CHAPTER

8

Solutes in HPLC

OUTLINE

8.1. NATURE OF THE SOLUTE

General Aspects

The chemical nature of solutes in a sample is a critical factor in selecting and performing an HPLC analysis. The classification of analytes based on their chemical nature can involve several criteria. One such classification can be based on the functional groups of the analytes (e.g., similar to that of an organic textbook). Another classification can be based on the role of the analyte in everyday life. This section presents a short description of these two classifications.

Essential in Modern HPLC Seperations
http://dx.doi.org/10.1016/B978-0-12-385013-3.00008-2

Solutes Classification Based on Their Chemical Structure

Depending on the molecular weight, the solutes/analytes are initially divided into small molecules and macromolecules. Small molecules are further classified based on their chemical structure. First, the organic and inorganic compounds are separately classified. For organic compounds, the simplest class is that of saturated hydrocarbons (linear or branched). Following this class are saturated cyclic hydrocarbons, unsaturated hydrocarbons with one or more double bonds, unsaturated hydrocarbons with triple bonds, aromatic hydrocarbons, and so on. Combinations of all these structures are possible.

Various functional groups can be attached on the hydrocarbon backbone. They can be classified based on the nature of atoms in the functional group or other criteria (such as monofunctional and bifunctional). This procedure will differentiate halogenated compounds, alcohols, enols, phenols, ethers, peroxy compounds, thiols, sulfides, amines, imines, a variety of other nitrogenous compounds (nitro, oximes, etc.), aldehydes, ketones, carboxylic acids, various derivatives of organic acids (such as esters, lactones, and acyl chlorides), derivatives of carbonic acid (such as ureas and cyanates), sulfonic acids, and other less common types. More than one type of functional group can be attached on the hydrocarbon backbone. Specific classes of compounds are generated from this large variety of combinations, such as carbohydrates, amino acids, and lipids.

A special group of compounds is that of heterocycles (both aromatic and nonaromatic). Heterocycles can be classified based on the heteroatoms in the cycle such as oxygen (furans, pyrans), nitrogen (pyrole, pyrazole, imidazole, triazole, pyridine, pyrazines, etc.), sulfur (thyophene), or different heteroatoms (oxazoles, thiazole, oxadiazoles, etc.).

In addition to all the classes of compounds containing C, H, O, N, or S, other types of organic compounds must be included, such as those containing boron, silicon, arsenic, stibium, and metallic elements (organometallic compounds).

For macromolecules, the chemical structure also can be an important criterion of classification. Specific groups such as polymeric carbohydrates, lignins, tannins, Maillard browning polymers, proteins, and nucleic acids, as well as various types of synthetic polymers, can be differentiated based on their structure.

The information regarding the chemical nature of the analytes (solutes) is very important since the selection of the type of separation, the best type of stationary phase to be used, the nature and pH of the mobile phase, and so on can be determined by it.

Classification Based on the Role of the Solute in Everyday Life

Depending on the role/function in everyday life, analytes are frequently classified in various categories. No specific criteria are set for such classification, and only some of the most important categories are indicated here. It should also be mentioned that the classification based on the role in everyday life is frequently combined with that based on the analyte structure.

One common classification of samples is that of biological and nonbiological. The two groups are frequently subdivided into multiple subgroups. Another classification is based on criteria related to environmental issues, and groups such as environmental pollutants, pesticides, herbicides, and fungicides are considered. Several large groups of compounds are classified on the basis of their role in health issues. These are the groups of pharmaceuticals, metabolites, and biodegradation products [1]. Numerous classes of pharmaceuticals are further differentiated such as antibiotics, anticancer drugs, antiepileptics, steroids, analgesics, and vitamins. Metabolites and biodegradation products are also further classified based

on their source, mechanism of production, and so on. A classification based on the compound's role in toxicology or in forensic science is also utilized, with groups such as carcinogens, illicit drugs, poisons, and other toxins.

Analysis of food and agricultural products plays an important role in everyday life, and as a result, classifications in groups related to these fields are also common. Analyte groups such as nutrients, flavors, toxins in food, and main food components are frequently differentiated. Numerous other groups of compounds not listed previously are common. Among these can be mentioned flavors, polymer additives, solvents, surfactants, and dyes.

Although such classifications as those based on the role of the analyte in everyday life may appear to be unrelated to the HPLC separation technique, this classification is in fact extremely important. Most real-life samples are complex mixtures that must be separated and their components detected. Depending on the group to which they belong, the analytes may be associated in specific ways. The matrix that must be separated from the analytes depends very frequently on the group of the sample classified based on its everyday life criteria. For example, biological samples such as blood/plasma will contain proteins, while food products may contain large amounts of lipids or of carbohydrates. The selection of a separation frequently depends on the matrix that is described by the information regarding the class of sample. Also, the specificity of the separation and sensitivity of the analysis is frequently determined by the type of sample as classified based on its role in everyday life.

8.2. PARAMETERS FOR SOLUTE CHARACTERIZATION IN THE SEPARATION PROCESS

General Aspects

The chemical nature of the solute determines a number of physicochemical parameters that are important for the HPLC separation. These parameters must be known in order to help in the choice of type of chromatography, column, mobile phase, and separation conditions. It should be noticed that solute parameters are predetermined (depending on the sample), while those of the column and mobile phase can be selected to fit the analysis needs. Solute properties are important not only for the choice of column and mobile phase, but also for decisions regarding the whole process of analysis: sampling, sample preparation, selection of the type of detection, and the quantitation procedure (see Section 1.3). It should be underlined that the solutes in a sample include all the species that must be separated, and not only the analytes. For this reason, besides the properties of the analytes, the properties of the matrix components must be considered in an HPLC analysis.

Molecular Weight

The molecular mass of one molecule is its mass expressed in unified atomic mass units (1/12 the mass of one atom of the isotope carbon-12, sometimes named dalton or Da). The relative molecular mass is frequently (and not correctly) indicated as molecular weight (MW) of a molecule. MW is the ratio of the mass of the molecule to 1/12 of the mass of isotope carbon-12 (MW is dimensionless). MW and molecular mass are numerically equal, but they are not the same parameter, although the terms are frequently used interchangeably.

The molecular weight (relative molecular mass) represents an important parameter in molecular characterization. It is related to many other molecular properties, as shown in Chapter 4. Also, a common purpose of MW in HPLC is its use in the differentiation between small molecules and macromolecules. In fact, as a common definition, macromolecules are chemical compounds formed from at least 1000 atoms linked by covalent bonds. However,

instead of number of atoms, a MW higher than about 5000 is commonly used to indicate a macromolecule, and a MW lower than about 2000 to indicate a small molecule. Between these two limits is a gray area where, among others, the molecules known as oligomers are placed.

Acidic or Basic Character of Solutes

A considerable number of molecules contain some specific functional groups likely to lose or gain protons (e.g. in aqueous solutions). As shown in Section 3.5, basic compounds are protonated as positive ions, while acidic compounds are protonated as neutral molecules, the process depending on the pH of the solvation medium. The nonionizable, acidic, basic, or amphoteric character of molecules is important for their separation. Tables with acidity constants are readily available in the literature for common compounds. For more complex molecules, various techniques are available for pK_a calculation. A common procedure for this purpose is based on the calculation of partial charges of atoms in the molecule [2,3]. Computer programs are also available for the calculation of pK_a (e.g., MarvinSketch [4]). For macromolecules with multiple ionizable groups, the individual pK_a values for each functionality are not anymore a relevant parameter.

Isoelectric Point

The isoelectric point (pI) is the pH value at which the molecule carries no electrical charge (in a solution). The concept is particularly important for zwitterionic molecules such as amino acids, peptides, and proteins. For an amino acid, the isoelectric point is the average of pK_a values for the amine and the carboxilic group. In the case of amino acids with multiple ionizable groups (e.g., lysine with two amino groups or aspartic acid with two acid groups), the isoelectric point is given by the average of the two pK_a of the acid and base that lose/gain a proton from the neutral form of the amino acid. This can be extended to the definition of pI of peptides and proteins. The pI value can be used to indicate the global basic or acidic character of a zwitterionic molecule; compounds with pI > 7 can be considered basic, and those with pI < 7 can be considered acidic. For complex molecules such as proteins, isoelectric point is useful in the description of acidic or basic character, where individual pK_a values are not relevant

Molecular Polarity and Octanol/Water Partition Constant

The asymmetrical charge distribution in a molecule (separation of the center of positive charges from that of negative charges), which causes the molecule to act as an electric dipole, is defined as polarity (see Sections 1.1 and 4.1). However, the term *polarity* is frequently used with a wider meaning, including polarity caused by the presence of a dipole moment m_i in the molecule, as well as the molecule polarizability α_i. These two parameters are involved in the calculation of various interactions between solutes, solutes and stationary phase, and solutes and mobile phase molecules. Besides a quantitative characterization of molecules based on m_i and α_i values, it is common to qualitatively assess a polar character by the presence in the molecules of polar functional groups such as -OH, -COOH, -NH_2, >NH, and -SO_3H since these groups bring both dipole moments and polarizability to molecules.

Molecules can also be characterized by the opposite of polar character, which is the hydrophobic character. Octanol/water partition constant K_{ow} is commonly used for characterizing the hydrophobicity of a compound (see rel. 1.1.1). For this reason, K_{ow} values for solutes are important parameters in HPLC, in particular for RP-HPLC, but also for other chromatographic types, such as HILIC. Octanol/water parameters have, besides chromatography,

a widespread utilization in other important fields of science, such as drug design and environmental studies. For this reason, values for K_{ow} for a large body of compounds are experimentally available [5,6], can be calculated using computer programs, for example, MarvinSketch 5.4.0.1, ChemAxon Ltd. [4], EPI Suite [7], and can be calculated using an additive fragment methodology [8]. Various calculation procedures for log K_{ow} are described in the literature [9–13]. The polar and/or hydrophobic character of a molecule can be related to its K_{ow} value. Positive values for log K_{ow} indicate some hydrophobic character, and larger values show more hydrophobicity. Molecules with low or negative values for log K_{ow} are frequently indicated as polar, although there is not a direct relation between K_{ow} and the charge distribution in the molecule.

Experimental K_{ow} (or log K_{ow}) as reported in the literature [6] may be listed for the same compound as having several values (that are close or relatively close). Also, different estimation methods may not always provide identical values or values identical to the experimental ones for K_{ow}. This should not create any significant confusion regarding K_{ow} since large variations in its value for a unique compound are uncommon. The dependence between calculated K_{ow} values using two different computer programs (MarvinSketch [4] and EPI Suite [7]) and K_{ow} experimental values is shown in Figure 8.2.1. For allowing a uniform comparison between compounds, the K_{ow} values used in this book were the calculated ones based on the MarvinSketch 5.4.0.1 program.

Another aspect regarding octanol/water partition is related to the characterization of ionic compounds using a distribution coefficient D_{ow} instead of K_{ow}, as defined in Section 3.5. The distribution coefficient D_{ow} for ionizable species depends on pH, and only for nonionizable compounds $D_{ow} = K_{ow}$. Variation of D_{ow} with the pH can also be obtained using computing programs (e.g., MarvinSketch 5.4.0.1). For the partition processes, the value of D_{ow} better

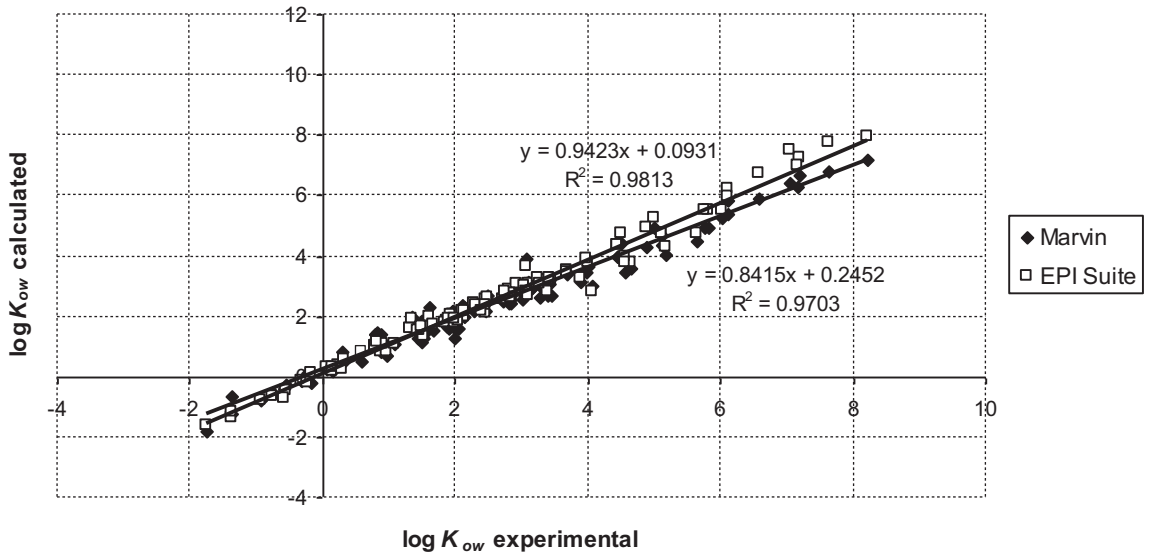

FIGURE 8.2.1 The dependence between the calculated K_{ow} (MarvinSketch [4] and EPI Suite [7]) and experimental K_{ow} values.

describes the affinity of a compound to one of the phases involved in partition.

The values of log K_{ow} are very useful for characterization of polarity in small molecules. However, the same technique cannot be applied for characterization of the polarity of polymers, where the log K_{ow} values lose their meaning. Proteins, for example, have the capability of folding, and the polar groups have the tendency to congregate in such a manner as to maximize electrostatic interactions. In a polar solvent like water or aqueous solutions of acids, the protein may change its tertiary and quaternary structure and expose polar side chains toward the solvent, the hydrophobic moieties being congregated toward a more hydrophobic core. The opposite effect may take place in the presence of organic solvents. This behavior would lead to a variable octanol/water partition.

Solubility in Water from Octanol/Water Partition Constant

Water solubility is an important parameter since mobile phases usually contain water. The evaluation of solubility of different organic compounds in water based on thermodynamic considerations is typically much less precise compared to the estimation based on the values of parameter K_{ow}. Values for the octanol/water partition constant K_{ow} are readily available. The estimation of solubility of organics in water from K_{ow} values is typically done using regression equations of the form:

$$\log C_i = a \log K_{ow} + b \tag{8.2.1}$$

or of the form:

$$\log (1/C_i) = a' \log K_{ow} - b' \tag{8.2.2}$$

where C_i is the solubility in water and a, a', b, b' are empirical parameters. Some values for these parameters are given in Table 8.2.1 [14].

Van der Waals Molecular Volume and Surface

Van der Waals molecular volume $\mathcal{V}$ and van der Waals surface $\mathcal{A}$ are important parameters for understanding molecular interactions

TABLE 8.2.1 Regression Equation for the Estimation of Water Solubility Based on log K_{ow} [14]

Equation	Units for C_i	Used for chemical class
$\log C_i = -1.37 \log K_{ow} + 7.26$	μM/L	Aromatics, chlorinated hydrocarbons
$\log (1/C_i) = 1.113 \log K_{ow} - 0.926$	mol/L	Alcohols
$\log (1/C_i) = 1.229 \log K_{ow} - 0.720$	mol/L	Ketones
$\log (1/C_i) = 1.013 \log K_{ow} - 0.520$	mol/L	Esters
$\log (1/C_i) = 1.182 \log K_{ow} - 0.935$	mol/L	Ethers
$\log (1/C_i) = 1.221 \log K_{ow} - 0.832$	mol/L	Alkyl halides
$\log (1/C_i) = 1.294 \log K_{ow} - 1.043$	mol/L	Alkynes
$\log (1/C_i) = 1.294 \log K_{ow} - 0.248$	mol/L	Alkenes
$\log (1/C_i) = 0.966 \log K_{ow} - 0.339$	mol/L	Aromatics
$\log (1/C_i) = 1.214 \log K_{ow} - 0.850$	mol/L	Various
$\log (1/C_i) = 1.237 \log K_{ow} + 0.248$	mol/L	Alkanes

TABLE 8.2.2 Van der Waals Radii r for Several Elements in Å [15]

H 1.20						
Li 1.82	Be 1.53	B 1.92	C 1.70	N 1.55	O 1.52	F 1.47
Na 2.27	Mg 1.73	Al 1.84	Si 2.10	P 1.80	S 1.80	Cl 1.75
K 2.75	Ca 2.31	Ga 1.87	Ge 2.11	As 1.85	Se 1.90	Br 1.85
Rb 3.03	Sr 2.49	In 1.93	Sn 2.17	Sb 2.06	Te 2.06	I 1.98

during a separation. In RP-HPLC in particular, van der Waals molecular area $\mathcal{A}$ plays an important role in the expression of the free energy for the partition equilibrium and directly to the value of capacity factor for an analyte, as shown in Section 5.1 (see rel. 5.1.26). The calculation of both van der Waals molecular volume and area starts with the concept of van der Waals radius of an atom. This is the radius of an imaginary sphere used to model the atoms describing its finite size. The radius is obtained based on results from gas kinetic collision cross sections, gas critical volumes, crystal densities (extrapolated at 0 K), liquid state properties, X-ray diffraction data, and the like [15,16]. Several "mean" van der Waals radii (in Å) are indicated in Table 8.2.2.

For most organic molecules, the atoms are placed at covalent-bond distance, which is shorter than the sum of the van der Waals radii of the connected atoms. For this reason, the van der Waals molecular volume is smaller than the sum of the volume of each component atom. This is also true for the molecular surface. For a diatomic molecule with two atoms with r_1 and r_2 van der Waals radii, and the covalent-bond distance l, the calculation of the molecular volume is obtained starting with the two interlocked spheres as shown in Figure 8.2.2.

The volume of one sphere $\mathcal{V}_2$ can be calculated as $\mathcal{V}_2 = (4/3)\pi\,(r_2)^3$. The volume of a spherical segment $\Delta\mathcal{V}_1 = \pi\, h_1^2(r_1 - h_1/3)$ must be added to $\mathcal{V}_2$, and the volume of another spherical segment with $\Delta\mathcal{V}_{2\text{-}1} = \pi\, h_2^2(r_2 - h_2/3)$ must be subtracted from $\mathcal{V}_2$ to obtain the total volume. The values for h_1 and h_2 are obtained as $h_1 = r_1 + l - m$, $h_2 = r_2 - m$ and $m = (r_2^2 - r_1^2 + l^2)\ /\ (2l)$. The total volume will be $\mathcal{V} = \mathcal{V}_2 + \Delta\mathcal{V}_1 - \Delta\mathcal{V}_{2\text{-}1}$. *Note:* The volume of a molecule $v_j = V_j/\mathcal{N}$ (where V is the molar volume $V = \mathrm{MW}/\rho$) is not the same as van der Waals volume $\mathcal{V}$ since the volume of voids and the changes in the volume due to interactions are included in the value for v.

The area of the molecule can be obtained as the sum of two spherical segments with $\mathcal{A}_1 = 2\,\pi\, r_1\, h_1$, and $\mathcal{A}_2 = 2\,\pi\, r_2\,(2r_2 - h_2)$. Since direct calculations of van der Waals molecular volumes and areas are difficult for larger molecules, various approximations were developed for this purpose [17]. Computer programs are available for calculating van der Waals volume and areas (e.g., MarvinSketch [4]). (Molar volume V can also be obtained using computer programs [18])

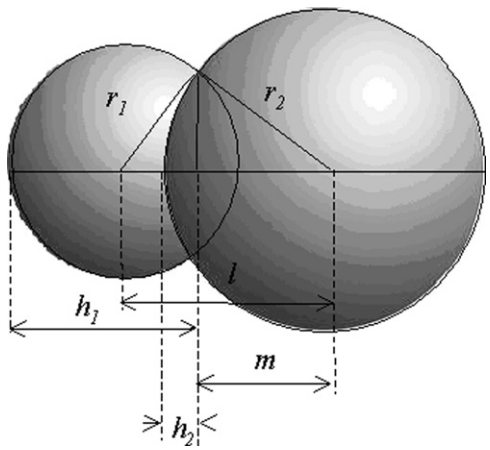

FIGURE 8.2.2 Description of the calculation of the volume of a diatomic molecule.

Correlation between van der Waals Molecular Surface and Octanol/Water Partition Constant

An expression for the octanol/water partition constant K_{ow} can be obtained based on the solvophobic theory developed for RP-HPLC in Section 5.1. For liquid-liquid partition the compound j is distributed between two nonmiscible liquid phases A and B in an equilibrium of the type:

$$j_B \rightleftharpoons j_A \tag{8.2.5}$$

When equilibrium is attained for "distributing" the compound j between phases A and B, the difference between the chemical potentials $\mu_{i,A}$ and $\mu_{i,B}$ of the component j in each of the two phases A and B must be zero. As shown in Section 3.1, this condition (expressed by rel. 3.1.2) is equivalent to the following expression for the equilibrium constant (see rel. 3.1.12):

$$K_j = \exp[-\Delta G_j^0/(RT)] \tag{8.2.6}$$

Assuming that no volume changes occur during the process, the free enthalpy ΔG^0 is taken as equal to the free energy of the process ΔA^0, and rel. 8.2.6 can be written in the form:

$$\begin{aligned} K_j &= \exp\left[-\Delta A_j^0/(RT)\right] \\ &= \exp\left[-(A_{j,A} - A_{j,B})/(RT)\right] \end{aligned} \tag{8.2.7}$$

In rel. 8.2.7, the value for the free energy $A_{j,S}$ (where $S = A$ or B) is given by the change in standard free energy ${A^{sol}}_{j,S}$ necessary for placing a molecular species j into a solution formed by molecules S (symbol Δ and the index "^{0}" for standard expressions are omitted). This energy is given by rel. 4.1.58. In Section 4.1 it was shown that the expression for ${A^{sol}}_{j,S}$ is the following:

$$\begin{aligned} {A^{sol}}_{j,S} &= {A^{cav}}_{j,S} + {A^{es}}_{j,S} + {A^{disp}}_{j,S} \\ &\quad + RT \ln(RT/p_0 V_S) \end{aligned} \tag{8.2.8}$$

where the expressions for the terms in rel. 8.2.8 are given by rel. 4.1.74, 4.1.75, and 4.1.76, respectively. With rel. 8.2.8 used in the expression for ln K given by rel. 8.2.7, the following formula is obtained:

$$\begin{aligned} RT \ln K_j &= ({A^{cav}}_{j,B} - {A^{cav}}_{j,A}) + ({A^{es}}_{j,B} - {A^{es}}_{j,A}) \\ &\quad + ({A^{disp}}_{j,B} - {A^{disp}}_{j,A}) \\ &\quad + RT \ln (V_A/V_B) \end{aligned} \tag{8.2.9}$$

Each set of terms in rel. 8.2.9 can be further estimated based on the solvophobic theory developed in Section 4.1. For the difference in the free energy for the cavity formation, making the assumption that $W_S \approx 1$, the following formula is obtained:

$$\begin{aligned} {A^{cav}}_{j,B} - {A^{cav}}_{j,A} &= \mathcal{N}\,(\gamma\prime_B - \gamma\prime_A)\mathcal{A}_j + \mathcal{N}\Big[(\kappa^e{}_B \\ &\quad - 1)\,\gamma\prime_B (V_B)^{2/3} - (\kappa^e{}_A \\ &\quad - 1)\gamma\prime_A (V_A)^{2/3}\Big]\mathcal{A}_j/(V_j)^{2/3} \end{aligned} \tag{8.2.10}$$

The estimation in the difference in electrostatic forces based on rel. 4.1.64 gives the expression:

$${A^{es}}_{j,B} - {A^{es}}_{j,A} = -\frac{\mathcal{N} m_j^2}{2\nu_j}(\mathcal{D}_B \mathcal{P}_{j,B} - \mathcal{D}_A \mathcal{P}_{j,A}) \tag{8.2.11}$$

The estimation of the differences in the dispersion forces obtained from rel. 4.1.66 using the assumption that $Q''_{j,S} \approx 0.1\ Q'_{j,S}$ gives:

$$\begin{aligned} {A^{disp}}_{j,B} - {A^{disp}}_{j,A} &= -\frac{16.75\ D_j}{8\pi}(Q\prime_{j,B} \Upsilon_{j,B} D_B \\ &\quad - Q\prime_{j,A} \Upsilon_{j,A} D_A) \end{aligned} \tag{8.2.12}$$

With the terms given by rel. 8.2.10, 8.2.11, and 8.2.12, it can be seen that the expression of the equilibrium constant should have the following form:

$$\log K_j = a\,\mathcal{A}_j + b(V_j)^{-2/3}\mathcal{A}_j + c \tag{8.2.13}$$

where the values for a and b depend only on the solvents A and B, while the value for c depends on the solvents and the solute as well. For the two solvents A = octanol and B = water, it can be concluded that a and b are the same for all analytes j. Since in this case $\gamma'_B - \gamma'_A > 0$, and also $(\kappa^e_B - 1)\,\gamma'_B\,(V_B)^{2/3} - (\kappa^e_A - 1)\,\gamma'_A\,(V_A)^{2/3} > 0$, from rel. 8.2.10 it can be seen that a larger $\mathcal{A}_j$ will generate a larger K_{ow}.

The term c in rel. 8.2.13 includes the contribution of electrostatic and the dispersion forces to the free energy. It is therefore expected that more polar analytes will have a negative c, larger in absolute value. Predictions of this theory were very clearly verified when plotting log K_{ow} values as a function of $\mathcal{A}_j$ for a number of compounds. This type of plot is exemplified in Figure 8.2.3 for several classes of mono-functional compounds. The results from Figure 8.2.3 show that rel. 8.2.13 can be well approximated by a relation of the form:

$$\log K_{j,ow} = a'\mathcal{A}_j + c \tag{8.2.14}$$

where a' is a constant (for $\mathcal{A}_j$ in Å^2, $a' \approx 1.46\ 10^{-2}$ with K_{ow} calculated based on MarinSketch; other values are $a' \approx 1.57\ 10^{-2}$ with K_{ow} calculated based on EPI Suite, and $a' \approx 1.63\ 10^{-2}$ with experimental K_{ow}). The value for c depends only on the nature of the substituent (polar or nonpolar) on the hydrophobic moiety of the compound j. Further investigation of rel. 8.2.14 indicates that it can be extended with excellent agreement to the following formula:

$$\log K_{i,ow} = a'\mathcal{A}_i + \sum_n c_n \tag{8.2.15}$$

where c_n are constants for each substituent. Values for c_n for different organic groups attached to an aliphatic hydrocarbon chain are given in Table 8.2.3 [19].

The calculation of log K_{ow} based on rel. 8.2.15 with the values for c indicated in Table 8.2.3, for a set of 147 compounds, including mono, bi, tri,

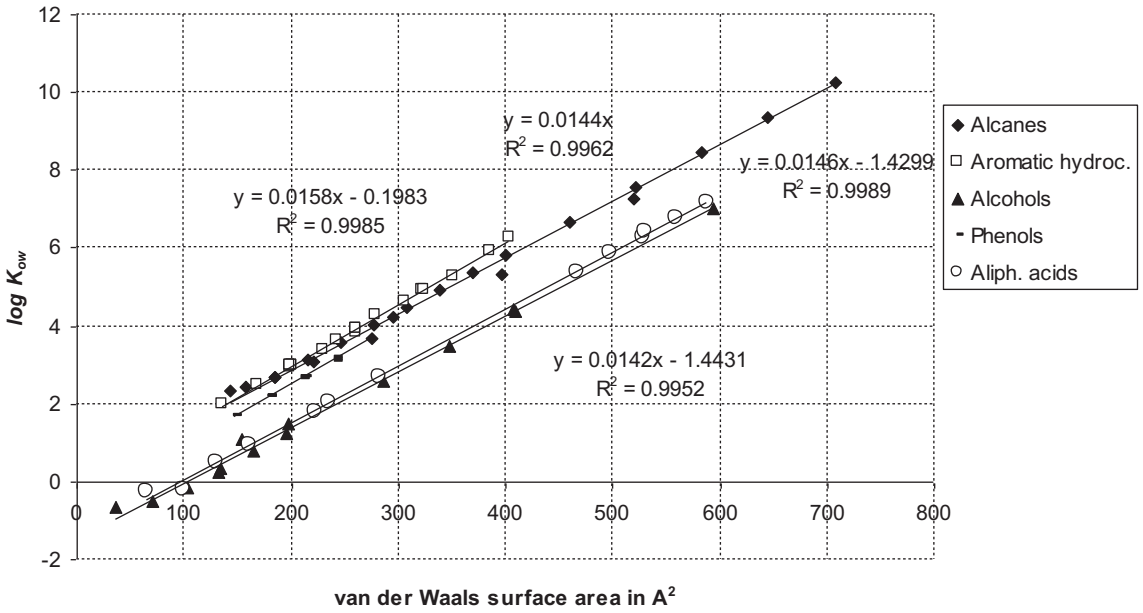

FIGURE 8.2.3 Dependence of log K_{ow} on van der Waals molecular surface $\mathcal{A}_j$ for different classes of mono-functional compounds.

TABLE 8.2.3 Values for Constant c in the Calculation of log K_{ow} (MarvinSketch Values) from van der Waals Surface Area of the Molecule* ($a' \approx 1.46\ 10^{-2}$).

Group	c_n	Group	c_n
Aromatic ring**	0.055	Aliphatic secondary amine	−1.897
Alcohol	−1.444	Aliphatic tertiary amine	−2.200
Phenol	−0.470	Aromatic primary amine	−0.998
Aliphatic ether	−1.581	Aromatic secondary amine	−1.320
Aromatic/aliphatic ether	−0.825	Ketone	−1.512
Aliphatic acid	−1.375	Nitro aromatic	−0.612
Aromatic acid	−0.852	Chloro aromatic	0.357
Aliphatic primary amine	−1.650	Bromo aromatic	0.481
Aromatic ester	−1.100	Nitrile	−1.072

** Note: each group is counted as many times as present.*
*** Note: Separated rings are counted individually, condensed rings are counted as one.*

tetra, and pentafunctional ones, with identical or different functional groups, produced an excellent agreement with log K_{ow} target values obtained using the MarvinSketch program (for which parameter c was optimized). This is shown in Figure 8.2.4 where the excellent correlation with $R^2 = 0.9961$ is displayed.

Optimization of c_n paramaters based on log K_{ow} values calculated with EPI Suite or with experimental values leads to values close (but

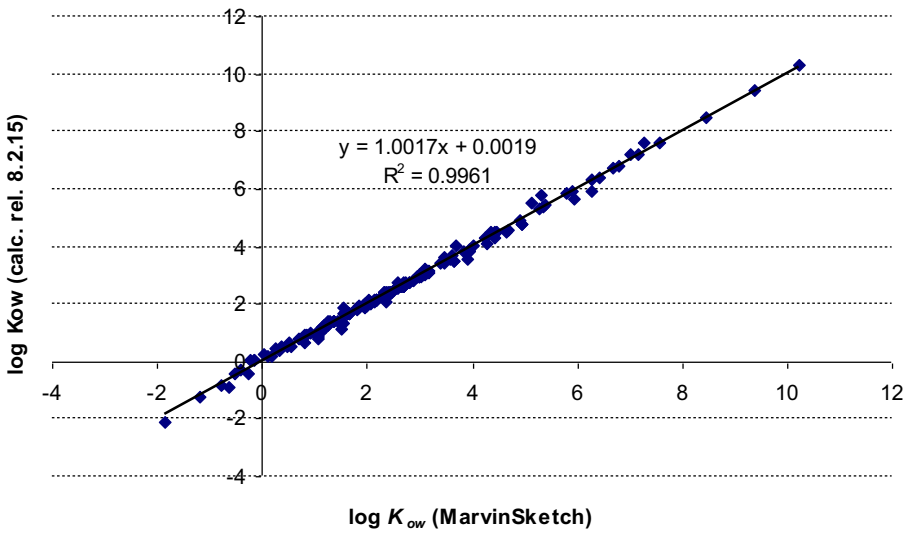

FIGURE 8.2.4 Correlation between log K_{ow} calculated and target values obtained using MarvinSketch. Calculation based on rel. 8.2.15 with $a' = 1.46\ 10^{-2}$ and the values for c_n indicated in Table 8.2.3.

not identical) to those from Table 8.2.3, and also to very good correlations between calculated values and the target values.

Based on the theory developed for the evaluation of log K_{ow} the expression for c_n should be given by:

$$c_n = [(A^{es}{}_{j,B} - A^{es}{}_{j,A}) + (A^{disp}{}_{j,B} - A^{disp}{}_{j,A})]/(2.303\,RT) + \log\,(V_A/V_B) \quad (8.2.16)$$

where A = octanol and B = water. Relation 8.2.16 indicates that the value for c can provide useful information regarding the electrostatic and dispersion forces between the molecule of the analyte and the two solvents.

Partial Charge Distribution

Analysis of the electronic population is a subject of considerable interest for understanding molecular properties (see, e.g., [20,21]). The partial charge distribution determines physicochemical properties such as dipole moment m, ionization constants, and reactivity. Various approaches and procedures are reported in the literature for calculating partial charges [21–23]. Computer programs are also available for partial charges calculation (e.g,. MarvinSketch [4,24,25]). The values for partial charges (point charges) are useful for understanding molecular polarity, although they are not used frequently in specific calculations related to separations. The potential application of this parameter could be in those LC separations based on electrostatic interactions, such as HILIC, NP, ANP, and the chiral mechanism.

Parameters for Solute Hydrophobicity, Steric Effects, Hydrogen Bonding, and Ion-Exchange Character

In the development of certain models for the characterization of stationary phases, it was necessary to include parameters for characterization of solutes (analytes), with which the stationary phase interacts. For example, in Section 6.4 it was shown that the capacity factor for a compound j is given by the expression (see rel. 6.4.15):

$$\log k(j) = \log k_{EB} + \eta'(j)H_{c1} - \sigma'(j)S^*_{c1} + \beta'(j)A_{c1} + \alpha'(j)B_{c1} + \kappa'(j)C_{c1} \quad (8.2.17)$$

where $\log k_{EB}$ is the capacity factor for ethylbenzene, parameters H_{c1}, S^*_{c1}, A_{c1}, B_{c1}, and C_{c1} describe the column characteristics. The corresponding parameters $\eta'(j)$, $\sigma'(j)$, $\beta'(j)$, $\alpha'(j)$, and $\kappa'(j)$ describe solute properties: $\eta'(j)$-hydrophobicity, $\sigma'(j)$-steric interactions, $\beta'(j)$-hydrogen-bonding acceptance, $\alpha'(j)$-hydrogen-bonding donation, and $\kappa'(j)$-cation exchange or ion-ion interactions. The values of these parameters would be very useful for solute characterization in HPLC, and some were reported in the literature for selected analytes [26–29]. However, these parameters were obtained for a specific mobile phase and were established with the specific purpose of column characterization. They were obtained from best-fit regression lines and are not readily available for solutes that were not in the "test" group.

Molar Volume, Hildebrand Solubility Parameter

In several expressions for the calculation or estimation of capacity factor k and therefore of α, the molar volume V and the solubility parameter δ for the solvent, the stationary phase, and the analyte were involved. The discussion regarding these parameters as given for the solvent molecule is applicable for the solute molecule. The calculation of the molar volume V and of δ was discussed in Section 7.1.

Capacity factor k can be calculated in principle by knowing Ψ for the stationary phase

and K_i for the solute i ($k_i = K_i\,\Psi$). Relation 7.1.17 shows, for example, how the equilibrium constant K_i depends on δ_i and V_i, and on solvent and stationary-phase properties. However, the direct calculation of a specific retention factor k based on that formula or of a selectivity α based on rel. 7.1.18 does not generate accurate values and can be useful only as a directional information.

Evaluation of Solubility of Nonpolar Compounds from Hildebrand Solubility Parameter

The dissolution process of nonelectrolytes can be viewed hypothetically as being formed from two steps. The first step is melting of the compound i to form a supercooled liquid (a liquid cooled below its freezing point without solidifying). The second step is the mixing of this liquid with the solvent S. For a transformation at constant pressure and temperature at equilibrium, the free enthalpy $\Delta G = 0$, and this holds true for melting at fusion temperature T_f. Therefore, the free enthalpy of melting (fusion) is $\Delta G_f = 0$ and $\Delta \boldsymbol{H}_f = T_f\,\Delta S_f$, where $\Delta \boldsymbol{H}_f$ is the heat of fusion and ΔS_f is the entropy of fusion. At a given temperature T, the free energy of melting will be given by the expression:

$$\Delta G_f \;=\; \Delta \boldsymbol{H}_f - T\Delta S_f \;=\; \Delta \boldsymbol{H}_f - (T/T_f)\Delta \boldsymbol{H}_f \tag{8.2.18}$$

The free energy of mixing for ideal solutions does not take place with any heat change, and therefore:

$$\Delta G^{mix} \;=\; -T\,\Delta S^{mix} \tag{8.2.19}$$

The ideal behavior of a solution assumes that each component has its molecules behaving as if it is surrounded by molecules of the same kind (to follow Raoult law). The variation in the entropy of mixing ΔS^{mix} of ideal solutions is given by the expression [30]:

$$\Delta S^{mix} \;=\; -R\sum_j n_j \ln x_j \qquad \left(\text{with } x_i \;=\; n_i \Big/ \sum_{j=1}^{r} n_j\right) \tag{8.2.20}$$

For the dissolution process of i in S, it can be assumed that $x_S = 1$ for the solvent, and therefore $\Delta G^{mix} = -\,RT\;n_i\;\ln\;x_i$, (the index S from $x_{i,S}$ and $n_{i,S}$ is omitted), or for $n_i = 1$ the expression for ΔG^{mix} becomes:

$$\Delta G^{mix} \;=\; -RT\ln x_i \tag{8.2.21}$$

The total free enthalpy of dissolution $\Delta G = -\,\Delta G_f + \Delta G^{mix}$, and at the equilibrium $\Delta G = 0$. Including expressions 8.2.18 and 8.2.19 in the formula for ΔG leads to the expression:

$$(1 - T/T_f)\Delta \boldsymbol{H}_f + RT\ln x_i \;=\; 0, \tag{8.2.22}$$

which can be rearranged in the form:

$$\ln x_i \;=\; [\Delta \boldsymbol{H}_f(1/T_f - 1/T)]/R \tag{8.2.23}$$

Relation 8.2.23 gives the formula for solubility (expressed as the maximum mole fraction of a solute in the solvent) for the formation of an ideal solution. As shown by rel. 8.2.23, the solubility increases with the temperature T. This is true for many compounds, including both nonelectrolytes and electrolytes, although there are exceptions.

In rel. 8.2.23 the mole fraction x_i can be changed into molar concentration c_i using the approximation:

$$x_i \;=\; \frac{n_i}{n_i + n_S} \approx \frac{n_i}{n_S} \;=\; \gamma_i c_i \frac{M_S}{1000\rho_S} \tag{8.2.24}$$

where γ_i is the activity coefficient included to correct for the deviation from ideal solutions, M_S is the molecular mass (weight) MW of the

solvent, and ρ_S is its density. In this case, rel. 8.2.23 can be written in the form [31]:

$$\ln \frac{\gamma_i c_i M_S}{1000\rho_S} = \frac{\Delta H_f}{RT}\frac{T - T_f}{T_f} \tag{8.2.25}$$

From rel. 8.2.25, the solubility for nonelectrolytes can be estimated based on solubility parameters δ, using for γ_i its expression given by rel. 7.1.12. Rearranging rel. 8.2.25 and using rel 7.1.12 for ln γ_i, the expression for ln c_i is given by the formula:

$$\ln c_i = \frac{\Delta H_f}{RT}\frac{T - T_f}{T_f} - \frac{V_i}{RT}(\delta_i - \delta_S)^2 + \ln \frac{1000\rho_S}{M_S} \tag{8.2.26}$$

Considering $\Delta H_f \approx T_f \Delta S_f$, and taking $\Delta S_f \approx 13$ cal mole^{-1} deg^{-1} (the fusion entropy being relatively constant for many compounds), rel. 8.2.26 can be estimated by the formula:

$$\ln c_i = 6.54\frac{T - T_f}{T_f} - \frac{V_i}{RT}(\delta_i - \delta_S)^2 + \ln \frac{1000\rho_S}{M_S} \tag{8.2.27}$$

Relation 8.2.27 is useful for evaluation of a solid nonelectrolyte solubility into a solvent.

Solvatochromic Parameters for the Solute

Solvatochromic parameters for the solute can be developed in the same manner as for a solvent (see Section 7.1). Parameters α, β, and π^* for a number of compounds were listed in Table 7.4.1, and more such parameters are available in the literature [32,33]. However, except for a limited number of analytes, estimation of these parameters is difficult or impossible. This reduces the utility of formulas using such parameters to only directional information and comparison for analogous cases where the solvent is changed while the analytes and the stationary phase remain the same. Accurate calculations, for example, of the capacity factor for a specific analyte, are not typically successful.

Other Solute Properties Affecting Separation

Depending on the type of HPLC analysis, specific physical properties of the solute play various roles in the separation process. A short enumeration of some of these properties used for calculating the capacity factor is as follows for the main HPLC types.

1) For RP-HPLC, the solvophobic theory shows that the capacity factor for an analyte is expressed by rel. 5.1.26. In that expression, all terms depend on analyte characteristics except for log Ψ and the term that accounts for the change in the free volume of the system during the retention process. Rel 5.1.26 includes the term representing the interactions in an ideal gas system consisting of solute molecule j and the ligand L of the stationary phase $E_T(j, L)$, the terms describing the dispersion and electrostatic forces between the molecule j and the molecules of the surrounding solvent S, a term involving polar/polarizability interactions, and two terms related to the energy necessary for cavity formation by the analyte in the mobile phase. The interaction between the solute j and ligand L in the ideal gas phase is related to the energy of intermolecular interactions (as given by rel. 4.1.46) where the dipole moment m_j, polarizability $\alpha_{0,j}$, and ionization potential I_j are implicated. Indirectly, the molecular dimensions are also a factor in the expression of $E_T(j, L)$. The same parameters of the solute m_j, $\alpha_{0,j}$, and I_j, plus molar volume V_j and molecular diameter d_j are important for the term describing the dispersion and electrostatic forces with the surrounding solvent. The terms related to the energy necessary for cavity formation in

the mobile phase are related in particular to the van der Waals solute area $\mathcal{A}_j$.

The detailed contribution of each of these parameters is rather difficult to assess, since most interactions take place competitively between the solute and the stationary phase on one hand, and between the solute and the mobile phase on the other. Only general comments can be made indicating that higher dipole moment and polarizability are conducing to stronger polar interactions, and that a larger van der Waals area $\mathcal{A}_j$ of the solute molecule is an indication of stronger hydrophobic interactions.

2) In ion-pair chromatography (IP) the theory of separation is basically expressed by rel. 5.2.30 (see Section 5.2). According to this formula, the capacity factor (and therefore the separation) in IP is determined by a capacity factor $k_j(0)$ in the absence of IPA (hetaeron) and by K_{IPA}, and by the charge and concentration of IPA. The values for $k_j(0)$ depend on the same parameters as in RP-HPLC.

3) In polarity-based separations, basically the same factors as in RP-HPLC determine the capacity factor of analyte. As discussed in Section 5.3, the weight of each term in the expression of the change in free energy of the separation process, and therefore in the expression of the capacity factor, is different from that in RP-HPLC, with the polar interactions playing a more important role. For this reason, the more polar compounds are more strongly retained than the less polar ones. Similar to the case of RP-HPLC, the direct calculation of capacity factors is not a practical way for assessing solute (analyte) behavior in the separation process.

4) In ion-exchange chromatography (IC), the separation process is described by equation 5.4.7, which indicates that the exchange equilibrium depends on an osmotic term, the charges of the ions, and the activity coefficients of each species. Again, since evaluation of the terms involved in equation 5.4.7 is difficult (e.g., the activity coefficient of an ion in the resin), practical observations are used for predicting the separation, such as the knowledge that cations typically separate in the order single charge (Na^+, K^+, NH_4^+, etc.) followed by double charge (Mg^{2+}, Ca^{2+}, etc.), while small anions typically follow the sequence F^-, Cl^-, Br^-, NO_3^-, SO_4^{2-}, $Cr_2O_7^{2-}$. Organic anions such as small organic acids can also be separated using IC, specific methods for separation being reported in the literature (see, e.g., [34]).

5) In size exclusion, the physical dimensions of the molecule are important for characterization of the separation (see rel. 5.6.1 and 5.6.2). In this separation technique, the hydrodynamic volume of the solute (or the effective mean radius) is involved in calculating separation parameters.

8.3. OTHER PARAMETERS FOR SOLUTE CHARACTERIZATION

General Aspects

Various properties of the solute molecule are involved in the chromatographic process, some of them, in particular being those that are directly related to the separation process (discussed in Section 8.1). In addition to those, other properties are of importance. One such property is the diffusion coefficient D_j of the analyte. As shown in rel. 2.2.2, the longitudinal diffusion in an HPLC column is proportional with D_j. Different theories and also empirical relations were developed for estimation of diffusion coefficients. For example, Stokes theory empirically modified for better prediction gives the following formula for the diffusion coefficient for nonelectrolytes j in liquids B:

$$D_{j,B} = 7.4 \cdot 10^{-8} \frac{(\psi_B M_B)^{0.5} T}{\eta V_j^{0.6}} \qquad (8.3.1)$$

where V_j is the molar volume of solute j (in cm^3 $mole^{-1}$), M_B is the molecular weight of solvent B, T is temperature in Kelvin degrees, η the viscosity of the solution (in 10^{-4} g cm^{-1} s^{-1}, or centipoise), and ψ_B an "association" factor for the solvent (ψ_B is 1 for nonpolar solvents, 1.5 for ethanol, 1.9 for methanol, 2.6 for water).

Other formulas describing properties of the solutes involve parameters such as the refractive index (see rel. 4.1.68), diameter of the molecule, solvent-accessible area SASA, Kihara parameter [35], and rate of hydrolysis. These parameters can be either found in the literature or estimated [14].

Physical Properties of the Analyte Determining Detection

The choice of detection in HPLC is determined by a particular physicochemical property of the analyte that allows its measurement at very low levels. The selected property must have a value significantly different from that of the mobile phase. Examples of such properties include UV absorption, refractive index, fluorescence, molecular mass, and fragmentation in a mass spectrometer. The subject of properties determining the choice of detection in HPLC is vast and beyond the purpose of this book. Such properties are typically discussed in detail in each individual method of analysis.

References

[1] Kowalska S, Krupczyńska K, Buszewski B. Some remarks on characterization of application of stationary phases for RP-HPLC determination of biologically important compounds. Biomed. Chromatogr. 2006;20:4–22.

[2] Clarke FH, Cahoon NM. Ionization constants by curve fitting: Determination of partition and distribution coefficients of acids and bases and their ions. J. Pharm. Sci. 1987;76:611–20.

[3] Dixon SL, Jurs PC. Estimation of pK_a for organic oxyacids using calculated atomic charges. J. Comput. Chem. 1993;14:1460–7.

[4] http://www.chemaxon.com

[5] Hansch C, Leo A. Exploring QSAR, Fundamentals and Applications in Chemistry and Biology. Washington, DC: ACS; 1995.

[6] Hansch C, Leo A, Hoekman D. Exploring QSAR, Hydrophobic, Electronic and Steric Constants. Washington, DC: ACS; 1995.

[7] http://www.epa.gov/oppt/exposure/pubs/episuite.htm

[8] Moldoveanu SC, David V. Sample Preparation in Chromatography. Amsterdam: Elsevier; 2002.

[9] Klopman G, Li J-Y, Wang S, Dimayuga M. Computer automated log P calculations based on an extended group of contribution approach. J. Chem. Inf. Comput. Sci. 1994;34:752–81.

[10] Viswanadhan VN, Ghose AK, Revankar GR, Robins RK. Atomic physicochemical parameters for three dimensional structure directed quantitative structure-activity relationships IV. Additional parameters for hydrophobic and dispersive interactions and their application for an automated superposition of certain naturally occurring nucleoside antibiotics. J. Chem. Inf. Comput. Sci. 1989;29:163–72.

[11] Csizmadia F, Tsantili-Kakoulidou A, Panderi I, Darvas F. Prediction of distribution coefficient from structure. 1. Estimation method. J. Pharm. Sci. 1997;86:865–71.

[12] Erős D, Kövesdi I, Őrfi L, Takács-Novák K, Acsády G, Kéri G. Reliability of log P predictions based on calculated molecular descriptors. A critical review,. Curr. Med. Chem. 2002;9:1819–29.

[13] Iwase K, Komatsu K, Hirono S, Nakagawa S, Moriguchi I. Estimation of hydrophobicity based on the solvent-accessible surface area of molecules. Chem. Pharm. Bull. 1985;33:2114–21.

[14] Lyman WJ, Reehl WF, Rosenblatt DH. Handbook of Chemical Property Estimation Methods. Washington, DC: ACS; 1990.

[15] Bondi A. Van der Waals volumes and radii. J. Phys. Chem. 1964;68:441–51.

[16] http://www.webelements.com/periodicity/van_der_waals_radius/

[17] Zhao YH, Abraham MH, Zissimos AM. Fast calculation of van der Waals volume as a sum of atomic and bond contributions and its application to drug compounds. J. Org. Chem. 2003;68:7368–73.

[18] http://www.aim.env.uea.ac.uk/aim/density/density.php

[19] Moldoveanu SC, David V. unpublished results

[20] Mulliken RS. Electronic population analysis on LCAO MO molecular wave function. I, J. Chem. Phys. 1955;23:1833–41.

[21] Moldoveanu SC, Savin A. Aplicatii in Chimie ale Metodelor Semiempirice de Orbitali Moleculari, Edit. Academiei RSR, Bucuresti 1980.
[22] Heinz H, Suter UW. Atomic charges for classical simulations of polar systems. J. Phys. Chem. B 2004;108:18341–52.
[23] Rappe AK, Goddard III WA. Charge equilibration for molecular dynamics simulations. J. Phys. Chem. 1991;95:3358–63.
[24] Stewart JJP. MOPAC-7, QCPE 113. Bloomington: Indiana University; 1994.
[25] Frisch MJ, Frisch A, Foresman JB. Gaussian 94. Pittsburgh: Gaussian Inc.; 1995.
[26] Gilroy JL, Dolan JW, Snyder LR. Column selectivity in reversed-phase liquid chromatography IV. Type-B alkyl-silica columns. J. Chromatogr. A 2003;1000: 757–78.
[27] Marchand DH, Snyder LR, Dolan JW. Characterization and applications of reversed-phase column selectivity based on the hydrophobic-subtraction model. J. Chromatogr. A 2008;1191:2–20.
[28] Marchand DH, Croes K, Dolan JW, Snyder LR, Henry RA, Kallury KMR, Waite S, Carr PW. Column selectivity in reversed-phase liquid chromatography. VIII. Phenylalkyl and fluoro-substituted columns. J. Chromatogr. A 2005;1062:65–78.
[29] Bidlingmeyer B, Chan CC, Fastino P, Henry R, Koerner P, Maule AT, Marques MRC, Neue U, Ng L, Pappa H, Sander L, Santasania C, Snyder L, Woznyak T. HPLC column classification. Pharmacopeial Forum 2005;31:637–45.
[30] Hildebrand JH, Scott RI. The Solubility of Non-Electrolytes. York: Dover, New; 1964.
[31] Gmehling JG, Anderson TF, Prausnitz JM. Solid-liquid equilibria using UNIFAC. Ind. Eng. Chem. Fundam. 1978;17:269–73.
[32] Kamlet MJ, Abboud JLM, Abraham MH, Taft RW. Linear solvation energy relationships. 23. A comprehensive collection of the solvatochromic parameters, π^*, α and β, and some methods for simplifying the generalized solvatochromic equation. J. Org. Chem. 1983;48:2877–87.
[33] http://www.stenutz.eu/chem/solv26.php
[34] Roger WJ, Michaux S, Bastin M, Bucheli P. Changes to the content of sugars, sugar alcohols, myo-inositol, carboxylic acids and inorganic anions in developing grains from different varieties of Robusta (*Coffea canephora*) and Arabica (*C. arabica*) coffees. Plant Sci. 1999;149:115–23.
[35] Tee LS, Gotoh S, Stewart WE. Molecular parameters for normal fluids. Ind. Eng. Chem. Fundamen. 1966;5:363–7.