Journal of Chromatography A, 656 (1993) 417–435
Elsevier Science Publishers B.V., Amsterdam

CHROM. 25 193

Review

Quantitative structure–retention relationships applied to reversed-phase high-performance liquid chromatography

Roman Kaliszan

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

ABSTRACT

Quantitative structure (reversed-phase)–retention relationships (QSRRs) derived by means of various statistical procedures are reviewed from the viewpoint of identifying retention affecting structural factors and understanding the mechanism of separations. A brief summary of the theoretical background of QSRRs is followed by presentation of reversed-phase high-performance liquid chromatographic (RP-HPLC) separation theories of relevance to the reported QSRRs. Hydrophobicity parameters derived by RP-HPLC are discussed in detail in relation to liquid–liquid partition coefficients with special emphasis on properties of new reversed-phase materials. Reported QSRR equations are critically reviewed, bearing in mind their statistical significance and physical meaning. Information on retention mechanism, as extracted by factorial methods of data analysis, is briefly analysed. It is concluded that QSRRs employing molecular descriptors expected to model fundamental intermolecular interactions and the QSRRs employing solvatochromic parameters are of similar potency for retention prediction and provide rationalization of the RP-HPLC retention mechanism. QSRR studies are demonstrated to be of value in the search for new reliable and precise descriptors of the structures of solutes of relevance to their properties, including properties other than chromatographic, *e.g.*, bioactivity.

CONTENTS

0021-9673/93/$24.00

1. INTRODUCTION

Quantitative structure–retention relationships (QSRRs) are one of the most extensively studied manifestations of linear free-energy relationships (LFERs) [1]. LFERs are extrathermodynamic relationships, *i.e.*, they are not necessarily a consequence of thermodynamics. Extrathermodynamic approaches combine detailed models of processes with certain concepts of thermodynamics. It is well known that the thermodynamic properties of a given system are bulk properties reflecting just the net interactive effects in that system. The magnitude of thermodynamic parameters represents the combination of individual interactions that may take place at the molecular level. Thus, from the chemical point of view classical thermodynamics is inadequate [2–4]. However, it is believed that the demonstration of LFERs suggests the presence of a real connection between some correlated quantities, and that the nature of this connection can subsequently be identified. One asumes that correlations among specific quantities are attributable to some physico-chemical relationships. The statistically derived correlations encourage attempts to identify the relationships behind them.

The following goals of QSRR studies can be identified during their 15-year-old history (Fig. 1) [4,5]: (i) prediction of retention for a new solute; (ii) identification of the most informative (regarding properties) structural descriptors; (iii) elucidation of the molecular mechanism of separation operating in a given chromatographic system; (iv) evaluation of complex physicochemical properties of solutes (other than chromatographic); and (v) estimation of relative biological activities within a set of solute xenobiotic compounds.

In previous publications [1,4–7], individual applications of QSRRs were reviewed. Here the coverage is restricted to reversed-phase HPLC and attention is focused on recent findings regarding items (i)–(iv) above, with special emphasis on the molecular mechanism of retention. Readers interested in item (v) are referred to a recent paper entitled "information potential of chromatographic data for pharmacological classification and drug design" [8].

2. BACKGROUND AND METHODOLOGY OF QSRRs

Chromatographic retention data must be some function of temperature, of chemical structure of the solute, of the stationary phase and of the

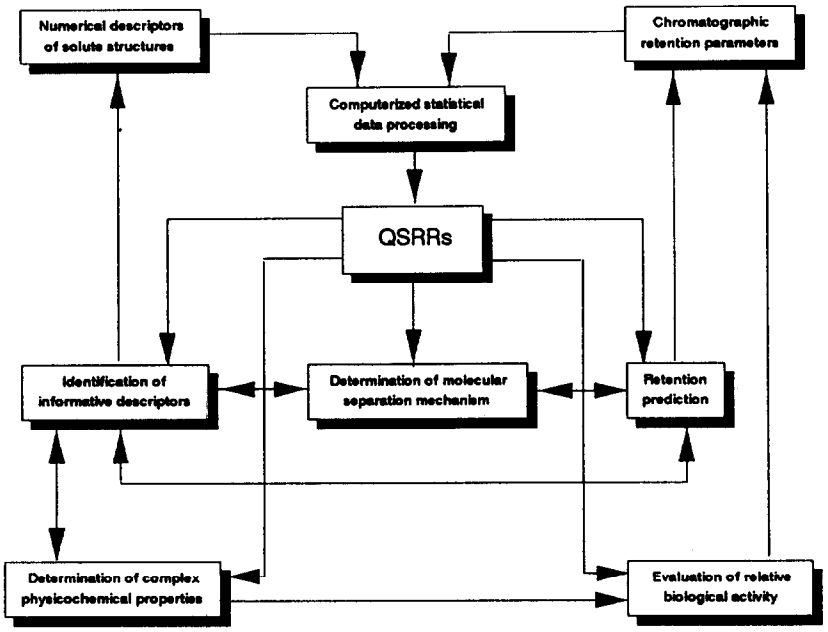

Fig. 1. Methodology and goals of QSRR studies.

mobile phase, all of them mutually interacting. However, there is no general, strict, unequivocally verifiable canonical equation relating retention to the four main chromatographic variables, that is, the temperature, the structure of the solute, the stationary phase and the mobile phase. Even if the stationary and mobile phases applied in a given chromatographic system remain constant, still a precise, quantitative description of the retention of a series of solutes appears problematic, the more so the more diverse are the solutes considered, although the problem is by no means trivial for homologues. The plots of log k' *versus* carbon number of a homologue are usually linear but only for some limited range of aliphatic chain length. In general, relationships between empirical or theoretically calculated molecular structural parameters require statistical evaluation in order to check the significance of the resulting correlations [1].

Basically, the research strategy now applied in QSRR studies was transferred from studies initiated in 1960s on quantitative structure–(biological) activity relationships (QSARs) [9]. It remains similar but, naturally, the emphasis in QSRRs is placed on specifically chromatographic aspects of structure–property relationships.

One needs two kinds of input data to undertake QSRR studies (Fig. 1): a set of quantitatively comparable retention data (dependent variable) for a sufficiently large set of solutes and a set of quantities (independent variables) assumed to reflect structural features of the solutes being studied. Through the use of chemometric computational techniques, retention parameters are characterized in terms of various combinations of solute descriptors or in terms of systematic knowledge extracted (learned) from these descriptors. Good QSRRs can provide otherwise inaccessible information on solutes and on the chromatographic system involved. That information may be of use in other structure–property relationship studies.

To obtain statistically significant and physically meaningful QSRRs, reliable input data must be provided and stringent mathematical analysis must be carried out. The great advantage of QSRRs over other quantitative structure–property relationship studies is that chromatography can readily yield a great amount of relatively precise and reproducible data. In a chromatographic process all conditions may be kept constant. Thus, solute structure becomes the single independent variable in the system.

The first-reported QSRRs were derived by multiple regression analysis of retention data against a set of structural descriptors. Another early approach to QSRRs was based on the assumption of additive substituent effects on retention analogously to the *de novo* non-parameter method of correlation analysis applied by Free and Wilson [10] in medicinal chemistry. Later, factorial methods of data analysis were occasionally employed in QSRR studies. Very recently reports have appeared [11,12] on the application of artificial neural systems (neural networks) in retention prediction.

From the viewpoint of employing QSRRs to gain insight into molecular mechanisms of chromatographic separations, the most valuable and common approach remains that based on multiparameter regression. There are also several reports on the application of factor analysis for this purpose. Most applications of the *de novo* non-parameter method, factorial methods and artificial neural networks are aimed at retention prediction (including so-called expert systems) and optimization of chromatographic separation conditions. Readers interested in these applications are referred to recently published specialistic books [13,14].

One has to be cognizant of the fact that not every QSRR equation provides meaningful and reliable information. Some published equations are statistically invalid [1] and sometimes formally valid correlations are developed for chemically invalid principles. Formal requirements regarding reporting of results of regression [1,15] and factorial [16] analysis may be found elsewhere.

3. RP-HPLC SEPARATION THEORIES AT THE BASIS OF QSRRs

One can attempt to derive QSRRs making no mention of any existing chromatographic theory. A typical strategy is to generate a multitude of solute descriptors that are next regressed against retention data. Observing all the statistical rules, one selects the minimum number of descriptors needed to produce an equation yielding the

calculated retention data in satisfactory agreement with the observed values. The number of descriptors that can be assigned to an individual solute is virtually unlimited. In Table 1 the structural descriptors that are more commonly used in QSRR studies are listed. However, are reported also numerous rare, sometimes *ad hoc* designed solute descriptors. It is often difficult to assign any physical sense to such parameters. It is even more difficult to interpret QSRR equations consisting of terms produced by various transformations and combinations of such descriptors, *e.g.*, their square roots, cubes, reciprocals or products. If QSRRs result from the analysis of tens or hundreds of descriptors, then most likely several equations with similar predictive abilities but consisting of different sets of variables can be derived. From the point of view of prediction of retention, all this does not matter (as long as the QSRR is not fortuitous). However, QSRRs that are not interpretable in physical terms are not very informative regarding the mechanism of retention.

A more promising QSRR strategy is to start from the existing theories of chromatographic separations and to attempt to quantify the abilities of solutes to take part in the postulated intermolecular interactions [6]. These fundamental intermolecular interactions involving solute molecules, molecules forming mobile phase and molecules of stationary phase are as follows: (i) dipole–dipole (Keesom); (ii) dipole–induced dipole (Debye); (iii) instantaneous dipole–induced dipole (London); (iv) hydrogen bonding; (v) electron pair donor–electron pair acceptor; and, possibly, (vi) solvophobic interactions. The potential energy, E, of the first three types of interactions is approximated by

$$E = -W^2\varepsilon^{-1}r^{-6}[2\mu_1^2\mu_2^2/3kT + \alpha_2\mu_1^2 + \alpha_1\mu_2^2 + 3I_1I_2\alpha_1\alpha_2/2(I_1 + I_2)] \quad (1)$$

where W and k are constants, ε is relative electric permittivity of the medium, r is distance between the interacting molecules, T is the absolute temperature and μ, α and I are the dipole moments, polarizabilities and ionization potentials, respectively, of the interacting molecules.

Eqn. 1 substantiates the assumption that, within a set of solutes of similar hydrogen-bonding and charge-transfer properties, chromatographed under identical conditions, the retention parameters can be approximated by a combina-

TABLE 1

STRUCTURAL DESCRIPTORS OF SOLUTES USED IN QSRR ANALYSIS

Bulkiness-related (non-specific) parameters	*Polarity-related (electronic) parameters*
Molecular mass	Dipole moments
Refractivity	Atomic excess charges
Molecular volume	Orbital energies
Total energy	Superdelocalizabilities
Solvent-accessible area	Partially charged surfaces
Geometry-related (shape) parameters	*Molecular graph-derived (topological) parameters*
Moments of inertia	Adjacency matrix indices
Length-to-breadth ratio	Distance matrix indices
Angle strain energy	Information content indices
Physico-chemical parameters	*Indicator variables*
Hydrophobic constants	
Hammett constants	
Solubility parameters	
Boiling points	
Solvatochromic parameters	

tion of polarizabilities, ionization potentials and squares of dipole moments. In pre-QSRR days, attempts were made to select solutes either with similar dipole moments and varying polarizability [17] or with similar polarizability and varying dipole moments [18], and to relate retention to the variable. Those first correlations were moderately successful but clearly illustrated the trends implied in eqn. 1.

Tijssen *et al.* [19] considered three types of interactions: dispersion, orientation and the so-called acid–base interactions. The ability of an individual compound to take part in the respective interactions is reflected by its specific partial solubility parameter. The problem encountered when testing the predictive potency of the approach was to determine precisely the solubility parameters. Similarly, Horváth *et al.*'s [20] solvophobic theory, Martire and Boehm's [21] molecular statistical theory and several other early theoretical approaches to RP-HPLC required a knowledge of a number of physicochemical parameters that were mostly not available for individual solutes. Thus, absolute, strict verification of these theories was extremely difficult. However, individual solute properties affecting retention were identified which, in turn, suggested the choice of the most informative structural descriptors for QSRRs.

More recently, Carr and co-workers [22–24], in studies on the nature of RP-HPLC separations, proposed an approach based on the solvatochromic comparison method and linear solvation energy relationships (LSERs). They considered a general equation to examine the chemical and physical characteristics of a solute that determine retention, with the following form:

$$\log k' = \text{constant} + M(\delta_m^2 - \delta_s^2)V_2/100 + S(\pi_s^* - \pi_m^*)\pi_2^* + A(\beta_s - \beta_m)\alpha_2 + B(\alpha_s - \alpha_m)\beta_2 \quad (2)$$

where the subscript 2 designates a solute property such as molar volume (V_2), polarizability-dipolarity (π_2^*), hydrogen bond acidity (α_2) and hydrogen bond basicity (β_2). Each solute property is multiplied by a term that represents the difference in complementary "solvent" properties of the mobile (subscript m) and the stationary (subscript s) phases. Thus, α_m and α_s are the abilities of the phases (bulk or bonded) to donate a hydrogen bond. These properties complement the solute's ability to accept a hydrogen bond (β_2). Similarly, δ_m^2 and δ_s^2, the squares of the Hildebrand solubility parameter or cohesive energies of the two phases, complement the solute molar volume.

Another recent theory that had an impact on QSRR studies is the mean-field statistical theory of Dill [25,26], applied to RP-HPLC by Dorsey and co-workers [27–29]. According to this theory, two driving forces dominate retention: (i) the free-energy change resulting from contact interactions of the solute and neighbouring molecules of the stationary and mobile phases and (ii) ordering of the stationary phase hydrocarbon chains leading (at higher hydrocarbon bonding density) to an entropic exclusion of solute from the stationary phase relative to that which would be expected in amorphous hydrocarbon–water partition system.

There are two important consequences of this theory for retention prediction and other QSRR studies. One is that retention in RP-HPLC increases with the grafted stationary phase chain density up to a density value of about 3.0 μmol/m, where the retention reaches a plateau. Another conclusion derived from the theory concerns the nature of the slope and intercept of the rectilinear relationship between logarithm of capacity factor, log k', and composition of binary organic–water eluent. The slope was postulated [30,31] to be directly proportional to the size of the solutes, although measures of solute size such as Van der Waals volume and molecular connectivity indices did not confirm the theoretical expectations [30].

It has been argued [32] that the RP-HPLC distribution coefficient could be calculated from known values of the activity coefficients of the substance of interest in both chromatographic phases. A means for the assessment of activity coefficients is the UNIFAC group contribution method according to Fredenslund *et al.* [33]. The UNIFAC method transforms a solution of molecules into a solution of groups. The magnitude of

a given group contribution to the activity coefficient depends on the Van der Waals group volume and surface area. The number of distinct groups is limited but is not so small as to neglect significant effects of molecular structure on physical properties. The parameters were tabulated for immediate reference [33].

There are theoretical approaches aimed at the prediction of RP-HPLC retention using the substituent and/or fragmental contribution to retention parameters. Recent papers by Smith and Burr [34,35], Hindriks *et al.* [36] and by Wells and Clark [37] report successful predictions of retention or retention-related parameters for variously substituted solutes. Apart from the purely predictive application of these methods (which form the basis of expert systems), there are interesting attempts to identify and quantify mutual interactions between substituents (fragments). For instance, Smith and Burr [35] described the RP-HPLC retention parameter, I, of disubstituted (X, Y) aromatic solutes employing the following equation:

$$I = I_P + I_{S,R} + \sum I_{S,Al-X} + \sum I_{S,Ar-X} + \sum I_{I,X-Y} \quad (3)$$

where I_P represents the retention parameter of a parent unsubstituted compound, $I_{S,R}$ is a contribution for saturated alkyl chains, $I_{S,Al-X}$ are contributions for substituents on saturated aliphatic carbons, $I_{S,Ar-X}$ are contributions for aromatic substituents and $I_{I,X-Y}$ are terms accounting for any interactions between substituents caused by electronic, hydrogen bonding and steric effects. The interaction terms are calculated by the following equation:

$$I_{I,X-Y} = (\sigma_X \rho_Y^* + \sigma_Y \rho_X^*) + F_{HB}^* + F_0^* \quad (4)$$

where ρ^*, F_{HB}^* and F_0^* are expressed in units of the retention parameter; ρ^* are the suceptibilities of X and Y to the modifying effects of Y and X on the Hammett constants of the substituents, σ_X and σ_Y, F_{HB}^* is a term accounting for hydrogen bonding and F_0^* is a term reflecting the *ortho* effect. It should be mentioned that the σ/ρ correction values, along with *ortho* effects, were demonstrated earlier [38] to be of limited value for the description of the RP-HPLC retention parameters of substituted phenols and anilines.

4. HYDROPHOBICITY CONCEPT IN CHROMATOGRAPHY

The nature of hydrophobicity [39] and the calculation of octanol–water partition coefficients from chromatographic data [40] are subjects of separate reviews in this volume.

4.1. Hydrophobicity parametrization by RP-HPLC

Since Boyce and Millborrow [41] extrapolated retention parameters determined at various organic–water eluent compositions to a pure water eluent, it became a common practice to employ extrapolated data as measures of hydrophobicity. The extrapolation is based on the assumption of the linear Soczewiński–Wachtmeister relationship [42] between log k' and the volume fraction of the organic modifier in a binary aqueous eluent. It has been demonstrated that the rectilinear relationship in RP-HPLC applies only over a limited solvent composition range that varies depending on the solute and the chromatographic system employed [31,43,44]. In effect, the values of the logarithm of the capacity factor extrapolated to a pure aqueous eluent (the intercepts in the Soczewiński–Wachtmeister equation denoted commonly by log k'_w) are usually different from those determined experimentally and depend on the organic modifier employed. Because of this observation, some workers are inclined to believe that the extrapolation of capacity factors to 0% organic modifier is a manipulation and the value of log k'_w itself has no physical meaning [45,46].

Interpretation of log k'_w as the logarithm of capacity factor corresponding to a pure water (buffer) eluent might be misleading, especially if the extrapolation is carried out over a considerable eluent composition range with a fitting function that is basically unreliable [31,47,48]. The parameter is not devoid of merits, however, as it may be regarded [42,49] as a means of normalizing retention.

If extrapolation to pure water is a normalization of hydrophobicity measures, then the question arises of what the most appropriate description of the dependence of retention parameters on the composition of the mobile phase is. Much effort has been devoted to solving this problem and individual models are discussed in detail elsewhere in this volume [31].

Isocratic capacity factors determined with various organic modifiers naturally depend on the properties of the modifier. One could expect the values extrapolated to pure water (log k'_w), however, to be independent of the organic modifier used. Unfortunately, this is not usually the case although Michels and Dorsey [49,50] reported common ln k'_w values for methanol–water, ethanol–water and acetonitrile–water eluents if ln k' was extrapolated against $E_T(30)$ (defined in ref. 3). Different modifiers yield different chromatographic measures of solute hydrophobicity. There is no reason to assume that one modifier provides a better measure of hydrophobicity than another. If the reference hydrophobicity scale is that of the 1-octanol–water partition system (log P), then individual organic modifiers appear advantageous. Braumann *et al.* [51] strongly advocated the view that a general relationship between log P and log k'_w can only be expected for capacity factors determined in methanol–water eluents. According to them, similar solute–solvent interactions operate in methanol–water and 1-octanol–water systems, whereas other organic modifiers (acetonitrile, tetrahydrofuran) introduce interactions that are not present in the 1-octanol–water system. Although Braumann *et al.* [51] postulated the identity of log k'_w and log P, there is evidence that even with non-polar solutes, such as chlorobiphenyls and alkylbenzenes, separate regressions of log P *versus* log k'_w have to be developed for each class of compound [52].

4.2. Correlations between RP-HPLC retention parameters and 1-octanol–water partition coefficients

One can certainly expect close correlations between log P and RP-HPLC retention parameters if the chromatographic system closely resembles the conventional 1-octanol–water partition system. Several procedures for obtaining such a system have been reported, based on dynamically coating a stationary phase with 1-octanol and using a 1-octanol-saturated aqueous eluent [53–55]. The correlations were satisfactory but serious technical problems made the approach impractical.

To achieve high correlations between retention parameters determined in stable RP-HPLC systems and log P, different workers have recommended a specific treatment of the stationary phase before use and the presence of various additives in the mobile phase [56–59]. Nonetheless, the correlations of such determined chromatographic data with log P were good only as long as the solutes being analysed were more or less closely related (congeneric).

In contrast to earlier tendencies, in recent publications on relationships between log P and reversed-phase liquid chromatographic parameters (from both HPLC and TLC), only moderate correlations were reported [52,60–62]. For a series of congeneric solutes the correlation between log k'_w and log P reported by Clark *et al.* [52], $r = 0.953$, appears reliable and realistic. However, poor correlations have been reported even for congeneric pyrazine derivatives [63] and oxazoline derivatives [64].

To retain log P as the RP-HPLC retention descriptor, some workers have introduced empirical corrections to log P or hydrophobic substituent (fragmental) constants [65,66]. Although such correlations may be of use for retention prediction, they are of little help in understanding the mechanism of separations. The same holds true if the correlation between log k' and log P is improved by the introduction of indicator variables [63] or molecular refractivity [67] into the regression equations. In the latter instance the statistical significance of the refractivity term in the equation reported may be questioned.

Patel *et al.* [68], in a QSRR study, employed 1-octanol–water partition coefficients to account simultaneously for changes in solute structure and mobile phase composition. They modelled ln k' from RP-HPLC by the equations

$$\ln k' = A + B(\log P/P_{sm}) + C(1/P_{sm}^2) \quad (6)$$

or

$$\ln k' = A' + B'(\log P/P_{sm}) + C'(\log P/P_{sm}^2) \quad (7)$$

where A, B, C, A', B' and C' are regression coefficients, P is the 1-octanol–water partition coefficient of the solute and P_{sm} is the calculated 1-octanol–water partition coefficient of a solvent mixture containing water and either methanol, acetonitrile or tetrahydrofuran. P_{sm} is calculated from the equation

$$\log P_{sm} = \sum_{i}^{n} (x_i \log P_{s,i}) \quad (8)$$

where x_i is the mole fraction of the ith solvent component, $P_{s,i}$ is the 1-octanol–water partition coefficient of the ith solvent component and n is the total number of pure solvents present in the solvent mixture.

The limited performance of log P, determined in a liquid–liquid partition system, in modelling RP-HPLC retention suggests differences in the two types of partition processes. There is evidence that with hydrocarbonaceous stationary phases the solute molecule can penetrate vertically into the bonded hydrocarbon layer [69]. In addition, retention on such phases is affected by the surface density of the bonded alkyl chains [28]. In such a situation the chromatographic process cannot be directly modelled by bulk organic–water partitioning processes, because the non-polar stationary phase is an interphase (immobilized at one end) and not a bulk medium.

The above discussion should be borne in mind when applying chromatographic data as a substitute for log P. Probably the least disputable way to obtain the octanol–water log P value is by the centrifugal partition chromatographic technique [70]. The problem with this method, however, is that it requires sophisticated laboratory equipment.

4.3. New RP-HPLC stationary phase materials recommended for the determination of chromatographic measures of hydrophobicity

In spite of, or due to, differences between bulk liquid–liquid and chromatographic partitioning, RP-HPLC provides a means of quantitative characterization of hydrophobicity. Different partition chromatographic systems can produce different hydrophobicity measures, each of them highlighting specific aspects of a complex property such as hydrophobicity. Hydrophobicity is as much a "phobia" against an aqueous environment as a "philia" towards non-polar species (lipophilicity). Hence the chemistry of the contact of the solute with the stationary phase cannot be neglected.

For many years octadecyl-bonded silica (ODS) stationary phases were commonly employed in hydrophobicity studies. However, the retention data obtained with nominally the same type of reversed-phase columns under identical mobile phase conditions are hardly comparable [71,72].

A serious disadvantage of silica-based reversed-phase materials is their chemical instability at pH > 8. The log P values are determined (or calculated) for neutral, non-ionized forms of solutes. The chromatographic determination of the hydrophobicity of non-ionized forms of organic bases cannot be performed directly on silica-based materials.

In attempts to provide a universal, continuous chromatographic hydrophobicity scale (not necessarily mimicking log P), several RP-HPLC materials have recently been tested. These materials are claimed to be devoid of the major problems of alkyl-bonded silicas, *i.e.*, they have no accessible free silanols and they are chemically stable over a wide pH range [73].

Poly(styrene–divinylbenzene) (PS–DVB) copolymers are stable over the pH range 1–14. They are reported to promote moderate correlations with log P, which hold usually only within subgroups of congeneric solutes [74–77]. However, columns packed with PS–DVB are characterized by low efficiency and the material suffers from excessive shrinkage and swelling [78]. Recently, several polymeric phases having a chemically bonded octadecyl moiety have been tested in hydrophobicity determinations. Phases such as octadecylpolyvinyl copolymer or rigid macroporous polyacrylamide with bonded octadecyls do not undergo swelling or shrinkage and offer the possibility of having reasonable flow-rates without undesirable pressure increases at the column inlet [78–80]. Depending on the specific phase used, the reported correlations with log P of test

solutes are low or at best as good as those obtained with the ODS phase. There is evidence, however, that individual polymeric phases provide a specific input to retention. For example, the octadecylpolyvinyl copolymer was reported to be less hydrophobic than alkylsilicas but strongly retained some specific compounds [81].

In recent years, great progress has been achieved in the technology of silica-based reversed-phase materials. Owing to the high C_{18} bonding densities, significant protection of ODS phases against hydrolysis was attained [82]. Hydrocarbonaceous silica phases exhibiting a high level of silanol deactivation became commercially available [83]. These new phases proved valuable for hydrophobicity determinations of drug solutes [84].

When alumina-based reversed-phase materials appeared, there was interest in them from the viewpoint of hydrophobicity parametrization [85]. Alumina is stable over a wide pH range and possesses no interferring silanol groups. Polybutadiene chemically encapsulated alumina (PBA) reversed-phase material was introduced by Bien-Vogelsang *et al.* [86]. Owing to chemical stability of PBA, the non-ionized forms of acids, bases and neutral species can be analysed in the same HPLC system operated at an appropriately adjusted pH. Hence a continuous hydrophobicity scale may be obtained in an easier, faster and more reproducible manner than is the case with the octanol–water system.

Carbon supports for RP-HPLC [87,88] may also appear interesting for comparative hydrophobicity studies. They have good chemical stability over a wide pH range, they do not exhibit peak tailing for amines, as do silica-based materials, and they do not adsorb phosphates or carboxylates, as do alumina and zirconia polymer-coated phases [89]. On the other hand, their high selectivity could provide information on specific features of the hydrophobicity of the solutes. This information could be of value for structure–activity studies and thus the reported [62,74] lack of correlation of log k' as determined on graphitic carbon with log P should not be discouraging.

The hydrophobic effect is assumed to be one of the "driving forces" for passive diffusion of xenobiotics through biological membranes and drug–receptor binding. If the hydrophobicity measuring system is to model a given biological phenomenon, then close similarity of the component entities is a prerequisite. Hence the partition system expected to model the transport through biological membranes should be composed of an aqueous phase and an organized phospholipid layer (bilayer). Miyake *et al.* [90] derived HPLC hydrophobicity parameters employing a column of silica gel coated physically with dipalmitoylphosphatidylcholine (DPPC). Leaving aside the inconveniences regarding their preparation and stability, the systems with DPPC adsorbed on silica probably do not emulate the lipid dynamics of biological membranes because the adsorbed lipids are not organized in a similar manner to natural (or artificial) membranes.

The recently introduced immobilized artificial membranes (IAM) as chromatographic packing materials, appear to be more reliable and convenient models of natural membranes [91,92]. The IAM surfaces are synthesized by covalently binding of the membrane-forming phospholipids to solid surfaces. They are cofluent monolayers of immobilized membrane lipids, wherein each lipid molecule is covalently bound to the surface. Membrane lipids possess polar head groups and two non-polar alkyl chains. One of the alkyl chains is linked to the solid surface. The immobilized lipid head groups protrude away from stationary phase surface and are the first contact site between solutes and IAM (Fig. 2).

Correlations between log k' data determined on IAM-type columns and experimental log P values are not high [93]. With non-end-capped columns the log k' determined with an eluent composed of sodium phosphate buffer (pH 7.00)–acetonitrile (80:20, v/v) correlates with log P with a correlation coefficient $r = 0.81$; for a column end-capped with methylglycolate (IAM.PC.MG) operated with an eluent composed of buffer–acetonitrile (75:25, v/v) the corresponding value was $r = 0.76$.

There was a very weak correlation between log k' from an IAM column and log k'_w determined on a deactivated hydrocarbonaceous silica column. This means that retention data determined on IAM columns contain information on the properties of solutes which are

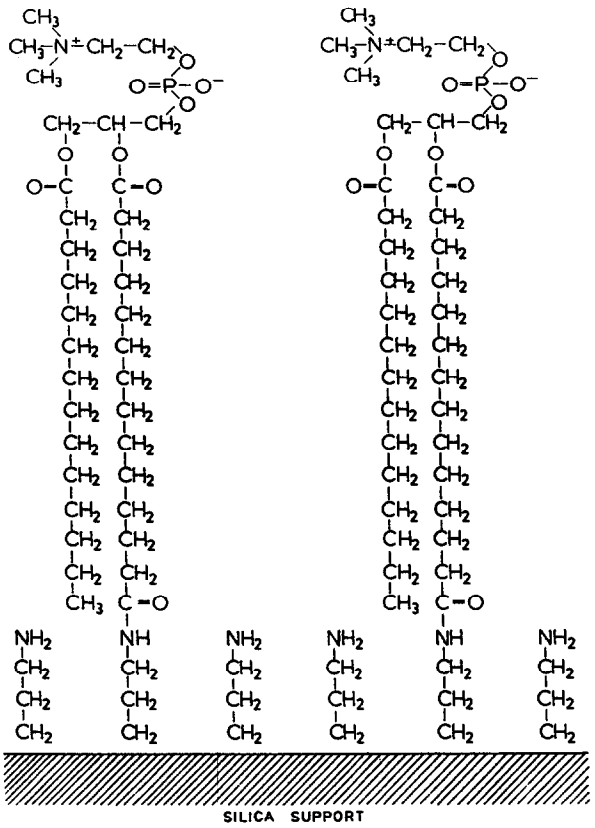

Fig. 2. Basic chemical structure of immobilized artificial membrane (IAM) stationary phases. In the case of the end-capped phase the residual propylamine groups are chemically bound to methylglycolate.

distinct from those provided by hydrocarbonaceous silica reversed-phase columns and by the 1-octanol–water slow equilibrium systems.

The IAM columns are easy to operate although there is a problem with column stability. The recommendation by the producer of avoiding alcohols as mobile phase components also causes some inconveniences.

It can be assumed that the IAM columns facilitate the hydrophobicity characteristics, which is most suitable for modelling of pharmacokinetics of drugs.

5. PHYSICO-CHEMICAL MEANING OF QSRR EQUATIONS REPORTED FOR RP-HPLC

Hundreds of QSRR equations have already been reported. This section covers the most important studies with emphasis on the more recent publications.

There are QSRR equations, aimed mostly at retention prediction, which contain terms (descriptors) of obscure physico-chemical meaning. One can argue that good retention prediction proves the validity of the descriptors present in the QSRR equation and that one should try to discover the physical sense hidden in an effective structural descriptor. This is often difficult.

A representative QSRR equation developed from the multi-parameter approach describing RP-HPLC retention indices, I_R, of polyhalogenated biphenyls is [94]

$$I_R = -66511(\pm 3646)\ [\text{fraction of positively charged surface area}] - 2469(\pm 455)\ [\text{fraction of negatively charged surface area}] - 72.9(\pm 18.8)\ [\text{number of } \textit{ortho} \text{ substituents}] + 3351(\pm 954)\ [\text{relative positive charge}] - 15.8(\pm 7.0)\ [\text{path-3 kappa Kier index}]^3 + 840.2 \quad (9)$$

$$n = 53,\ R = 0.968,\ s = 55,\ F_{6,48} = 285$$

where n is the number of solutes used to derive the regression equation, R is the multiple correlation coefficient, s is the standard estimate error, F is the value of the statistical significance test (F-test) for the model and the numbers in parentheses represent 95% confidence limits.

Eqn. 9 predicts relative retention on an ODS column, with pure methanol as the mobile phase, within the series of polyhalogenated biphenyls. However, it is difficult to assign physical meanings to the descriptors selected. The first two descriptors are defined as the surface areas of either positively or negatively charged portions of the molecule divided by the total surface area. It is not clear what might represent the relative positive charge descriptor, defined as the charge of the most positive atom in the molecule divided by the total charge of the molecule. Among the descriptors in eqn. 9, the least significant is a molecular graph-derived index, path-3 kappa, proposed by Kier [95]. Kappa indices are calculated by an algorithm that uses the number of atoms and the number of edge (bond) paths connecting the atoms in the graph.

The path-3 kappa Kier index might encode the "general shape of the molecules", but it is difficult to decide what would be that "shape" raised to the third power.

Regression equations containing indices derived from hydrogen-suppressed molecular graphs form a separate family of QSRRs. However numerous is this family, it is of little informative value regarding retention mechanisms, even though a high predictive potency in RP-HPLC of several topological indices has often been claimed in the past [1,5,6] and is still reported occasionally [96–99]. In spite of the impressive imagination of designers of the myriad of molecular graph-derived indices, it is difficult to assign a definite physical sense to individual indices (not to mention their various transformations, such as squares, square roots and reciprocals). Probably the only systematic structural information that can be extracted from molecular graph-derived indices is that concerning the bulkiness of a series of solutes [100].

Solute bulkiness certainly strongly affects RP-HPLC retention. The molecular bulkiness parameters used in QSRR studies may be considered reliable descriptors of dispersive interactions. This is evidenced by excellent correlations among these parameters and retention data determined in systems in which dispersive interactions are decisive, *i.e.*, when polar interactions are either meaningless or constant. Numerous such correlations have been reported for RP-HPLC retentions of homologous or non-polar solutes [1,5,6]. Among the more recent examples of such QSRRs, that reported for a series of triazine derivatives [101] deserves discussion. The authors described the intercept of the linear function of log k' *versus* volume fraction of acetonitrile in the mobile phase in terms of the water-accessible non-polar surface area of solutes for energy minimized structures ($r = 0.942$). The slope of the function was poorly correlated ($r = 0.878$) with the difference between the non-polar surface area minus water-accessible non-polar surface area. In fact, the authors gave a two-parameter equation using both parameters independently (instead of their difference), but the high intercorrelation between them ($r = 0.99$) invalidates that regression equation.

For solutes of equal molecular size and polarity, the differences in RP-HPLC retention may arise from steric effects. To prove steric effects on retention by QSRRs one needs numerical measures of molecular shape. Obtaining one-dimensional shape descriptors is possible only for specific solutes. Such a group of solutes are planar polycyclic aromatic hydrocarbons (PAHs). QSRR equations describing the RP-HPLC retention of PAHs have been derived [102–106]. The shape parameter employed was the molecular length-to-breadth ratio introduced previously [107] to describe the retention of PAHs on nematic phases in gas chromatography. Based on the QSRRs derived for isomeric PAHs, Sander and Wise [102] postulated a "slot" model of retention on hydrocarbon-bonded silica phases. According to this model, solute molecules would immerse in a hydrocarbon layer of the stationary phase.

Various structural descriptors, more or less directly bound to molecular size, have been tested in QSRRs describing RP-HPLC retention [1,5,6]. However, none was able to account at the same time for differences in structurally specific, polar properties of solutes. In such a situation, multi-parameter QSRRs were studied in which a bulkiness descriptor was accompanied by polarity parameters. The first to be studied, the total dipole moment, performed poorly. It has long been known [17,18] that for molecules such as 1,4-dioxane with an overall dipole moment of zero, a better explanation of relative retention was given by the assumption of the effective dipole moment as twice that of diethyl ether. The two dipoles in 1,4-dioxane are in opposition and, therefore, cancel each other. In chromatography, however, single dipoles interact at close range with molecules forming the RP-HPLC system.

Looking for an effective polarity parameter, Kaliszan *et al.* [108,109] applied a submolecular polarity parameter, Δ, defined as the largest difference in electron excess charges on a pair of atoms in a given solute molecule. The test series of solutes was twelve mono- and disubstituted benzene derivatives with a range of functional groups. The compounds were chromatographed on three octadecylsilica stationary phases with

different hydrocarbon coverages. For each stationary phase, four or five compositions of methanol–water eluents were employed. To describe the retention of individual solutes in specific RP-HPLC systems, a molecular size-related structural descriptor was used together with the polarity parameter Δ. As the size descriptor, the total energy, E_T, from quantum chemical calculations was applied.

Assuming a linear dependence of log k' on the fraction of methanol in the water–methanol eluent, X, and on an octadecyl coverage of the stationary phase, C, the following equation was derived describing the retention of solute i on phase j:

$$\log k'_{i,j} = [0.045(\pm 0.007)E_{T,i} - 2.649(\pm 0.919)\Delta_i - 0.105(\pm 0.067)C_j - 0.495(\pm 0.583)]X + [-0.038(\pm 0.004)E_{T,i} + 2.166(\pm 0.492)\Delta_i + 0.170(\pm 0.036)C_j + 1.296(\pm 0.312)] \quad (10)$$

The correlation between the 144 pairs of log k' values determined experimentally and calculated by eqn. 10 was $r = 0.986$. Replacing Δ with the total dipole moment of the solutes, μ, makes any retention prediction unreliable. This is not surprising because the correlation between μ and Δ is only $r = 0.77$.

The advantage of the submolecular polarity descriptor, Δ, over the total dipole moment in predicting RP-HPLC retention was also demonstrated in QSRR studies of RP-HPLC data determined on a polybutadiene-encapsulated alumina stationary phase [110]. The set of solutes consisted of selected rigid and planar compounds. In this way, the possibility that the conformation of a solute interacting with the components of the chromatographic phases differs from the conformation for which structural descriptors are determined was eliminated. Multiple regression analysis in which sixteen various size-related, molecular graph-derived and quantum chemical descriptors were considered, yielded the following equation:

$$\log k'_w = 0.089MR - 2.505\Delta - 1.62 \quad (11)$$

$$n = 21,\ R = 0.909,\ s = 0.50,\ F = 43$$

where MR was calculated as the sum of the bond refractivities for all pairs of connected atoms according to Vogel *et al.* [111] and the other symbols were as explained earlier.

If the parameter MR is interpreted as reflecting the ability of a solute to take part in non-specific dispersive interactions with components of the chromatographic system, then Δ can be assumed to reflect the ability of a solute to participate in specific polar interactions. Eqn. 11 indicates that the net effect of attractive dispersive interactions of a solute with the stationary phase on the one hand and with the mobile phase on the other provides a positive input for retention parameters. The net effect on the retention of attractive polar interactions between a solute and molecules of the mobile phase and between a solute and the stationary phase is negative, *i.e.*, the more polar the solute, the less it is retained.

Eqn. 11 rationalizes the mechanism of RP-HPLC separation on polybutadiene-coated alumina phases and is in agreement with the report by Arenas and Foley [112] on the nature of processes that determine RP-HPLC retention on these stationary phase materials. However, the equation does not allow for the precise prediction of log k'_w for a given solute. The parameters MR and Δ are rough measures of the non-polar and polar properties of solutes and cannot be claimed to be universal for all QSRR studies. Structural specificity of an individual set of solutes may give rise to the application of better performing retention descriptors. The submolecular polarity parameter Δ performs better than the total dipole moment as a retention descriptor. This polarity descriptor resembles the "bond dipole moment" or the "effective dipole moment" previously suggested [17,18] to explain differences in retention among solutes of similar size.

There is a report [113] in which a successful description of RP-HPLC data by a three-parameter regression equation containing the Van der Waals volume, the square of the dipole moment and the hydrogen bond energy as independent variables was claimed. Unfortunately, no in-

formation on the significance of individual terms in the QSRR equations was provided. Similar statistical objections concern a paper [114] in which ln k'_w is described in terms of substituent constants: Hansch π, Hammett σ, Taft E_s and Swain–Lupton F.

Recently the RP-HPLC retention of a series of benzodiazepine derivatives chromatographed on specially deactivated hydrocarbonaceous silica with methanol–buffer eluents was related to various additive–constitutive molecular descriptors and to the descriptors derived by molecular modelling [115]. The most informative equation obtained was

$$\log k'_w = 1.823 + 0.444(\pm 0.046) f_{X+Y} - 1.187(\pm 0.312) C_3^* - 0.010(\pm 0.003)\mu^2 + 0.012(\pm 0.004) MR \quad (12)$$

$$n = 21,\ R = 0.930,\ F = 25.7,\ p < 1.0 \cdot 10^{-6}$$

where the statistical terms are as explained earlier, except p, which denotes the significance levels of the regression equation. The variables used in eqn. 12 denote the following (Fig. 3): f_{X+Y} is the sum of hydrophobic fragmental constants of substituents at the aromatic part of the molecule according to Taylor [116], MR is the molecular refractivity of the whole molecule calculated according to Vogel [117], μ is the total dipole moment and C_3^* is the electron excess charge on carbon atom C-3 of the 1,4-benzodiazepine system.

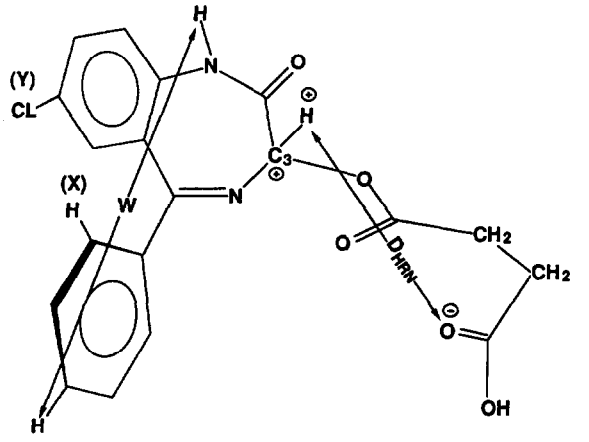

Fig. 3. Determination of structural descriptors of benzodiazepine derivatives used in QSRR analysis. X and Y are substituents in aromatic rings; C_3^* is the excess charge on carbon C-3; W is the width of the molecule along the phenyl substituent; D_{HRN} is the distance between hydrogen at C-3 and the most negatively charged atom of another substituent at C-3.

Eqn. 12 shows that the hydrophobicity of the aromatic part of solutes as quantified by f_{X+Y} and solute bulkiness as described by MR provide net positive inputs to RP-HPLC retention. On the other hand, the structural descriptors that may be related to solute polarity, C_3^* and μ^2, account for retention-decreasing effects.

QSRRs based on linear solvation energy relationships (LSERs) and the solvatochromic comparison method have been the subject of several publications. Carr *et al.* [23] analysed log k' data from RP-HPLC determined for a set of benzene derivatives in five acetonitrile–water systems at four different temperatures. By regression analysis they derived coefficients c, m, s, a and b of twenty equations of the form

$$\log k'_i = c + mV_2/100 + s\pi_2^* + a\alpha_2 + b\beta_2 \quad (13)$$

where i denotes individual eluent composition and temperature, V_2 is the solute molar volume, π_2^* is its polarizability–dipolarity, α_2 represents hydrogen bond acidity and β_2 is the hydrogen bond basicity of the solutes. With the data set considered, not every variable of eqn. 13 appeared statistically significant. The meaningful equations obtained contained the $V_2/100$, β_2 and either π_2^* or α_2 terms. Carr *et al.* [23] opted for π_2^* as the third regression parameter. The predictive efficiency of the three-variable regression equations obtained was very high. It should be noted, however, that the authors excluded several outlier solutes (for reasons not explained) when deriving their individual regression equations. The highest number of solutes considered in the regression was 21 out of a total of 26. For nine of the reported equations, the number of solutes taken into consideration was only 15–17 out of a total of 20.

Aplying the same approach to log k' determined for eight small aromatic solutes on ODS with methanol–water (60:40, v/v) as the eluent, Park *et al.* [118] obtained QSRRs relating retention to the V_2 and β_2 parameters. The small size

of the solute series limited the possibility of demonstrating the significance for RP-HPLC of other variables in eqn. 13.

Recently, Chinese workers employed solvatochromic parameters in QSRR equations describing the slope of the relationship of log k' *versus* volume fraction of organic solvent in a binary aqueous mobile phase [119,120]. From the analysis of the significance of the coefficients of the individual regression variables, V_2, π_2^*, β_2 and α_2, it appears that the term representing hydrogen bonding acidity, α_2, is statistically insignificant. This is in spite of using a sufficiently large group of solutes ($n = 49$) for the QSRR study. Thus, the authors obtained basically the same form of relationship as that reported by Carr *et al.* [23].

In view of the reported QSRRs, it is difficult to judge whether solvatochromic parameters are more suitable and convenient for the description of the RP-HPLC retention of chemical compounds than the parameters obtained theoretically by molecular modelling and quantum chemical calculations. Solvatochromic parameters are empirical. Respective data are available for an increasing number of compounds [121,122]. On the other hand, theoretical descriptors can be obtained readily for any structure owing to the common access to calculation chemistry software.

Modelling of RP-HPLC retention according to the UNIFAC group contribution method [33] has been attempted only occasionally [32,123]. The prediction of retention was moderately good. An interesting observation was that if hexane, cumene and octanol were used to model the ODS stationary phase then the best prediction of retention was obtained with octanol [32].

6. STRUCTURAL AND MECHANISTIC INFORMATION ON RP-HPLC RETENTION FROM FACTORIAL METHODS OF DATA ANALYSIS

Factor analysis is applied in chemistry to determine the "intrinsic dimensionality" of certain experimentally determined chemical properties, that is, the number of "fundamental factors" required to account for the variance [124]. Once the number of factors has been determined, the next step is to try to identify these abstract factors with physically significant parameters. The advantage that factor analysis has over regression analysis is that individual factors can be tested for possible identification with the abstract factors without simultaneously identifying all the other fundamental factors.

Factor analysis of chromatographic data has most often been used as a clustering technique. The proximities of points representing solutes or chromatographic systems suggest similarities of properties. Information extracted from factor analysis is mostly exploited for stationary and mobile phase classification, for optimization of separation conditions from the viewpoint of retentivity and selectivity [125]. In RP-HPLC, factor analysis is seldom focused on the physico-chemical exploitation of the data. The extracted factors most often remain abstract. There are examples, however, in which one of the factorial axes can be attributed to the contribution of solute properties such as the partition coefficient [126] or electronic characteristics [127]. For a series of monosaccharides and polyols, one of two independent factors could be directly connected to the number of accessible hydroxyl groups on the solutes and the other, well quantified, remained physico-chemically not interpreted [128].

Another example of the application of factorial methods of data analysis to obtain some insight into the mechanism of RP-HPLC retention may be found in papers by Forgács and co-workers [129,130]. They analysed RP-HPLC data determined for a series of phenol and aniline derivatives on a graphitized carbon column. The most important finding was that the retention of solutes was not governed by their lipophilicity and bulkiness, as would normally be expected in RP-HPLC. This observation is consistent with reports on the poor correlation of log k' from RP-HPLC on graphitized carbon with log P [74] and the reported importance for retention of charge-transfer interactions [62]. In a subsequent paper, Forgács and Cserháti [131] demonstrated by means of factorial analysis that the retention of anilines on graphitized carbon was mainly dependent on electronic parameters

and on the hydrogen acceptor capacity of substituents.

Typically, factorial methods of chromatographic data analysis support only very general concepts regarding the mechanism of RP-HPLC retention, *e.g.*, that "predominantly hydrophobic and polar forces affect the separation process" [132]. More instructive in this respect appears to be factor analysis of structural data of solutes and subsequent application of the extracted systematic information to describe retention. An example of such an approach in QSRRs concerns RP-HPLC retention data determined on polybutadiene-encapsulated alumina (PBA) for a set of non-ionized organic bases and neutral species [110]. For 21 solutes, sixteen structural descriptors were determined along with $\log k'_w$ parameters. Among the descriptors considered were size-related parameters, such as molecular mass, refractivity, total energy, molecular connectivity indices and Wiener index. There were also descriptors related to molecular polarity, such as dipole moments, maximum electron excess charge differences and the energies of the highest occupied and lowest unoccupied molecular orbitals. Separate groups of descriptors formed parameters claimed to encode molecular shape and the indices calculated from the probabilities of finding equivalent atoms or patterns of atoms in a given structural formula. Several of the descriptors considered were strongly mutually intercorrelated, which excluded their simultaneous use in a multiple regression equation.

By principal component analysis (PCA) of a matrix of sixteen structural descriptors of 21 solutes, systematic information dispersed over many variables was extracted. The first principal component, PC1, accounted for 48.6% of the variance in the structural data considered and the second principal component, PC2, for 25.2%. PC1 basically condensed information on the molecular size of the solutes. PC2 was influenced mainly by structural descriptors, such as maximum charge difference or dipole moment. The equation describing $\log k'_w$ in terms of inputs ("scores") to PC1 and PC2 by individual solutes is

$$\log k'_w = 0.594(\mathrm{PC1}) - 0.902(\mathrm{PC2}) + 0.885 \quad (14)$$

$$n = 21,\ R = 0.948,\ s = 0.380,\ F = 80.6$$

where the symbols are as explained earlier.

Eqn. 14 provides qualitative information similar to eqn. 11 concerning the mechanism of RP-HPLC retention on PBA columns. However, owing to the exploitation of information from many descriptors, the statistical value of eqn. 14 is higher and it provides a better prediction of $\log k'_w$ than eqn. 11. Both eqns. 11 and 14 show evidence that the mechanism of RP-HPLC retention on PBA is similar to that on ODS (see eqn. 10), *i.e.*, that a net positive input due to non-specific dispersive interactions and a net negative input due to polar interactions of a solute with the stationary and moblie phase determine retention.

7. CONCLUSIONS

During the last 15 years, many reports have been published on QSRRs which specifically concern RP-HPLC. In general, the physical significance, predictive potency and statistical quality of these QSRRs are lower than those in gas–liquid chromatography on non-polar stationary phases [1]. On the other hand, QSRRs reported for RP-HPLC are, as a rule, better than those in normal-phase liquid chromatography. This observation is a consequence of increasing system complexity, which is a consequence of the increasing importance of structurally specific, directional, polar properties of solutes for their retention. The identification and quantitative characteristics of various aspects of the polarity of chemical entities form a problem in all types of structure–property relationship studies and the knowledge gained in one type of study can be exploited in another. RP-HPLC can readily yield a great amount of well measurable data concerning a property (retention) for diverse chemical structures. In such a situation, QSRR analysis may be a convenient means of selecting the most promising structural descriptors.

The descriptors used in QSRR studies may be of empirical, additive–constitutive semi-empirical and theoretical nature. A typical empirical descriptor of solutes is the logarithm of the partition coefficient in 1-octanol–water liquid–

liquid partition systems, log *P*. In general, there is some correlation between RP-HPLC retention data and log *P* for a given series of solutes. This relationship proves certain common features of the liquid–liquid and chromatographic partition systems. It supports the assumption that in RP-HPLC the partition, rather than adsorption, processes are decisive for retention. However, each RP-HPLC system provides a distinct individual hydrophobicity measure of solutes. Information on specific aspects of solute hydrophobicity obtained from different RP-HPLC systems may be unsuitable for the prediction of log *P* but may be of use for the evaluation of other hydrophobicity-related properties, *e.g.*, bioactivity.

Solvatochromic structural parameters provided progress in the prediction of RP-HPLC and also other partition data. QSRRs employing these structural descriptors of solutes, stationary phases and mobile phases can be of help in rationalizing the mechanism of RP-HPLC retention. General validity of the parameters cannot be expected yet, however, and individual solutes deviate from the relationships postulated. In addition, solvatochromic parameters are not readily available for every set of solutes of potential interest in QSRR analysis.

QSRR models of RP-HPLC based on theoretically calculated structural parameters, which are assumed to reflect the abilities of solutes to participate in fundamental intermolecular interactions with components of chromatographic systems, allow the interpretation of retention mechanisms in simple rational terms. The models are of comparable predictive potency to those based on solvatochromic parameters and linear solvation energy relationships. Their advantage is that a variety of structural descriptors can easily be generated by standard molecular mechanics and quantum chemistry calculations. It is expected that in this way structural parameters will be identified that account better for physico-chemical and biological properties.

Finally, a chemometric approach can help to systematize our knowledge of factors that affect retention in RP-HPLC. One can attempt to list some more straightforward advantages resulting from QSRR studies of RP-HPLC data (Table 2). However, it must be emphasized that QSRR is the best training object for elaborating the strategy and methodology of all structure-prop-

TABLE 2

BASIC APPLICATIONS AND ACHIEVEMENTS OF QSRR STUDIES OF RP-HPLC RETENTION DATA

Prediction of retention
Expert systems: deriving and testing. Optimization of separation conditions. Identification of individual solutes in mixtures of compounds (congeneric). Tutorials on chromatography and structural chemistry
Generation and testing of structural descriptors
Molecular refractivity, volume, surface area and total energy as "bulkiness" parameters. Localized dipoles as polarity descriptors. Orbital energies as charge-transfer parameters. Length-to-breadth ratio as a shape parameter of planar solutes. Validation of descriptive potency of the substituent and fragmental constants: electronic, hydrophobic, steric, UNIFAC and the fragment contributions to retention. Testing the applicability of individual solubility parameters and solvatochromic parameters
Identification of principles determining separation in individual chromatographic systems
Comparison of RP-HPLC systems with the standard, reference, slow-equilibrium partition systems. Estimation of inputs to retention due to non-polar (non-specific) and polar (structurally specific) interactions. Distinguishing partition and adsorption retention mechanisms
Prediction of complex physico-chemical and biological properties
Selection of RP-HPLC systems providing the required hydrophobicity data for a given class of solutes (suppression of ionization). Identification of chromatographic systems mimicking the properties of the biophase. Correlation of chromatographic with bioactivity data and rationalization of drug development processes with reduction of animal use. Testing the performance of multivariate methods of data analysis in extracting systematic, bioactivity or other property-relevant information from diverse RP-HPLC data

erty relationship studies. Thus, based on QSRR, the ultimate goal of chemistry, the design and production of entities of any required property, might be attained sooner.

8. ACKNOWLEDGEMENTS

This work was supported by Grant No. 408319101 from the Komitet Badań Naukowych, Warsaw, Poland.

REFERENCES

1 R. Kaliszan, *Quantitative Structure–Chromatographic Retention Relationships*, Wiley, New York, 1987.
2 J.M. Prausnitz, *Science*, 205 (1979) 759.
3 C. Reichardt, *Solvent Effects in Organic Chemistry*, Verlag Chemie, Weinheim, 1979.
4 R. Kaliszan, *Anal. Chem.*, 64 (1992) 619A.
5 R. Kaliszan, *CRC Crit. Rev. Anal. Chem.*, 16 (1986) 323.
6 R. Kaliszan, in P.R. Brown and R.A. Hartwick (Editors), *High Performance Liquid Chromatography*, Wiley, New York, 1989, Ch. 14.
7 R. Kaliszan, *Quant. Struct.–Act. Relat.*, 9 (1990) 83.
8 R. Kaliszan, *Adv. Chromatogr.*, 33 (1993) 147.
9 C. Hansch and T. Fujita, *J. Am. Chem. Soc.*, 86 (1964) 1616.
10 S.M. Free, Jr., and J.W. Wilson, *J. Med. Chem.*, 7 (1964) 395.
11 K.L. Peterson, *Anal. Chem.*, 64 (1992) 379.
12 R.C. Glen, V.S. Rose, J.C. Lindon, R.J. Ruane, I.D. Wilson and J.K. Nicholson, *J. Planar Chromatogr.–Mod. TLC*, 4 (1991) 432.
13 K. Jinno, *A Computer-Assisted Chromatography System*, Hüthig, Heidelberg, 1990.
14 J.L. Glajch and L.R. Snyder (Editors), *Computer-Assisted Method Development for High-Performance Liquid Chromatography*, Elsevier, Amsterdam, 1990; *J. Chromatogr.*, Vol. 485 (1989).
15 M.S. Charton, S. Clementi, S. Ehrenson, O. Exner, J. Shorter and S. Wold, *Quant. Struct.–Act. Relat.*, 4 (1985) 29.
16 J.-F. Gal, P.-C. Maria, M. Chastrette, D. Zakarya, O. Exner, U. Haldna, M. Sjostrom, S. Wold and R.I. Zalewski, *Quant. Struct.–Act. Relat.*, 9 (1990) 132.
17 R.P.W. Scott, *J. Chromatogr.*, 122 (1976) 35.
18 B.L. Karger, L.R. Snyder and C. Eon, *J. Chromatogr.*, 125 (1976) 71.
19 R. Tijssen, H.A.H. Billiet and P.J. Schoenmakers, *J. Chromatogr.*, 122 (1976) 185.
20 C. Horváth, W. Melander and J. Molnar, *J. Chromatogr.*, 125 (1976) 129.
21 D.E. Martire and R.E. Boehm, *J. Liq. Chromatogr.*, 3 (1980) 753.
22 P.C. Sadek, P.W. Carr, R.M. Doherty, M.J. Kamlet, R.W. Taft and M.H. Abraham, *Anal. Chem.*, 57 (1985) 2971.
23 P.W. Carr, R.M. Doherty, M.J. Kamlet, R.W. Taft, W. Melander and C. Horváth, *Anal. Chem.*, 58 (1986) 2674.
24 L.R. Snyder, P.W. Carr and S.C. Rutan, *J. Chromatogr. A*, 656 (1993) 537.
25 K.A. Dill, *J. Phys. Chem.*, 91 (1987) 1980.
26 K.A. Dill, J. Naghizadeh and J.A. Marqusee, *Annu. Rev. Phys. Chem.*, 39 (1988) 425.
27 P.T. Ying, J.G. Dorsey and K.A. Dill, *Anal. Chem.*, 61 (1989) 2540.
28 K.B. Sentell and J.G. Dorsey, *Anal. Chem.*, 61 (1989) 930.
29 P.T. Ying and J.G. Dorsey, *Talanta*, 38 (1991) 237.
30 B.P. Johnson, M.G. Khaledi and J.G. Dorsey, *Anal. Chem.*, 58 (1986) 2354.
31 K. Valko, L.R. Snyder and J.L. Glajch, *J. Chromatogr. A*, 656 (1993) 501.
32 L. Dasko, *J. Chromatogr.*, 543 (1991) 267.
33 A. Fredenslund, R.L. Jones and J.M. Prausnitz, *AIChE J.*, 21 (1975) 1086.
34 R.M. Smith and C.M. Burr, *J. Chromatogr.*, 485 (1989) 325.
35 R.M. Smith and C.M. Burr, *J. Chromatogr.*, 550 (1991) 335.
36 R. Hindriks, F. Maris, J. Vink, A. Peeters, M. De Smet, D. Massart and L. Buydens, *J. Chromatogr.*, 485 (1989) 255.
37 M.J.M. Wells and C.R. Clark, *Anal. Chem.*, 64 (1992) 1660.
38 A. Tsantili-Kakoulidou, N. El Tayar, H. van de Waterbeemd and B. Testa, *J. Chromatogr.*, 389 (1987) 33.
39 P.W. Carr, J. Li, A.J. Dallas, D.I. Eikens and L.C. Tan, *J. Chromatogr. A*, 656 (1993) 113.
40 W.J. Lambert, *J. Chromatogr. A*, 656 (1993) 469.
41 C.B.C. Boyce and B.V. Millborrow, *Nature*, 208 (1965) 537.
42 E. Soczewiński and C.A. Wachtmeister, *J. Chromatogr.*, 7 (1962) 311.
43 P.J. Schoenmakers, H.A.H. Billiet, R. Tijssen and L. de Galan, *J. Chromatogr.*, 149 (1978) 519.
44 L.R. Snyder, J.W. Dolan and J.R. Gant, *J. Chromatogr.*, 165 (1979) 3.
45 K. Miyake, N. Mizuna and H. Terada, *J. Chromatogr.*, 439 (1988) 227.
46 P.M. Sherblom and R.P. Eganhouse, *J. Chromatogr.*, 454 (1988) 37.
47 T. Braumann, *J. Chromatogr.*, 373 (1986) 191.
48 T. Braumann and B. Jastorff, *J. Chromatogr.*, 350 (1985) 105.
49 J.J. Michels and J.D. Dorsey, *J. Chromatogr.*, 499 (1990) 435.
50 J.J. Michels and J.D. Dorsey, *J. Chromatogr.*, 457 (1988) 85.
51 T. Braumann, H.-G. Genieser, C. Lullmann and B. Jastorff, *Chromatographia*, 24 (1987) 777.
52 C.R. Clark, J.M. Barksdale, C.A. Mayfield, W.R. Ravis and J. DeRuiter, *J. Chromatogr. Sci.*, 28 (1990) 83.

53 D. Henry, J.H. Block, J.L. Anderson and G.R. Carlson, *J. Med. Chem.*, 19 (1976) 619.
54 S.H. Unger and G.H. Chiang, *J. Med. Chem.*, 24 (1981) 262.
55 S.J. Lewis, M.S. Mirrlees and P.J. Taylor, *Quant. Struct.–Act. Relat.*, 2 (1983) 1.
56 J.E. Garst, *J. Pharm. Sci.*, 73 (1984) 1616.
57 J.J. Sabatka, D.J. Minick, T.K. Shumaker, G.L. Hodgson, Jr., and D.A. Brent, *J. Chromatogr.*, 384 (1987) 349.
58 F. Gago, J. Alvarez-Builla and J. Elguero, *J. Liq. Chromatogr.*, 10 (1987) 1031.
59 M.C. Pietrogrande, F. Dondi, G. Blo, P.A. Borea and C. Bighi, *J. Liq. Chromatogr.*, 10 (1987) 1065.
60 K.P. Dross, R. Mannhold and R.F. Rekker, *Quant. Struct.–Act. Relat.*, 11 (1992) 36.
61 G.L. Biagi, M. Recanatini, A.M. Barbaro, M.C. Guerra, A. Sapone, P.A. Borea and M.C. Pietrogrande, in C. Silipo and A. Vittoria (Editors), *QSAR: Rational Approaches to the Design of Bioactive Compounds*, Elsevier, Amsterdam, 1991, p. 83.
62 N. Tanaka, T. Tanigawa, K. Kimata, K. Hosoya and T. Araki, *J. Chromatogr.*, 549 (1991) 29.
63 C. Yamagami and N. Takao, *Chem. Pharm. Bull.*, 39 (1991) 1217.
64 F. Demotes-Mainard, C. Jarry, J. Thomas and P. Dallet, *J. Liq. Chromatogr.*, 14 (1991) 795.
65 K. Jinno and Y. Yokoyama, *J. Chromatogr.*, 550 (1991) 325.
66 A.A. Petrauskas and V.K. Svedas, *J. Chromatogr.*, 585 (1991) 3.
67 A.L. Pereira, E.J.L. Barreiro, A.C.C. Freitas, C.J.C. Correa and L.N.L.F. Gomes, *J. Liq. Chromatogr.*, 14 (1991) 1161.
68 H.B. Patel, D.N. King and T.M. Jefferies, *J. Chromatogr.*, 555 (1991) 21.
69 A. Tchapla, H. Colin and G. Guiochon, *Anal. Chem.*, 56 (1984) 621.
70 N. El Tayar, R.-S. Tsai, P. Vallat, C. Altomare and B. Testa, *J. Chromatogr.*, 556 (1991) 181.
71 P.E. Antle, A.P. Goldberg and L.P. Snyder, *J. Chromatogr.*, 321 (1985) 1.
72 J. Nawrocki, *Chromatographia*, 31 (1991) 177.
73 R. Kaliszan, *Quant. Struct.–Act. Relat.*, 9 (1990) 83.
74 V. De Biasi, W.J. Lough and M.B. Evans, *J. Chromatogr.*, 353 (1986) 279.
75 S. Bitteur and R. Rosset, *J. Chromatogr.*, 394 (1987) 279.
76 K. Miyake, F. Kitaura, N. Mizuno and H. Terada, *Chem. Pharm. Bull.*, 35 (1987) 377.
77 W.J. Lambert, L.A. Wright and J.K. Stevens, *Pharm. Res.*, 7 (1990) 577.
78 A. Bechalany, T. Rothlisberger, N. El Tayar and B. Testa, *J. Chromatogr.*, 473 (1989) 115.
79 C. Altomare, R.S. Tsai, N. El Tayar, B. Testa, A. Carotti, S. Cellamare and P.G. De Benedetti, *J. Pharm. Pharmacol.*, 43 (1991) 191.
80 J.V. Dawkins, N. Gaggot, L.L. Lloyd, J.A. McConville and F.P. Warner, *J. Chromatogr.*, 452 (1988) 145.
81 J. Yamaguchi and T. Hanai, *Chromatographia*, 27 (1989) 371.
82 K.B. Sentell, K.W. Bornes and J.G. Dorsey, *J. Chromatogr.*, 455 (1988) 95.
83 T.L. Ascah and B. Feibush, *J. Chromatogr.*, 506 (1990) 357.
84 R. Gami-Yilinkou and R. Kaliszan, *Chromatographia*, 30 (1990) 277.
85 R. Kaliszan, R.W. Blain and R.A. Hartwick, *Chromatographia*, 25 (1988) 5.
86 U. Bien-Vogelsang, A. Deege, H. Figge, J. Kohler and G. Schomburg, *Chromatographia*, 19 (1984) 170.
87 J.H. Knox and B. Kaur, in P.B. Brown and R.A. Hartwick (Editors), *High Performance Liquid Chromatography*, Wiley, New York, 1989, Ch. 4.
88 T.P. Weber and P.W. Carr, *Anal. Chem.*, 62 (1990) 2620.
89 M.P. Rigney, T.P. Weber and P.W. Carr, *J. Chromatogr.*, 484 (1989) 273.
90 K. Miyake, F. Kitaura, N. Mizuno and H. Terada, *J. Chromatogr.*, 389 (1987) 47.
91 C. Pidgeon and U.V. Venkatarum, *Anal. Biochem.*, 176 (1989) 36.
92 H. Thurnhofer, J. Schnabel, M. Betz, G. Lipka, C. Pidgeon and H. Hauser, *Biochim. Biophys. Acta*, 1064 (1991) 275.
93 R. Kaliszan, A. Kaliszan and I.W. Wainer, *J. Pharm. Biomed. Anal.*, (1993) in press.
94 M.N. Hasan and P.C. Jurs, *Anal. Chem.*, 62 (1990) 2318.
95 L.B. Kier, *Med. Res. Rev.*, 7 (1987) 417.
96 N. Adler, N. Rak and K. Sertie-Bionda, *Fresenius' Z. Anal. Chem.*, 334 (1989) 136.
97 J. Mokrosz and B. Duszyńska, *Quant. Struct.–Act. Relat.*, 9 (1990) 33.
98 G. Bazylak, *J. Planar Chromatogr.*, 5 (1992) 275.
99 H. Vuorela, P. Lehtonen and R. Hiltunen, *J. Chromatogr.*, 507 (1990) 367.
100 K. Ośmiałowski and R. Kaliszan, *Quant. Struct.–Act. Relat.*, 10 (1991) 125.
101 K. Valko and P. Slegel, *J. Chromatogr.*, 592 (1992) 59.
102 L.C. Sander and S.A. Wise, *CRC Crit. Rev. Anal. Chem.*, 18 (1987) 299.
103 K. Jinno, *Adv. Chromatogr.*, 30 (1989) 123.
104 S.A. Wise, L.C. Sander, R. Lapouyade and P. Garrigues, *J. Chromatogr.*, 514 (1990) 111.
105 S.F.Y. Li, H.K. Lee and C.P. Ong, *J. Liq. Chromatogr.*, 12 (1989) 3251.
106 K. Jinno, S. Shimura, N. Tanaka, K. Kimata, J.C. Fetzer and W.R. Biggs, *Chromatographia*, 27 (1989) 285.
107 A. Radecki, H. Lamparczyk and R. Kaliszan, *Chromatographia*, 12 (1979) 595.
108 R. Kaliszan, K. Ośmiałowski, S.A. Tomellini, S.-H. Hsu, S.D. Fazio and R.A. Hartwick, *Chromatographia*, 20 (1985) 705.
109 R. Kaliszan, K. Ośmiałowski, S.A. Tomellini, S.-H. Hsu, S.D. Fazio and R.A. Hartwick, *J. Chromatogr.*, 352 (1986) 141.

110 R. Kaliszan and K. Ośmiałowski, *J. Chromatogr.*, 506 (1990) 3.
111 A.J. Vogel, W.T. Cresswell, G.H. Jeffery and J. Leicester, *J. Chem. Soc.*, (1952) 514.
112 R.V. Arenas and J.P. Foley, *Anal. Chim. Acta*, 246 (1991) 113.
113 P. Lu, H. Zou and Y. Zhang, *J. Chromatogr.*, 509 (1990) 171.
114 H.F. Zou, Q.S. Wang, R.Y. Gao, H.Z. Yang, B.W. Yang, Y.K. Zhang and P.C. Lu, *Chromatographia*, 31 (1991) 143.
115 R. Kaliszan, A. Kaliszan, T.G. Noctor, W.P. Purcell and I.W. Wainer, *J. Chromatogr.*, 609 (1992) 69.
116 P.J. Taylor, in C. Hansch, P.G. Sammes and J.B. Taylor (Editors), *Comprehensive Medicinal Chemistry*, Vol. 4, Pergamon Press, Oxford, 1990, p. 241.
117 A.I. Vogel, *Textbook of Practical Organic Chemistry*, Chaucer, London, 1977, p. 1034.
118 J.H. Park, M.D. Jang and M.J. Shin, *J. Chromatogr.*, 595 (1992) 45.
119 H. Zou, Y. Zhang and P. Lu, *J. Chromatogr.*, 522 (1990) 49.
120 N. Chen, Y. Zhang and P. Lu, *J. Chromatogr.*, 606 (1992) 1.
121 M.J. Kamlet, R.M. Doherty, M.H. Abraham, Y. Marcus and R.W. Taft, *J. Phys. Chem.*, 92 (1988) 5244.
122 D.E. Leahy, J.J. Morris, P.J. Taylor and A.R. Wait, *J. Chem. Soc. Perkin Trans. 2*, 1992 (1992) 705.
123 J. Petrovic, S. Lomic and I. Sefer, *J. Chromatogr.*, 348 (1985) 49.
124 E.R. Malinowski and G.D. Howery, *Factor Analysis in Chemistry*, Wiley, New York, 1980.
125 M. Righezza and J.R. Chretien, *J. Chromatogr.*, 556 (1991) 169.
126 B. Walczak, M. Dreux, J.R. Chretien, K. Szymoniak, M. Lafosse, L. Morin-Allory and J.P. Doucet, *J. Chromatogr.*, 353 (1986) 109.
127 B. Walczak, L. Morin-Allory, J.R. Chretien, M. Lafosse and M. Dreux, *Chemometr. Intell. Lab. Syst.*, 1 (1986) 79.
128 L. Morin-Allory and B. Herbreteau, *J. Chromatogr.*, 590 (1992) 203.
129 E. Forgács, T. Cserháti and K. Valko, *J. Chromatogr.*, 592 (1992) 75.
130 E. Forgács and T. Cserháti, *J. Chromatogr.*, 600 (1992) 43.
131 E. Forgács and T. Cserháti, *Chromatographia*, 33 (1992) 356.
132 S.J. Schmitz, H. Zwanziger and H. Engelhardt, *J. Chromatogr.*, 544 (1991) 381.