# Insights into chiral recognition mechanisms in supercritical fluid chromatography. I. Non-enantiospecific interactions contributing to the retention on tris-(3,5-dimethylphenylcarbamate) amylose and cellulose stationary phases 

Caroline West ${ }^{\text {a }, *}$, Yingru Zhang ${ }^{\mathrm{b}}$, Luc Morin-Allory ${ }^{\text {a }}$<br>${ }^{\text {a }}$ Institut de Chimie Organique et Analytique (ICOA), Université d'Orléans, CNRS UMR 6005, B.P. 6759, rue de Chartres, 45067 Orléans Cedex 2, France<br>${ }^{\mathrm{b}}$ Bioanalytical and Discovery Analytical Sciences, Bristol-Myers Squibb Company, PO Box 4000, Princeton, NJ 08648-4000, USA

## A R T I C L E I N F O

## Article history:

Available online 8 December 2010

## Keywords:

Chiral stationary phases
Supercritical fluid chromatography
Polysaccharide
Enantiomer separation
Retention mechanism
Solvation parameter model
Quantitative structure-retention relationships


#### Abstract

In this series of papers, we use a systematic approach to investigate the factors responsible for enantiorecognition in supercritical fluid chromatography (SFC) on chiral stationary phases (CSPs). In this first part, the interactions contributing to the retentions of the achiral solutes are measured with a modified version of the solvation parameter model. Since stereospecific interactions were not accounted for in the classical linear solvation energy relationship using Abraham descriptors, we introduce two additional descriptors, flexibility and globularity, to rationally quantify the stereochemical properties that may significantly affect enantiomeric resolutions. Two polysaccharide stationary phases presenting identical bonded groups on different polysaccharide backbones, namely tris-(3,5-dimethylphenylcarbamate) on amylose and on cellulose, are compared using 230 achiral and structurally diverse solutes. The experimental results are evaluated based on statistics and the chemical intuition of the chromatographic systems.


© 2010 Elsevier B.V. All rights reserved.

## 1. Introduction

The chromatographic separation of enantiomers is one of the most difficult challenges in analytical chemistry. In recent years, the rapidly growing technique of supercritical fluid chromatography with packed stationary phases ( pSFC ) has attracted great interests, as evidenced by several recent papers [1-3] and the steady rise in the number of industrial users, as the renewed interest of major instrument manufacturers. SFC , which uses a supercritical or near critical $\mathrm{CO}_{2}$-based mobile phase, often provides a faster separation and higher efficiency due to its higher diffusivity (fast mass transfer) and lower viscosity (low pressure drop) comparing to HPLC [4]. In addition, pSFC replaced water or hexane in HPLC with $\mathrm{CO}_{2}$ as the bulk mobile phase that is not only greener (environmentally friendly) but also less costly to use and to remove [5,6]. Such advances in speed and cost are extremely important not only for analytical separation but even more so for preparative enantioseparations, particularly in the pharmaceutical industry. As a result, pSFC has been most widely used for chiral separations, and is now becoming the first choice for enantioseparation and purification in drug discovery environment [7-11].

[^0]Most HPLC analytical columns can be used for pSFC, as they work equally well in supercritical fluids. The addition of organic modifiers (such as alcohols) and additives (such as diethylamine, ammonium acetate or water) to $\mathrm{CO}_{2}$ extends the utility of pSFC to polar and even ionic compounds. However, because the mobile phase plays a crucial role in chiral separation, HPLC and pSFC enantioseparations are often not equivalent, and their mobile phase conditions are often not transferable [11-13].

Consequently, method development in chiral SFC, as in chiral HPLC, generally relies on a systematic screening of chiral stationary phases (CSPs) and mobile phases, in a preferential order or in parallel based on personal experience of the chromatographer. Despite the large body of experimental data [1] that has been reported since the first chiral SFC separations presented by Mourier et al. in 1985 [14], most publications cover applications or research related to one particular analyte family (see for example [15-19]), and no clear guideline for choosing a stationary phase or mobile phase is available for a new analyte. Therefore, developing a chiral resolution is still largely a time-consuming "trial-and-error" process with multiple columns and mobile phases.

A searchable database, ChirBase [20], that comprises a large number of reported chiral separation conditions from the literature, is aimed at helping chromatographers to select suitable separation conditions [21]. However, since the literature data is not verified, and the compiled data is sometimes incomplete, the included separations are often not the latest and the best possible for the
given analyte. Moreover, no clear association between particular CSPs and analyte types can be extracted from the database, understandably due to the complexity of the intermolecular interaction between the solutes and various CSPs. Nevertheless, while the utility of the database as guidance for method development is limited in our experience, ChirBase suggests that derivatized polysaccharide CSPs are the most popular for SFC enantioseparations.

Polysaccharide CSPs were pioneered by Hesse and Hagel [22] and particularly developed by Okamoto and co-workers [23]. Although fundamental studies of the chiral recognition mechanisms on polysaccharide CSPs have been reported [24-29], a clear description of the chromatographic processes operating on these CSPs is still missing: reliable prediction of enantioselectivity for a new chiral molecule on any polysaccharide CSP is not currently available.

In order to understand the intermolecular interactions and ultimately be able to predict pSFC chiral resolution, a standardized method for characterizing and comparing CSPs under a pSFC condition is required. Therefore, we decided to investigate a number of columns containing various stationary phase chemistries and obtained from different manufacturers.

Both the stationary phases and the mobile phases should be compared, since the chromatographic behaviour of a stationary phase and solute depends on the mobile phase used. As solvation occurs for a solute and stationary phase, the three-dimensional structure and solvation state may change depending on the composition of the mobile phase. To study such an extremely complex system, we divided the problems and tackled them individually. This first paper focuses on our study of different stationary phases under a fixed mobile phase condition. The study of different mobile phases will be the subject of future work. Nevertheless, the chromatographic properties of the stationary phase equilibrated with the given mobile phase were considered when we chose the testing procedure.

We used quantitative structure-retention relationships (QSRRs), to obtain precise information on the interactions contributing to chiral retention and separations. By improving the knowledge and understanding of enantioselective SFC separation, we aim to help chromatographers in choosing the best chromatographic system when facing a new separation problem.

Two types of data are necessary to construct a QSRR [30]: chromatographic retention data for a sufficiently large group of compounds, and a data set reflecting the physico-chemical properties of the said solutes. Describing molecular structures based on the physico-chemical properties of the solutes is the very heart of QSRRs. Indeed, we use solute descriptors to convert the molecular structure information into mathematical data.

There are generally two ways to establish a QSRR. The first one is to perform a multilinear regression on a great number of solute descriptors, then establish a relationship between the chromatographic retention data and the analytes functionalities that are selected by eliminating the insignificant descriptors. This approach is generally chosen when one seeks the best correlation, for prediction purposes for instance. However, the physico-chemical meaning associated with each term is sometimes difficult to understand. Moreover, the selected descriptors are not necessarily the same when another chromatographic system is considered, rendering all comparisons impossible.

The second approach consists of postulating the possible factors influencing the chromatographic retention, as we used in a number of studies on achiral SFC systems [31,32]. Thus, the solute descriptors are limited to a small selection. In this approach, as long as the initially selected interaction terms provide a sufficiently complete description of the possible retention mechanism, high correlation can still be obtained. Moreover, in this case, the meaning
of each term is more straightforward, and a comparison between chromatographic systems is possible.

One of the most widely used QSRR for chromatographic processes is the linear solvation energy relationship (LSER) that uses Abraham descriptors [33,34], also known as the solvation parameter model. The retention factors $(k)$ of selected probes can be related through this relationship to specific interactions by the following equation:
$\log k=c+e E+s S+a A+b B+v V$
In this equation, capital letters represent the solute descriptors, related to particular interaction properties, while lower case letters represent the system constants, related to the complementary effect of the stationary and mobile phases on these interactions. $c$ is the model intercept term and is dominated by the phase ratio. $E$ is the excess molar refraction (calculated from the refractive index of the molecule) and models polarizability contributions from $n$ and $\pi$ electrons; $S$ is the solute dipolarity/polarizability; $A$ and $B$ are the solute overall hydrogen-bond acidity and basicity; $V$ is the McGowan characteristic volume in units of $\mathrm{cm}^{3} \mathrm{~mol}^{-1} / 100$. The system constants ( $e, s, a, b, v$ ), obtained through a multilinear regression of the retention data for a certain number of solutes with known descriptors, reflect the magnitude of difference for that particular property between the mobile and stationary phases. Thus, if a particular coefficient is numerically large, then any solute having the complementary property will interact very strongly with either the mobile phase (if the coefficient is negative) or the stationary phase (if the coefficient is positive).

In the past, the solvation parameter model has proven to be successful for the classification of non-enantioselective stationary phases of various natures in SFC [35].

In general, the solvation parameter model is not considered to be applicable to highly specific phenomena in which stereochemical factors play a critical role [36]. Although it has been reported that the model could provide some information on the intermolecular interactions involved in these processes [37-41], we were concerned that the five usual descriptors may not be sufficient for a precise description of enantio-recognition mechanisms, since no term in the solvation parameter model is related to the shape or stereochemistry of the solute that certainly plays a significant role in enantio-recognition processes.

In this series of papers, a modified version of the solvation parameter model will be used in a systematic study of various types of CSPs with varied mobile phase compositions in SFC. The work presented in this first paper is intended first to evaluate the interactions contributing to retention on polysaccharide CSPs in pSFC. The interactions contributing to enantioseparation will be evaluated in a subsequent paper. This split approach, evaluating the retention and enantioseparation separately, is derived from the perception that a combination of enantiospecific and non-enantiospecific interactions contributes to an enantiomeric resolution where non-enantiospectific interactions would contribute mostly to the retention, and a difference in enantiospecific interactions with a CSP between two enantiomers would drive the separation.

As mentioned above, polysaccharide CSPs are the most widely used stationary phases for SFC enantioseparations [42]. Among all polysaccharide CSPs, tris-(3,5-dimethylphenylcarbamate) on amylose (ADMPC) and cellulose (CDMPC) (Fig. 1) are the two most successful ones for chiral resolutions of a wide variety of pharmaceutical compounds. Additionally, these stationary phases are known to be uniform, with little inter-column variability.

We initially selected these two CSPs to evaluate how shape affects the intermolecular interactions that drive retention, because the two CSPs only differ by the orientation of the polysaccharide backbone. Both polysaccharides are based on D-glucose units, mod-
a


b


Fig. 1. Structures of the stationary phases investigated in this study (a) ADMPC and (b) CDMPC. The polysaccharide derivatives are coated onto a silica substrate.
ified with the same aromatic groups, but the primary difference between the two phases is the nature of the glycosidic linkage. In cellulose, it is a $\beta-1,4$ linkage, and in amylose it is an $\alpha-1,4$ linkage. While this is a small difference at the microscopic level, it has a major impact on the shapes and macromolecular structures of the polysaccharides. Cellulose and amylose strands both assume a helical structure, with a left-handed $3 / 2$ helix and lefthanded $4 / 3$ helix respectively [43]. Thus a chiral helical groove surrounded by polar ligands exists along the main chains. However, while cellulose strands lie side by side, amylose strands adopt a supramolecular helical structure [44,45]. The difference in this supramolecular structure results in different spatial arrangements of the chiral cavities. Fig. 2 shows possible structures of these CSPs.

The work presented herein focuses on (i) searching for additional solute descriptors, which could help extract additional information from the experimental data to more accurately


Fig. 2. Molecular models of tris-3,5-dimethylphenylcarbamates of (a) cellulose and (b) amylose, perpendicular to (top) and along (bottom) the helix axis. Copyright: 1999 The Chemical Society of Japan [44] and 2002 American Chemical Society [45].
describe the chromatography interactions for stereoisomers, (ii) building an appropriate solute set, which would provide precise and complete information on all possible interactions established between the analytes and the chromatographic system, and (iii) the establishment of a relationship between solute descriptors and retention on ADMPC and CDMPC CSPs.

## 2. Experimental

### 2.1. Stationary phases

The columns used in this study were Chiralcel OD-H (CDMPC) and Chiralpak AD-H (ADMPC) ( $150 \mathrm{~mm} \times 4.6 \mathrm{~mm}, 5 \mu \mathrm{~m}$ ) from Daicel (Tokyo, Japan). Their structures are presented in Figs. 1 and 2.

For the purpose of comparison, two non-enantioselective stationary phases were also used: Synergi Polar RP ( $250 \mathrm{~mm} \times 4.6 \mathrm{~mm}$, $4 \mu \mathrm{~m}$ ) from Phenomenex (Le Pecq, France), and XBridge Shield C18 $(250 \mathrm{~mm} \times 4.6 \mathrm{~mm}, 5 \mu \mathrm{~m})$ from Waters (Guyancourt, France).

All columns were new at the start of this study to eliminate any concern with respect to the changes of column properties as a result of their prior use under different mobile phase conditions.

### 2.2. Chemicals

The 230 solutes used in this study are presented in Table 1, together with their solute descriptors. The chemicals were obtained from several different manufacturers. All solutions were prepared in methanol.

The HPLC grade methanol ( MeOH ) used in the studies was provided by SDS Carlo Erba (Val-de-Reuil, France). Carbon dioxide was provided by Messer (Puteaux, France).

### 2.3. Apparatus and operating conditions

Chromatographic separations were carried out using equipment manufactured by Jasco (Tokyo, Japan). Two model 980-PU pumps were used, one for carbon dioxide and a second for the modifier. Control of the mobile phase composition was performed by the modifier pump. The pump head used for pumping the carbon dioxide was cooled to $-5^{\circ} \mathrm{C}$ by a cryostat (Julabo F10c, Seelbach, Germany). When the two solvents (methanol and $\mathrm{CO}_{2}$ ) were mixed, the fluid was introduced into a dynamic mixing chamber PU 4046 (Pye Unicam, Cambridge, UK) connected to a pulsation damper (Sedere, Orleans, France). The injector valve was supplied with a $5 \mu \mathrm{~L}$ loop (model 7125 Rheodyne, Cotati, CA, USA).

The columns were thermostated by an oven (Jetstream 2 Plus, Hewlett-Packard, Palo Alto, USA), regulated by a cryostat (Haake D8 GH, Karlsruhe, Germany). The detector was a UV-vis. HP 1050 (Hewlett-Packard, CA), with a high-pressure resistant cell. After the detector, the outlet column pressure was controlled by a Jasco 88081 pressure regulator. The outlet regulator tube (internal diameter 0.25 mm ) was heated to $60^{\circ} \mathrm{C}$ to avoid ice formation during the $\mathrm{CO}_{2}$ depressurization.

UV detection was carried out at 254 nm , or 210 nm for nonaromatic solutes. Injection volumes were $1-5 \mu \mathrm{~L}$. Chromatograms were recorded using the Azur software (Datalys, France).

Operating conditions were as follows: carbon dioxide-methanol 90:10 (v/v), $3 \mathrm{~mL} / \mathrm{min}, 25^{\circ} \mathrm{C}$ (controlled with an oven), outlet pressure 150 bar. These conditions were chosen so as to allow suitable retention of the probe solutes on all stationary phases. Sufficient retentions are necessary to ensure that the retention factors are statistically significant. On the other hand, too much retention would extend the analysis beyond a reasonable time.

Since chiral and achiral SFC separations can be significantly and variably affected by the presence of mobile phase additives depending on the extent to which they are adsorbed onto the stationary

Table 1
Chromatographic solutes and molecular descriptors.

| $\mathrm{N}^{\circ}$ | Compound | E | $S$ | A | B | V | $F$ | G |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 1 | Benzene | 0.610 | 0.52 | 0.00 | 0.14 | 0.7164 | 0.00 | 2.04 |
| 2 | Toluene | 0.601 | 0.52 | 0.00 | 0.14 | 0.8573 | 0.00 | 1.89 |
| 3 | Ethylbenzene | 0.613 | 0.51 | 0.00 | 0.15 | 0.9982 | 0.63 | 1.77 |
| 4 | Propylbenzene | 0.604 | 0.50 | 0.00 | 0.15 | 1.1391 | 1.11 | 1.60 |
| 5 | Butylbenzene | 0.600 | 0.51 | 0.00 | 0.15 | 1.2800 | 1.50 | 1.43 |
| 6 | Pentylbenzene | 0.594 | 0.51 | 0.00 | 0.15 | 1.4209 | 1.82 | 1.27 |
| 7 | Hexylbenzene | 0.591 | 0.50 | 0.00 | 0.15 | 1.5620 | 2.08 | 1.21 |
| 8 | Heptylbenzene | 0.577 | 0.48 | 0.00 | 0.15 | 1.7029 | 2.31 | 0.98 |
| 9 | Octylbenzene | 0.579 | 0.48 | 0.00 | 0.15 | 1.8438 | 2.50 | 0.88 |
| 10 | Nonylbenzene | 0.578 | 0.48 | 0.00 | 0.15 | 1.9847 | 2.67 | 0.73 |
| 11 | Decylbenzene | 0.579 | 0.47 | 0.00 | 0.15 | 2.1254 | 2.81 | 0.61 |
| 12 | Undecylbenzene | 0.579 | 0.47 | 0.00 | 0.15 | 2.2665 | 2.94 | 0.49 |
| 13 | Dodecylbenzene | 0.571 | 0.47 | 0.00 | 0.15 | 2.4074 | 3.06 | 0.38 |
| 14 | Tridecylbenzene | 0.560 | 0.46 | 0.00 | 0.15 | 2.5483 | 3.16 | 0.27 |
| 15 | Tetradecylbenzene | 0.560 | 0.46 | 0.00 | 0.16 | 2.6890 | 3.25 | 0.18 |
| 16 | Allylbenzene | 0.717 | 0.60 | 0.00 | 0.22 | 1.0961 | 1.11 | 1.64 |
| 17 | Cumene | 0.602 | 0.49 | 0.00 | 0.16 | 1.1391 | 0.56 | 1.81 |
| 18 | p-Cymene | 0.607 | 0.49 | 0.00 | 0.19 | 1.2800 | 0.50 | 1.63 |
| 19 | t-Butylbenzene | 0.619 | 0.49 | 0.00 | 0.18 | 1.2800 | 0.00 | 1.85 |
| 20 | Naphthalene | 1.340 | 0.92 | 0.00 | 0.20 | 1.0854 | 0.00 | 1.81 |
| 21 | 1-Methylnaphthalene | 1.344 | 0.92 | 0.00 | 0.20 | 1.2263 | 0.00 | 1.72 |
| 22 | 2-Methylnaphthalene | 1.304 | 0.92 | 0.00 | 0.20 | 1.2263 | 0.00 | 1.64 |
| 23 | 1-Ethylnaphthalene | 1.371 | 0.88 | 0.00 | 0.20 | 1.3672 | 0.38 | 1.61 |
| 24 | 2-Ethylnaphthalene | 1.331 | 0.90 | 0.00 | 0.20 | 1.3672 | 0.38 | 1.51 |
| 25 | Aniline | 0.955 | 0.96 | 0.26 | 0.41 | 0.8162 | 0.00 | 1.97 |
| 26 | N -methylaniline | 0.948 | 0.90 | 0.17 | 0.43 | 0.9571 | 0.63 | 1.78 |
| 27 | $\mathrm{N}, \mathrm{N}$-Dimethylaniline | 0.957 | 0.81 | 0.00 | 0.41 | 1.0980 | 0.56 | 1.74 |
| 28 | N -Ethylaniline | 0.945 | 0.85 | 0.17 | 0.43 | 1.0980 | 1.11 | 1.56 |
| 29 | $\mathrm{N}, \mathrm{N}$-Diethylaniline | 0.953 | 0.80 | 0.00 | 0.41 | 1.3798 | 1.36 | 1.57 |
| 30 | Benzylamine | 0.829 | 0.77 | 0.15 | 0.72 | 0.9571 | 0.63 | 1.83 |
| 31 | N -Methylbenzylamine | 0.797 | 0.74 | 0.13 | 0.74 | 1.0980 | 1.11 | 1.58 |
| 32 | $\mathrm{N}, \mathrm{N}$-Dimethylbenzylamine | 0.668 | 0.80 | 0.00 | 0.69 | 1.2389 | 1.00 | 1.63 |
| 33 | o-Phenylenediamine | 1.260 | 1.40 | 0.24 | 0.73 | 0.9160 | 0.00 | 1.93 |
| 34 | 2-Phenylethylamine | 0.824 | 0.87 | 0.32 | 0.72 | 1.0980 | 1.11 | 1.68 |
| 35 | Pyridine | 0.631 | 0.84 | 0.00 | 0.52 | 0.6753 | 0.00 | 2.08 |
| 36 | 2-Ethylpyridine | 0.613 | 0.71 | 0.00 | 0.59 | 0.9571 | 0.63 | 1.76 |
| 37 | Benzamide | 0.990 | 1.50 | 0.49 | 0.67 | 0.9728 | 0.56 | 1.85 |
| 38 | Phenylurea | 1.110 | 1.33 | 0.79 | 0.79 | 1.0726 | 0.50 | 1.69 |
| 39 | Benzenesulfonamide | 1.130 | 1.55 | 0.55 | 0.80 | 1.0971 | 0.50 | 1.74 |
| 40 | p-Toluenesulfonamide | 1.100 | 1.55 | 0.55 | 0.87 | 1.2380 | 0.45 | 1.58 |
| 41 | Caffeine | 1.500 | 1.72 | 0.00 | 1.28 | 1.3632 | 0.00 | 1.54 |
| 42 | Theophylline | 1.500 | 1.60 | 0.54 | 1.34 | 1.2223 | 0.00 | 1.66 |
| 43 | Nicotinamide | 1.010 | 1.09 | 0.63 | 1.00 | 0.9317 | 0.56 | 1.89 |
| 44 | Nicotinic acid | 0.790 | 1.21 | 0.57 | 0.73 | 0.8906 | 0.56 | 1.92 |
| 45 | $o$-Picolinic acid | 0.790 | 1.21 | 0.57 | 0.73 | 0.8906 | 0.56 | 1.87 |
| 46 | Indazole | 1.180 | 1.22 | 0.53 | 0.35 | 0.9053 | 0.00 | 1.96 |
| 47 | Carbazole | 1.787 | 2.01 | 0.18 | 0.08 | 1.3154 | 0.00 | 1.59 |
| 48 | Acridine | 2.356 | 1.32 | 0.00 | 0.58 | 1.4133 | 0.00 | 1.52 |
| 49 | Quinoline | 1.268 | 0.97 | 0.00 | 0.54 | 1.0443 | 0.00 | 1.83 |
| 50 | 1-Naphtylamine | 1.670 | 1.26 | 0.20 | 0.57 | 1.1852 | 0.00 | 1.77 |
| 51 | 2-Naphtylamine | 1.670 | 1.28 | 0.22 | 0.55 | 1.1852 | 0.00 | 1.73 |
| 52 | 1-Naphthalenemethylamine | 1.566 | 1.25 | 0.10 | 0.73 | 1.3917 | 0.38 | 1.67 |
| 53 | Aminodiphenylmethane | 1.360 | 1.40 | 0.21 | 0.78 | 1.5649 | 0.71 | 1.43 |
| 54 | N -Benzylbenzamide | 1.460 | 1.97 | 0.26 | 0.76 | 1.7215 | 0.63 | 1.04 |
| 55 | 2,2'-Bipyridyl | 1.384 | 1.33 | 0.00 | 0.81 | 1.2420 | 0.38 | 1.55 |
| 56 | 4,4'-Bipyridyl | 1.470 | 1.42 | 0.00 | 0.78 | 1.2420 | 0.38 | 1.58 |
| 57 | $o$-Toluidine | 0.966 | 0.92 | 0.23 | 0.45 | 0.9571 | 0.00 | 1.87 |
| 58 | $m$-Toluidine | 0.946 | 0.95 | 0.23 | 0.45 | 0.9571 | 0.00 | 1.81 |
| 59 | $p$-Toluidine | 0.923 | 0.95 | 0.23 | 0.45 | 0.9571 | 0.00 | 1.81 |
| 60 | Benzoic acid | 0.730 | 0.90 | 0.59 | 0.40 | 0.9317 | 0.56 | 1.88 |
| 61 | Isophthalic acid | 0.940 | 1.46 | 1.14 | 0.77 | 1.1470 | 0.83 | 1.69 |
| 62 | Trimesic acid | 1.140 | 1.84 | 1.71 | 1.10 | 1.3623 | 1.00 | 1.51 |
| 63 | 1-Naphtoic acid | 1.460 | 1.20 | 0.65 | 0.46 | 1.3007 | 0.36 | 1.68 |
| 64 | 2-Naphtoic acid | 1.460 | 1.15 | 0.61 | 0.44 | 1.3007 | 0.36 | 1.62 |
| 65 | 1-Naphtylacetic acid | 1.460 | 1.55 | 0.60 | 0.67 | 1.4416 | 0.67 | 1.54 |
| 66 | 2-Naphtylacetic acid | 1.470 | 1.40 | 0.57 | 0.50 | 1.4416 | 0.67 | 1.44 |
| 67 | 2-Biphenylcarboxylic acid | 1.580 | 1.55 | 0.57 | 0.54 | 1.5395 | 0.63 | 1.46 |
| 68 | 4-Biphenylcarboxylic acid | 1.480 | 1.34 | 0.61 | 0.51 | 1.5395 | 0.63 | 1.33 |
| 69 | 4-Biphenylacetic acid | 1.580 | 1.55 | 0.57 | 0.54 | 1.6804 | 0.88 | 1.18 |
| 70 | Diphenylacetic acid | 1.330 | 1.53 | 0.57 | 0.56 | 1.6804 | 0.88 | 1.46 |
| 71 | 3,3-Diphenylpropionic acid | 1.330 | 1.53 | 0.57 | 0.57 | 1.8213 | 1.11 | 1.58 |
| 72 | Benzilic acid | 1.460 | 1.44 | 0.74 | 1.03 | 1.7391 | 0.83 | 1.39 |
| 73 | Benzyl alcohol | 0.803 | 0.87 | 0.39 | 0.56 | 0.9160 | 0.63 | 1.83 |
| 74 | 3-Phenyl-1-propanol | 0.821 | 0.94 | 0.31 | 0.65 | 1.1978 | 1.50 | 1.76 |
| 75 | 4-Phenyl-1-butanol | 0.811 | 0.90 | 0.33 | 0.70 | 1.3387 | 1.82 | 1.49 |
| 76 | 1-Phenyl-1-cyclohexanol | 0.990 | 0.88 | 0.31 | 0.56 | 1.5119 | 0.36 | 1.57 |

Table 1 (Continued )

| $\mathrm{N}^{\circ}$ | Compound | E | $S$ | A | B | V | F | G |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 77 | R,S-1,2-Diphenyl-1,2-ethanediol | 1.610 | 1.47 | 0.54 | 1.09 | 1.7234 | 0.88 | 1.71 |
| 78 | Naphthalene methanol | 1.530 | 1.27 | 0.39 | 0.62 | 1.2850 | 0.38 | 1.68 |
| 79 | Naphthalene ethanol | 1.540 | 1.31 | 0.30 | 0.70 | 1.4259 | 0.71 | 1.75 |
| 80 | Cyclopropyldiphenylcarbinol | 1.580 | 1.30 | 0.31 | 0.71 | 1.8379 | 1.00 | 1.35 |
| 81 | Methylbenzilate | 1.350 | 1.40 | 0.17 | 1.03 | 1.8800 | 0.79 | 1.23 |
| 82 | Anisole | 0.708 | 0.75 | 0.00 | 0.29 | 0.9160 | 0.63 | 1.96 |
| 83 | Benzaldehyde | 0.820 | 1.00 | 0.00 | 0.39 | 0.8730 | 0.63 | 1.91 |
| 84 | Acetophenone | 0.818 | 1.01 | 0.00 | 0.48 | 1.0139 | 0.56 | 1.81 |
| 85 | Propiophenone | 0.804 | 0.95 | 0.00 | 0.51 | 1.1548 | 1.00 | 1.62 |
| 86 | Valerophenone | 0.797 | 0.95 | 0.00 | 0.51 | 1.4366 | 1.67 | 1.29 |
| 87 | Coumarin | 1.060 | 1.76 | 0.00 | 0.43 | 1.0619 | 0.00 | 1.92 |
| 88 | Benzonitrile | 0.742 | 1.11 | 0.00 | 0.33 | 0.8711 | 0.00 | 1.79 |
| 89 | Nitrobenzene | 0.871 | 1.11 | 0.00 | 0.28 | 0.8906 | 0.56 | 1.87 |
| 90 | p-Dinitrobenzene | 1.130 | 1.63 | 0.00 | 0.46 | 1.0648 | 0.83 | 1.65 |
| 91 | Fluorobenzene | 0.477 | 0.57 | 0.00 | 0.10 | 0.7340 | 0.00 | 1.99 |
| 92 | Chlorobenzene | 0.718 | 0.65 | 0.00 | 0.07 | 0.8388 | 0.00 | 1.92 |
| 93 | Bromobenzene | 0.882 | 0.73 | 0.00 | 0.09 | 0.8914 | 0.00 | 1.89 |
| 94 | Iodobenzene | 1.188 | 0.82 | 0.00 | 0.12 | 0.9746 | 0.00 | 1.85 |
| 95 | Naphtylaldehyde | 1.549 | 1.40 | 0.00 | 0.45 | 1.2420 | 0.38 | 1.73 |
| 96 | Naphtylacetate | 1.391 | 1.53 | 0.00 | 0.59 | 1.4416 | 0.67 | 1.52 |
| 97 | Cyanonaphthalene | 1.472 | 1.51 | 0.00 | 0.39 | 1.2401 | 0.00 | 1.63 |
| 98 | Naphtylacetonitrile | 1.481 | 1.55 | 0.00 | 0.51 | 1.3810 | 0.71 | 1.52 |
| 99 | Nitronaphthalene | 1.600 | 1.51 | 0.00 | 0.29 | 1.2596 | 0.36 | 1.64 |
| 100 | Fluoronaphthalene | 1.144 | 0.97 | 0.00 | 0.13 | 1.1030 | 0.00 | 1.78 |
| 101 | Chloronaphthalene | 1.417 | 1.00 | 0.00 | 0.14 | 1.2078 | 0.00 | 1.73 |
| 102 | Bromonaphthalene | 1.598 | 1.13 | 0.00 | 0.13 | 1.2604 | 0.00 | 1.71 |
| 103 | Iodonaphthalene | 1.928 | 1.22 | 0.00 | 0.16 | 1.3436 | 0.00 | 1.68 |
| 104 | Benzophenone | 1.447 | 1.50 | 0.00 | 0.50 | 1.4808 | 0.67 | 1.44 |
| 105 | Deoxybenzoin | 1.360 | 1.60 | 0.00 | 0.51 | 1.6217 | 0.94 | 1.40 |
| 106 | Benzil | 1.445 | 1.59 | 0.00 | 0.62 | 1.6374 | 0.88 | 1.32 |
| 107 | Benzyl benzoate | 1.330 | 1.42 | 0.00 | 0.47 | 1.5395 | 0.63 | 1.23 |
| 108 | Phenol | 0.805 | 0.89 | 0.60 | 0.30 | 0.7751 | 0.00 | 1.98 |
| 109 | Eugenol | 0.946 | 0.99 | 0.22 | 0.51 | 1.3544 | 1.25 | 1.55 |
| 110 | Vanillin | 1.040 | 1.33 | 0.29 | 0.69 | 1.1313 | 0.91 | 1.64 |
| 111 | Thymol | 0.822 | 0.79 | 0.52 | 0.44 | 1.3387 | 0.45 | 1.49 |
| 112 | Pyrocatechol | 0.970 | 1.10 | 0.88 | 0.47 | 0.8338 | 0.00 | 1.94 |
| 113 | Resorcinol | 0.980 | 1.00 | 1.09 | 0.52 | 0.8338 | 0.00 | 1.91 |
| 114 | Hydroquinone | 1.063 | 1.27 | 1.06 | 0.57 | 0.8337 | 0.00 | 1.92 |
| 115 | Pyrogallol | 1.165 | 1.35 | 1.35 | 0.62 | 0.8925 | 0.00 | 1.90 |
| 116 | Phloroglucinol | 1.355 | 1.12 | 1.40 | 0.82 | 0.8925 | 0.00 | 1.86 |
| 117 | $\alpha$-Naphtol | 1.520 | 1.05 | 0.60 | 0.37 | 1.1441 | 0.00 | 1.78 |
| 118 | $\beta$-Naphtol | 1.520 | 1.08 | 0.61 | 0.40 | 1.1440 | 0.00 | 1.75 |
| 119 | 2-Phenylphenol | 1.550 | 1.40 | 0.56 | 0.49 | 1.3829 | 0.36 | 1.48 |
| 120 | 4-Phenylphenol | 1.560 | 1.41 | 0.59 | 0.40 | 1.3829 | 0.36 | 1.46 |
| 121 | Thiophene | 0.687 | 0.57 | 0.00 | 0.15 | 0.6411 | 0.00 | 2.12 |
| 122 | Thiophenol | 1.000 | 0.80 | 0.09 | 0.16 | 0.8799 | 0.00 | 1.91 |
| 123 | Thioanisole | 1.063 | 0.68 | 0.00 | 0.32 | 1.0208 | 0.63 | 1.75 |
| 124 | Methylphenylsulfone | 1.080 | 1.85 | 0.00 | 0.76 | 1.1382 | 0.50 | 1.71 |
| 125 | $o$-Xylene | 0.663 | 0.56 | 0.00 | 0.16 | 0.9982 | 0.00 | 1.86 |
| 126 | $m$-Xylene | 0.623 | 0.52 | 0.00 | 0.16 | 0.9982 | 0.00 | 1.72 |
| 127 | $p$-Xylene | 0.613 | 0.52 | 0.00 | 0.16 | 0.9982 | 0.00 | 1.72 |
| 128 | o-Cresol | 0.840 | 0.86 | 0.52 | 0.30 | 0.9160 | 0.00 | 1.88 |
| 129 | m-Cresol | 0.822 | 0.88 | 0.57 | 0.34 | 0.9160 | 0.00 | 1.83 |
| 130 | p-Cresol | 0.820 | 0.87 | 0.57 | 0.31 | 0.9160 | 0.00 | 1.83 |
| 131 | 2,3-Dimethylphenol | 0.850 | 0.85 | 0.52 | 0.36 | 1.0569 | 0.00 | 1.92 |
| 132 | 2,4-Dimethylphenol | 0.843 | 0.80 | 0.53 | 0.39 | 1.0569 | 0.00 | 1.70 |
| 133 | 2,5-Dimethylphenol | 0.840 | 0.79 | 0.54 | 0.37 | 1.0569 | 0.00 | 1.70 |
| 134 | 2,6-Dimethylphenol | 0.860 | 0.79 | 0.39 | 0.39 | 1.0569 | 0.00 | 1.75 |
| 135 | 3,4-Dimethylphenol | 0.830 | 0.86 | 0.56 | 0.39 | 1.0569 | 0.00 | 1.79 |
| 136 | 3,5-Dimethylphenol | 0.820 | 0.84 | 0.57 | 0.36 | 1.0569 | 0.00 | 1.66 |
| 137 | o-Isopropylphenol | 0.842 | 0.88 | 0.52 | 0.38 | 1.1978 | 0.50 | 1.69 |
| 138 | m-Isopropylphenol | 0.811 | 0.92 | 0.55 | 0.38 | 1.1978 | 0.50 | 1.65 |
| 139 | p-Isopropylphenol | 0.791 | 0.89 | 0.55 | 0.38 | 1.1978 | 0.50 | 1.75 |
| 140 | o-Chlorophenol | 0.853 | 0.88 | 0.32 | 0.31 | 0.8975 | 0.00 | 1.89 |
| 141 | m-Chlorophenol | 0.909 | 1.06 | 0.69 | 0.15 | 0.8975 | 0.00 | 1.85 |
| 142 | p-Chlorophenol | 0.915 | 1.08 | 0.67 | 0.20 | 0.8975 | 0.00 | 1.85 |
| 143 | $o$-Nitrophenol | 1.015 | 1.05 | 0.05 | 0.37 | 0.9493 | 0.50 | 1.87 |
| 144 | $m$-Nitrophenol | 1.050 | 1.57 | 0.79 | 0.23 | 0.9493 | 0.50 | 1.79 |
| 145 | p-Nitrophenol | 1.070 | 1.72 | 0.82 | 0.26 | 0.9493 | 0.50 | 1.79 |
| 146 | o-Nitrobenzylalcohol | 1.064 | 1.42 | 0.35 | 0.70 | 1.0902 | 0.91 | 1.81 |
| 147 | m-Nitrobenzylalcohol | 1.064 | 1.35 | 0.44 | 0.64 | 1.0902 | 0.91 | 1.67 |
| 148 | p-Nitrobenzylalcohol | 1.064 | 1.39 | 0.44 | 0.62 | 1.0902 | 0.91 | 1.66 |
| 149 | $o$-Nitrotoluene | 0.866 | 1.11 | 0.00 | 0.28 | 1.0315 | 0.50 | 1.77 |
| 150 | $m$-Nitrotoluene | 0.874 | 1.10 | 0.00 | 0.25 | 1.0315 | 0.50 | 1.69 |
| 151 | $p$-Nitrotoluene | 0.870 | 1.11 | 0.00 | 0.28 | 1.0315 | 0.50 | 1.69 |
| 152 | 2,4-Dinitrotoluene | 1.150 | 1.61 | 0.00 | 0.49 | 1.2057 | 0.77 | 1.60 |
| 153 | 2,6-Dinitrotoluene | 1.150 | 1.60 | 0.00 | 0.45 | 1.2057 | 0.77 | 1.69 |

Table 1 (Continued )

| $\mathrm{N}^{\circ}$ | Compound | E | $S$ | A | B | V | F | G |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 154 | o-Methylacetophenone | 0.780 | 1.00 | 0.00 | 0.51 | 1.1548 | 0.50 | 1.72 |
| 155 | $m$-Methylacetophenone | 0.806 | 1.00 | 0.00 | 0.51 | 1.1548 | 0.50 | 1.63 |
| 156 | p-Methylacetophenone | 0.842 | 1.00 | 0.00 | 0.52 | 1.1548 | 0.50 | 1.63 |
| 157 | Methylbenzoate | 0.733 | 0.85 | 0.00 | 0.46 | 1.0726 | 0.50 | 1.60 |
| 158 | Ethylbenzoate | 0.689 | 0.85 | 0.00 | 0.46 | 1.2135 | 0.91 | 1.40 |
| 159 | Propylbenzoate | 0.675 | 0.80 | 0.00 | 0.46 | 1.3544 | 1.25 | 1.25 |
| 160 | Butylbenzoate | 0.668 | 0.80 | 0.00 | 0.46 | 1.4953 | 1.54 | 1.10 |
| 161 | Dimethylphthalate | 0.780 | 1.40 | 0.00 | 0.84 | 1.4288 | 0.71 | 1.55 |
| 162 | Diethylphthalate | 0.729 | 1.40 | 0.00 | 0.86 | 1.7106 | 1.25 | 1.40 |
| 163 | Dipropylphthalate | 0.713 | 1.40 | 0.00 | 0.86 | 1.9924 | 1.67 | 1.53 |
| 164 | Dibutylphthalate | 0.700 | 1.40 | 0.00 | 0.86 | 2.2742 | 2.00 | 1.18 |
| 165 | Methylparaben | 0.900 | 1.37 | 0.69 | 0.45 | 1.1313 | 0.45 | 1.54 |
| 166 | Ethylparaben | 0.860 | 1.35 | 0.69 | 0.45 | 1.2722 | 0.83 | 1.37 |
| 167 | Propylparaben | 0.860 | 1.35 | 0.69 | 0.45 | 1.4131 | 1.15 | 1.22 |
| 168 | Butylparaben | 0.860 | 1.33 | 0.71 | 0.46 | 1.5540 | 1.43 | 1.08 |
| 169 | Biphenyl | 1.360 | 0.99 | 0.00 | 0.26 | 1.3242 | 0.38 | 1.51 |
| 170 | 1-Phenyl-naphthalene | 1.950 | 1.20 | 0.00 | 0.34 | 1.6932 | 0.28 | 1.33 |
| 171 | Diphenylmethane | 1.220 | 1.04 | 0.00 | 0.33 | 1.4651 | 0.71 | 1.62 |
| 172 | Acenaphthene | 1.604 | 1.05 | 0.00 | 0.22 | 1.2586 | 0.00 | 1.68 |
| 173 | Acenaphtylene | 1.750 | 1.14 | 0.00 | 0.26 | 1.2156 | 0.00 | 1.72 |
| 174 | Fluorene | 1.588 | 1.06 | 0.00 | 0.25 | 1.3565 | 0.00 | 1.59 |
| 175 | Phenanthrene | 2.055 | 1.29 | 0.00 | 0.29 | 1.4544 | 0.00 | 1.58 |
| 176 | Anthracene | 2.290 | 1.34 | 0.00 | 0.28 | 1.4544 | 0.00 | 1.54 |
| 177 | 9-Methylanthracene | 2.020 | 1.28 | 0.00 | 0.23 | 1.5953 | 0.00 | 1.52 |
| 178 | Fluoranthene | 2.600 | 1.52 | 0.00 | 0.25 | 1.5846 | 0.00 | 1.51 |
| 179 | Pyrene | 2.600 | 1.52 | 0.00 | 0.25 | 1.5846 | 0.00 | 1.62 |
| 180 | Chrysene | 2.710 | 1.66 | 0.00 | 0.29 | 1.8234 | 0.00 | 1.34 |
| 181 | Benz[a]anthracene | 2.710 | 1.66 | 0.00 | 0.29 | 1.8234 | 0.00 | 1.32 |
| 182 | Tetracene | 2.710 | 1.66 | 0.00 | 0.29 | 1.8234 | 0.00 | 1.27 |
| 183 | Benzo[a]pyrene | 3.320 | 1.84 | 0.00 | 0.31 | 1.9536 | 0.00 | 1.38 |
| 184 | Perylene | 3.320 | 1.84 | 0.00 | 0.31 | 1.9536 | 0.00 | 1.41 |
| 185 | Binaphthyl | 2.820 | 1.81 | 0.00 | 0.33 | 2.0622 | 0.22 | 1.17 |
| 186 | Triphenylene | 2.710 | 1.66 | 0.00 | 0.29 | 1.8234 | 0.00 | 1.42 |
| 187 | $o$-Terphenyl | 2.000 | 1.18 | 0.00 | 0.30 | 1.9320 | 0.50 | 1.28 |
| 188 | p-Terphenyl | 2.040 | 1.48 | 0.00 | 0.30 | 1.9320 | 0.50 | 1.00 |
| 189 | Acetanilide | 0.900 | 1.39 | 0.48 | 0.67 | 1.1137 | 0.50 | 1.56 |
| 190 | Ascorbic acid | 1.230 | 1.68 | 1.12 | 1.65 | 1.1116 | 0.83 | 1.85 |
| 191 | Salicylic acid | 0.890 | 0.84 | 0.71 | 0.38 | 0.9904 | 0.50 | 1.87 |
| 192 | Amitriptyline | 2.246 | 1.78 | 0.00 | 1.00 | 2.3996 | 0.65 | 1.13 |
| 193 | Antipyrine | 1.320 | 1.50 | 0.00 | 1.48 | 1.4846 | 0.33 | 1.46 |
| 194 | Aspirin | 0.781 | 0.80 | 0.49 | 1.00 | 1.2879 | 0.77 | 1.72 |
| 195 | p-Chloroacetanilide | 0.980 | 1.47 | 0.64 | 0.51 | 1.2361 | 0.45 | 1.44 |
| 196 | Codeine | 1.780 | 1.95 | 0.33 | 1.78 | 2.2057 | 0.19 | 1.54 |
| 197 | Cyclobarbital | 1.440 | 1.35 | 0.49 | 1.45 | 1.7859 | 0.56 | 2.00 |
| 198 | Diphenhydramine | 1.360 | 1.43 | 0.00 | 0.95 | 2.1872 | 1.50 | 1.00 |
| 199 | Dithranol | 1.980 | 1.68 | 0.41 | 0.52 | 1.6305 | 0.00 | 1.49 |
| 200 | Harmaline | 1.710 | 1.05 | 0.31 | 0.68 | 1.6578 | 0.28 | 1.32 |
| 201 | Lidocaine | 1.010 | 1.50 | 0.12 | 1.21 | 2.0589 | 1.47 | 1.59 |
| 202 | Mephentermine | 0.710 | 0.76 | 0.13 | 0.60 | 1.5207 | 1.25 | 1.67 |
| 203 | Metacetamol | 1.050 | 1.70 | 1.09 | 0.78 | 1.1724 | 0.91 | 1.65 |
| 204 | Noscapine | 2.390 | 3.09 | 0.00 | 2.09 | 2.8751 | 0.59 | 1.74 |
| 205 | Papaverine | 2.190 | 2.76 | 0.00 | 1.47 | 2.5914 | 1.11 | 1.06 |
| 206 | Paracetamol | 1.060 | 1.63 | 1.04 | 0.86 | 1.1724 | 0.45 | 1.50 |
| 207 | Phenobarbital | 1.630 | 1.80 | 0.73 | 1.15 | 1.6999 | 0.56 | 1.97 |
| 208 | Phenylbutazone | 1.846 | 2.62 | 0.00 | 1.28 | 2.4239 | 1.00 | 1.07 |
| 209 | Phenyltoloxamine | 1.380 | 1.39 | 0.00 | 0.92 | 2.1872 | 1.50 | 1.16 |
| 210 | Primidone | 1.510 | 2.08 | 0.51 | 1.45 | 1.6842 | 0.59 | 1.58 |
| 211 | Procaine | 1.135 | 1.68 | 0.44 | 1.23 | 1.9767 | 1.76 | 0.95 |
| 212 | Quinine | 2.469 | 1.23 | 0.37 | 1.97 | 2.5512 | 0.74 | 1.20 |
| 213 | 2-Pentanone | 0.143 | 0.68 | 0.00 | 0.51 | 0.8288 | 0.20 | 1.78 |
| 214 | 3-Pentanone | 0.154 | 0.66 | 0.00 | 0.51 | 0.8288 | 0.20 | 1.76 |
| 215 | $\gamma$-Terpinene | 0.497 | 0.32 | 0.00 | 0.20 | 1.3230 | 0.50 | 1.55 |
| 216 | Terpinolene | 0.593 | 0.31 | 0.00 | 0.20 | 1.3230 | 0.00 | 1.60 |
| 217 | Myrcene | 0.483 | 0.29 | 0.00 | 0.21 | 1.3886 | 2.22 | 1.44 |
| 218 | Citral | 0.490 | 0.84 | 0.00 | 0.50 | 1.4473 | 2.00 | 1.76 |
| 219 | $\beta$-Ionone | 0.875 | 0.90 | 0.00 | 0.50 | 1.7614 | 0.71 | 1.58 |
| 220 | $\beta$-Terpineol | 0.500 | 0.49 | 0.31 | 0.44 | 1.4247 | 0.00 | 1.74 |
| 221 | $\boldsymbol{\gamma}$-Terpineol | 0.500 | 0.49 | 0.31 | 0.44 | 1.4247 | 0.45 | 1.72 |
| 222 | Nerol | 0.498 | 0.61 | 0.27 | 0.66 | 1.4903 | 2.00 | 1.78 |
| 223 | Geraniol | 0.513 | 0.63 | 0.39 | 0.66 | 1.4903 | 2.00 | 1.79 |
| 224 | 1,8-Cineole | 0.383 | 0.33 | 0.00 | 0.76 | 1.3591 | 0.00 | 1.89 |
| 225 | Squalene | 0.760 | 0.55 | 0.00 | 0.59 | 4.0776 | 2.59 | 1.40 |
| 226 | Myristic acid | 0.150 | 0.67 | 0.57 | 0.39 | 2.1556 | 4.00 | 0.48 |
| 227 | Palmitic acid | 0.150 | 0.68 | 0.57 | 0.40 | 2.4374 | 4.12 | 0.27 |
| 228 | Stearic acid | 0.150 | 0.68 | 0.57 | 0.41 | 2.7192 | 4.21 | 0.08 |
| 229 | Oleic acid | 0.252 | 0.68 | 0.60 | 0.52 | 2.6762 | 3.95 | 1.37 |
| 230 | Linoleic acid | 0.384 | 0.76 | 0.60 | 0.59 | 2.6332 | 3.68 | 0.21 |

E: Excess molar refraction; S: dipolarity/polarizability; $A$ : hydrogen bond acidity; $B$ : hydrogen bond basicity; $V$ : McGowan's characteristic volume; $F$ : flexibility; $G$ : globularity.
phases [37,46,47], including additives would further complicate the understanding of the already complex phenomena. Thus, we used a simple carbon dioxide-methanol mobile phase in this study. Naturally, this can cause some peak tailing, particularly for the highly basic, possibly ionized solutes. Precautions were taken while interpreting the results from the retention data of these possibly ionized solutes.

Other effects of mobile phase composition will be discussed in future papers.

### 2.4. Data analysis and molecular modelling

Retention factors ( $k$ ) were calculated based on the retention time $t_{\mathrm{R}}$, determined using the peak maximum (even when tailing did occur) and the hold-up time $t_{0}$ measured on the first negative peak due to the unretained dilution solvent (always visible in these conditions) [48].

Abraham descriptors and $\log P$ values were determined with the Absolv Webboxes program, based on ADME Boxes version 3.5 (Pharma Algorithms, ACD Labs, Toronto, Canada). Whenever an exact match was found in the Absolv database, the experimental values were preferred. When no exact match could be found, the descriptors calculated by Absolv were used.

Extra descriptors (flexibility and globularity) were computed using MOE 2008.10 and 2009.10 (Chemical Computing Group, Montreal, Canada), and QikProp 2009/08/20 (Schrödinger). A stochastic conformational analysis was performed with MOE 2009.10 with the following parameters: maximum number of iterations 10000; RMS gradient 0.005; maximum number of rejected structures 50; MM iteration limit 500; RMSD limit 0.25 ; strain cut off 1 ; conformation limit 1 (so as to retrieve only the conformation with the minimum energy).

Multiple linear regression analyses and principal component analyses were performed using XLStat 7.5 software (Addinsoft, New York, NY).

## 3. Results and discussion

### 3.1. Solute descriptors

There were two important criteria in searching for the new descriptors. First of all, they must be relevant to and carry the information of the sterochemical interactions. For instance, although widely used, $\log P$ is rarely relevant to SFC processes, as will be discussed in a following section. Secondly, they must be as independent to the existing Abraham descriptors as possible. It is essential that each descriptor is encoding different properties, particularly for a multilinear regression analysis.

We identified flexibility and shape as two important properties for enantioselective chromatography, which were not included in the Abraham descriptors $E, S, A, B$ and $V$.

Flexibility (or rather non-flexibility or rigidity) is important for chiral resolution because flexible molecules have more conformers, thus more ways for intermolecular interactions, which dilutes the enantioselective interactions [15].

Globularity, on the other hand, can be related to steric impedance to insertion into the stationary phase, which is totally absent from the LSER approach. This essentially reflects the extra difficulty in inserting a "bulky" solute into a region of space of the stationary phase. In their hydrophobic subtraction model for the characterization of RP-HPLC stationary phases, Dolan and co-workers introduced a term related to steric hindrance ( $\sigma^{\prime} S^{*}$ ) [49-51]. This description term was quite effective in comparing stationary phases with different resistance to insertion, but it is unfortunately available only for a small number of solutes,
thus other comparable descriptors needed to be found. While the compounds used in the hydrophobic subtraction model are quite flexible, other authors have used rigid molecules with different three-dimensional structures to assess "shape selectivity" of the stationary phases: for example triphenylene - o-terphenyl in Tanaka and Euerby tests [52,53], tetrabenzonaphthalene benzo[a]pyrene by Sander and Wise [54], and isomers of $\beta$-carotene in Lesellier's carotenoid test [55], all used for ODS-type stationary phases. The property measured with flexible molecules could be different from that with rigid ones. A globularity term would be related to steric selectivity, more to rigid molecules than to flexible ones because globularity of flexible molecules is generally low, as they can be more easily flattened or extended in a rod-like manner.

Eight possible descriptors, which could reflect flexibility and globularity properties, were computed with MOE and QikProp software programs, as detailed in the "supplementary material" section.

The eight descriptors were calculated for all solutes in Table 1. Prior to the descriptors calculation, the solute structure needs to be energy minimized. To do so, a molecular mechanics force field must be selected based on the state and environment of the molecule. We rely on a gaseous state force field (MMFF94(s), which is appropriate for drug-like molecules [56]), on a neutral molecule. This means that the uncharged molecule is considered to be alone in space, not influenced by any other, to optimize its conformation.

As usually observed in molecular modelling, any simplification and assumption to facilitate a practical computation do come with a number of unanswered questions or issues. First of all, we have no clear idea of the cybotactic region in the vicinity of the molecule. Studies [57-59] suggested that the solute in a carbon dioxide-methanol mobile phase would be in a predominantly methanolic environment, but specific solvation of 1,2 -amino-alcohols by carbon dioxide was also reported [60]. Furthermore, other studies indicate that, if some aggregation of methanol molecules occurs in the supercritical mixtures, a homogeneous mixing of carbon dioxide and methanol exists in some regions [61]. As a result, the gaseous state force field might not perfectly model the molecule's environment in the supercritical mobile phase used in this study.

The very flexible molecules are obviously the most difficult to apprehend, as the complexity due to the possible conformational changes can expand exponentially beyond the range of a reasonable research possibility. Moreover, as a molecule approaches the stationary phase, its conformation may vary because intermolecular interactions between the solute and stationary phase could provide sufficient energy to rotate single bonds. Furthermore, solvation could result in conformational changes that affect the interaction between solutes and stationary phases.

Another important concern is the effect of temperature. In LSER studies, it is generally assumed that the solute parameters do not change with temperature. However, temperature could affect the states of flexibility and ionization of the molecule.

Since acidity of the supercritical mobile phase is unknown, as well as $\mathrm{pKa}^{*}$ of the ionisable solutes in this environment, the protonation state of ionisable species is difficult to evaluate. Moreover, $\mathrm{pH}^{*}$ in the stationary phase, depending on the solvation of the ligands, might be different from $\mathrm{pH}^{*}$ in the bulk mobile phase. This point will be further discussed in a following section.

Then, to investigate the possible correlations between the five Abraham descriptors and the eight suggested flexibility and shape descriptors, a principal component analysis (PCA) was performed using all 13 descriptors for the solutes in Table 1. The correlation of the new terms with Abraham descriptor V (molecular volume), was our major concern.

The PCA loading plot can be found as "supplementary material" (Figure S1).


Fig. 3. Principle of the augmented solvation parameter model: molecular properties and interactions related to each solute descriptor and coefficient.

Based on this analysis, we selected and introduced two additional descriptors in the solvation parameter model: the $G$ descriptor, representing globularity of the solute and calculated based on glob(Maestro), and the $F$ descriptor, representing solute flexibility based on b1rotR. A new solvation equation is thus proposed:
$\log k=c+e E+s S+a A+b B+v V+f F+g G$
The different interaction capabilities represented by the seven selected descriptors are represented in Fig. 3.

The possible use of the above equation naturally depends on the availability of Abraham descriptors and our additional descriptors.

Abraham descriptors are now available for a wide range of solutes in the literature, but for the model to have practical utility it will always be necessary to determine them for new solutes. This can be achieved through experiments, or simply with the help of a software program, as indicated in Section 2.

The $F$ descriptor can be simply calculated by hand for any new molecule, taking into account the ratio between rotatable bonds and total number of bonds. As b1rotR values are essentially comprised between 0 and 0.8 for a large majority of our test solutes, we multiply it by 5 , which brings the $F$ values in the range $0-4$ that is comparable to the Abraham descriptors, $E, S, A, B$ and $V$. It is important to ensure that the coefficient amplitude and errors are comparable in order to facilitate the proper interpretation of the coefficients.

The $G$ descriptor can also be easily calculated with any molecular modelling program providing solvent accessible surface area $(S)$ and the corresponding volume $(V)$, according to the following equation:
$\operatorname{glob}($ maestro $)=\frac{4 \pi(3 V / 4 \pi)^{2 / 3}}{S}$
Similarly to the $F$ descriptor, the glob value was scaled as follows: to the glob value calculated by Maestro, we deduce 0.75 then multiply the rest by 10 . This was done to obtain a more homogeneous scale with respect to the other descriptors. This way, flat and rod-like molecules have a $G$ value close to 0 , while most globular species have a $G$ value larger than 2 .

It must be pointed out that, since we have a highly diverse collection of probe solutes with limited cross-correlation, the correlations observed among the different descriptors is not
incriminated to the choice of the solutes. The building of the solute set and study on molecular descriptors were performed in a joint fashion, with the solute set being constructed bit by bit to ensure limited covariance between the descriptors.

### 3.2. Solute set

The procedure used for the characterization of stationary phases with the solvation parameter model, as for any quantitative structure-retention relationship, involves the analysis of a large set of test solutes. A solute set of about a hundred compounds was successfully used to characterize the achiral stationary phases in SFC in our previous works [35]. For the added complexity of the CSPs in this study, we doubled the size and diversity of the solute set in order to characterize enantioselective stationary phases with varied interaction capabilities. To generate retention data for such a large number of solutes, the speed advantage of SFC was very helpful. Only the solutes that eluted within an hour were retained in the final data set that contains 230 solutes, as shown in Table 1. The retention factors of the solutes are widely distributed with $\log k$ ranging from -1 to 2 .

Beside the number of solutes in the data set, the diversity and balance of the solutes with wide variety of chemical structures are essential. Although simple benzenic and naphthalenic monofunctional species used in our achiral stationary phase characterization were included, we added many solutes including drug-like molecules that differed in physico-chemical properties and in three-dimensional structures, as evidenced by the wide array of globularities. The final solute set covers a wide range of functional groups likely to be found in chiral analytes of pharmaceutical interest. With such a large and diverse data set, introduction of additional solutes would not significantly modify the results of the multilinear regression. For detection reasons, most of the chosen solutes contain an aromatic group, which may dominate the behaviour of the molecule, especially as the CSPs contain an aromatic moiety. Nevertheless, non-aromatic solutes are not negligible, not only because they are important types of solutes (amino-acids, steroids and terpenes), but also because they can be used to assess the contributions of aromaticity to retention.

First of all, it was observed on the score plot of the principal component analysis mentioned above that all compounds were
reasonably scattered in the PCA space. No outliers, who would indicate extreme characteristics, seemed to be present. Only the long-chain compounds as alkylbenzenes and fatty acids somewhat stood out of the distribution due to their large volume and flexibility. Special attention was thus given to the behaviour of these molecules in the following. It appeared in the models that were built later that none of them behaved as outliers.

The 230 solutes provide a uniform distribution of each solvation descriptor within a wide enough space and each descriptor covers a wide range. (Repartition of the solutes in each descriptor space can be observed on Figure S2 in the "supplementary material" section.) While clustering was avoided as much as possible, an exception was found in the $A$ descriptor for which a large proportion of solutes showed a value of zero. This is because of the very definition of this parameter; and because the number of solutes with significant hydrogen-bonding acidity is limited. In addition, because monofunctional benzenic and naphthalenic species that are not flexible represented a large part of the solute set, the proportion of rigid molecules (low $F$ value) was also relatively large.

Minima, maxima, average and standard deviation values for each descriptor can be found in Table S1 in the "supplementary material" section. Each descriptor covers a wide range that defines the applicability domain of the models to be established, which in turn will ensure the predictability from the models.

Moreover, cross-correlation should be avoided for meaningful data interpretation with multiple linear regression analysis. We relied on the orthogonal definition of the descriptors and the size and diversity of the solutes to ensure that the least possible covariance exist between the solute descriptors.

The Abraham descriptors were designed to provide complementary information and, as we have shown in the above paragraph, our newly introduced descriptors are orthogonal to the five older ones. Absence of cross-correlation inherent to the choice of solutes can be checked in the covariance matrix, presenting the determination coefficients between the descriptor values (Table S2 in the "supplementary material" section). Each descriptor was also plotted against another, and non-correlation was reflected by the random scatter of the data. The $E$ and $S$ descriptors appear to present some correlation; this was not unexpected as both $E$ and $S$ reflect some of the polarizability characteristics of the solute. We know certain correlation exists between the $E$ and $S$ descriptors, particularly when only aromatic solutes are used [36]. Thus a number of positional aromatic isomers and aliphatic solutes were included in the solute set, as both were helpful in breaking the covariance. All aliphatic solutes contain at least one double-bond, for UV detection at 210 nm . Structurally more complex solutes of pharmaceutical interest also clearly help in limiting the covariance.
$G$ also appears to be somewhat correlated to $V$ and $F$. Judging from the plots of one descriptor against the others, these high correlation coefficients are essentially due to a minority of solutes acting as levers, namely the long chain alkylbenzenes (compounds 6-15 in Table 1) and fatty acids (compounds 226-230 in Table 1). Indeed, covariance estimated through the correlation coefficient can be somewhat overestimated, because this coefficient can be strongly influenced by a few points acting as levers, while the rest of the points would be scattered. A better covariance coefficient, expressing total covariance that would be more representative of the whole distribution and less susceptible to levers would be useful, but we have not found anything satisfying so far.

### 3.3. Comparison of ADMPC and CDMPC retention to octanol-water partition

The polysaccharide CSPs are multimodal, meaning they can be used in normal-phase, or reversed-phase, or polar organic mode. A comparison of retention factors with octanol-water partition coef-


Fig. 4. Relationship between retention on the ADMPC phase $\left(\log k_{\mathrm{AD}}\right)$ and the octanol-water partition coefficient $(\log P)$. The red line is drawn across the alkylbenzene homologous series (solutes 2-15 in Table 1); the blue line is drawn across the PAH solutes (solutes 172-188 in Table 1); the green line is drawn across phenols (solutes 108, 113 and 116 in Table 1, green squares); the orange line is drawn across nitrotoluenes (solutes 2, 149 and 152 in Table 1, orange triangles). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)
ficients was carried out to assess the normal and reversed-phase character of SFC separations on ADMPC and CDMPC. The reason for this study is because it would be useful to have a method to transfer a HPLC separation to SFC.

The relation between $\log P$ and $\log k$ on ADMPC is shown (Fig. 4), whose trend is similar on CDMPC. The observed behaviour is typical of aromatic achiral stationary phases in SFC [31], with retention of non-polar homologous series (like alkylbenzenes, alkylphenones or phthalates) and PAHs showing positive correlations to $\log P$ (but with different slopes). A close examination of the retention of some polar solute families shows negative trends relating retention to $\log P$. For instance, homologous series with either hydroxyl group increment (phenol, resorcinol, phloroglucinol) or nitro group increment (toluene, o-nitrotoluene, 2,4-dinitrotoluene) showed a good correlation between retention factors and octanol-water partition with negative slopes. Therefore, the results suggest that some normal-phase trends exist for polar solutes, and reversedphase trends exist for non-polar solutes in pSFC. This is a first indication of why chiral HPLC and chiral SFC may not be equivalent, and method transfer between the techniques is often non-practical.

### 3.4. Comparison of ADMPC and CDMPC retention to non-enantioselective stationary phases

To examine the effect of the aromatic moieties on the retention process without the stereospecific contribution from the polysaccharide backbones, we compared ADMPC and CDMPC with achiral stationary phases, bonded with similar ligands. Among the commercially available columns, we found two having ligands resembling the 3,5-dimethylphenyl-carbamate moieties, a phenyl-oxy-propyl-bonded silica phase (Phenomenex Synergi Polar RP, named OPHE in the following), and an ODS phase with a carbamateembedded group (Waters XBridge Shield C18). OPHE allowed the comparison of the effects of the aromatic group, with a little contribution of a polar functional group in the spacer arm. The polar


Fig. 5. $\kappa-\kappa$ plots comparing the retention on the ADMPC phase ( $\log k_{\mathrm{AD}}$ ) to the retention on two achiral stationary phases (a) Synergi Polar RP (log $k_{\text {OPHE }}$ ) and (b) XBridge Shield C18 $\left(\log k_{\mathrm{C} 18}\right)$.
embedded C18 phase was for the effects of a carbamate function on retention.

The $\kappa-\kappa$ plot (Fig. 5a) that compares logarithms of retention factors on ADMPC to that on OPHE displays a decent correlation coefficient ( $R^{2}=0.69$ ). The correlation with CDMPC (not shown) is similar ( $R^{2}=0.67$ ). Retention of most compounds was higher on the polysaccharide phases than on the phenyl-ether bonded phase. Methylene selectivity was very close on both phases, while the alkylbenzene series appeared to be closely matching the first bisector. This suggests that density of the accessible aromatic functions in both columns is similar. It is also noted that the large, polyfunctional solutes (essentially among drug molecules 192-212) are scattered at a larger distance from the regression line, while smaller solutes with only one polar function correlate better. This suggests that solutes possessing multiple functional groups tend to have more stereo-induced interactions with the polysaccharide phases than simpler molecules. Stereo-induced interactions occur when the three-dimensional functionality of the molecule allows joint interactions with different parts of the CSP, causing more retention when attractive interactions occur (cooperative adsorption), or less retention when repulsive interactions occur.

The correlation with the carbamate-embedded ODS phase (Fig. 5b) was much worse ( $R^{2}=0.36$ and 0.42 respectively). Again the points representing the alkylbenzene solutes were close to the first bisector and aligned in a parallel fashion to this line, indicating identical methylene selectivity. All other points are below the first bisector, indicating that the polysaccharide phases allowed more retention than the ODS phase for polar solutes.

In summary, the data suggest that the aromatic ring is the principal function participating in retention on the ADMPC and CDMPC phases, while the carbamate group is of secondary importance. However, we must point out that the carbamate-embedded C18 phase may not be an adequate mimic to the carbamate function in the CSPs, because the carbamate function in the achiral stationary phase is buried under the long alkyl chain and thus far less accessible as compared to the carbamate linkage in the CSPs.

In addition, the scattering of the points from the larger and structurally more complicated drug-like molecules suggests that shape does play an important role in the chromatographic retention process on the polysaccharide CSPs. The orientation of the polysaccharide CSPs likely creates stereo-induced interactions, leading to
different retentions compared to the non-stereoselective stationary phases.

### 3.5. Comparison of ADMPC and CDMPC based on $\kappa-\kappa$ plots

The $\kappa-\kappa$ plot (Fig. 6) comparing logarithms of retention factors on CDMPC to ADMPC displayed a high correlation coefficient ( $R^{2}=0.73$ ). Since CDMPC and ADMPC have identical bonding ligands on the polysaccharide backbone, retention correlation was expected especially as the solutes included many monofunctional molecules. The fact that the two columns did not exhibit perfect correlation for nearly $25 \%$ of the solutes indicates the significant


Fig. 6. $\kappa-\kappa$ plot comparing the retention on the ADMPC phase $\left(\log k_{\mathrm{AD}}\right)$ to the retention on the CDMPC phase $\left(\log k_{\mathrm{OD}}\right)$. The continuous black line is the linear regression line; the interrupted red line is the first bisector; the dotted blue line is a parallel to the first bisector relating the retention of the alkylbenzene homologous series (solutes 2-15 in Table 1). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)
difference in stereochemical interactions resulted from the very different macromolecular structures of the two CSPs.

A large proportion of the solute points were close to the first bisector (interrupted red line), which would, on first glance, indicate generally similar retention behaviours. However, this observation does not resist long to closer examination of the scatter of points. Indeed, the alkylbenzene homologous series appear to be aligned in a parallel fashion to the first bisector but along a line that is above the first bisector. This indicates that the sites available to interact with the solutes are more on the CDMPC phase than on the ADMPC. In other words, phase ratios are different between the two columns. To compensate that, the retentions of solutes were compared based on the line drawn through the alkylbenzene solutes (dotted blue line). Most points of the polar solutes are situated below the blue line, indicating that polar solutes are more retained on the ADMPC phase than on the CDMPC phase. This suggests that polar interactions are stronger with the ADMPC phase than with the CDMPC phase. This difference is unlikely a result of the differential density of the accessible phenylcarbamate ligands on the two columns, because, on the contrary, more interaction sites seem to be present in the CDMPC phase as it provides more retention for non-polar solutes. Thus it is reasonable to propose that different supramolecular structures or polymorphisms of the two phases cause the different intermolecular interactions between the solutes and the stationary phases. More thorough study using multiple linear regressions was conducted and is discussed in the following section.

Under a closer look at the isomeric species (for instance compounds 125-156 in Table 1), it appeared that the elution orders within each isomer family were completely different on the two phases, and the elution order was not consistent from one isomer family to another. For instance, ADMPC showed total co-elution of ethylbenzene and the three xylene isomers, while all four of them were separated on CDMPC. Other groups of solutes showed similar trends: phenanthrene and anthracene, both three-ringed polynuclear aromatic solutes with different three-dimensional shape, were not separated on ADMPC while they were separated on CDMPC. Although these are relatively simple molecules, the data suggest that the cellulose-based phase (CDMPC) was more sensitive to solute shape when no other solute property (as hydrogenbonding or polarity) was affected by this change of shape. Actually, this is not entirely true as electron repartition in the molecule is different for solutes of different shape, so this particular behaviour could be related either to steric effects or to dipole-dipole type interactions.

The CDMPC phase, thus, seems to display specific shape recognition towards aromatic non-polar solutes, which was not observed on ADMPC. From a practical point of view, this could be interesting for the separation of hydrocarbon species with little or no polar functionalities, like terpenes.

For polar isomeric species, the situation was much more complicated. For example, isopropylphenols had the same elution order on the two phases, while cresols had different elution orders. Moreover, on CDMPC, the nitrotoluenes, nitrobenzylalcohols and nitrophenol isomer families all eluted in the ortho, meta, para order. However, the elution orders of the three isomer families were all different on ADMPC. Simple considerations on the polarity and hydrogen-bonding capabilities of the different isomers are not sufficient to explain these seemingly erratic elution order differences. The supramolecular structures of the CSPs that influence (either enhance or eliminate) the intermolecular interactions such as hydrogen-bonding with the solutes must play a key role in the elution order.

While limited conclusions can be drawn from the $\kappa-\kappa$ plot, a more detailed study based on the solvation parameter model was conducted to probe the differences between the two phases.
3.6. Comparison of ADMPC and CDMPC based on the modified solvation parameter model

Non-enantiospecific interactions that are evaluated in this study using achiral solutes are important, because some of them may actively participate in the enantio-recognition process, while others might be "in excess" of those required for the chiral recognition process. They may also influence the magnitude of enantioselectivity of the chiral resolution by increasing retention: long retention with small separation factors can generally be attributed to strong non-enantioselective interactions.

The system constants for each chromatographic system, presented in Table 2 and Fig. 7, are obtained by multiple linear regression analysis on the logarithm of the measured retention factors $(\log k)$, regressed against the seven molecular descriptors, according to Eq. (2). The presented data were obtained after elimination of outliers (as further detailed below) from the remaining 208 and 200 experimental retention factors on ADMPC and CDMPC respectively

The quality of the fits on ADMPC and CDMPC was estimated using the adjusted determination coefficient ( $R_{\text {adj }}^{2}=$ 0.84 and 0.89 respectively), standard error in the estimate ( 0.20 and 0.14 respectively) and Fischer $F$ statistic (279 and 373 respectively). Statistical significance of each individual coefficient to explain retention was assessed using the $t$-ratio, which is defined as the ratio of the regression coefficient to its standard error. Coefficients that were not significant, at the $95 \%$ confidence level, were eliminated from the model.

Graphs of the residuals (difference between the experimental and predicted $\log k$ values) plotted against $\log k$ or against the values of each individual descriptor showed no correlation.

Since some complex pharmaceutical compounds are included in the solute set, the fit quality of the regression is not as good as those generally observed on achiral stationary phases.

The limited quality of the regression may also be related to the $\mathrm{pH}^{*}$ of the supercritical mobile phase. Some studies [62-64] indicated that the $\mathrm{pH}^{*}$ of supercritical carbon dioxide-methanol mixtures is most probably acidic. Indeed the $\mathrm{pH}^{*}$ of the carbon dioxide-methanol mixture used in this study may be between 3 and 4 [59]. Thus, it is highly probable that very acidic solutes such as trimesic acid are in their anionic form, while basic solutes are in their protonated form. Some basic solutes (among solutes 25-59 and among the drugs, 192-212) were found to be extreme outliers from the model calculation evidenced on Figure S3 in the supplementary material, that shows the relationship between the experimental and the predicted retention factors on ADMPC and CDMPC. For instance, isophthalic acid ( $\mathrm{n}^{\circ} 61$ ), trimesic acid ( $\mathrm{n}^{\circ} 62$ ), benzylamine ( $\mathrm{n}^{\circ} 30$ ), o-phenylenediamine ( $\mathrm{n}^{\circ} 33$ ), 2-phenylethylamine ( $\mathrm{n}^{\circ} 34$ ), nicotinamide ( $\mathrm{n}^{\circ} 43$ ), amitriptyline ( $\mathrm{n}^{\circ} 192$ ), mephentermine ( $\mathrm{n}^{\circ} 202$ ), noscapine ( $\mathrm{n}^{\circ} 204$ ), procaine ( $\mathrm{n}^{\circ} 211$ ) or quinine ( $\mathrm{n}^{\circ} 212$ ) are all quite far from the first bisector. These are not "chance" outliers as they are all N -containing bases except the two acids that are most likely in their anionic forms. Since oxygen-containing compounds with similar capacity for H-bond and dipole-type interactions were not influenced to the same extent, we presume that the additional retention may be attributed to interactions that are not considered by the model, such as electrostatic interactions with partially dissociated residual silanol groups under the polysaccharide coating. Ion-dipole interactions between an ionic solute and the polar carbamate moiety may also be possible.

We had already encountered this problem in achiral SFC [31]. One solution would be to use additives to suppress ionization. However, the mechanism of interaction could be largely dictated by the type and composition of the mobile phase as it changes both the solute and the stationary phase. Thus, it is desirable to keep the

Table 2
System constants and statistics for both columns.

| Stationary phase | $c$ | $e$ | $a$ | $b$ | $g$ | $n$ | $R_{\text {adj }}^{2}$ |  |
| :--- | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :--- |
| ADMPC | -0.759 | 0.731 | 0.718 | 0.338 | -0.164 | 208 | 0.843 |  |
|  | 0.070 | 0.029 | 0.045 | 0.048 | 0.039 |  | 0.20 |  |
| CDMPC | -0.543 | 0.694 | 0.535 | 0.175 | -0.181 | 200 | 0.882 |  |
|  | 0.050 | 0.021 | 0.033 | 0.037 | 0.028 |  | 0.14 |  |
|  |  |  |  |  |  |  |  |  |

$n$ is the number of solutes considered in the regression, $R_{\mathrm{adj}}^{2}$ is the adjusted correlation coefficient, SE in the standard error in the estimate, $F$ is Fischer's statistic and the numbers in italics represent $95 \%$ confidence limits.
mobile phase simple and constant while characterizing the stationary phases.

Since all current descriptors are calculated for neutral solutes without considering electrostatic interactions, the solvation parameter model is not expected to provide accurate predictions of chromatographic properties of solutes in a fully or partially ionized form. Using additional terms for ionisable solutes was suggested by some researchers [65-70], but these descriptors require exact knowledge of the $\mathrm{pH}^{*}$ and $\mathrm{pKa}^{*}$ of all the species. This will be addressed in future works.

Despite the issue with some ionic solutes, the results are surprisingly good, considering the complexity of the chromatographic systems. In addition, the large data set included in the final regression analysis (more than 200 on both columns) gives us the confidence in the results. In fact, the sign and magnitude of each regression coefficient obtained are well in accordance with the intuition for the chemical nature of the chromatographic systems.

The vast majority of the variance in the retention information can be explained with the seven selected descriptors.

The interaction coefficients or system constants obtained on both CSPs are represented in Fig. 7. Each coefficient represents the difference, for a particular type of interaction, between the solute-stationary phase interactions and the solute-mobile phase interaction. Thus, when a particular coefficient is zero, it does not mean that this particular type of interaction does not occur, but rather that they are of the same magnitude between solute and stationary phase on one hand, solute and mobile phase on the other hand. As a result, the coefficient is equal to zero as the overall interaction is cancelled and becomes insignificant in explaining the retention in the chromatographic system.

The $c$ intercept (not shown in the figure), which is solute independent and not related to a specific interaction, is significantly
more negative on ADMPC than on CDMPC, indicating that phase ratio are somewhat different on the two phases, with more stationary phase available for interaction in the CDMPC phase than in the ADMPC phase. This is consistent with the above conclusions based on the $\kappa-\kappa$ plot.

With respect to the two additional descriptors, it appears that globularity is significant in both ADMPC and CDMPC retention models, while flexibility is not significant to explain retention. Although the $g$ coefficient values are relatively small, both the $t$-tests and error bars indicate that they are statistically significant. We continue to study the retention and separation process using these new descriptors. It will be shown in our subsequent papers that the relevance of both the $F$ and $G$ descriptors is significant to enantioselective separations. Relevance should be evaluated from long-term empirical success, and not from one failure to produce a QSRR model.

The $s$ coefficient was not significant to explain retention on these two CSPs. We were concerned that it might be related to crosscorrelation between the $E$ and $S$ descriptors. However, our further work with other CSPs proves $s$ to be significant, as will also be discussed in our subsequent papers. Similarly, non-significance of the $v$ coefficient in this case cannot be attributed to cross-correlation between the $V$ and $G$ descriptors, as will appear for different chromatographic systems.

The most significant coefficients appeared to be the $e, a$ and $b$ on both CSPs. This is very explainable and consistent with the chemical nature of the ligand, 3,5-dimethylphenyl-carbamate, on these two CSPs.

We must point out that comparison of the normalised coefficients yields identical conclusions to those based on nonnormalised coefficients. This is the result of an appropriate scaling of descriptors, as indicated by the comparable range, average and


Fig. 7. System constants issued from the multiple linear regression analysis of retention factors on (a) ADMPC and (b) CDMPC, based on Eq. (2).
standard deviation values (Table S1 in the "supplementary material" section). Consequently, only non-normalised coefficients are discussed here.

The $e$ coefficient is the largest contributor, which indicates that $\pi-\pi$ interactions between solute and the phenyl groups on the CSPs principally contribute to retention. As in the above comparisons to achiral stationary phases, the same conclusion that the aromatic ring is the major contributor to retention on these CSPs is derived. The interactions with the carbonyl group of the carbamate function might also participate in the $e$ term. Furthermore, not only $\pi-\pi$ interactions but also induced dipole-dipole interactions of Keesom type may contribute to the large $e$ value. This is also in accordance with the chemical intuition of these CSPs: aromatic groups are known to be located on the surface of the polysaccharide helices, while the less polar carbamate groups are located in the interior [71].

The large $a$ coefficient is certainly related to strong hydrogenbonding between acidic solutes and the carbonyl group of the carbamate spacer arm. Interaction with accessible silanol groups of the base silica might also contribute to this term. Methanol molecules adsorbed in the stationary phase could also promote hydrogen-bonding, which will be investigated in our further study for the effects of different mobile phase modifiers.

The positive $b$ coefficient is most likely related to hydrogenbonding between basic solutes and the proton-donor, - NH- group of the carbamate function, on CSPs. Again, accessible silica and adsorbed mobile phase components might also contribute to this term.

The smaller values of the $a$ and $b$ coefficients compared to the $e$ coefficient (especially on CDMPC) are also consistent with the above observation that interaction with the carbamate group is of secondary importance to retention compared to the aromatic group.

While the $e$ coefficient was not significantly different between the two CSPs, both $a$ and $b$ coefficients were somewhat smaller on CDMPC than on ADMPC. This indicates that acidic and basic solutes are less retained on CDMPC than on ADMPC. In other words, ADMPC creates more opportunities for hydrogen-bonding than CDMPC. This is consistent with the finding that polar interactions are stronger on ADMPC than CDMPC, from the $\kappa-\kappa$ plots discussed previously.

As the $a / b$ ratios on both phases are similar, more carbamate groups must be accessible on ADMPC, because the carbamate function participates in both hydrogen bonding modes, as a donor and an acceptor through the carbonyl and NH groups respectively. As shown in Fig. 2, representing possible structures of amylose and cellulose, the spatial arrangement of the chiral grooves is quite different between the two CSPs. Thus the difference in supramolecular structure of ADMPC and CDMPC must be responsible for these differences in the intermolecular interactions in terms of retention and elution orders as discussed above.

The negative $g$ coefficients on both ADMPC and CDMPC indicate that bulky solutes are less retained than flat solutes. In other words, flat and rod-like solutes are preferentially retained on both stationary phases, compared to more compact solutes of identical volume. This makes sense as compact molecules may have fewer possibilities for interacting with the stationary phase, while a more "extended" molecule would have more possibilities, thus should be more retained. Also, it seems reasonable that a spherical molecule would have more difficulties in entering the stationary phase cavities, thus would have less accessible stationary phase to interact with and would be less retained.

Removal of this coefficient from the LSER model does not strongly affect the statistical significance of the whole equation, but we will show in the following part of this study that it is very significant to enantioseparation.

All in all, the reasonable agreement between the experimental results and chemical intuition provides us the confidence in the chemical interpretation of the coefficients.

## 4. Conclusions

In the first paper of this series, we have demonstrated the usefulness of the solvation parameter model to measure nonenantiospecific interactions contributing to the retention of achiral solutes on two polysaccharide CSPs (ADMPC and CDMPC) in supercritical fluid chromatography.

We have also defined two additional descriptors based on flexibility and globularity of the solutes. These individual descriptors were found to be only weakly correlated to each other and to the classical five Abraham descriptors over a suitably large collection of compounds.

Reasonably good QSRR models of retention factors of over 200 achiral solutes were obtained on both ADMPC and CDMPC phases, where the new globularity descriptor appeared to be significant.

CDMPC and ADMPC generally show similar capabilities for intermolecular interactions, which is consistent with the fact that they possess the same bonded ligand (3,5-dimethylphenylcarbamate). On both stationary phases, $\pi-\pi$ interactions with the aromatic groups and carbonyl groups of the carbamate moiety were the major contributor to retention, immediately followed by donor and acceptor hydrogen-bonding with the carbamate group. Some differences appeared in the amplitude of hydrogen-bonding coefficients (larger on ADMPC than on CDMPC) between the two CSPs.

Since achiral solutes were used, the study described in this paper is more relevant to the interactions contributing to retention than to the separation of a given enantiomer. Nevertheless, the study based on the modified solvation parameter provided a good starting point for further assessment of interactions contributing to enantiomeric separations. That is the topic of the second part of this series.

## Acknowledgments

Nathalie Percina is acknowledged for highly appreciated technical assistance with chromatographic analyses. Guillaume Guenegou is acknowledged for assistance with molecular modelling and chemometrics.

## Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.chroma.2010.11.084

## References

[1] D. Mangelings, Y. Vander Heyden, J. Sep. Sci. 31 (2008) 1252.
[2] R. Mukhopadhyay, Anal. Chem. May (2008) 3091.
[3] L.T. Taylor, Anal. Chem. 82 (2010) 4925.
[4] K.W. Phinney, Anal. Chem. 72 (5) (2002) 204A.
[5] C.J. Welch, W.R. Leonard Jr., J.O. DaSilva, M. Biba, J. Albaneze-Walker, D.W. Henderson, B. Laing, D.J. Mathre, LC-GC North Am. 23 (1) (2005) 16.
[6] L. Miller, M. Potter, J. Chromatogr. B 875 (2008) 230.
[7] E. Abbott, T.D. Veenstra, H.J. Issaq, J. Sep. Sci. 31 (2008) 1223.
[8] T.Q. Yan, C. Orihuela, J. Chromatogr. A 1156 (2007) 220.
[9] M. Maftouh, C. Granier-Loyaux, E. Chavana, J. Marini, A. Pradines, Y. Vander Heyden, C. Picard, J. Chromatogr. A 1088 (2005) 67.
[10] C. White, J. Burnett, J. Chromatogr. A 1074 (2005) 163.
[11] Y. Zhao, W.A. Pritts, S. Zhang, J. Chromatogr. A 1189 (2008) 245.
[12] A. Van Overbeke, P. Sandra, A. Medvedovici, W. Baeyens, H.Y. Aboul-Enein, Chirality 9 (1997) 126.
[13] P. Borman, B. Boughtflower, K. Cattanach, K. Crane, K. Freebairn, G. Jonas, I. Mutton, A. Patel, M. Sanders, D. Thompson, Chirality 15 (2003) S1.
[14] P.A. Mourier, E. Eliot, M.H. Caude, R.H. Rosset, A.G. Tambute, Anal. Chem. 57 (1985) 2819.
[15] S. Svensson, A. Karlsson, O. Gyllenhaal, J. Vessman, Chromatographia 51 (2000) 283.
[16] C.M. Kraml, D. Zhou, N. Byrne, O. McCornell, J. Chromatogr. A 1100 (2005) 108.
[17] L. Toribio, M. Jesús del Nozal, J.L. Bernal, C. Alonso, J.J. Jiménez, J. Chromatogr. A 1144 (2007) 255.
[18] J. Qian-Cutrone, B. Dasgupta, E.S. Kozlowski, R. Dalterio, D. Wang-Iverson, V.M. Vrudhula, J. Pharm. Biomed. Anal. 48 (2008) 1120.
[19] C. Wenda, A. Rajendran, J. Chromatogr. A 1216 (2009) 8750.
[20] ChirBase Project, ENSSPICAM, University of Aix-Marseille III, Marseille, France.
[21] P. Piras, C. Roussel, J. Pierrot-Sanders, J. Chromatogr. A 906 (2001) 443.
[22] G. Hesse, R. Hagel, Chromatographia 6 (1973) 277.
[23] T. Shibala, K. Mori, Y. Okamoto, in: A.M. Krstulovic (Ed.), Polysaccharide Phases in Chiral Separation by HPLC, Ellis Horwood, New York, 1989, p. 336.
[24] T.D. Booth, I.W. Wainer, J. Chromatogr. A 737 (1996) 157.
[25] T.D. Booth, I.W. Wainer, J. Chromatogr. A 741 (1996) 205.
[26] C. Roussel, C. Suteu, J. Chromatogr. A 761 (1997) 129.
[27] C. Roussel, B. Bonnet, A. Piederriere, C. Suteu, Chirality 13 (2001) 56.
[28] R.B. Kasat, N.-H.L. Wang, E.I. Franses, J. Chromatogr. A 1190 (2008) 110.
[29] S. Ma, S. Shen, H. Lee, M. Eriksson, X. Zeng, J. Xu, K. Fandrick, N. Yee, C. Senanayake, N. Grinberg, J. Chromatogr. A 1216 (2009) 3784.
[30] K. Héberger, J. Chromatogr. A 1158 (2007) 273.
[31] C. West, E. Lesellier, J. Chromatogr. A 1191 (2008) 21.
[32] C. West, L. Fougère, E. Lesellier, J. Chromatogr. A 1189 (2008) 227.
[33] M.H. Abraham, A. Ibrahim, A.M. Zissimos, J. Chromatogr. A 1037 (2004) 29.
[34] C.F. Poole, S.K. Poole, J. Chromatogr. A 965 (2002) 263.
[35] C. West, E. Lesellier, in: E. Grushka, N. Grinberg (Eds.), Advances in Chromatography, vol. 48, CRC Press, Taylor and Francis, 2010, p. 195.
[36] M.F. Vitha, P.W. Carr, J. Chromatogr. A 1126 (2006) 143
[37] J.A. Blackwell, R.W. Stringham, Chirality 11 (1999) 98.
[38] J. Lokajová, E. Tesařová, D.W. Armstrong, J. Chromatogr. A 1088 (2005) 57.
[39] A. Berthod, C.R. Mitchell, D.W. Armstrong, J. Chromatogr. A 1166 (2007) 61.
[40] C.R. Mitchell, D.W. Armstrong, A. Berthod, J. Chromatogr. A 1166 (2007) 70.
[41] C.R. Mitchell, N.J. Benz, S. Zhang, J. Chromatogr. B 875 (2008) 65.
[42] G. Felix, A. Berthod, P. Piras, C. Roussel, Sep. Purif. Rev. 37 (2008) 229.
[43] X. Chen, C. Yamamoto, Y. Okamoto, Pure Appl. Chem. 79 (9) (2007) 1561.
[44] C. Yamamoto, E. Yashima, Y. Okamoto, Bull. Chem. Soc. Jpn. 72 (1999) 1815.
[45] C. Yamamoto, E. Yashima, Y. Okamoto, J. Am. Chem. Soc. 124 (2002) 12583.
[46] J.A. Blackwell, Chirality 10 (1998) 338.
[47] Y.K. Ye, K.G. Lynam, R.W. Stringham, J. Chromatogr. A 1041 (2004) 211.
[48] K. Gurdale, E. Lesellier, A. Tchapla, J. Chromatogr. A 866 (2000) 241.
[49] N.S. Wilson, M.D. Nelson, J.W. Dolan, L.R. Snyder, R.G. Wolcott, P.W. Carr, J. Chromatogr. A 961 (2002) 171.
[50] N.S. Wilson, M.D. Nelson, J.W. Dolan, L.R. Snyder, P.W. Carr, J. Chromatogr. A 961 (2002) 195.
[51] N.S. Wilson, J.W. Dolan, L.R. Snyder, P.W. Carr, L.C. Sander, J. Chromatogr. A 961 (2002) 217.
[52] M. Euerby, P. Petersson, J. Chromatogr. A 994 (2003) 13.
[53] K. Kimata, K. Iwaguchi, S. Onishi, K. Jinno, R. Eksteen, K. Hosoya, M. Araki, N. Tanaka, J. Chromatogr. Sci. 27 (1989) 721.
[54] L.C. Sander, S.A. Wise, Anal. Chem. 56 (1984) 504.
[55] E. Lesellier, C. West, A. Tchapla, J. Chromatogr. A 1111 (2006) 62.
[56] T.A. Halgren, J. Comput. Chem. 17 (1996) 490.
[57] A.P. Abbott, E.G. Hope, D.J. Palmer, J. Phys. Chem. B 111 (2007) 8119.
[58] C.R. Yonker, R.D. Smith, J. Phys. Chem. 92 (1988) 2374.
[59] C. West, M. Mengue-Metogo, E. Lesellier, Poster at the HPLC 2009 Symposium in Dresden, 2009.
[60] N. Bargmann-Leyder, C. Sella, D. Bauer, A. Tambute, M. Caude, Anal. Chem. 67 (1995) 952.
[61] R.L. Smith Jr., C. Saito, S. Suzuki, S.-B. Lee, H. Inomata, K. Aria, Fluid Phase Equilibr. 194-197 (2002) 869.
[62] D. Wen, S.V. Olesik, Anal. Chem. 72 (2000) 475
[63] K.L. Toews, R.M. Shroll, C.M. Wai, N.G. Smart, Anal. Chem. 67 (1995) 4040.
[64] J. Zheng, Supercritical Fluid Chromatography of Ionic Compounds, PhD thesis, Virginia Polytechnic Institute, Blacksburg, VI, 2005.
[65] J. Li, J. Chromatogr. A 982 (2002) 209.
[66] J. Li, Anal. Chim. Acta 522 (2004) 113.
[67] M. Rosés, D. Bolliet, C.F. Poole, J. Chromatogr. A 829 (1998) 29.
[68] M.H. Abraham, Y.H. Zhao, J. Org. Chem. 69 (2004) 4677.
[69] Y.H. Zhao, M.H. Abraham, A.M. Zissimos, J. Chem. Inf. Comput. Sci. 43 (2003) 1848.
[70] J. Li, J. Sun, S. Cui, Z. He, J. Chromatogr. A 1132 (2006) 174.
[71] R. Thompson, J. Liq. Chromatogr. Rel. Technol. 28 (2005) 1215.


[^0]:    * Corresponding author. Tel.: +33 238 494778; fax: +33 238417281.

    E-mail address: caroline.west@univ-orleans.fr (C. West).

