.7	322.1	284.9	1	2.635	−0.318	−0.0044
42	Phenylacetic acid	1.56	458.0	306.0	229.0	2	1.641	−0.360	−0.0018
43	4-Phenylbutyric acid	2.56	566.6	366.7	283.0	2	1.802	−0.360	−0.0016
44	2-Phenylethyl acetate	2.57	577.1	378.8	342.8	2	1.909	−0.287	−0.0018
45	Propiophenone	2.2	488.7	320.8	296.4	1	2.835	−0.296	−0.0050
46	p-Quaterphenyl	6.97	948.5	552.6	552.6	0	0.046	−0.133	0.0000
47	p-Terphenyl	5.23	744.6	450.0	450.0	0	0.000	−0.131	0.0000
48	1,2,4,5-Tetrachlorobenzene	4.33	496.3	327.5	327.5	0	0.000	−0.108	0.0000
49	Toluene	2.29	384.6	265.6	265.6	0	0.264	−0.179	−0.0001
50	m-Toluic acid	2.23	446.3	295.6	227.3	2	4.553	−0.308	−0.0146
51	Tri-o-tolyl phosphate	5.73	992.3	569.9	539.4	4	3.156	−1.058	−0.0021
52	Valeric acid	1.39	404.3	283.9	200.6	2	2.060	−0.368	−0.0034

in Eq. (24) for s) as compared to respective values in the combined thermodynamic/extrathermodynamic models (0.0280 in Eq. (32) and −0.0207 in Eq. (33)). On the other hand, there is a very good agreement of coefficients at ΔA_{WAS} parameters in Eqs. (23) and (24) (0.0196 and −0.0169, respectively) and in QSRR Eqs. (34) and (35) (0.0279 and −0.0231, respectively).

Third set of reversed-phase HPLC data used here to test retention prediction performance of the solvophobic theory-based model in comparison to QSRR model was taken from Guillot et al. [23] and is presented in Table 3, along with the analyte structure descriptors. The log k_w values listed were determined on a Acquity BEH Shield RP18 stationary phase in gradient mode with methanol as organic modifier. We excluded four nitro compounds of total 52 reported by the original authors from our molecular studies.

As the s data were not given, the correlation between log k_w and s could not be tested here. The theoretical values $B' = 10.549\ Å^3/D^2$ and $C' = 0.0816\ Å^{-2}$ correspond to the values in an equation of the form of Eq. (7): 78.7 and 0.0182, respectively, and analogously to 45.67 and 0.0167, respectively, in the equations analogous to Eq. (8). Thus, agreement of the theoretical B' and C' with those estimated theoretically is again limited (to the order of magnitude).

A combined thermodynamic/extrothermodynamic model, with the correction of analytes hydrophobic contact area for hydrogen-bonding atoms, resulted in the below given equation describing log k_w data from Table 3:

$$\log k_w = -1.66(\pm 0.33) + 0.0162(\pm 0.0008)\Delta A_{WAS} - 0.358(\pm 0.089)N_H \quad (36)$$
$$r^2 = 0.922 \quad SD = 0.526 \quad F = 265 \quad p < 0.0001 \quad n = 48$$

The retention modeling potency of Eq. (36) is illustrated in Fig. 9. It is evidently very good.

The QSRR modeling of data from Table 3 provided the following regression equation:

$$\log k_w = -2.56(\pm 0.44) - 0.0194(\pm 0.0187)\mu^2 + 3.66(0.62)\delta_{min} + 0.0194(\pm 0.0011)\Delta A_{WAS} \quad (37)$$
$$r^2 = 0.898 \quad SD = 0.607 \quad F = 128 \quad p < 0.0001 \quad n = 48$$

The retention predicting capacity of Eq. (37) is illustrated in Fig. 10.

A better fitting of the data from Table 3 to the solvophobic theory based model is also reflected by a smaller difference of the

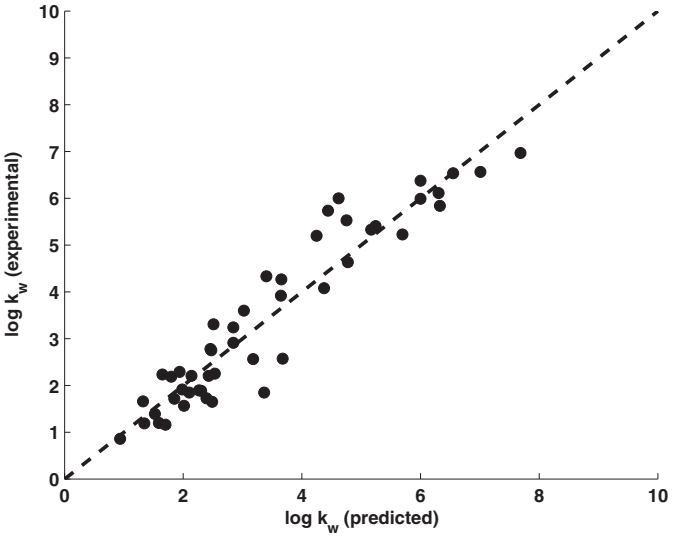

Fig. 10. The parameter log k_w, experimental from Table 3 vs. predicted by Eq. (37).

coefficients at the contact area parameters in the thermodynamic-based Eq. (23) (0.0196) as compared to the respective value in the combined thermodynamic/extrathermodynamic model Eq. (36) (0.0162). There is also an excellent agreement of the theoretical/thermodynamic coefficient in Eq. (23) (0.0196) with that at ΔA_{WAS} in Eq. (37) (0.0194).

One more set of retention data used to test the validity of the proposed approaches: a thermodynamic and a statistical one, to the modeling of reversed phase HPLC behavior of structurally defined analytes was taken from one of our previous publications [24]. The analytes listed in Table 4 have originally been chosen for the sake of comparison of separation properties of individual HPLC columns. Therefore, that set showed a highly regular retention behavior as regards the dependence of log k_w on organic modifier content in the mobile phase. The more so that the column employed was a modern highly reproducible Alltima C18.

The relationship between log k_w and s is illustrated in Fig. 11. There is a nearly perfect agreement of the slope of that relationship (−1.29) with analogous parameters for data presented in Fig. 1 (−1.31) and Fig. 6 (−1.30).

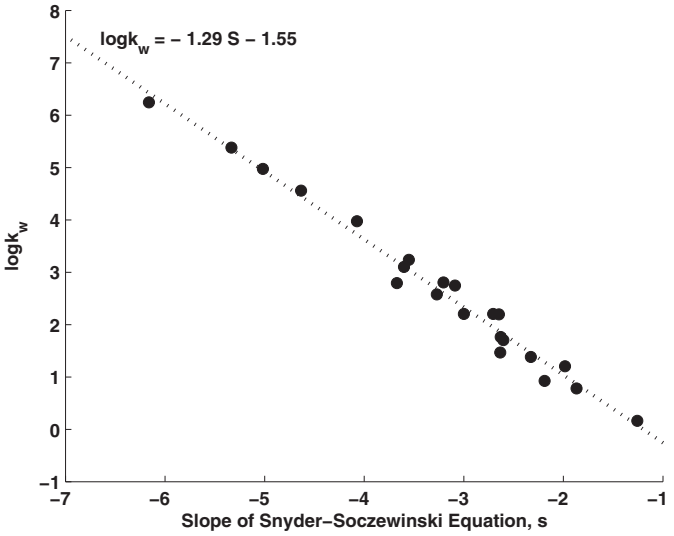

Fig. 11. The relationship between intercepts, log k_w, and slopes, s, of the relationship between logarithms of retention coefficients, k, and percent of methanol, φ, in methanol–water eluent for 22 test analytes chromatographed [24] on a Alltima C18 stationary phase.

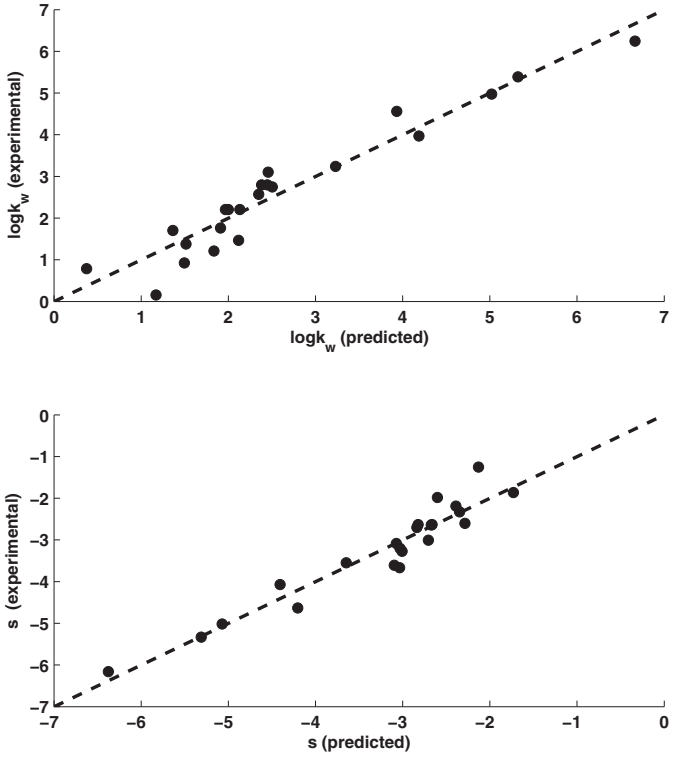

Fig. 12. The parameter log k_w and s, experimental from Table 3 vs. predicted by Eqs. (38) and (39), respectively.

The agreement of the theoretical values of B' and C' from the Horváth equation (2): $B' = 10.549\,\text{Å}^3/\text{D}^2$ and $C' = 0.0816\,\text{Å}^{-2}$, with the corresponding coefficients of analogous Eqs. (7) and (8) was limited, as in the case of previously considered data. Specifically, for example, the coefficients at ΔA_{WAS} in equation of the kind of Eq. (7) was 0.0159 (as compared to $C' = 0.0816$) and that of B' was 219 (as compared to 10.549).

However, very good combined thermodynamic/extrathermodynamic models were obtained with the correction of analyte hydrophobic contact area for hydrogen-bonding atoms for the data listed in Table 4:

$$\log k_w = -2.71(\pm0.53) + 0.0196(\pm0.0017)\Delta A_{WAS} - 0.690(\pm0.076)N_H$$
$$r^2 = 0.922 \quad SD = 0.461 \quad F = 112 \quad p < 0.0001 \quad n = 22 \tag{38}$$

$$s = 1.07(\pm0.42) - 0.0156(\pm0.0013)\Delta A_{WAS} + 0.489(\pm0.060)N_H$$
$$r^2 = 0.916 \quad SD = 0.367 \quad F = 103 \quad p < 0.0001 \quad n = 22 \tag{39}$$

The good retention modeling efficiency of Eqs. (38) and (39) is illustrated in Fig. 12.

The QSRR modeling of data in Table 4 produced the following equations:

$$\log k_w = -1.29(\pm0.59) - 0.136(\pm0.027)\mu^2 + 3.95(1.51)\delta_{\min} + 0.0176(\pm0.00174)\Delta A_{WAS}$$
$$r^2 = 0.928 \quad SD = 0.454 \quad F = 77.8 \quad p < 0.0001 \quad n = 22 \tag{40}$$

$$s = 0.131(\pm0.540) + 0.0925(\pm0.0246)\mu^2 - 2.67(1.39)\delta_{\min} - 0.0142(\pm0.0016)\Delta A_{WAS}$$
$$r^2 = 0.896 \quad SD = 0.418 \quad F = 51.8 \quad p < 0.0001 \quad n = 22 \tag{41}$$

The retention prediction performance of Eqs. (40) and (41) is illustrated in Fig. 13.

There is an excellent agreement of the coefficients at the contact area parameters in the thermodynamics-based equation

Table 4
Intercepts, log k_w, and slopes, s, of the linear relationship between logarithms of retention coefficients determined on an octadecylsilica column (Alltima C18) and percent of methanol in aqueous eluent for a series of test analytes [24], along with their structural parameters from molecular modeling: ΔA_{WAS} – water accessible surface area; ΔA_{HSA}, hydrophobic surface area; N_H – number of hydrogen-bonding atoms; μ_A – total dipole moment; δ_{min} – electron excess charge of the most negatively charged atom; V_A – molecular volume of analyte, β is analyte parameter calculated from Eq. (6).

No.	Name	log k_w	s	V_A	ΔA_{WAS}	ΔA_{HSA}	N_H	μ	δ_{min}	β
1	*n*-Hexylbenzene	5.38	−5.33	646.6	409.2	409.2	0	0.349	−0.211	−0.0059
2	1,3,5-Triisopropylbenzene	6.25	−6.16	777.1	477.6	477.5	0	0.00419	−0.206	−0.0010
3	3-Trifluoromethylphenol	2.79	−3.67	432.4	294.6	198.5	1	2.09	−0.245	−0.0021
4	3,5-Dichlorophenol	3.10	−3.60	440.1	298.4	273.9	1	1.41	−0.243	−0.0012
5	4-Cyanophenol	1.46	−2.63	409.9	281.1	281.1	1	3.31	−0.244	0.0000
6	4-Iodophenol	2.58	−3.27	434.3	292.8	243.6	1	1.58	−0.302	0.0000
7	Anisole	2.21	−2.71	407.9	281.8	269.5	1	1.25	−0.212	−0.0017
8	Benzamide	0.92	−2.19	418.2	284.6	194.4	1	3.58	−0.433	−0.0018
9	Benzene	2.20	−2.65	331.8	240.0	240.0	0	0.000	−0.130	−0.0060
10	Chlorobenzene	2.80	−3.20	375.3	262.8	262.8	0	1.31	−0.129	0.0000
11	Cyclohexanone	1.21	−1.98	384.0	266.8	228.1	1	2.97	−0.294	0.0000
12	Dibenzothiophene	3.97	−4.07	556.0	351.2	351.2	0	0.522	−0.271	−0.0023
13	Phenol	1.38	−2.33	353.1	250.3	201.3	1	1.23	−0.253	−0.0023
14	Hexachlorobutadiene	4.56	−4.63	516.8	338.4	338.4	0	0.000	−0.073	−0.0044
15	Indazole	1.70	−2.60	405.0	277.9	241.3	1	1.55	−0.203	−0.0018
16	Caffeine	0.78	−1.87	569.4	368.0	305.2	1	3.71	−0.362	−0.0004
17	1-Methyl-2-pyrrolidinone	0.16	−1.26	381.5	268.0	216.3	1	3.59	−0.353	−0.0001
18	Naphthalene	3.24	−3.55	459.0	302.6	302.6	0	0	−0.128	0.0000
19	4-Chlorophenol	2.20	−3.00	396.2	273.5	273.5	1	1.48	−0.248	−0.0003
20	Toluene	2.75	−3.09	384.7	265.6	265.6	0	0.264	−0.179	0.0000
21	Benzonitrile	1.76	−2.63	389.7	270.7	270.7	0	3.336	−0.135	−0.0012
22	Benzoic acid	1.83	−2.83	408.7	279.5	200.9	1	2.418	−0.365	−0.0094
23	Isopropylbenzene	4.97	−5.02	629.0	393.9	393.9	0	0.2312	−0.206	−0.0190

(0.0196 in Eq. (23) describing log k_w and −0.0169 in Eq. (24) for s) as compared to respective values in the combined thermodynamic/extrathermodynamic models (0.0196 in Eq. (38) and −0.0156 in Eq. (39)). There is also a very good agreement of coefficients at ΔA_{WAS} parameters in Eqs. (23) and (24) (0.0196 and −0.0169, respectively), and in the QSRR Eqs. (40) and (41) (0.0176 and −0.0142, respectively).

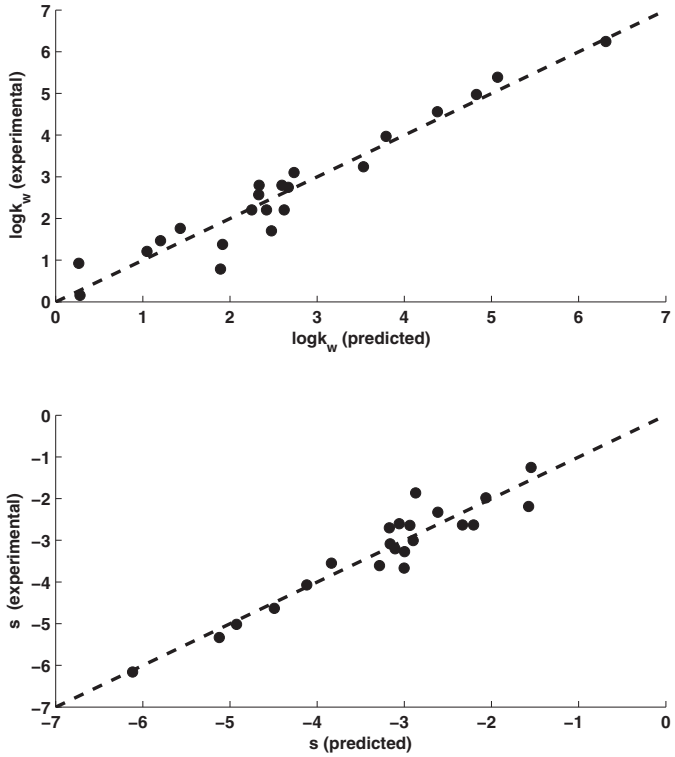

Fig. 13. The parameter log k_w and s, experimental from Table 3 vs. predicted by QSRR Eqs. (40) and (41), respectively.

3. Conclusions

Chemical reactivity, in a sense of breaking of the existing or forming the new chemical bonds, seems to emerge as an innate feature of a molecule which has been coded in its chemical formula. However, compound's property strongly depends on the environment in which it is actually placed. Unlike chemical reactions, the interactions of molecules placed in the environment cause neither the breaking of existing bonds nor the formation of new bonds. These interactions with molecules forming the environment determine compound's properties: physicochemical or biological [2].

To predict a given property, and hence, to design requested chemical product, the relationships must be identified between the chemical structure and the desired property. Optimally, these relationships should be described quantitatively. For such studies chromatography appears to be especially suitable because it can yield a large amount of quantitatively comparable, precise and reproducible property data (retention parameters) for larger sets of structurally diversified compounds (analytes). Therefore, studies of structure–retention relationships can be considered a model approach to establish strategy and methods of property predictions.

In this work the retention prediction performance of the comperhensive solvophobic theory of reversed-phase liquid chromatography of Horvath and co-workers [11,16,17,21] has been tested. The standard chromatographic parameters, log k_w, and, s, were modeled with regard to the chemical structure of four series of analytes when strictly observing the rules of classical thermodynamics, when applying an extrathermodynamically driven correction to the model based on the thermodynamic hermeneutics and when using Quantitative Structure–Retention Relationships (QSRR). It was proved that describing molecular equilibrium between phases by means of approximate thermodynamic equations is of limited precision, however rationally interpretably in clear physicochemical terms. The combined thermodynamicals/extrathermodynamic model with an empirical correction accounting for the number of polar atoms in analyte molecule resulted in an improved agreement between the observed and the model-predicted retention data. However, a purely extrathermodynamic QSRR model, employing analyte descriptors from

calculation chemistry, which were selected *a priori* on the basis of chemical intuition, produced retention predictions only little worse than those based on the rather sophisticated thermodynamic formalistics. The structural descriptors of analytes used in QSRR were: water-accessible molecular surface area, A_{WAS}, square of a total dipole moment, μ^2, and the largest atomic electron charge deficiency in the molecule, $\delta_{\min}$. The molecular parameters presented present in our thermodynamics/extrathermodynamic model were A_{WAS}, as above and the number of hydrogen bond-forming atoms, N_H. Thus, the intuitive QSRR model resembles the thermodynamical model with the pair of descriptors μ^2 and $\delta_{\min}$ corresponding to N_H. The three parameters, μ^2, $\delta_{\min}$ and N_H, account for polarity of the molecule. Whereas the molecular "bulkiness" parameter, A_{WAS}, can be considered as a descriptor of ability of an analyte to participate in the nonspecific dispersive intermolecular interactions which are additive, the polar interactions are mostly constitutive and thus difficult to estimate from structural formulas. Further studies combing thermodynamics and QSRR may help to precise define and determine polar properties of compounds, which are poorly accounted for by the conventional chemical notation.

Acknowledgements

This work was supported by Grants No. NN405G30038 and No. NN405376037 from the Polish Ministry of Science and Higher Education.

References

[1] J.M. Prausnitz, Science 205 (1979) 759.
[2] R. Kaliszan, Chem. Rev. 107 (2007) 3212.
[3] J.E. Leffler, E. Grunwald, Rates and Equilibria of Organic Reactions as Treated by Statistical, Thermodynamic, and Extrathermodynamic Methods, Wiley, New York, 1963.
[4] R. Kaliszan, in: P.R. Brown, R.A. Hartwick (Eds.), High Performance Liquid Chromatography, Wiley, New York, 1989, p. 563.
[5] J.C. Giddings, Unified Separation Science, Wiley, New York, 1991.
[6] R. Kaliszan, Quantitative Structure–Chromatographic Retention Relationships, Wiley, New York, 1987.
[7] M.N. Hasan, P.C. Jurs, Anal. Chem. 62 (1990) 2318.
[8] M. Karelson, V.S. Lobanov, A.R. Katritzky, Chem. Rev. 96 (1996) 1027.
[9] V. Consonni, R. Todeschini, Handbook of Molecular Descriptors, Wiley-VCH, Weinheim, New York, 2000.
[10] K.A. Dill, J. Phys. Chem. 91 (1987) 1980.
[11] C. Horváth, W. Melander, I. Molnár, J. Chromatogr. 125 (1976) 129.
[12] B.L. Karger, L.R. Snyder, C. Eon, J. Chromatogr. 125 (1976) 71.
[13] D.E. Martire, J. Phys. Chem. 87 (1983) 2425.
[14] R.P.W. Scott, J. Chromatogr. 122 (1976) 35.
[15] R. Tijssen, H.A.H. Billiet, P.J. Schoenmakers, J. Chromatogr. 122 (1976) 185.
[16] C. Horvath, W. Melander, I. Molnar, Anal. Chem. 49 (1977) 142.
[17] A. Vailaya, C. Horvath, J. Phys. Chem. B 101 (1997) 5875.
[18] I. Molnár, Chromatographia 62 (2005) s7.
[19] M.A. Al-Haj, R. Kaliszan, A. Nasal, Anal. Chem. 71 (1999) 2976.
[20] M.A. Al-Haj, R. Kaliszan, B. Buszewski, J. Chromatogr. Sci. 39 (2001) 29.
[21] A. Vailaya, C. Horvath, J. Chromatogr. A 829 (1998) 1.
[22] P.J. Shoenmakers, H.A.H. Billiet, L. De Galan, J. Chromatogr. 185 (1979) 179.
[23] A. Guillot, Y. Henchoz, C. Moccand, D. Guillarme, J.L. Veuthey, P.A. Carrupt, S. Martel, Chem. Biodivers. 6 (2009) 1828.
[24] R. Kaliszan, M.A. van Straten, M. Markuszewski, C.A. Cramers, H.A. Claessens, J. Chromatogr. A 855 (1999) 455.