ELSEVIER

Journal of Chromatography A, 962 (2002) 41–55

JOURNAL OF CHROMATOGRAPHY A

www.elsevier.com/locate/chroma

Combination of linear solvent strength model and quantitative structure–retention relationships as a comprehensive procedure of approximate prediction of retention in gradient liquid chromatography

Tomasz Bączek, Roman Kaliszan*

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

Received 12 February 2002; received in revised form 2 May 2002; accepted 3 May 2002

Abstract

Quantitative structure–retention relationships (QSRR) combined with the linear solvent strength (LSS) model are demonstrated to provide approximate predictions of gradient reversed-phase high-performance liquid chromatography (HPLC) retention time for any structurally defined analyte on a once characterized column. The approach requires at first the determination of retention times for a predesigned model series of 15 analytes in two gradient runs. Then by employing the LSS theory a given HPLC system of interest is quantitatively characterized. Structure of the model analytes is next described quantitatively by means of three structural descriptors from standard molecular modeling: total dipole moment, electron excess charge of the most negatively charged atom and water-accessible molecular surface area. With those data the general QSRR equations are derived which describe gradient retention times of the model analytes in the specific column/eluent system. Having now the structural descriptors for any analyte to be chromatographed in such a characterized HPLC system, one employs respective general QSRR equations to calculate its expected gradient retention time at given gradient conditions by means of appropriate LSS equations. Additionally, the chromatographic parameters log k_w and S can be calculated and retention coefficients corresponding to chosen isocratic conditions evaluated. The approach provides retention predictions which can be treated as a first approximation of actual data. Predictions are not yet precise enough for practical separation purposes but can be of use in rational modification of analytical conditions aimed at optimization of separations.

Keywords: Gradient elution; Linear solvent strength; Multiple regression analysis; Retention prediction; Quantitative structure–retention relationships

1. Introduction

However disappointing it may appear, the fact is that valid predictions from chemical formulae of even the most simple properties, like boiling point, chromatographic retention or anaesthetic potency, can only be demonstrated within series of homologues or otherwise closely congeneric compounds by extrapolation or interpolation of the measured properties of several representatives of the series.

*Corresponding author. Tel.: +48-58-349-3260; fax: +48-58-349-3262.

E-mail address: romankal@farmacja.amg.gda.pl (R. Kaliszan).

0021-9673/02/$ – see front matter
PII: S0021-9673(02)00557-5

Most of the claimed quantitative property predictions for structurally diverse compounds have as a rule been derived retrospectively, probably often with some data ''smoothing''. That is because of the lack of adequate means of translation of the conventional atom and bond notation of chemical entities into their physicochemical (biological) properties. However, fundamental works of Pople and co-workers in the early 1980s [1,2] allowed for an estimation of changes in physicochemical properties of compounds accompanying the changes in their chemical structure.

Properties of chemical compounds are believed to be encoded in their structure. However, a specific property is actually revealed by the environment in which the molecules of a given type are placed. Observed physicochemical and/or biological properties are in fact net effects of intermolecular interactions between the molecules of an individual chemical entity and the molecules forming its environment. Therefore, properties other than the standard chemical reactivity (in a synthetic sense) are consequences of supramolecular chemistry [3].

Prediction and quantitative evaluation of properties is the main task of chemometrics [4,5]. To test the predictive potency of chemometric models, chromatography appears a unique and most suitable physicochemical system [6]. That is because in chromatographic systems all the measurement conditions can be kept constant for a large, statistically representative series of structurally diverse analytes, and the chromatographic retention parameters, which can easily be collected in a precise and reproducible manner, can be treated as arising solely from the differences in chemical structure of individual analytes.

Gradient elution is used in HPLC to overcome problems arising at isocratic conditions. In general, isocratic mode often fails when dealing with samples which contain analytes that either cover a wide range of polarities or of molecular mass or both. Therefore, for separation of complex samples gradient elution should be considered as the first step in the HPLC method development [7,8]. It applies to small molecular mass analytes [9,10] as well as to peptides and proteins [11,12].

Evaluation of retention in terms of chemical structure of analytes and of physicochemical properties of both the mobile and the stationary phase is known under the acronym QSRR: quantitative structure–retention relationships. Successful prediction of isocratic retention parameters by means of QSRR equations has been reported by several authors, e.g., Kaliszan [5,13], Carr et al. [14,15], Forgacs and Cserhati [16], Valko et al. [17], Park et al. [18] and others.

QSRR are statistically derived relationships between chromatographic parameters and the quantities (descriptors) characterizing molecular structure of the analytes. QSRR found application to: (i) get insight into the molecular mechanism of separation operating in a given chromatographic system; (ii) identify the most informative structural descriptors of analytes; (iii) evaluate complex physicochemical properties of analytes, e.g., lipophilicity; (iv) evaluate properties of stationary phases; (v) predict relative differences in biological activities within a set of congeneric drugs or other xenobiotics; (vi) predict retention for a new analyte [13,19].

By QSRR analysis it was found [20] that both intercept, log k_w, and slope, S, in the linear semilogarithmic equations relating logarithm of retention coefficient, k, in reversed-phase HPLC to the concentration of organic modifier in the binary aqueous eluent, can be described by regression equations comprising the same structural theoretically assessed descriptors of analytes, however accompanied by different regression coefficients. Statistically significant QSRR equations employing analytes' total energy, E_T, and the maximum excess charge difference between a pair of atoms in the molecule, Δ, were derived. The two structural descriptors could readily be obtained by quantum chemical calculations for any given structural formula. Evaluation of isocratic retention of individual analytes chromatographed at defined concentration of organic modifier in the eluent was demonstrated.

Galushko et al. [21] proposed another QSRR model for calculation of isocratic retention based on structural formulas of analytes. In that model the molecular bulkiness-dependent interactions of analytes with the components of a chromatographic system are assumed to be accounted for by the partial molar volume descriptor, V. That descriptor appears to be a fairly reliable measure of structurally nonspecific inputs to analyte retention. Unfortuna-

tely, less reliable and rather obscure is another structural descriptor proposed by Galushko et al. [21], namely ΔG, which is to reflect differences in the so-called electrostatic intermolecular interactions involving the analytes.

Marked interest suggested empirical solvatochromic parameters of analytes as retention predictors. In 1976, Kamlet and Taft [22,23] introduced the solvatochromic method to evaluate relative polarity of solvents. The original theory was adapted to chromatography by Abraham and co-workers [24,25] under the term linear solvation energy relationships (LSER). The general LSER equation in HPLC is of the form:

$$\log k = \log k_0 + rR_2 + vV_x + s\pi_2^{\mathrm{H}} + a\sum\alpha_2^{\mathrm{H}} + b\sum\beta_2^{\mathrm{H}} \quad (1)$$

where R_2 is the excess molar refraction of the analyte, V_x is its molecular volume produced by the McGowan algorithm, π_2^{H} is dipolarity/polarizability descriptor, $\Sigma\alpha_2^{\mathrm{H}}$ is a measure of the ability of the analyte to donate a hydrogen bond, $\Sigma\beta_2^{\mathrm{H}}$ is an analogous parameter corresponding to the hydrogen bond accepting potency, $\log k_0$ is a constant and r, v, s, a and b are regression coefficients accounting for the net complementary properties of the chromatographic system formed by a given stationary and mobile phase [26].

Solvatochromic parameters of analytes have been determined spectroscopically from the shifts of specific absorption bands of indicator compounds or by gas-chromatographic measurements [27]. Appropriate data are now available for about 4000 compounds. However, they are obviously not available for all the analytes of potential chromatographic interest.

Until recently the QSRR were applied to isocratic retention data only. However, in 2001, Li and Cai [28] employed LSER combined with the linear solvent strength model (LSS) to evaluate retention of a series of test analytes in linear gradient elution. The approach appears to work but its application is naturally limited to those analytes for which the necessary LSER parameters are available.

Certainly, of a wider analytical interest could only be the QSRR approach based on the structural descriptors of analytes which can readily be assessed for any existing or hypothetical compound. Such are the structural descriptors which are provided by molecular modeling. A multitude of individual descriptors produced by calculation chemistry have been reported to produce retention predictions of specific sets of chromatographic data [4,13,16]. Unfortunately, the QSRR models derived retrospectively seldom performed satisfactorily when applied to sets of retention data not used in deriving the models. The most robust general QSRR model resulting from previous extensive studies of this laboratory [29–31] employs the following analyte descriptors: (i) total dipole moment, μ; (ii) electron excess charge of the most negatively charged atom, δ_{Min}; (iii) water-accessible molecular surface area, A_{WAS}. The general QSRR equation has the form:

$$\text{retention parameter} = k_1 + k_2\mu + k_3\delta_{\mathrm{Min}} + k_4A_{\mathrm{WAS}} \quad (2)$$

where k_1–k_4 are regression coefficients.

By equations of the form of Eq. (2), retention parameters for a structurally representative and sufficiently large (for meaningful statistics) model series of analytes chromatographed in a given HPLC system can be described. A model series of 18 test analytes has previously been designed [29,30]. It was found, when completing the present project, that the model series could have been shortened to 15 compounds without meaningful loss of statistical significance of the resulting general QSRR equations. In effect, retention of all the compounds forming the reduced series of model analytes does not depend on the pH of the mobile phase. The values of the molecular descriptors of the analytes used to derive QSRR equations characterizing individual HPLC columns, along with those of the analytes serving to evaluate the prediction potency of the QSRR equations (test analytes), are presented in Table 1.

Experimentally observed changes in retention parameters accompanying the changes in composition of the mobile phase are generally in a good agreement with the linear solvent strength (LSS) model [32]. In that model the logarithm of retention coefficient for a given analyte, $\log k$, is linearly related to the volume fraction of organic modifier in a binary aqueous eluent:

Table 1
Molecular descriptors of the analytes used to derive model QSRR and of the test analytes of several chemical classes used to check retention prediction potency of the proposed approach

No.	Analyte	μ (D)	δ_{Min} (electron)	A_{WAS} (Å^2)
QSRR model analytes				
1	Benzamide	3.583	−0.4333	293.46
2	4-Cyanophenol	3.311	−0.2440	290.90
3	Indazole	1.547	−0.2034	284.44
4	Benzonitrile	3.336	−0.1349	279.14
5	Indole	1.883	−0.2194	292.38
6	2-Naphthol	1.460	−0.2518	323.16
7	Anisole	1.249	−0.2116	288.94
8	Benzene	0.000	−0.1301	245.21
9	1-Naphthylacetonitrile	3.031	−0.1381	364.26
10	Benzyl chloride	1.494	−0.1279	296.17
11	Naphthalene	0.000	−0.1277	311.58
12	Biphenyl	0.000	−0.1315	358.08
13	Phenanthrene	0.020	−0.1279	374.73
14	Pyrene	0.000	−0.1273	392.41
15	2,2′-Dinaphthyl ether	1.463	−0.1606	510.36
Test analytes				
Benzene derivatives				
16	Toluene	0.264	−0.1792	274.37
17	Ethylbenzene	0.320	−0.2105	300.40
18	*n*-Propylbenzene	0.336	−0.2118	329.97
19	*n*-Butylbenzene	0.341	−0.2107	360.86
20	*n*-Amylbenzene	0.349	−0.2107	391.43
21	*n*-Hexylbenzene	0.349	−0.2106	421.46
22	Cumene	0.247	−0.2057	322.15
23	2-Ethyltoluene	0.468	−0.2106	323.11
24	1,2,3-Trimethylbenzene	0.487	−0.1807	319.85
25	1,3,5-Trimethylbenzene	0.000	−0.1786	332.12
26	Anthracene	0.000	−0.1267	379.15
27	1-Methylnaphthalene	0.274	−0.1811	335.17
28	1-Bromonaphthalene	1.414	−0.1540	340.71
29	*o*-Xylene	0.437	−0.1804	297.13
30	*m*-Xylene	0.258	−0.1790	302.97
31	*p*-Xylene	0.000	−0.1780	303.57
Organic acid derivatives				
32	3-Cyanobenzoic acid	3.907	−0.3554	322.24
33	3-Fluorobenzoic acid	2.759	−0.3568	293.89
34	3-Nitrobenzoic acid	5.518	−0.3553	321.90
35	*o*-Toluic acid	2.077	−0.3695	308.71
36	*p*-Toluic acid	2.809	−0.3670	316.80
37	4-Ethylbenzoic acid	2.889	−0.3672	343.94
38	3-Hydroxybenzoic acid	3.496	−0.3584	299.84
39	4-Hydroxybenzoic acid	3.010	−0.3682	300.43
40	Benzoic acid	2.418	−0.3651	287.97
41	1-Naphthylacetic acid	2.028	−0.3742	376.83
42	Acetylsalicylic acid	5.816	−0.3321	353.76
43	Naproxen	2.346	−0.3584	446.68
44	Piroxicam	2.067	−0.9370	490.37
45	Ketoprofen	2.779	−0.3581	481.35
46	Fenbufen	4.028	−0.3584	490.29
47	Diclofenac	1.783	−0.3754	462.21

Table 1. Continued

No.	Analyte	μ (D)	δ_{Min} (electron)	A_{WAS} (Å^2)
Aniline derivatives				
48	2-Chloroaniline	1.676	−0.4010	284.98
49	2-Methoxyaniline	0.802	−0.4035	306.41
50	3,4-Dichloroaniline	3.707	−0.4025	309.72
51	3,5-Dichloroaniline	2.989	−0.4026	312.41
52	3,5-Dimethylaniline	1.274	−0.4137	322.95
53	3-Chloroaniline	2.603	−0.4073	288.71
54	3-Methylaniline	1.469	−0.4131	293.50
55	4-Chloroaniline	3.086	−0.4066	289.22
56	*N*-Ethylaniline	1.867	−0.3605	327.54
57	4-Methoxyaniline	1.966	−0.4157	309.35
Miscellaneous				
58	Coumarin	4.818	−0.2880	310.80
59	Phthalimide	3.348	−0.4025	306.23
60	Phthalonitrile	5.298	−0.1134	308.61
61	1,4-Naphthoquinone	1.332	−0.2698	324.50
62	Phenylacetylene	0.257	−0.1964	290.81
63	Carbazole	1.206	−0.2449	361.17
64	9,10-Anthraquinone	0.003	−0.2863	388.67
65	Xanthene	1.146	−0.1523	376.35
66	Hexachlorobutadiene	0.000	−0.0730	340.60
67	1,3,5-Triisopropylbenzene	0.080	−0.2056	478.59

$$\log k = \log k_w - S\varphi \quad (3)$$

where $\log k_w$ is the value of $\log k$ extrapolated to 100% water as the mobile phase ($\varphi = 0$); S is a constant characteristic for a given analyte and a given chromatographic system; φ is a volume fraction of organic solvent (solvent B) in the mobile phase ($\varphi = \%B/100$).

Based on Eq. (3), $\log k_w$ value is derived by extrapolation of the relationship $\log k$ vs. % organic modifier. The procedure involves several chromatographic measurements and therefore is time-consuming. In such a situation, it would be advantageous to calculate $\log k_w$ and S from retention data obtained in two gradient experiments. Such an opportunity is offered by the LSS model as demonstrated by Snyder and co-workers [32–34]. The calculations can be done by commercially distributed chromatographic softwares. Based on results of two gradient runs carried at different gradient times, t_G, one can calculate not only gradient retention time, t_R, at predesigned gradient conditions but also parameters $\log k_w$ and S, and consequently, retention coefficients corresponding to selected isocratic conditions. That is possible according to the following relationships [32–35]:

$$t_R = (t_0/b)\log(2.3k_0 b + 1) + t_0 + t_D \quad (4)$$

where t_0 is dead time, k_0 is isocratic value of k at the start of gradient elution, t_D is gradient delay time (hold-up or dwell time) and b is gradient steepness parameter which is described by the formula:

$$b = t_0 \Delta\varphi\, S/t_G = V_m \Delta\varphi\, S/t_G\, F \quad (5)$$

where V_m is column dead volume, $\Delta\varphi$ is the change in the mobile phase composition, t_G is gradient time, F is flow-rate, and S is the slope of the $\log k$ vs. φ relationship.

In that work, pairs of general QSRR equations (Eq. (2)) were derived to describe two gradient retention times of model analytes (compounds no. 1–15 in Table 1) for each column and organic modifier of the eluent studied. The QSRR models served to calculate the predicted gradient retention times of test analytes (compounds no. 16–67 in Table 1), as well as the parameters $\log k_w$ and S, characterizing isocratic retention. Calculated and

experimental retention times were next employed to determine the relative errors in the gradient retention coefficient k^*.

2. Experimental

2.1. Equipment

Chromatographic measurements were made with a HPLC apparatus (Waters Corporation, Milford, MA, USA) equipped with a pump, variable wavelength UV–Vis detector, autosampler and thermostat. Data were collected using the Waters Millennium 2.15 software. The following columns were employed: Inertsil ODS-3, 15.0×0.46 cm I.D., particle size 5 μm (GL Sciences, Shinjuku-ku, Tokyo, Japan), XTerra MS, 15.0×0.46 cm I.D., particle size 5 μm (Waters Corporation), Supelcosil LC-18 column, 15.0×0.46 cm I.D., particle size 5 μm (Supelco, Bellefonte, PA, USA), all three packed with octadecyl-bonded silica, Aluspher 100 RP-select B, 12.5×0.40 cm I.D., particle size 5 μm (Merck KGaA, Darmstadt, Germany), packed with polybutadiene-coated alumina, PRP-1, 15.0×0.41 cm I.D., particle size 5 μm (Hamilton Company, Reno, NV, USA), made of cross-linked polystyrene(divinylbenzene) and Discovery HS F5, 15×0.46 cm I.D., particle size 3 μm (Supelco) packed with a pentafluorophenylpropyl-terminated reversed-phase.

The injected sample volume was 20 μl. All the chromatographic measurements were done at 35 °C with eluent flow-rate of 1 ml/min.

Mobile phases contained either methanol or acetonitrile as organic modifiers. Water or buffers of pH 2.5, 3.0, 7.0 and 9.5, necessary for suppression of dissociation of individual analytes, formed the aqueous component of the eluents.

A wide pH range universal buffer (Britton–Robinson buffer) [36] consisted of two parts: A and B. Part A was a mixture of 0.04 *M* acetic acid, 0.04 *M* phosphoric acid and 0.04 *M* boric acid. Part B, 0.2 *M* sodium hydroxide, was added to part A at amounts providing the requested pH. Parts A and B of the buffer were prepared by dissolving individual components purchased from Fluka Chemie AG (Buchs, Switzerland) in water.

Phosphate buffer (0.02 *M*) of pH 3.0 was prepared by dissolving sodium dihydrogen phosphate dihydrate (Merck KGaA) in water and adjusting the pH with hydrochloric acid 37% (Fluka Chemie).

The pH of the buffers was measured at 21 °C before mixing with the organic modifiers. The pH measurements were done with a HI 9017 pH-meter (Hanna Instruments, Bedfordshire, UK).

2.2. Chemicals

Methanol was from P.C. Odczynniki (Gliwice, Poland) and acetonitrile from Merck KGaA. Water was prepared with a Milli-Q Water Purification System (Millipore Corporation, Bedford, MA, USA).

The following analytes listed in Table 1 were selected to derive model QSRR: benzamide, indazole, benzonitrile, 2-naphthol, anisole, 1-naphthylacetonitrile, benzyl chloride, naphthalene, biphenyl, pyrene, 2,2′-dinaphthyl ether, all from Lancaster (Newgate, UK); indole and benzene, both from P.C. Odczynniki; 4-cyanophenol from Aldrich Chemical (Gillingham, UK) and phenanthrene from Koch-Light Laboratories (Koinbrook Bucks, UK).

The following analytes from Table 1 were used to test the retention prediction potency of the approach proposed in this study: ethylbenzene, cumene, *n*-propylbenzene, anthracene, *n*-butylbenzene, *n*-amylbenzene, 2-ethyltoluene, 1,3,5-trimethylbenzene, 1,2,3-trimethylbenzene, 1-methylnaphthalene, *o*-xylene, *m*-xylene, *o*-toluic acid, *p*-toluic acid, 4-ethylbenzoic acid, 3-hydroxybenzoic acid, 4-hydroxybenzoic acid, 1-naphthylbenzoic acid, phenylacetylene, carbazole, xanthene, 9,10-anthraquinone, hexachlorobutadiene, 1,4-naphthoquinone, coumarin, phthalimide, phthalonitrile, aniline, phenol, 2-chloropyridine, all from Lancaster; toluene from Fluka; 2-fluorobenzoic acid, 3-cyanobenzoic acid, 3-fluorobenzoic acid, 3-nitrobenzoic acid, 2-chloroaniline, 3,4-dichloroaniline, 3,5-dichloroaniline, 3,5-dimethylaniline, 3-chloroaniline, 3-methylaniline, 4-chloroaniline, *N*-ethylaniline, 4-methoxyaniline, all obtained from LC Resources (Walnut Creek, CA, USA); *n*-hexylbenzene and 1,3,5-triisopropylbenzene, both from Aldrich Chemicals; *p*-

xylene from Romil Chemicals (Shepshed, UK); benzoic acid from Merck KGaA; 2-methoxyaniline from P.C. Odczynniki; acetylsalicylic acid, ketoprofen, diclofenac, fenbufen, naproxen, piroxicam, all from the drug and reagent collection of the Medical University of Gdańsk (Gdańsk, Poland).

2.3. Determination of retention parameters for QSRR studies

Retention times, t_R, of the model series of analytes were measured on six columns washed with two linear gradients of 5–100% B at two gradient times, t_G. The data from these two gradient experiments were used to derive model QSRR and also served to calculate appropriate log k_w and S values for individual column/eluent systems. The calculations of log k_w and S were performed with the DryLab program (LC Resources, Walnut Creek, CA, USA). The t_R values, along with the calculated log k_w and S parameters for model analytes (analytes no. 1–15 from Table 1) corresponding to individual columns and eluent systems, are collected in Table 2.

2.4. Structural descriptors of analytes

Molecular structure descriptors of the analytes employed in QSRR analysis, i.e. total dipole moment, μ, electron excess charge of the most negatively charged atom, δ_{Min}, and water-accessible van der Waals surface area, A_{WAS}, were calculated by standard molecular modeling and are collected in Table 1. HyperChem program for personal computers with the extension ChemPlus (HyperCube, Waterloo, Canada) was used. The software performed geometry optimization by the molecular mechanics MM+ force field method followed by quantum chemical calculations according to the semi-empirical AM1 method [37,38].

2.5. QSRR analysis

Multiple regression analysis (MRA) equations were derived employing Microsoft Excel software (Microsoft Co., Redmond, WA, USA) and Statistica (StatSoft, Tulsa, OK, USA) run on a personal computer. Regression coefficients (±SD), multiple correlation coefficients, R, standard errors of estimate, s, significance levels of each term and of the whole equations, P, and values of the F-test of significance, F, were calculated and are subsequently reported.

Gradient retention times, t_R, were regressed against the three structural descriptors obtained from molecular modeling: total dipole moment, μ, electron excess charge of the most negatively charged atom, δ_{Min}, and water-accessible molecular surface area, A_{WAS}. The resulting QSRR equations (Eqs. (6)–(21)), characterizing all the HPLC systems studied, are collected in Table 3.

Having QSRR equations for a given HPLC system and employing the three above indicated structural descriptors the t_R, log k_w and S parameters could be calculated for any analyte to be chromatographed in a once characterized HPLC system.

Here, the experimental, $t_{R\,exp}$, and calculated, $t_{R\,calc}$, gradient retention times corresponding to gradient time $t_G = 20$ min, along with the calculated parameters log k_w and S, are given in Table 4 for a subseries of test analytes from Table 1 subjected to HPLC analysis on an Inertsil ODS-3 column with both methanol and acetonitrile as organic modifier of the eluent. In Table 5 the experimental and the calculated data are collected for the whole series of test analytes chromatographed at two gradient times on an XTerra MS column with the methanol-modified eluent.

In both Tables 4 and 5 the relative error in gradient retention coefficient, k^*, is given to quantify the prediction potency of the QSRR/LSS approach here proposed. The calculations of the error were done according to the following equation [32]:

$$\text{relative error in } k^* = (t_0/2.3b)(\delta k/k) \quad (22)$$

where δk is the absolute difference between the experimental and the calculated gradient retention coefficient, k is the experimental gradient retention coefficient and the remaining symbols are as explained earlier.

To illustrate the actual gradient retention prediction capabilities of all QSRR models specified in Table 3 the respective experimental and calculated t_R data are plotted in Figs 1–6.

Table 2
Gradient retention times, $t_{R\ exp}$ (min), at indicated gradient time, t_G (min), along with the log k_w and S values for model QSRR analytes from Table 1 determined on six HPLC columns with methanol (MeOH) or acetonitrile (ACN) as modifiers of the aqueous eluents

No.	Model analytes	Inertsil ODS-3								Supelcosil LC-18							
		MeOH				ACN				MeOH				ACN			
		log k_w	S	$t_{R\ exp}$		log k_w	S	$t_{R\ exp}$		log k_w	S	$t_{R\ exp}$		log k_w	S	$t_{R\ exp}$	
				$t_G=20$	$t_G=60$			$t_G=20$	$t_G=60$			$t_G=10$	$t_G=30$			$t_G=10$	$t_G=30$
1	Benzamide	1.22	3.66	9.63	14.03	1.09	5.28	7.63	10.27	1.10	3.19	6.84	9.45	0.88	4.75	5.37	6.57
2	4-Cyanophenol	1.64	3.39	12.53	21.28	1.50	4.25	10.40	16.67	1.35	2.95	7.93	12.06	1.05	2.43	7.36	10.07
3	Benzonitrile	1.92	3.35	13.95	25.71	1.76	3.61	10.53	17.79	1.82	3.32	9.05	15.62	1.47	4.24	7.14	10.93
4	Indazole	1.95	3.60	14.35	26.35	1.73	5.00	12.80	22.43	1.68	3.04	9.01	15.14	1.52	3.18	8.27	13.22
5	Indole	2.13	3.36	15.63	30.08	2.30	4.49	13.84	26.83	1.81	2.86	9.77	17.08	1.89	3.71	8.77	15.20
6	Benzene	2.25	3.09	16.80	35.28	1.74	2.72	13.73	26.91	2.50	3.81	10.35	19.97	2.13	4.37	8.67	15.44
7	Anisole	2.36	3.33	17.17	34.53	2.12	3.69	14.67	27.97	1.99	2.77	10.62	19.39	1.85	3.28	9.23	16.07
8	2-Naphthol	2.80	4.26	17.36	34.43	2.41	4.81	14.91	26.35	1.84	2.48	10.69	19.04	1.63	2.61	9.53	16.00
9	Benzyl chloride	3.00	4.04	17.41	37.76	2.62	4.21	15.41	31.84	2.97	4.43	10.60	21.24	2.55	4.59	9.41	17.81
10	1-Naphthylacetonitrile	3.19	4.71	18.48	39.92	2.76	4.75	16.03	32.91	2.58	3.50	11.12	21.91	2.35	3.89	9.81	18.44
11	Naphthalene	3.66	4.48	20.24	45.92	2.82	4.01	17.60	37.25	3.03	3.64	12.14	25.13	2.57	3.76	10.61	20.70
12	Biphenyl	4.39	5.12	21.31	49.97	3.09	4.08	18.80	40.96	3.54	4.07	12.65	27.08	3.03	4.15	11.18	22.75
13	Phenanthrene	4.70	5.30	22.03	52.32	3.16	3.93	19.71	43.39	3.67	4.00	13.12	28.43	2.86	3.57	11.81	24.06
14	Pyrene	5.01	5.37	22.99	55.28	3.23	3.69	21.01	46.80	4.02	4.18	13.61	30.08	2.88	3.23	12.65	26.14
15	2,2′-Dinaphthyl ether	6.35	6.65	23.68	58.43	3.94	4.35	22.03	51.09	5.05	5.23	13.79	31.27	3.39	3.69	13.07	27.98
		XTerra MS				Aluspher 100 RP-select B				PRP-1				Discovery HS F5			
		MeOH				MeOH				MeOH				MeOH			
		log k_w	S	$t_{R\ exp}$		log k_w	S	$t_{R\ exp}$		log k_w	S	$t_{R\ exp}$		log k_w	S	$t_{R\ exp}$	
				$t_G=20$	$t_G=40$			$t_G=10$	$t_G=30$			$t_G=10$	$t_G=30$			$t_G=10$	$t_G=30$
1	Benzamide	1.05	4.35	9.98	11.38	−0.19	3.37	2.17	2.20	1.78	4.05	10.47	15.84	0.86	1.08	9.86	12.52
2	4-Cyanophenol	1.34	3.67	12.32	15.27	0.21	3.37	2.92	3.47	2.30	3.70	12.36	21.1	1.55	1.96	11.90	19.95
3	Benzonitrile	1.68	3.53	14.53	19.62	0.37	3.37	5.58	5.70	3.01	3.37	13.65	24.95	1.57	2.16	11.44	19.19
4	Indazole	1.79	3.93	14.63	19.53	0.60	3.37	3.92	3.93	2.76	3.72	15.14	29	1.54	1.94	11.89	19.89
5	Indole	1.86	3.40	16.00	22.07	0.86	3.37	7.6	8.16	3.44	3.73	15.53	30.59	1.76	2.02	12.79	22.67
6	Benzene	1.88	2.84	17.57	25.87	0.86	3.37	9.76	13.71	3.33	3.35	15.44	30.6	2.10	2.43	13.10	24.65
7	Anisole	2.06	3.28	17.63	25.14	0.89	3.37	7.85	8.48	3.46	3.36	16.74	33.31	1.72	1.97	12.75	22.40
8	2-Naphthol	2.53	4.31	17.77	24.91	1.47	3.52	7.73	8.12	3.64	4.01	16.21	32.13	1.45	1.61	12.46	20.38
9	Benzyl chloride	2.55	3.72	18.33	27.68	1.38	2.48	10.17	15.34	3.71	3.48	17.77	35.38	2.26	2.57	13.31	25.58
10	1-Naphthylacetonitrile	2.95	4.81	19.28	28.83	1.80	4.13	10.62	14.91	3.79	3.48	17.16	34.59	1.98	2.22	13.18	24.36
11	Naphthalene	3.12	4.04	21.32	33.16	2.35	3.82	12.05	20.6	4.06	3.48	20.31	38.22	2.32	2.38	14.25	27.91
12	Biphenyl	3.64	4.43	22.60	35.99	3.05	4.46	12.89	23.67	6.23	5.59	22.48	40.85	2.58	2.65	14.36	28.85
13	Phenanthrene	3.87	4.52	23.33	37.51	3.46	4.67	13.53	25.72	6.84	6.10	24.46	42.96	2.77	2.65	15.41	31.21
14	Pyrene	4.09	4.54	24.33	39.51	3.06	3.42	14.87	28.57	7.31	6.58	24.44	43.05	2.86	2.56	16.73	33.04
15	2,2′-Dinaphthyl ether	5.29	5.73	25.12	41.68	4.65	5.34	14.95	30.4	6.83	6.08	24.78	43.28	2.94	2.56	17.62	34.06

3. Results and discussion

3.1. Retention properties of HPLC columns as revealed by QSRR

Details of QSRR methodology of evaluation of retention properties of chromatographic columns in quantitative terms were reported elsewhere [13,19].

In the case of an Inertsil ODS-3 column the QSRR were derived for retention parameters determined with both the methanol-containing and the acetonitrile-containing mobile phases. Neat water was used as the weaker solvent. Having gradient retention time

Table 3
Coefficients k_1–k_4 (±SD) with their significance levels, P (underneath in parentheses), and statistical parameters, R, s, F and P, of regression equations of the form: $t_R = k_1 + k_2\mu + k_3\delta_{Min} + k_4 A_{WAS}$, for the series of model analytes designed to derive general QSRR equations characterizing individual stationary/mobile phase HPLC systems

Gradient retention time (min)	k_1	k_2	k_3	k_4	R	s	F	P	Eq. no.
Inertsil ODS-3 (methanol-containing mobile phase)									
t_R	12.0137 (±1.2517)	−1.3869 (±0.1852)	16.3590 (±2.9991)	0.0326 (±0.0032)	0.9861	0.7515	129	7E−11	(6)
(t_G=20 min)		(P=1E−05)	(P=0.0002)	(P=6E−10)					
t_R	13.8983 (±3.9875)	−4.3315 (±0.5899)	45.9778 (±0.5899)	0.1176 (±0.0101)	0.9866	2.3941	134	6E−09	(7)
(t_G=60 min)		(P=1E−05)	(P=0.0005)	(P=2E−07)					
Inertsil ODS-3 (acetonitrile-containing mobile phase)									
t_R	8.6883 (±1.7156)	−1.1501 (±0.2538)	18.0807 (±4.1106)	0.0356 (±0.0044)	0.9746	1.0301	69	20E−07	(8)
(t_G=20 min)		(P=0.0004)	(P=0.0009)	(P=5E−6)					
t_R	6.7328 (±4.6899)	−3.1146 (±0.6938)	45.2122 (±11.2369)	0.1128 (±0.0119)	0.9767	2.8158	76	1E−07	(9)
(t_G=60 min)		(P=0.0009)	(P=0.002)	(P=1E−06)					
Supelcosil LC-18 (methanol-containing mobile phase)									
t_R	7.9076 (±0.6208)	−0.7723 (±0.0918)	7.5117 (±1.4875)	0.0165 (±0.0016)	0.9870	0.3727	138	5E−09	(10)
(t_G=10 min)		(P=4E−06)	(P=0.0004)	(P=5E−07)					
t_R	8.6244 (±1.9809)	−2.2422 (±0.2930)	21.5172 (±4.7461)	0.0598 (±0.005)	0.9869	1.1893	137	5E−09	(11)
(t_G=30 min)		(P=1E−05)	(P=0.0009)	(P=1E−07)					
Supelcosil LC-18 (acetonitrile-containing mobile phase)									
t_R	6.2794 (±0.8801)	−0.5901 (±0.1302)	9.6941 (±2.1086)	0.0181 (±0.0022)	0.9750	0.5284	71	2E−07	(12)
(t_G=10 min)		(P=−0.0009)	(P=0.0008)	(P=6E−06)					
t_R	5.7206 (±2.2609)	−1.7119 (±0.3345)	24.0791 (±5.4171)	0.0572 (±0.0058)	0.9802	1.3575	90	5E−08	(13)
(t_G=30 min)		(P=0.0003)	(P=0.001)	(P=8E−07)					
XTerra MS									
t_R	11.5023 (±1.3586)	−1.5589 (±0.2010)	16.7260 (±3.2551)	0.0374 (±0.0035)	0.9867	0.8157	135	6E−09	(14)
(t_G=20 min)		(P=9E−06)	(P=0.0003)	(P=3E−7)					
t_R	10.6460 (±2.7412)	−3.0372 (±0.4055)	30.7067 (±6.5680)	0.0818 (±0.0070)	0.9867	1.6458	135	6E−09	(15)
(t_G=40 min)		(P=1E−05)	(P=0.0007)	(P=1E−07)					
Aluspher 100 RP-select B									
t_R	1.3120 (±2.1932)	−1.5745 (±0.3244)	10.5352(±5.2548)	0.0370 (±0.0056)	0.9603	1.3168	43	2E−06	(16)
(t_G=10 min)		(P=0.0005)	(P=0.0702)	(P=4E−05)					
t_R	−15.1438 (±3.8436)	−3.6515 (±0.5317)	–	0.1064 (±0.0108)	0.9680	2.5891	89	6E−08	(17)
(t_G=30 min)		(P=2E−05)		(P=4E−07)					
PRP-1									
t_R	8.9891 (±1.9274)	−1.4034 (±0.2851)	15.6420 (±4.6180)	0.0423 (±0.0049)	0.9741	1.1572	68	2E−07	(18)
(t_G=10 min)		(P=0.0005)	(P=0.0056)	(P=3E−06)					
t_R	25.4337 (±3.8068)	−2.2862 (±0.5632)	44.7120 (±9.1211)	0.0590 (±0.0097)	0.9687	2.2856	56	6E−07	(19)
(t_G=30 min)		(P=0.0019)	(P=0.0055)	(P=8E−05)					
Discovery HS F5									
t_R	7.9635 (±0.8818)	−0.4653 (±0.1305)	6.2602 (±2.1129)	0.0223 (±0.0022)	0.9727	0.5295	64	3E−07	(20)
(t_G=10 min)		(P=0.0044)	(P=0.0129)	(P=8E−07)					
t_R	11.8072 (±2.7204)	−1.3506 (±0.4024)	23.8696 (±6.5181)	0.0583 (±0.0069)	0.9689	1.6333	56	6E−07	(21)
(t_G=30 min)		(P=0.0064)	(P=0.0037)	(P=4E−06)					

data determined for a series of predesigned model analytes (compounds nos. 1–15 in Table 1) the corresponding multiple regression QSRR equations were derived for individual columns and eluent modifier. For Inertsil ODS-3 the coefficients and the statistical parameters of the equations relating the t_R data determined at two t_G values to the three selected molecular descriptors are presented in Table 3.

Table 4
Experimental and calculated gradient retention times, along with the calculated, from t_R, isocratic retention parameters log k_w and S, and the relative error in gradient retention coefficient k^* for a subseries of test analytes from Table 1 chromatographed on Inertsil ODS-3 column with linear gradient 5–100% B at gradient time $t_G = 20$ min

No.	Analyte	Methanol-containing mobile phase					Acetonitrile-containing mobile phase				
		log k_w	S	$t_{R\ exp}$	$t_{R\ calc}$	Relative error in k^*	log k_w	S	$t_{R\ exp}$	$t_{R\ calc}$	Relative error in k^*
1	Coumarin	1.31	3.26	13.95	10.75	0.20	1.42	5.26	11.71	9.00	0.38
2	Phthalimide	1.41	3.62	11.95	10.77	0.12	1.39	5.87	9.28	8.46	0.22
3	Phthalonitrile	1.66	3.29	12.29	12.87	0.07	1.76	4.34	11.87	11.53	0.06
4	1,4-Naphthoquinone	2.71	4.27	15.33	16.33	0.23	2.36	4.63	13.09	13.83	0.19
5	Toluene	2.64	3.68	19.01	17.66	0.14	2.18	3.74	16.61	14.91	0.17
6	Phenylacetylene	2.79	3.87	17.76	17.93	0.03	2.29	3.87	15.57	15.20	0.05
7	Ethylbenzene	2.87	3.99	20.16	17.92	0.23	2.35	3.99	17.97	15.21	0.26
8	Carbazole	3.36	4.73	19.09	18.11	0.17	2.78	4.64	16.61	15.73	0.15
9	Cumene	3.20	4.24	20.85	18.81	0.24	2.56	4.05	19.01	16.15	0.27
10	1-Bromonaphthalene	3.17	4.25	21.71	18.64	0.30	2.62	4.07	19.87	16.41	0.29
11	*n*-Propylbenzene	3.27	4.34	21.17	18.84	0.27	2.61	4.13	19.36	16.22	0.29
12	*n*-Butylbenzene	3.77	4.76	21.97	19.86	0.29	2.90	4.25	20.99	17.33	0.32
13	9,10-Anthraquinone	4.33	5.49	19.97	20.00	0.01	3.20	4.76	17.09	17.34	0.06
14	Xanthene	3.85	4.76	21.91	20.20	0.26	2.99	4.16	19.47	18.01	0.19
15	*n*-Amylbenzene	4.39	5.28	22.59	20.84	0.31	3.20	4.37	21.65	18.41	0.32
16	*n*-Hexylbenzene	5.14	5.90	23.20	21.82	0.32	3.51	4.50	22.61	19.48	0.32
17	Hexachlorobutadiene	3.83	4.25	22.29	21.92	0.06	2.84	3.51	21.79	19.49	0.19
18	Anthracene	4.43	4.89	22.03	22.30	0.07	3.13	3.83	20.03	19.90	0.02
19	1,3,5-Triisopropylbenzene	7.42	7.63	23.65	24.14	0.36	4.13	4.63	23.33	21.92	0.22
					Mean:	0.19				Mean:	0.21

In the case of methanol-modified mobile phases the description of t_R by the set of applied structural parameters is excellent (Eqs. (6) and (7)). All the coefficients at the three parameters are statistically significant ($P \leq 0.0002$ at $t_G = 20$ min and $P \leq 0.0005$ at $t_G = 60$ min) as is the whole equation ($P = 7E-11$ at $t_G = 20$ min and $P = 6E-09$ at $t_G = 60$ min). Multiple correlation coefficients, R, standard errors of estimate, s, and the values of the F-test of significance, F, are all very good.

The chromatographic system formed by Inertsil ODS-3 column and acetonitrile-containing mobile phases has been similarly characterized. All the coefficients at the molecular descriptors in Eqs. (8) and (9) are significant and the parameters R, s, F and P prove the high statistical quality of the equations.

Similar observations regarding the description of the t_R parameters refer also to the Supelcosil LC-18 column (Eqs. (10)–(13) in Table 3).

Characterization of retention properties of the other columns studied (XTerra MS, Aluspher 100 RP-select B, PRP-1 and Discovery HS F5) has been carried out analogously as for Inertsil ODS-3. Successful QSRR analysis was completed only for the retention data obtained with mobile phases containing methanol as the stronger solvent and universal buffer of pH 7.0 as the weaker solvent.

The multiple regression equations describing t_R data determined on the XTerra MS column are Eqs. (14) and (15) in Table 3. All the coefficients at the molecular descriptors are statistically highly significant ($P \leq 0.0003$). High correlation coefficient, R, low value of standard error of estimate, s, and high value of F confirm the significance of the equations.

A specific situation was observed for the Aluspher 100 RP-select B column. Here, there was no problem with the equation describing t_R determined at $t_G = 10$ min (Eq. (16)). However, as regards t_R determined at $t_G = 30$ min only μ and A_{WAS} appeared statistically significant retention descriptors. Nonetheless, the resulting QSRR equation (Eq. (17)) is of a good statistical quality ($R = 0.9680$, $s = 2.5891$, $F = 89$, $P = 6E-08$).

In the case of the PRP-1 column, the description

Table 5
Experimental and calculated gradient retention times, along with the calculated, from t_R, isocratic retention parameters log k_w and S, and the relative error in gradient retention coefficient, k^*, for the complete set of test analytes from Table 1 chromatographed on the XTerra MS column with linear gradient of methanol 5–100% at the indicated gradient time

No.	Analyte	log k_w	S	$t_G = 20$ min			$t_G = 40$ min		
				$t_{R\,exp}$	$t_{R\,calc}$	Relative error in k^*	$t_{R\,exp}$	$t_{R\,calc}$	Relative error in k^*
1	Toluene	2.26	3.51	18.48	18.35	0.02	30.10	26.78	0.08
2	Ethylbenzene	2.48	3.80	21.48	18.72	0.24	33.48	27.78	0.12
3	1-Bromonaphthalene	2.77	4.08	23.32	19.47	0.30	37.13	29.49	0.15
4	Cumene	2.76	3.97	22.67	19.73	0.26	35.63	29.93	0.13
5	*n*-Propylbenzene	2.83	4.08	22.72	19.78	0.27	36.25	30.11	0.14
6	Anthracene	3.67	4.22	23.70	23.56	0.02	38.03	37.77	0.01
7	*n*-Hexylbenzene	4.32	5.17	25.28	23.20	0.32	37.37	37.60	0.02
8	*n*-Butylbenzene	3.27	4.41	23.37	20.94	0.28	38.47	32.66	0.15
9	*n*-Amylbenzene	3.77	4.78	24.47	22.07	0.31	37.38	35.14	0.10
10	2-Ethyltoluene	2.73	4.05	22.73	19.33	0.28	35.80	29.19	0.14
11	1,3,5-Trimethylbenzene	2.98	3.95	23.03	20.94	0.21	36.75	32.33	0.11
12	1,2,3-Trimethylbenzene	2.71	3.91	22.83	19.68	0.26	35.83	29.78	0.13
13	1-Methylnaphthalene	2.95	4.03	22.55	20.58	0.21	35.87	31.67	0.11
14	*o*-Xylene	2.46	3.69	21.60	18.92	0.22	33.33	28.08	0.11
15	*m*-Xylene	2.57	3.73	21.87	19.44	0.21	33.92	29.15	0.11
16	*p*-Xylene	2.64	3.71	21.93	19.88	0.19	33.93	30.01	0.09
17	3-Cyanobenzoic acid	1.37	4.89	13.22	11.52	0.27	17.07	14.23	0.12
18	3-Fluorobenzoic acid	1.41	4.27	15.78	12.23	0.31	21.82	15.35	0.15
19	3-Nitrobenzoic acid	0.64	1.66	14.58	9.00	0.10	19.45	9.31	0.05
20	*o*-Toluic acid	1.80	4.65	16.58	13.63	0.33	23.52	18.24	0.16
21	*p*-Toluic acid	1.67	4.85	16.95	12.83	0.39	24.43	16.76	0.19
22	4-Ethylbenzoic acid	2.03	5.36	18.68	13.72	0.46	28.08	18.73	0.23
23	3-Hydroxybenzoic acid	1.22	4.23	11.33	11.27	0.01	13.58	13.55	0.00
24	4-Hydroxybenzoic acid	1.37	4.41	9.90	11.89	0.66	11.15	14.77	0.29
25	Benzoic acid	1.44	4.21	14.67	12.40	0.25	19.78	15.65	0.12
26	1-Naphthylacetic acid	2.86	5.80	17.92	16.18	0.35	26.78	23.82	0.16
27	Acetylsalicylic acid	1.12	5.49	13.32	10.11	0.42	18.38	11.72	0.21
28	Naproxen	4.41	7.37	19.37	18.56	0.32	30.10	29.05	0.12
29	Piroxicam	2.81	7.22	16.97	10.95	0.62	24.97	15.71	0.31
30	Ketoprofen	5.53	8.87	18.98	19.18	0.18	29.30	30.58	0.36
31	Fenbufen	5.41	9.91	19.65	17.57	0.80	30.58	27.51	0.36
32	Diclofenac	5.17	7.92	21.42	19.73	0.55	34.12	31.51	0.24
33	2-Chloroaniline	1.55	4.33	15.73	12.84	0.29	21.52	16.55	0.14
34	2-Methoxyaniline	2.11	4.68	15.50	14.96	0.10	21.55	20.89	0.03
35	3,4-Dichloroaniline	1.12	4.49	18.02	10.57	0.39	26.77	12.36	0.20
36	3,5-Dichloroaniline	1.44	4.93	19.27	11.79	0.43	29.10	14.76	0.22
37	3,5-Dimethylaniline	2.15	5.00	16.52	14.68	0.29	23.55	20.49	0.13
38	3-Chloroaniline	1.26	4.27	15.52	11.43	0.33	21.32	13.85	0.16
39	3-Methylaniline	1.70	4.56	14.08	13.28	0.14	18.52	17.51	0.05
40	4-Chloroaniline	1.08	3.95	15.33	10.71	0.31	20.97	12.45	0.15
41	*N*-Ethylaniline	2.15	4.90	17.45	14.81	0.34	25.07	20.70	0.15
42	4-Methoxyaniline	1.74	4.92	10.85	13.05	1.05	18.23	17.22	0.05
					Mean:	0.32		Mean:	0.14

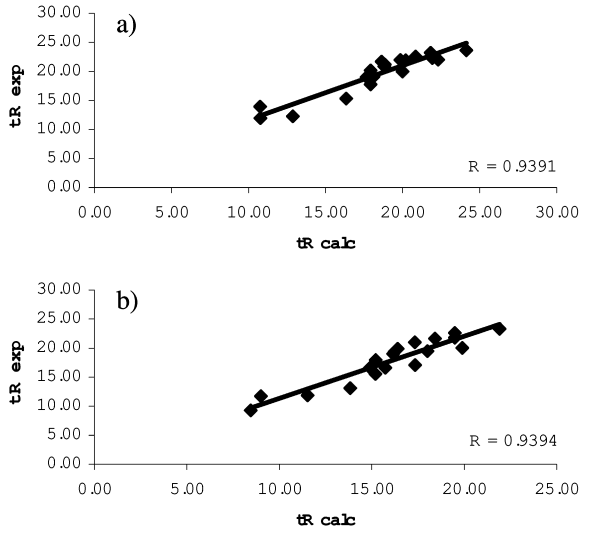

Fig. 1. Correlations between the calculated and the experimental gradient retention times for a subset of test analytes from Table 4 on Inertsil ODS-3 column: (a) methanol-containing mobile phase; (b) acetonitrile-containing mobile phase. Gradient time $t_G = 20$ min.

of both the t_R values is satisfactory with all the descriptors statistically significant in Eqs. (18) and (19).

Equations (20) and (21) describe t_R parameters

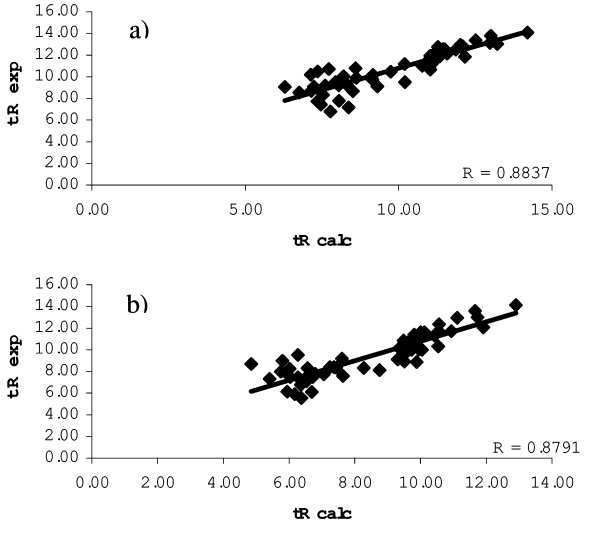

Fig. 2. Correlations between the calculated and the experimental gradient retention times for a set of test analytes from Table 5 on Supelcosil LC-18 column: (a) methanol-containing mobile phase; (b) acetonitrile-containing mobile phase. Gradient time $t_G = 10$ min.

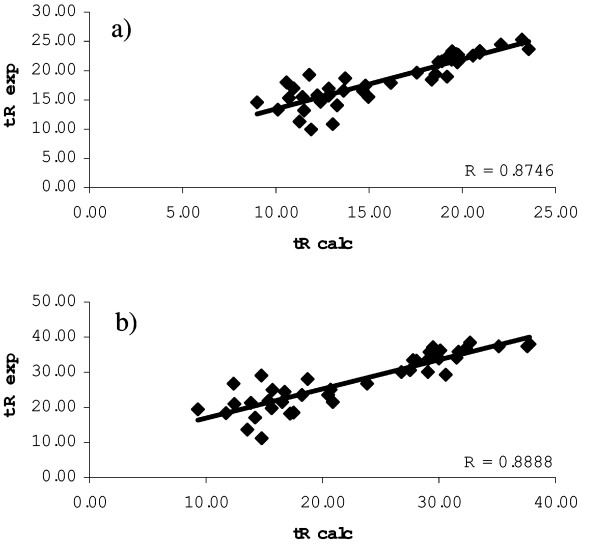

Fig. 3. Correlations between the calculated and the experimental gradient retention times obtained on XTerra MS column for a set of test analytes from Table 5: (a) $t_G = 20$ min; (b) $t_G = 40$ min.

determined on the Discovery HS F5 column. Quality of description of gradient retention times is closely similar to that obtained for PRP-1.

3.2. Testing of retention prediction potency of the derived QSRR

In Table 4 the retention parameters (calculated and experimental gradient retention times along with log k_w and S values) are collected for a structurally diverse set of analytes from Table 1 which had not been used to derive QSRR describing $t_{R\,exp}$ on Inertsil ODS-3 with neither methanol–water nor acetonitrile–water as mobile phases. For the methanol-containing mobile phases, $t_{R\,calc}$ was calculated by Eq. (6) from Table 3. For the acetonitrile-containing mobile phases Eq. (8) was used. Next, the log k_w and S parameters were calculated according to the LSS model employing t_R values from two gradient runs.

The differences between the calculated and the experimental gradient retention data for the 19 test analytes studied are reasonable, at least in case of the methanol-involving chromatographic system. For those systems the average difference between the calculated and the experimental retention time is 1.15 min. Goodness of prediction of gradient re-

tention time, t_R, on Inertsil ODS-3 washed with the methanol-containing eluent for a set of test analytes from Table 4 is illustrated in Fig. 1a. Correlation coefficient, R, between the predicted and the observed t_R equals 0.9391.

As can be calculated from Table 4, in the case of acetonitrile-containing mobile phases the differences between the calculated and the experimental gradient retention times are larger than those observed for the methanol-containing eluent with the mean value of 2.14 min. Although the QSRR equations (Eqs. (8) and (9)) used to calculate t_R were of lower statistical quality than Eqs. (6) and (7), the predictions of t_R made with the use of these equations are surprisingly accurate for most of the analytes. The correlation coefficient $R=0.9394$ for the relationship presented in Fig. 1b supports the validity of the general idea of the approach.

Standard deviations in the predicted gradient retention times were converted to relative errors in the gradient retention coefficient k^*. These errors are listed in Table 4 for a series of test analytes chromatographed on Inertsil ODS-3 with linear gradient 5–100% of methanol or acetonitrile at gradient time $t_G=20$ min. It has been assumed that to be practically useful for optimizing resolution, the errors in k^* should be no greater than about 5% [32]. As seen in Table 4 the present approach gives mean relative error in prediction of k^* on Inertsil ODS-3 at the conditions applied in this work of about 20%. Therefore, the retention predictions obtained by the approach developed here may be treated as a first approximation.

Studies of retention data determined on Supelcosil LC-18, XTerra MS, Aluspher 100 RP-select B, PRP-1 and Discovery HS F5 columns with methanol-buffer eluents and with acetonitrile-buffer eluents (Supelcosil LC-18 column) of the pH suppressing analyte dissociation, comprised a larger, chemically diverse series of test analytes formed by groups of simple benzene derivatives, organic acids (including nonsteroidal anti-inflammatory drugs) and aniline derivatives. For the sake of illustration, values of t_R, both experimental and calculated from model QSRR (Eqs. (14) and (15)) for the whole set of test analytes on the XTerra MS column, are given in Table 5, accompanied by the relative errors in gradient retention coefficients k^*. Satisfactory prediction of gradient retention times was obtained in the case of benzene derivatives and organic acids. The predictions in case of aniline derivatives were worse. That can be attributed to the known difficulties in liquid chromatographic separations of amines [7].

As can be seen in Table 5 the mean relative error in k^* in case of the Xterra MS column washed with a linear gradient 5–100% of methanol differs with the gradient time. In the case of $t_G=20$ min the mean error is about 32%, whereas with $t_G=40$ min it decreases significantly attaining the level of 14%.

Correlations between the observed and the calculated gradient retention times for all the HPLC systems studied are illustrated in Figs. 1–6.

Analyzing the data obtained on the Supelcosil LC-18 column (Fig. 2) one will note a general trend for almost all the analytes that the mean differences of the actual retention times and those calculated theoretically are very similar for methanol- and acetonitrile-containing mobile phases.

In the case of the XTerra MS columns the experiments were performed at gradient times, t_G, of 20 min and 40 min whereas with the Aluspher 100 RP-select B t_G was 10 min and 30 min. The agreement between the calculated retention times and the experimental data (Table 5) is confirmed by correlation analysis (Figs. 3 and 4). Scattering of data points in Fig. 4 is broader than in Figs. 1–3. The reason may be some specific (adsorption) inputs

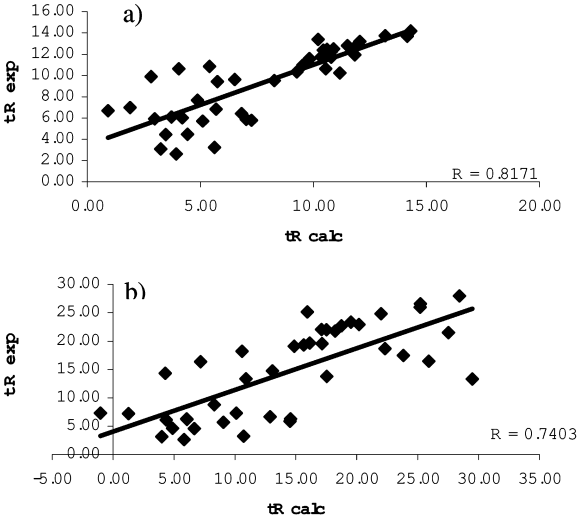

Fig. 4. Correlation between the calculated and the experimental gradient retention times obtained on Aluspher 100 RP-select B column for a set of test analytes from Table 5: (a) $t_G=10$ min; (b) $t_G=30$ min.

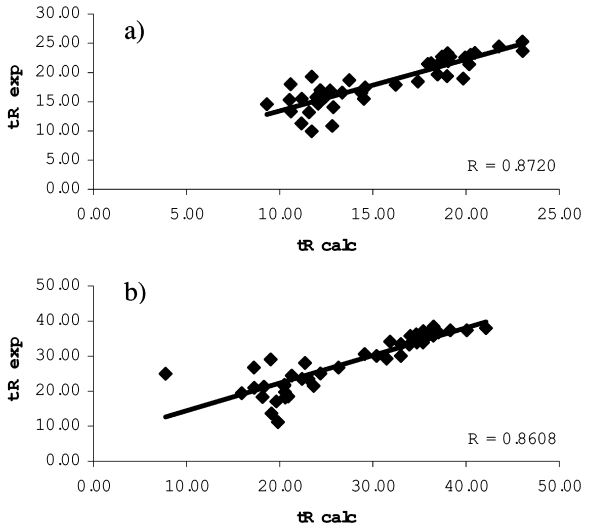

Fig. 5. Correlation between the calculated and the experimental gradient retention times obtained on PRP-1 column for a set of test analytes from Table 5: (a) $t_G = 10$ min; (b) $t_G = 30$ min.

to retention on the alumina-based Aluspher 100 RP-select B column which are not fully accounted for by the QSRR models employed.

Correlation analysis of the data determined on the PRP-1 and Discovery HS F5 columns also confirms general agreement between the calculated and the experimental retention parameters for structurally representative series of test analytes. The specific chemical nature of both the PRP-1 and Discovery HS F5 stationary phases, which certainly differ from the typical reversed-phase materials represented by Inertsil ODS-3, Supelcosil LC-18 and XTerra MS, manifests itself in a worse performance of the QSRR models. That is in agreement with our previous observations [19] on quantitative differentiation of retention properties of various reversed-phase materials by QSRR analysis.

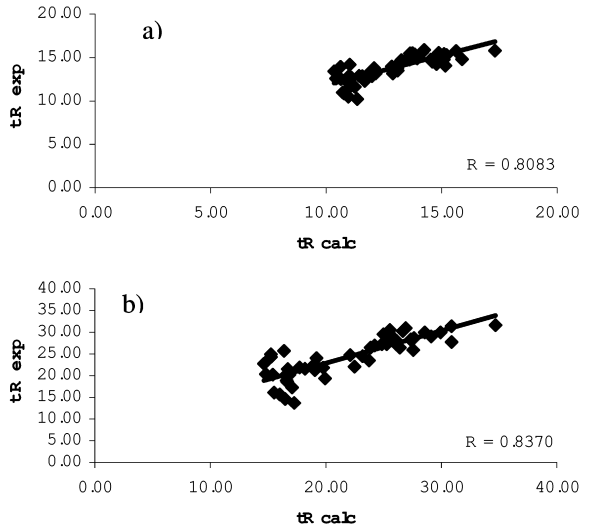

Fig. 6. Correlation between the calculated and the experimental gradient retention times obtained on Discovery HS F5 column for a set of test analytes from Table 5: (a) $t_G = 10$ min; (b) $t_G = 30$ min.

4. Conclusions

Quantitative structure–retention relationships (QSRR) and the linear solvent strength (LSS) model allow for approximate prediction of both gradient and isocratic reversed-phase HPLC retention time of any analyte on a once previously characterized column. Data (otherwise not attainable) which can guide further optimization of the analytical procedure can thus be obtained. Two gradient experiments carried out at different gradient times for an appropriately designed series of 15 model analytes serve to derive model QSRR equations. These equations, once established for a given HPLC system, are next used to evaluate retention parameters, t_R, log k_w and S, for any analyte of a known molecular structure to be chromatographed in a given HPLC system. Consequently, chromatographic conditions can be predetermined a priori which may help to optimize the separation of the analytes of interest.

The approach here described offers a general strategy of the computer-aided automatization of the chromatographic method development process. To that aim the commercially available softwares for molecular modeling (like HyperChem) and for optimization of gradient HPLC separations (like DryLab) can be combined with a statistics program (like Microsoft Excel or Statistica). The only experimental data required for a new HPLC column would be two gradient retention times for 15 model analytes. Having these data determined (or provided by the column manufacturer) one would draw the structure of the analytes to be separated on a personal computer screen and receive detailed instructions on the HPLC conditions to be applied as well as the expected gradient retention times. Also, a

rational choice of the presumably best column could be made for given set of analytes to be separated based on the QSRR equations which had previously been derived.

The approach proposed here allows for only a first approximation of retention for a new analyte. The relative error in k^* of 14–32% is larger than that (ca. 5%) allowing for routine applications to solve practical analytical problems. On the other hand, the predictions obtained are better than simple guessing of the retention from structural formulae which is the only alternative for most potential analytes for which no experimentally derived structural data are available. That can be checked by asking an expert on HPLC to guess the t_R data presented in Table 5 on the basis of the training data collected in Table 2.

Obviously, the limited accuracy of the prediction of gradient retention a priori from the computer-generated properties of analytes is due to errors in both the LSS model and in the property description. An improvement in prediction of retention changes, accompanying the changes in composition of the eluent, might probably be obtained by applying nonlinear models. However, the results of this work support the view that the main limitation of the a priori prediction of retention (as well as other physicochemical and biological properties) of chemical compounds from their structure is the inadequacy of translation of structural formulae into sets of numerical descriptors. There is a lack of such a translation that would reveal the properties encoded into the structure in a reliable manner. Hopefully, further progress in theoretical chemistry and molecular modeling will bring better means of characterization of chemical entities and make the predictions of their properties feasible. QSRR is a unique tool to test the performance of the proposed solutions.

References

[1] J.S. Binkley, J.A. Pople, W.J. Hehre, J. Am. Chem. Soc. 102 (1980) 939.
[2] W.J. Pietro, M.M. Francl, W.J. Hehre, D.J. Defrees, J.A. Pople, J.S. Binkley, J. Am. Chem. Soc. 104 (1982) 5039.
[3] J.M. Lehn, Chemia supramolekularna, PWN, Warszawa, 1998, in Polish.
[4] R. Kaliszan, Quantitative Structure–Chromatographic Retention Relationships, Wiley, New York, 1987.
[5] D.L. Massart, B.G.M. Vandeginste, S.N. Deming, Y. Michotte, L. Kaufman, Chemometrics: A Textbook, Elsevier, Amsterdam, 1988.
[6] R. Kaliszan, Anal. Chem. 64 (1992) 619.
[7] L.R. Snyder, J.J. Kirkland, J.L. Glajch, Practical HPLC Method Development, Wiley, New York, 1997.
[8] J.W. Dolan, L.R. Snyder, R.C. Wolcott, P. Haber, T. Bączek, R. Kaliszan, L.C. Sander, J. Chromatogr. A 857 (1999) 41.
[9] J.W. Dolan, L.R. Snyder, N.M. Djordjevic, D.W. Hill, T.J. Waeghe, J. Chromatogr. A 857 (1999) 1.
[10] K. Hayakawa, M. Hirano, K. Yoshikawa, N. Katsumata, T. Tanaka, J. Chromatogr. A 846 (1999) 73.
[11] T. Yoshida, T. Okada, J. Chromatogr. A 841 (1999) 19.
[12] A.W. Purcell, G.L. Zhao, M.I. Aguilar, M.T.W. Hearn, J. Chromatogr. A 852 (1999) 43.
[13] R. Kaliszan, Structure and Retention in Chromatography. A Chemometric Approach, Harwood Academic, Amsterdam, 1997.
[14] P.T. Jackson, M.R. Schure, T.P. Weber, P.W. Carr, Anal. Chem. 69 (1997) 416.
[15] J. Li, P.W. Carr, Anal. Chim. Acta 334 (1996) 239.
[16] E. Forgacs, T. Cserhati, Molecular Bases of Chromatographic Separations, CRC Press, Boca Raton, 1997.
[17] K. Valko, C. Bevan, D. Reynolds, Anal. Chem. 69 (1997) 2022.
[18] J.H. Park, M.D. Jang, D.S. Kim, P.W. Carr, J. Chromatogr. 513 (1990) 107.
[19] R. Kaliszan, M.A. van Straten, M. Markuszewski, C.A. Cramers, H.A. Claessens, J. Chromatogr. A 855 (1999) 455.
[20] R. Kaliszan, K. Ośmiałowski, S.A. Tomellini, S.-H. Hsu, S.D. Fazio, R.A. Hartwick, J. Chromatogr. 352 (1986) 141.
[21] S.V. Galushko, A.A. Kamenchuk, G.L. Pit, J. Chromatogr. A 660 (1994) 47.
[22] M.J. Kamlet, R.W. Taft, J. Am. Chem. Soc. 98 (1976) 377.
[23] R.W. Taft, M.J. Kamlet, J. Am. Chem. Soc. 98 (1976) 2886.
[24] P.C. Sadek, P.W. Carr, R.M. Doherty, M.J. Kamlet, R.W. Taft, M.H. Abraham, Anal. Chem. 57 (1985) 2971.
[25] M.H. Abraham, Chem. Soc. Rev. 23 (1993) 660.
[26] J.C. McGowan, M.H. Abraham, Chromatographia 23 (1987) 243.
[27] M.H. Abraham, University College London DataBase, 1997.
[28] J. Li, B. Cai, J. Chromatogr. A 905 (2001) 35.
[29] M.A. Al-Haj, R. Kaliszan, A. Nasal, Anal. Chem. 71 (1999) 2976.
[30] M.A. Al-Haj, R. Kaliszan, B. Buszewski, J. Chromatogr. Sci. 39 (2001) 29.
[31] R. Kaliszan, P. Haber, T. Bączek, D. Siluk, Pure Appl. Chem. 73 (2001) 1465.
[32] L.R. Snyder, J.W. Dolan, Adv. Chromatogr. 38 (1998) 115.
[33] M.A. Quarry, R.L. Grob, L.R. Snyder, Anal. Chem. 58 (1986) 907.
[34] L.R. Snyder, J.W. Dolan, J.R. Gant, J. Chromatogr. 165 (1979) 3.
[35] M.A. Quarry, R.L. Grob, L.R. Snyder, J. Chromatogr. 285 (1984) 19.
[36] Poradnik fizykochemiczny (Physicochemical Manual), Wydawnictwo Naukowo-Techniczne, Warszawa, 1974, in Polish.
[37] N.L. Allinger, J. Am. Chem. Soc. 99 (1977) 8127.
[38] M.J.S. Dewar, E.G. Zoebisch, E.F. Healy, J.J.P. Stewart, J. Am. Chem. Soc. 107 (1985) 3902.