



Insights into chiral recognition mechanisms in supercritical fluid chromatography. II. Factors contributing to enantiomer separation on *tris*-(3,5-dimethylphenylcarbamate) of amylose and cellulose stationary phases

Caroline West^{a,*}, Guillaume Guenegou^a, Yingru Zhang^b, Luc Morin-Allory^a

^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

^b Bioanalytical and Discovery Analytical Sciences, Bristol-Myers Squibb Company, PO Box 4000, Princeton, NJ 08648-4000, USA

ARTICLE INFO

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
Factorial discriminant analysis

ABSTRACT

In this second part of our work on enantioselective supercritical fluid chromatography (SFC), we investigate the factors participating in the chiral recognition process on *tris*-(3,5-dimethylphenylcarbamate) of amylose and cellulose chiral stationary phases (CSPs). 135 racemates with diverse structures were analysed under identical SFC conditions on both stationary phases. The possibility of identifying the differential interactions of an enantiomer pair within the chromatographic system is assessed using a modified version of the solvation parameter model and factorial discriminant analysis. It is illustrated that one relationship of intermolecular interactions is insufficient to express the enantioseparation of different groups of racemates. An innovative approach is used in unravelling the interactions taking part in the enantiorecognition process. Different intermolecular interactions participating in the enantiomeric separation are demonstrated between the two stationary phases.

© 2010 Elsevier B.V. All rights reserved.

1. Introduction

The chromatographic resolution of enantiomers is one of the most difficult challenges in analytical chemistry. Retention, efficiency and selectivity all constitute a chromatographic resolution. To achieve a resolution, a minimum retention is required; the column efficiency enhances the resolution following the square root of relationship, while resolution is directly proportional to the selectivity increase. Supercritical fluid chromatography (SFC), which uses supercritical or near critical CO₂ as the bulk mobile phase, provides higher efficiency than HPLC due to the higher diffusivity of CO₂. Thus choosing SFC for our study ensured that efficiency was maximized. Our effort to achieve chiral resolution was focused on enantioselectivity as selectivity has the largest impact on the chromatographic resolution of enantiomers.

Studies of enantioselectivity are of great interest to acquire a better understanding of chiral separations, as well as to provide an insight into developing more selective chiral stationary phases (CSPs). Although numerous studies have been published

for chiral separations on derivatized polysaccharide CSPs in HPLC and SFC, the chiral separation mechanism has not yet been fully elucidated. This is essentially due to the complexity of the macromolecular CSPs in which multiple interaction sites with different affinity exist, and the understanding of polysaccharide polymorph structures is still controversial. Thus, simple rules, such as the three-point interactions that were established primarily for Pirkle-type CSPs, are impossible to establish for the polysaccharide CSPs. As a result, despite the large body of experimental data [1], most publications cover applications or research related to one particular analyte or analyte family, and no clear guideline is available for choosing a stationary and mobile phase or predicting whether a separation is achievable for a given racemate.

Before predicting enantioresolution, we want to first improve our understanding of the chiral recognition mechanism on the most popular derivatized polysaccharide CSPs. A complete review on the use of chemoinformatic techniques to explore enantioselective recognition mechanisms was recently published by Del Rio [2]. Molecular modelling techniques can be helpful in elucidating the mechanism of enantioselective recognition. However, although some examples exist on polysaccharide CSPs [3,4], these techniques are most successful when the stationary phase structure is not too complex, to allow practical and reliable computation. In this respect, Pirkle-type phase can be more suitably studied using

* Corresponding author. Tel.: +33 238 494778; fax: +33 238 417281.
E-mail address: caroline.west@univ-orleans.fr (C. West).

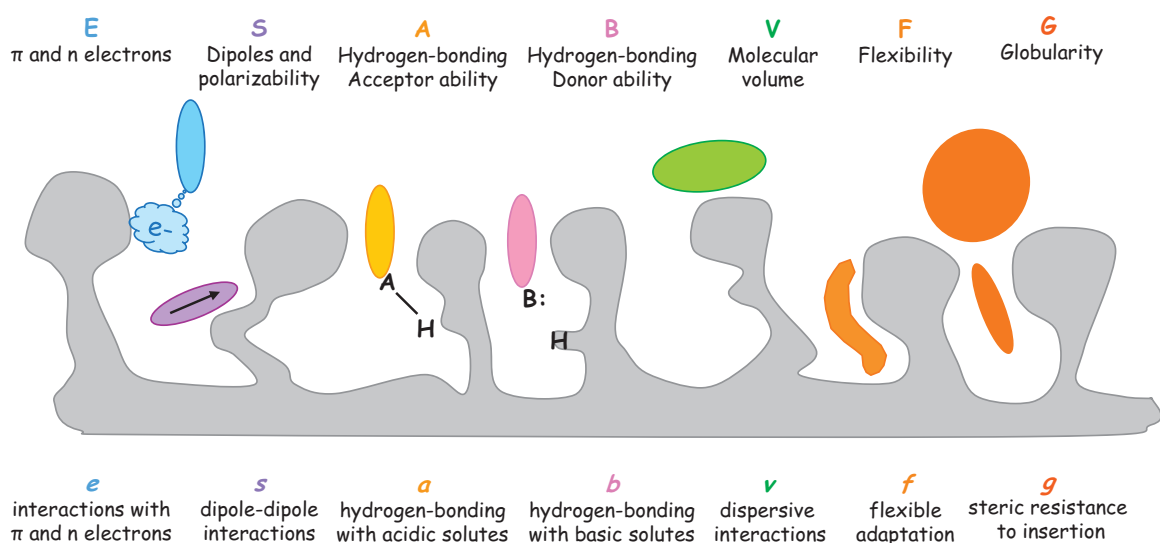


Fig. 1. Principle of the augmented solvation parameter model: molecular properties and interactions related to each solute descriptor and coefficient.

computational methods than the macromolecular polysaccharide phases.

Another approach is to try and relate the separation factor (α) to molecular descriptors, which can be called quantitative structure–separation relationships (QSSRs), or quantitative structure–enantioselective retention relationships (QSERRs). Such studies have been reported [5–7], but most of which focused on separation prediction potential rather than understanding the mechanism.

When considering the separation of enantiomers, one would naturally look for a descriptor that would differentiate the two enantiomers. Some rare examples of chirality-sensitive descriptors were reported [8–10]. However, as demonstrated by Del Rio and Gasteiger [11], the separation of an enantiomeric pair can be simply related to achiral molecular descriptors, because individual intermolecular interactions between the enantiomers and the CSP are achiral, while collectively they differentiate the two enantiomers. The intrinsic chirality of the solutes, while resulting in a rotation of polarized light, would not yield chromatographic separation. Moreover, numerous examples in the literature suggest that molecules sharing similar structure properties also offer a common mode of enantioselective recognition, thus the structural features would be critical to chiral interaction mechanisms, rather than the possible difference between the enantiomers.

Most existent molecular descriptors are not easily interpreted with respect to intermolecular interactions, thus are not very helpful in decrypting the enantio-recognition mechanisms. In the present study, we have privileged the use of descriptors, which have more straightforward meanings, based on our previous works. The aim of the present study was to improve the understanding of chiral recognition mechanisms, while the prediction capability will be studied later.

Recent studies of Mitchell et al. [12,13] showed that the solvation parameter model [14] can provide relevant information about the interactions contributing to retention and enantioselective separation on different macrocyclic glycopeptide and immobilized polysaccharide-based CSPs in HPLC.

In the first part of this series, we have shown how a modified version of the solvation parameter model could provide an understanding of the nature and strength of the intermolecular interactions controlling retention on two polysaccharide CSPs in SFC [15] (this issue).

The retention of selected probes is related to specific interactions by the following equation:

$$\log k = c + eE + sS + aA + bB + vV + fF + gG \quad (1)$$

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 phases on these interactions. k is the retention factor of the solute. c is the model intercept term, which when the retention factor is used as the dependent variable is dominated by the phase ratio. E , S , A , B and V are the usual Abraham descriptors [14], while F and G were introduced in the first part of this work and were intuitively chosen for their suspected role in the enantioseparation process. E is the excess molar refraction (calculated from the refractive index of the molecule) and models polarizability contributions from n and π 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 $\text{cm}^3 \text{mol}^{-1}/100$; F is the flexibility of the molecule, calculated as the fraction of rotatable bonds; G is the globularity, qualifying the compactness of the molecule. The system constants (e, s, a, b, v, f, g), 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 the 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). The different interaction capabilities reflected by the seven terms of the equation are presented in Fig. 1.

As selectivity or separation factor of two solutes is equal to the ratio of their retention factors, Eq. (2) is deduced from Eq. (1):

$$\log \alpha = e \Delta E + s \Delta S + a \Delta A + b \Delta B + v \Delta V + f \Delta F + g \Delta G \quad (2)$$

where α is the separation factor and ΔX represents the difference in the X molecular descriptor between the two solutes. Consequently, the coefficients also reflect the system's selectivity towards one particular molecular interaction.

However, the descriptors are considered identical for both enantiomeric solutes. Therefore Eq. (2) shows that no enantioseparation can occur. In the case of enantioseparation with CSPs, since the stationary phase is chiral, the possible interactions established between the two enantiomers and the CSP can be different. The

two enantiomers are no longer identical to the chiral selector (CSP) and each enantiomer “sees” the CSP differently with respect to their three-dimensional structures. Therefore, Eq. (3) can also be deduced from Eq. (1):

$$\log \alpha = \Delta eE + \Delta sS + \Delta aA + \Delta bB + \Delta vV + \Delta fF + \Delta gG \quad (3)$$

where Δx represents the difference in the x type of interactions, or the difference in the free energy of binding between the two enantiomers having the same X molecular descriptor with the CSP. This way, the enantiomers are considered seeing different CSP domains [12,16].

Since such a relationship is based on the intermolecular interactions contributing to the enantio-separation, a good understanding could lead to a predicting capability for whether a racemate is separable or not.

Since the descriptors are achiral, it must be pointed out that no absolute configuration of the enantiomers is considered. Therefore, the above equations could only model the retentions and separation factors irrespective to the elution order of the enantiomers.

In this paper, we investigate the factors contributing to enantiomer separation or selectivity on two polysaccharide CSPs, *tris*-(3,5-dimethylphenylcarbamate) of amylose and cellulose (ADMPC and CDMPC), on which we have previously determined the interactions contributing to retention [15] (this issue). We will use the modified version of the solvation parameter model to study the enantioselectivity. In addition, we will show how factorial discriminant analysis (FDA), a statistical method that is rarely used to explain analytical chemistry data, can be helpful in extracting the information from the multi-dimensional space of experimental results.

2. Experimental

2.1. Stationary phases

The columns used in this study were Chiralcel OD-H and Chiralpak AD-H (150 mm \times 4.6 mm, 5 μ m) from Daicel (Tokyo, Japan). 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 135 solutes used in this study are presented in Table 1, together with their solute descriptors. The majority of the products were from commercial sources. A small portion of them (solute 120–135 in Table 1) were in-house synthesized products, whose formulas are confidential. All solutions were prepared in methanol.

Solvent used was HPLC grade methanol (MeOH) 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°C by a cryostat (Julabo F10c, Seelbach, Germany). When the two solvents (methanol and 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 μ 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 880-81 pressure regulator. The outlet regulator tube (internal diameter 0.25 mm) was heated to 60°C to avoid ice formation during the CO_2 depressurization. UV detection was carried out at 254 nm, or 210 nm depending on solute structure. Injection volumes were 1–5 μ L. Chromatograms were recorded using the Azur software (Datalys, France).

Operating conditions were as follows: carbon dioxide–methanol 90:10 (v/v), 3 mL/min, 25°C (controlled with an oven), outlet pressure 170 bar. It was important to maintain identical conditions for all solutes, to ensure that separation factors could be compared on the same basis, because stereoselective interactions are markedly affected by the mobile phase conditions. The mobile phase conditions were chosen to ensure that reasonable elution times and separation factors could be measured, although they were not optimized for each racemate.

2.4. Data analysis

Retention factors (k) were calculated based on the retention time t_R , determined using the peak maximum (even when tailing occurred) and the hold-up time t_0 measured on the first negative peak due to the unretained dilution solvent (always visible in these conditions). Separation factors (α) were calculated as the ratio of retention factors of the second eluted enantiomer to the first eluted enantiomer k_2/k_1 .

Abraham descriptors were determined with the Absolv Web-boxes 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 2009.10 (Chemical Computing Group, Montreal, Canada) and QikProp 2009/08/20 (Schrödinger) respectively, based on the procedure described in the first part of this work [15].

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

3. Results and discussion

3.1. Solute set

A result can only be useful and valid within its application domain. A good predictive model should be able to extrapolate to new molecular structures. However, one limitation of most published QSRR approaches is that characterized solutes are usually structurally much simpler than the broad and complex array of pharmacologically active compounds that are targets for the study.

Therefore, to improve QSRR, the initial solute set must be as diverse as possible. As a result, we introduced compounds from diverse application areas in addition to simple and small molecules like phenylethanol, so that the results could be helpful in understanding how structural differences affected chiral recognition. The solute set comprised 135 racemates, as tabulated in Table 1. Compounds 1–26 were simple benzenic or naphthalenic solutes; compounds 27–92 were drug molecules; compounds 93–104 were pesticides; compounds 105–117 were natural products, terpenes and flavanones; compounds 118 and 119 were derivatized amino acids; compounds 120–135 were synthetic products of possible biological interest.

Table 1
Chromatographic solutes, their structures and molecular descriptors. *E*, excess molar refraction; *S*, dipolarity/polarizability; *A*, hydrogen bond acidity; *B*, hydrogen bond basicity; *V*, McGowan's characteristic volume; *F*, flexibility; *G*, globularity.

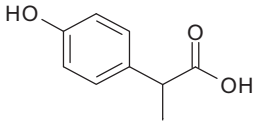
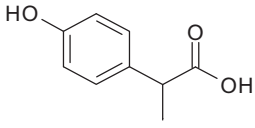
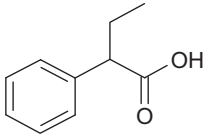
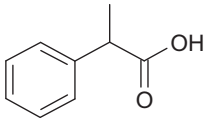
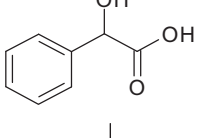
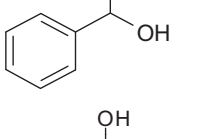
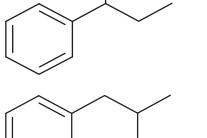
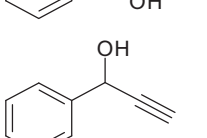
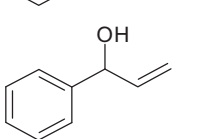
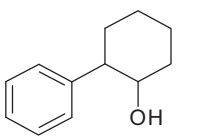
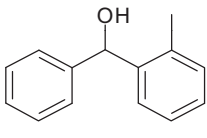
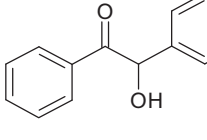

Nb	Compound		<i>E</i>	<i>S</i>	<i>A</i>	<i>B</i>	<i>V</i>	<i>F</i>	<i>G</i>	α_{ADMPC}	α_{CDMPC}
1	(4-Hydroxyphenylacetic)-2-propionic acid		0.908	1.460	0.590	0.680	1.272	1.667	1.590	1.21	1.31
2	2-Phenylbutyric acid		0.750	1.070	0.570	0.480	1.354	2.500	1.600	1.06	1.15
3	2-Phenylpropionic acid		0.730	0.970	0.570	0.680	1.214	1.818	1.658	1.09	1.03
4	Mandelic acid		0.900	1.050	0.740	0.890	1.131	1.818	1.700	1.37	1.61
5	1-Phenylethanol		0.784	0.830	0.300	0.660	1.057	1.100	1.741	1.11	1.26
6	1-Phenyl-1-propanol		0.775	0.830	0.300	0.660	1.198	2.000	1.679	1.11	1.26
7	1-Phenyl-2-propanol		0.787	0.900	0.300	0.720	1.198	2.000	1.599	1.04	1.09
8	1-Phenyl-2-propyn-1-ol		1.060	1.000	0.400	0.620	1.112	1.000	1.540	1.00	1.38
9	Phenylvinylcarbinol		0.930	0.890	0.310	0.610	1.155	2.000	1.694	1.05	1.24
10	2-Phenyl-1-cyclohexanol		1.000	0.980	0.310	0.440	1.512	0.714	1.575	2.06	1.54
11	2-Methylbenzhydrol		1.420	1.210	0.310	0.690	1.665	1.250	1.408	1.16	1.14
12	Benzoin		1.520	1.560	0.170	0.950	1.680	1.765	1.423	1.37	1.12

Table 1 (Continued)

Nb	Compound		<i>E</i>	<i>S</i>	<i>A</i>	<i>B</i>	<i>V</i>	<i>F</i>	<i>G</i>	α_{ADMPc}	α_{CDMPc}
13	1-(1-Naphthyl)-ethanol		1.530	1.150	0.310	0.620	1.426	0.714	1.608	1.21	1.55
14	1-(2-Naphthyl)-ethanol		1.530	1.150	0.310	0.620	1.426	0.714	1.494	1.03	1.14
15	Mandelic acid, methyl ester		0.800	1.010	0.170	0.890	1.272	1.667	1.441	1.12	1.67
16	Mandelic acid, ethyl ester		0.800	1.010	0.170	0.890	1.413	2.308	1.294	1.21	1.72
17	trans-Chlorostilbene oxide		1.570	1.390	0.000	0.450	1.679	1.111	1.332	1.35	1.08
18	(2,3-Epoxypropyl)-benzene		0.840	0.870	0.000	0.340	1.089	1.818	2.162	1.00	1.00
19	1,2-Epoxyethylbenzene		0.840	0.860	0.000	0.340	0.948	1.010	2.193	1.11	1.25
20	Methylphenylsulfoxide		1.104	1.730	0.000	0.880	1.080	1.111	1.764	1.02	1.00
21	Methyl- <i>p</i> -tolylsulfoxide		1.120	1.770	0.000	0.800	1.220	1.000	1.594	1.00	1.06
22	Phenylvinylsulfoxide		1.220	1.890	0.000	0.850	1.177	2.000	1.592	1.21	1.07
23	2-Aminobutane		0.170	0.320	0.160	0.630	0.772	2.500	1.962	1.00	1.00
24	2-Amino-1-phenylethanol		1.030	1.100	0.460	1.190	1.157	2.000	1.684	1.11	1.08
25	Phenylglycinol		1.010	1.170	0.460	1.060	1.157	2.000	1.842	1.92	1.26

Table 1 (Continued)

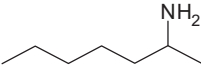
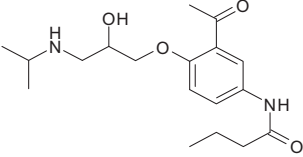
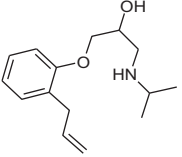
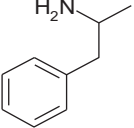
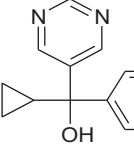
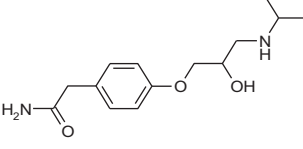
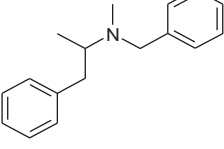
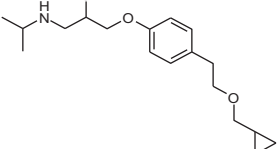
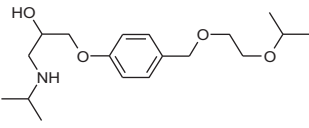
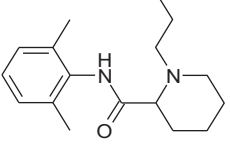
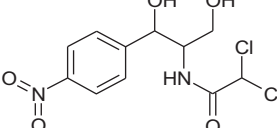
Nb	Compound	<i>E</i>	<i>S</i>	<i>A</i>	<i>B</i>	<i>V</i>	<i>F</i>	<i>G</i>	α_{ADMPC}	α_{CDMPC}	
26	Tuaminoheptane		0.210	0.500	0.210	0.620	1.195	5.714	1.542	1.00	2.35
27	Acebutolol		1.600	2.420	0.900	2.100	2.756	4.167	0.409	1.28	1.16
28	Alprenolol		1.250	1.090	0.150	1.440	2.159	4.444	0.891	1.22	1.00
29	Amphetamine		0.795	0.810	0.140	0.780	1.239	2.000	1.670	3.99	1.00
30	Ancymidol		1.720	1.580	0.310	1.330	1.955	1.905	1.287	2.51	1.33
31	Atenolol		1.450	1.880	0.690	2.000	2.176	4.211	0.722	1.23	1.90
32	Benzphetamine		1.320	1.280	0.000	0.740	2.129	2.623	1.284	1.75	1.00
33	Betaxolol		1.310	1.310	0.290	1.530	2.575	4.783	0.309	1.06	2.88
34	Bisoprolol		1.140	1.370	0.290	1.770	2.742	5.217	0.199	1.43	1.26
35	Bupivacaine		1.320	1.590	0.260	1.190	2.514	2.273	1.039	1.11	1.35
36	Chloramphenicol		1.850	1.700	0.700	1.750	2.073	3.000	1.253	1.36	1.00

Table 1 (Continued)

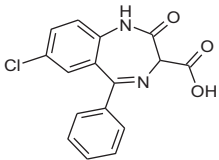
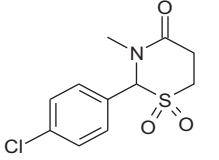
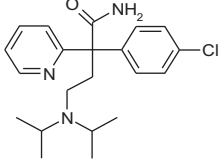
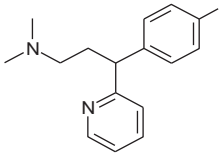
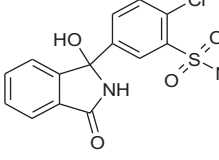
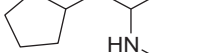
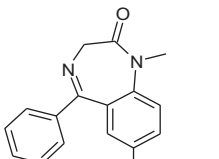
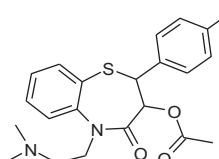
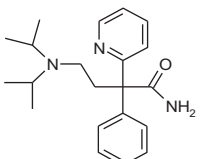
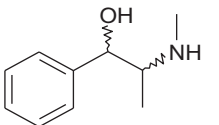
Nb	Compound	<i>E</i>	<i>S</i>	<i>A</i>	<i>B</i>	<i>V</i>	<i>F</i>	<i>G</i>	α_{ADMPC}	α_{CDMPC}	
37	Chlorazepate		2.270	2.140	1.040	1.340	2.148	0.833	1.005	2.36	1.42
38	Chlormezanone		1.440	2.400	0.000	1.390	1.831	0.556	1.374	1.52	1.27
39	<i>p</i> -Chlorodisopyramide		1.920	2.340	0.490	1.640	3.030	2.963	0.978	4.75	1.00
40	Chlorpheniramine		1.465	1.340	0.000	1.350	2.210	2.500	1.068	1.07	1.00
41	Chlorthalidone		2.640	3.050	1.010	1.980	2.175	0.833	1.112	1.20	1.28
42	Cyclopentamine		0.370	0.430	0.130	0.520	1.368	3.000	1.674	1.00	1.00
43	Diazepam		2.078	1.570	0.000	1.250	2.074	0.455	1.137	1.20	1.75
44	Diltiazem		2.420	2.550	0.000	2.120	3.137	1.935	0.664		1.00
45	Disopyramide		1.770	2.260	0.490	1.640	2.907	3.077	1.073	2.03	1.00
46	<i>psi</i> -Ephedrine		0.916	0.760	0.210	1.210	1.439	2.500	1.507	0.99	1.34

Table 1 (Continued)

Nb	Compound	<i>E</i>	<i>S</i>	<i>A</i>	<i>B</i>	<i>V</i>	<i>F</i>	<i>G</i>	α_{ADMPC}	α_{CDMPC}
47	Epinephrine	1.350	1.340	1.150	1.510	1.415	2.308	1.357	1.18	1.09
48	Ethyl-loflazepate	2.080	2.070	0.470	1.350	2.307	1.111	0.706	1.72	1.14
49	Fenopropfen	1.390	1.630	0.570	0.780	1.880	2.105	1.307	1.16	1.00
50	Flurbiprofen	1.500	1.510	0.570	0.580	1.839	1.579	1.190	2.04	1.00
51	Glafenine	2.830	2.580	0.580	1.680	2.630	2.143	0.558		1.08
52	Glutethimide	1.310	1.400	0.340	1.020	1.725	1.176	1.573	1.94	1.08
53	Hexobarbital	1.340	1.500	0.240	1.330	1.786	0.556	1.705	6.47	1.04
54	p-Hydroxyamphetamine	1.010	1.160	0.710	0.970	1.298	1.818	1.629	1.06	1.00
55	Ibuprofen	0.730	0.590	0.590	0.810	1.777	2.667	1.218	1.22	1.00
56	Imoxiterol	2.530	2.340	0.650	2.000	2.775	2.855	1.026		1.00

Table 1 (Continued)

Nb	Compound	<i>E</i>	<i>S</i>	<i>A</i>	<i>B</i>	<i>V</i>	<i>F</i>	<i>G</i>	α_{ADMPC}	α_{CDMPC}
57	Indoprofen	1.920	2.300	0.570	1.140	2.110	1.304	1.018	1.23	1.22
58	Ketamine	1.280	1.420	0.130	0.890	1.832	1.176	1.542	1.66	1.06
59	Ketoprofen	1.560	1.970	0.570	0.870	1.978	2.000	1.285	1.06	1.86
60	Labetalol	2.192	2.130	0.770	1.760	2.643	3.200	0.651		1.22
61	Lorazepam	2.510	1.280	0.450	1.630	2.114	0.435	1.101	1.25	1.21
62	Mepenzolate bromide	1.350	1.410	0.170	1.080	2.738	1.481	1.000		1.13
63	Mephobarbital	1.570	1.820	0.240	1.350	1.841	1.053	1.620	9.05	1.20
64	Metamphetamine	0.740	0.800	0.130	0.590	1.380	2.727	1.649	1.00	1.00
65	Methylphenidate	1.010	1.290	0.130	0.940	1.909	1.765	1.174	1.05	1.66
66	5-Methyl-5-phenylhydantoin	1.370	1.580	0.440	1.060	1.402	0.667	2.090	3.55	1.02

Table 1 (Continued)

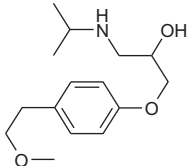
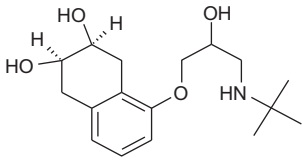
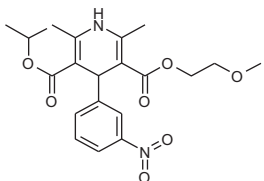
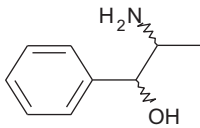
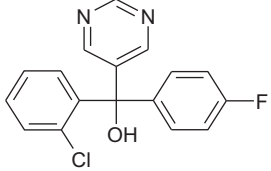
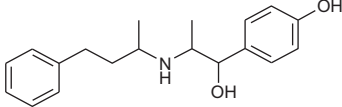
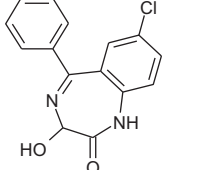
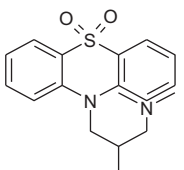
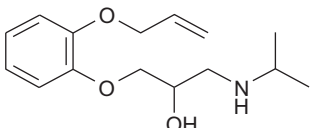
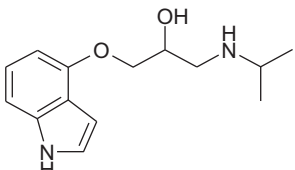
Nb	Compound	E	S	A	B	V	F	G	α_{ADMPC}	α_{CDMPC}	
67	Metoprolol		1.170	1.330	0.170	1.760	2.260	4.737	0.585	1.00	2.85
68	Nadolol		1.630	1.640	0.850	2.340	2.492	2.000	0.780	1.33	1.40
69	Nimodopine		1.600	2.410	0.130	1.790	3.117	2.581	1.101	1.00	1.00
70	Norephedrine		0.965	0.800	0.440	1.190	1.298	1.818	1.752	2.14	1.00
71	Nuarinol		2.090	1.890	0.310	1.240	2.189	1.250	1.277	1.45	1.06
72	Nylidrine		1.790	1.620	0.880	1.520	2.528	3.043	1.040	1.28	1.00
73	Oxazepam		2.350	1.100	0.450	1.600	1.992	0.455	1.047	1.19	1.65
74	Oxomemazine		1.890	2.480	0.000	1.500	2.542	1.600	1.150	1.29	1.00
75	Oxprenolol		1.310	1.490	0.170	1.620	2.217	4.737	1.261	2.15	2.21
76	Pindolol		1.700	1.650	0.300	1.480	2.009	3.158	1.095	1.00	1.00

Table 1 (Continued)

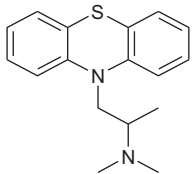
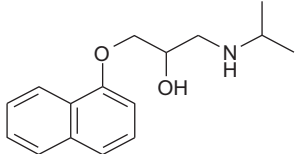
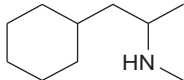
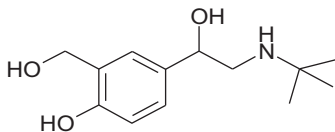
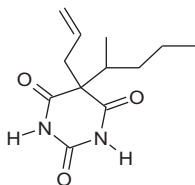
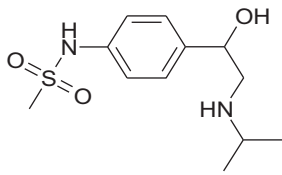
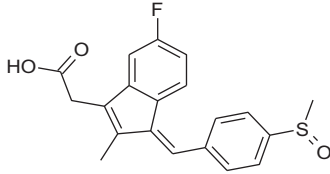
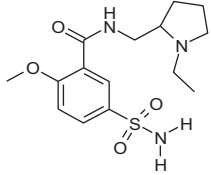
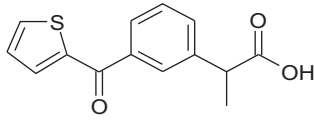
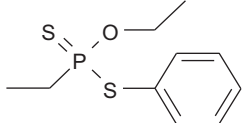
Nb	Compound	E	S	A	B	V	F	G	α_{ADMPC}	α_{CDMPC}
77	Promethazine 	2.140	1.720	0.000	1.090	2.283	1.364	1.303	1.19	1.03
78	Propranolol 	1.880	1.430	0.170	1.420	2.148	3.000	1.045	2.67	1.83
79	Propylhexedrine 	0.370	0.430	0.130	0.530	1.509	2.727	1.580		1.11
80	Salbutamol 	1.430	1.260	1.190	1.820	1.979	2.353	1.043	1.50	1.12
81	Secobarbital 	1.160	1.200	0.490	1.310	1.895	2.941	1.805	1.29	1.00
82	Sotalol 	1.520	1.860	0.740	1.750	2.101	3.333	0.875	1.32	1.16
83	Sulindac 	2.260	2.720	0.570	1.390	2.571	1.481	0.906	1.96	1.05
84	Sulpiride 	1.910	2.780	0.720	2.150	2.531	2.500	1.208	1.00	
85	Suprofen 	1.510	1.890	0.570	0.810	1.903	2.105	1.131	1.55	1.07
86	Terbutaline 	1.570	1.760	1.300	1.750	1.838	1.875	1.139	1.13	1.09

Table 1 (Continued)

Nb	Compound	E	S	A	B	V	F	G	α_{ADMPC}	α_{CDMPC}	
87	Terfenadine		2.550	2.040	0.630	1.800	4.013	2.105	0.266	2.49	
88	Thiopental		1.410	1.460	0.260	1.340	1.901	2.500	1.794	1.01	1.12
89	Verapamil		1.760	3.000	0.000	1.890	3.786	3.824	0.993	1.00	1.07
90	Viloxazine		1.050	1.330	0.170	1.200	1.870	2.778	1.441	1.17	1.18
91	Warfarine		2.300	2.180	0.350	1.490	2.308	1.600	1.101	3.00	2.47
92	Zopiclone		2.660	3.200	0.000	2.430	2.623	1.000	0.773	1.15	1.40
93	Cyanofenphos		1.890	1.820	0.000	0.950	2.246	2.381	0.838	1.08	1.00
94	EPN		1.990	1.910	0.000	0.910	2.266	2.727	0.911	1.19	1.00
95	Fenamifos		1.020	2.050	0.130	1.290	2.347	3.684	0.720	1.62	1.07

Table 1 (Continued)

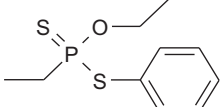
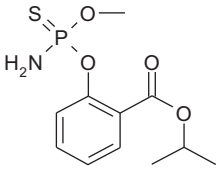
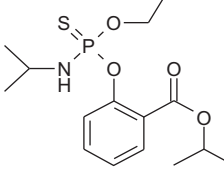
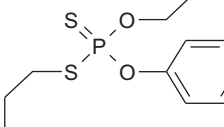
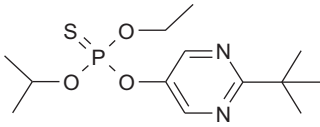
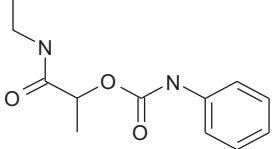
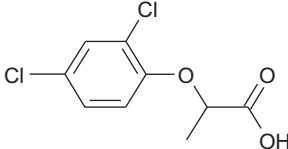
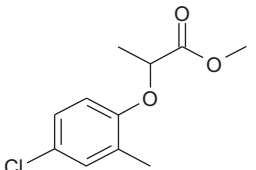
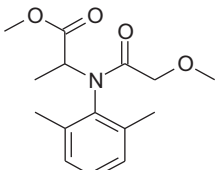
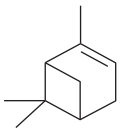
Nb	Compound	<i>E</i>	<i>S</i>	<i>A</i>	<i>B</i>	<i>V</i>	<i>F</i>	<i>G</i>	α_{ADMPC}	α_{CDMPC}
96	Fonofos 	1.440	0.880	0.000	0.640	1.870	3.571	1.405	1.00	1.00
97	Isofenphos 	1.390	1.380	0.130	1.560	2.644	3.636	1.150	1.00	1.00
98	Isocarbophos 	1.490	1.420	0.210	1.510	2.081	2.778	1.282	1.91	1.11
99	Sulfopros 	1.850	1.160	0.000	0.970	2.374	4.444	1.152	1.00	1.00
100	Tebupirimfos 	1.270	1.100	0.000	1.420	2.447	3.000	1.032	1.00	1.00
101	Carbetamide 	1.198	1.860	0.620	1.230	1.852	2.353	0.872	1.00	1.00
102	Dichlorprop 	1.050	1.400	0.570	0.610	1.517	2.143	1.560	1.00	1.00
103	Mecoprop methyl ester 	0.860	1.240	0.000	0.680	1.676	2.000	1.136	1.09	1.11
104	Metalaxyl 	0.850	1.960	0.000	1.600	2.233	3.000	1.796	1.12	1.66
105	α -Pinene 	0.446	0.140	0.000	0.120	1.257	0.000	1.789	1.90	

Table 1 (Continued)

Nb	Compound	<i>E</i>	<i>S</i>	<i>A</i>	<i>B</i>	<i>V</i>	<i>F</i>	<i>G</i>	α_{ADMPC}	α_{CDMPC}
106	β -Pinene	0.530	0.240	0.000	0.190	1.257	0.000	1.895	1.00	1.00
107	Limonene	0.488	0.280	0.000	0.210	1.323	1.000	1.666	1.68	2.85
108	Camphor	0.500	0.690	0.000	0.710	1.316	0.000	2.077	1.10	1.00
109	α -Terpineol	0.553	0.610	0.200	0.700	1.425	0.909	1.770	2.17	1.10
110	4-Terpineol	0.531	0.340	0.320	0.540	1.425	0.909	1.776	1.14	1.00
111	β -Citronellol	0.340	0.510	0.310	0.440	1.533	5.000	1.535	1.05	1.00
112	Linalool	0.398	0.550	0.200	0.670	1.490	4.000	1.386	1.99	1.00
113	Linalyl acetate	0.300	0.650	0.000	0.550	1.788	3.846	1.264	1.00	1.00
114	Methyl acetate	0.300	0.640	0.000	0.480	1.765	1.429	1.387	1.64	1.00
115	Flavanone	1.650	1.760	0.000	0.730	1.713	0.526	1.235	2.74	1.14
116	6-Methoxyflavanone	1.720	1.890	0.000	0.930	1.912	0.952	1.155	3.18	1.14

Table 1 (Continued)

Nb	Compound	<i>E</i>	<i>S</i>	<i>A</i>	<i>B</i>	<i>V</i>	<i>F</i>	<i>G</i>	α_{ADMPC}	α_{CDMPC}
117	4',5,7-Trihydroxyflavanone	2.230	2.190	1.300	1.140	1.889	0.455	1.114	1.42	1.19
118	Fmoc-Proline	2.310	2.360	0.570	1.150	2.483	1.429	1.008	1.53	1.02
119	Fmoc-Threonine (tBu)	2.090	2.370	0.800	1.400	3.073	1.935	0.908	1.05	1.11
120	Synthetic compound	1.140	1.510	0.260	0.710	1.428	0.714	1.645	1.25	1.54
121	Synthetic compound	2.660	2.520	0.750	1.100	2.041	0.417	1.281	1.81	3.42
122	Synthetic compound	2.610	2.460	0.260	1.090	2.279	1.154	1.019	2.26	2.14
123	Synthetic compound	1.770	2.080	0.260	0.980	1.813	1.000	1.309	3.15	1.12
124	Synthetic compound	2.430	2.310	0.260	0.820	1.982	0.435	1.305	2.91	4.07
125	Synthetic compound	1.950	2.200	0.500	1.090	1.813	0.500	1.308	1.13	1.40
126	Synthetic compound	2.570	2.600	0.310	1.580	2.479	1.071	1.046	5.28	1.57
127	Synthetic compound	1.530	1.610	0.000	0.750	1.771	0.500	1.513	1.96	1.26
128	Synthetic compound	2.360	1.730	0.630	1.160	2.092	1.364	1.213	2.53	1.13
129	Synthetic compound	2.120	1.530	0.310	1.230	2.611	1.111	1.112	1.90	1.16
130	Synthetic compound	1.740	3.520	0.510	1.510	2.846	3.704	0.758	1.00	1.33
131	Synthetic compound	2.440	2.130	0.290	1.460	2.533	2.692	0.652	1.15	
132	Synthetic compound	3.090	2.330	0.290	1.320	2.702	2.069	1.202	1.18	1.10
133	Synthetic compound	1.750	1.460	0.290	1.330	2.148	3.500	0.931	1.51	1.60
134	Synthetic compound	0.960	1.350	0.000	0.730	1.430	0.000	1.833	1.40	1.00
135	Synthetic compound	0.930	1.290	0.000	0.750	1.571	0.000	1.696	1.30	1.00

There are many origins of chirality in organic compounds. We were keen to introduce solutes with different stereogenic atoms. Compounds 20, 21, 22 and 83 possessed a chiral sulphur atom, while compounds 93–100 possessed a chiral phosphorous atom. Solute 43 was diazepam, possessing a chiral nitrogen. All others had chiral carbons, which is the most common chiral centre encountered. Racemates with more than one asymmetric centre were also included, such as labetalol or nadolol.

It must be pointed out that, while drug molecules with multiple functional groups are included for wider applicability, the fit quality of the statistical analyses is likely to be worse as a result.

As explained in the first part of this work, probe solutes should cover a large range of descriptor values and a variety of chemical functions, so that the introduction of additional solutes would not significantly modify the results of the statistical analyses. As can be seen in Figure S1 in the supplementary material section, showing the repartition of the solutes from Table 1 in each descriptor space, the 135 solutes selected indeed provided a uniform distribution of each solvation descriptor within a sufficiently wide space. Similar to the finding in the first part of the work, there were some crowding

in the lower values of *A*, but the clustering of values was limited for most descriptors. The final solute set provided adequate and proper retention (*i.e.* measurable with statistical significance without excessive retention), with $\log k$ values essentially comprised between -1 and $+2$. It was also important that the separation factors were widely spread. On the ADMPC phase, $\log \alpha$ values ranged from 0 to 1, while they ranged from 0 to 0.6 on the CDMPC phase.

Minima, maxima, average and standard deviation values for each descriptor can be found in Table S1 in the supplementary material section, and appear to be in the same range, which ensures possible comparability of the multiple linear regression coefficients.

The absence of cross-correlation inherent to the choice of solutes can be checked in the covariance matrix, with low determination coefficients between the descriptors (Table S2 in the supplementary material section). Each descriptor was plotted against another, and non-correlation was reflected by the random scatter of the data, without any particular compounds acting as levers. Only the *E* and *S* descriptors appeared to present some correlation, which was not unexpected as both *E* and *S* reflect some of the polarizability characteristics of the solutes. A significant proportion of non-aromatic

molecules, mainly small terpenes, was introduced to help breaking this covariance, as well as to ensure possible assessment of the contribution of the aromatic moiety to the chiral recognition mechanism. *V* and *G* similarly presented some correlation. The large proportion of aromatic compounds is also responsible for this correlation.

3.2. Factorial discriminant analyses based on two classes: the factors preventing separation

Factorial discriminant analysis (FDA) is a standard tool in modern data analysis. As principal component analysis (PCA), FDA provides a way to identify statistical patterns in data, and to present the data in such a way that their similarities and differences are highlighted. However, FDA is preferably used when data sets of classified data are available, and the common features to one classification group need to be determined. Thus *a priori* defined groups must be established, and the observation of the results indicates *a posteriori* classification. The discriminant rules are based on linear combinations of the observed variables, called discriminant factors (similar to the principal components obtained in PCA). The selection of factors is made in order to minimize the probability of mis-classification: it produces a score plot where compact groups should appear as spread as possible in the space.

In our case, the 135 racemates analysed on each CSP, represented by their seven molecular descriptors, were treated as variables. Two groups were defined: separated and non-separated racemates. This would have been a 135-dimensions problem for which no graphical representation is possible, since patterns in data can hardly be found in such a high dimension. The primary aim of FDA is to project all data points onto a space of reduced dimension, most practically on a plane. When only two classes are present, as is the case here, only one discriminant function *F1* is obtained, thus all points are represented as projections on this unique axis. The loading plot then allows identifying the factors contributing to the difference between the classes, in our case, the factors explaining why a given racemate could be resolved in the studied chromatographic system.

All racemates were divided in two groups for each column: the racemates that were separated and those that were not separated. Any racemate providing a separation factor above 1.0 was considered separated racemate and belonged to class 1 and only racemates with $\alpha = 1.0$ were considered non-separated and belonged to class 2, since we looked for separation mechanisms rather than prediction of resolution. Indeed, a separation with a non-zero (even very low) $\log \alpha$ under the common test conditions indicates partial enantio-recognition, which could likely be improved by optimizing the operating conditions during the course of method development. Visual examination of the chromatographic separation was relied upon in deciding whether a racemate was separated, particularly for those with bad peak shapes. Whenever a decision of separation was too difficult to make due to asymmetry or distortion of the chromatographic peak, the racemate was preferably eliminated from the data set.

The results are presented in Figs. 2 and 3 for ADMPC and CDMPC respectively.

The mathematical details of the FDA model can be found elsewhere. We only give a short review of the most essential resulting parameters here.

First of all, the statistics of the analyses must be observed, because statistically insignificant results should not be considered in drawing a conclusion. A complete statistical evaluation of the results can be found in the [supplementary material section](#). Judging from these analyses, both models showed statistical significance, thus were amenable to interpretation.

The utility of each variable in separating the samples into two classes using univariate procedures, were assessed by the Fischer weights (*F*) and the corresponding probabilities (*p*). Based on these statistics, *F* (flexibility) was the most significant parameter responsible for the class separation. The other parameters had different contributions on the two CSPs.

Molecular coordinates were represented on the *F1* axis in Figs. 2 and 3a for ADMPC and CDMPC, respectively. They were parted along a second axis according to the *a priori* class, so the position of all points and possible superposition between the classes would be clearer. The position of the barycentres of each class is shown in Figs. 2 and 3b.

Figs. 2 and 3c display the loading plots for ADMPC and CDMPC, respectively, where the contribution of each molecular descriptor to the *F1* axis is represented.

Thus, as appears on the score plots and was observed from the statistical analysis, a good discrimination was obtained on ADMPC, and not as good on CDMPC. The racemates that were separated are on the left-hand side of the *F1* axis (red points), while those that co-eluted are on the right-hand side (grey points). However, some molecules are plotted quite far from their original class. There is also a region of the axis where both classes are present. This is an uncertainty zone where accurate prediction might be difficult.

On both phases, flexibility was found to be the least favourable feature for enantioseparation (it points far to the right-hand side on the loading plots). This is understandable because flexible molecules have more conformers, thus more ways to interact with the stationary phase, which lessens the differences between the two enantiomers. This point had already been observed in past studies [18]. This is a first indication that interactions with minor contributions to total solute retention can still be essential to chiral recognition. Indeed, in the first part of this work, flexibility was shown not to be significant to retention on these CSPs. Thus, if flexibility has little effect on retention, this feature is however clearly not desirable for good resolution.

Enantioseparations are attributed to a combination of stereoselective interactions that usually cannot be reduced down to one single factor. For instance, on the ADMPC phase, globularity appears to favour enantioseparation (as it points to the left-hand side of the *F1* axis), but this is only true to small molecules because large volume, on the contrary, is a disadvantage for enantioseparation. It is understood because small globular molecules can have more possibilities for simultaneous close interactions with the stationary phase, while large globular molecules may simply not enter the chiral grooves, which would limit possibilities for enantio-specific interactions.

On the CDMPC phase, globularity was never found to be favourable for enantio-recognition. This is possibly because the chiral cavities in cellulosic phases are of different shapes, thus it might be more difficult for globular molecules to reach the chiral selectors. Indeed, the polar carbamate groups are preferably located inside the modified cellulose polymer helix [19], thus stereogenic fit is essential for the solutes to be able to establish polar interactions with the carbamate moiety. Literature data also indicate that cellulose CSPs tend to be more suitable for linear molecules [20].

On the other hand, molecular volume is of little significance to separation on CDMPC, as the *V* variable is very close to the centre of the loading plot. This might be related to the fact that the chiral cavities in this CSP are larger than that on amylosic phases [3], thus size of the solute is less critical to enantio-resolution than in the smaller cavities of ADMPC.

The *A* descriptor is pointed to the left hand-side of the *F1* axis for both CSPs, indicating identical influence of solute acidity. This showed that acidity or hydrogen-bond acceptor capability is a favourable feature for enantioseparation on ADMPC

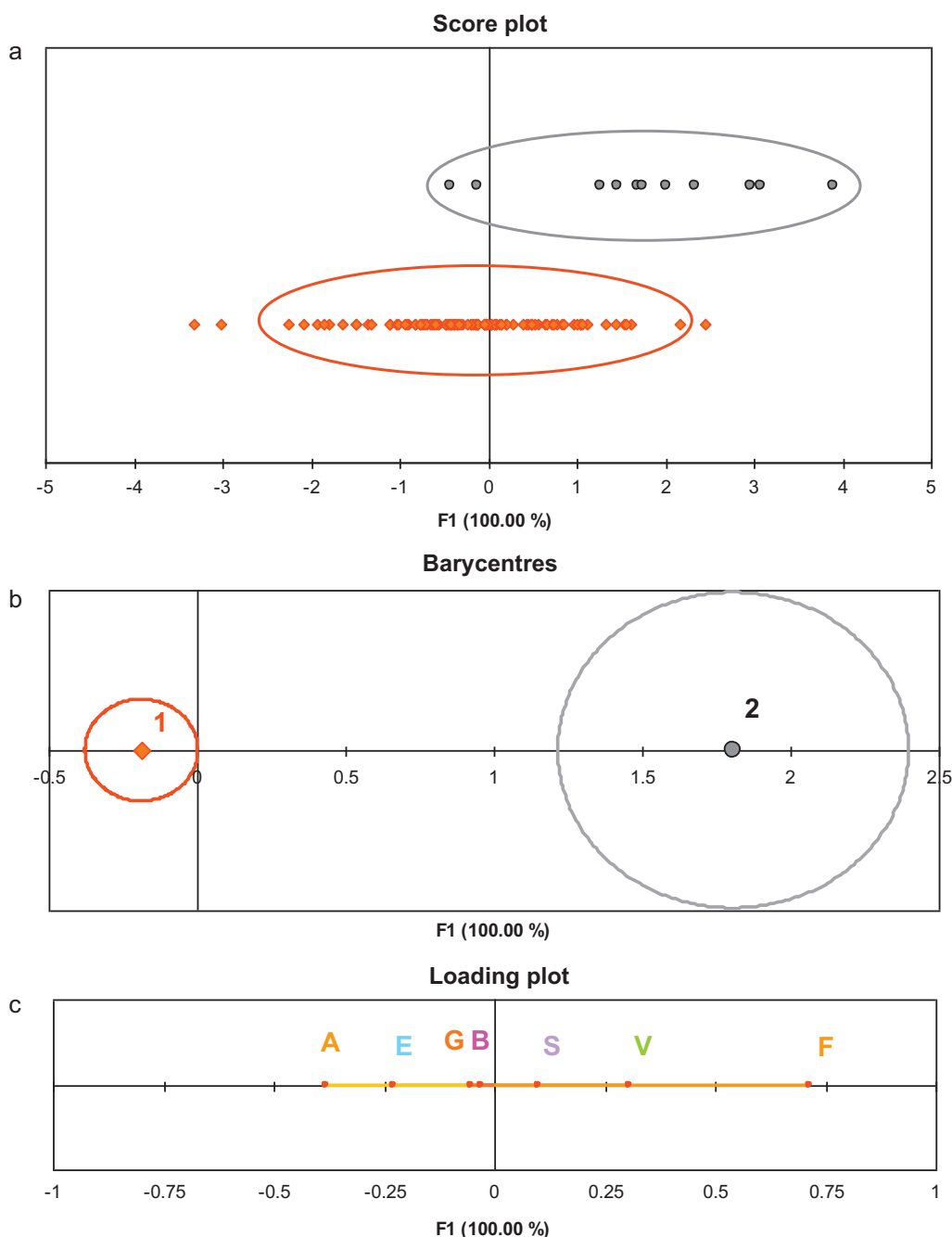


Fig. 2. Factorial discriminant analysis based on two classes of racemates (1 = experimentally separated, red diamonds; 2 = not-separated, black diamonds) on ADMPC, using the seven molecular descriptors of the modified solvation parameter model as variables. (a) Score plot, (b) barycentre plot and (c) loading plot. See text for details. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

and CDMPC. The hydrogen bond is most certainly established between the solute and the carbonyl group of the carbamate function.

The *E* descriptor is also pointed to the left hand-side of the F1 axis for both CSPs, indicating identical influence of solute aromaticity. This showed that π - π interaction capability is a favourable feature for enantioseparation on ADMPC and CDMPC.

Other features are all different between the two CSPs: the *S* and *B* variables point to the left hand-side for CDMPC, indicating favourable contribution to enantioselectivity, while they are zero or pointing to the right hand-side for ADMPC, indicating no or negative contribution to enantioseparation. Moreover, univariate analysis indicates that *S* and *B* are of little significance to separa-

tion on ADMPC: *F* (Fischer's statistic) is 0.3 and 0.0, while *p*-value is 0.6 and 0.9 respectively.

The different positions of the *B* variable is most surprising because this descriptor was found to be significant to explain retention using achiral solutes on both stationary phases. Also, the *S* descriptor was not significant to explain retention but appears to be significant to explain separation on CDMPC. This is another indication that the factors contributing to retention are not necessarily favourable for enantioselectivity, and vice versa. Another way of formulating this idea is that strong retention is not necessarily associated to high enantioselectivity, and vice versa. This was confirmed by studies from Kafri and Lancet [21] who reported a comprehensive data examination of enantioseparation mea-

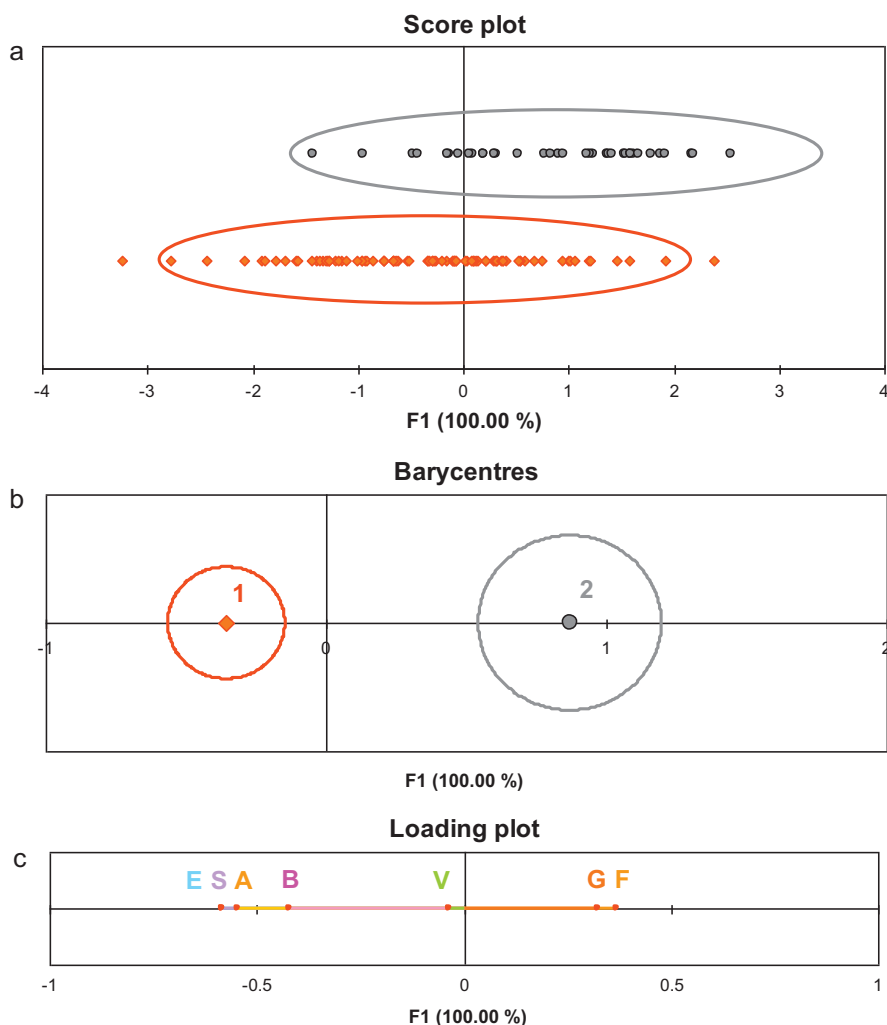


Fig. 3. Factorial discriminant analysis based on two classes of racemates (1 = experimentally separable, red diamonds; 2 = not separable, black diamonds) on CDMPC, using the seven molecular descriptors of the modified solvation parameter model as variables. (a) Score plot, (b) barycentre plot and (c) loading plot. See text for details. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

surements for over 72,000 chiral separations and concluded that no correlation existed between retention factors and separation factors.

Thus, while some features are common to ADMPC and CDMPC (*E*, *A* and *F* essentially), others (*S*, *B*, *V* and *G*) clearly have different influences on the enantiorecognition capability of the two CSPs. This naturally reflects in the experimental results, as shown in Fig. 4, where the logarithms of separation factors measured on both phases are compared. It is clear from this figure that those racemates that were well separated on one phase were often less well separated on the other. This is also in accordance with past observations on these stationary phases [22]. It is also apparent, as already mentioned above, that the success rate on ADMPC was greater than on CDMPC, with more data points being plotted below the first bisector, with several of them on the abscissa axis indicating no separation on CDMPC. This seems likely due to the more demanding requirements for interaction capabilities of solutes on the CDMPC phase, while the ADMPC phase, being more versatile and less demanding for solutes, would reach enantiomer resolution more easily. These results are in accordance with past studies [17,23].

Okamoto and co-workers suggested that the most important adsorbing sites for chiral discrimination on phenylcarbamate derivatives CSP were probably the polar carbamate groups, while

the π - π interaction of phenyl groups might be of secondary importance for the chiral recognition ability [24]. They also pointed out that contribution of steric effects of the derivatized polysaccharide should also be considered. Our results based on FDA are consistent with their conclusions, although π - π interactions appear to be significant. However, it is worth noting that the unbalance towards the number of non-separated racemates may have an impact on the choice of variables. However, good agreement between the observed coefficients and the chemical intuition increases one's confidence in the validity of the results.

3.3. Multiple linear regression analysis: an unsuccessful attempt for establishing the factors contributing to separation

To try to establish a relationship between the separation factors measured for each racemate and their molecular descriptors, multiple linear regressions were calculated based on Eq. (3). The use of α as a dependent variable was possible because all experimental data were acquired under identical operating conditions. However, this attempt was unsuccessful: there was no statistically significant relationship using the whole data set.

The implication of this failed attempt is important. It indicates that these CSPs should be considered as heterogeneous stationary phases, exhibiting mixed retention mechanisms.

Table 2

System constants and statistics for both columns. n is the number of solutes considered in the regression, R_{adj}^2 is the adjusted correlation coefficient, SD in the standard error in the estimate, F is Fischer's statistic and the numbers in italics represent 99.9% confidence limits.

Stationary phase	Case	c	e	s	a	b	ν	f	g	n	R_{adj}^2	SD	F
ADMPC	E1 in class 1	−0.911 <i>0.147</i>	0.838 <i>0.078</i>		0.391 <i>0.107</i>		−0.197 <i>0.095</i>	0.153 <i>0.044</i>		48	0.829	0.223	58
	Achiral	−0.759 <i>0.070</i>	0.731 <i>0.029</i>		0.718 <i>0.045</i>	0.338 <i>0.048</i>			−0.164 <i>0.039</i>	208	0.843	0.198	279
	E2 in class 2	−0.953 <i>0.127</i>	0.894 <i>0.079</i>		0.536 <i>0.149</i>	0.482 <i>0.138</i>				47	0.845	0.240	85
CDMPC	E1 in class 1		0.442 <i>0.079</i>	0.208 <i>0.086</i>		0.342 <i>0.095</i>	−0.445 <i>0.098</i>		−0.251 <i>0.141</i>	36	0.798	0.164	29
	Achiral	−0.543 <i>0.050</i>	0.694 <i>0.021</i>		0.535 <i>0.033</i>	0.175 <i>0.037</i>			−0.181 <i>0.028</i>	200	0.882	0.141	373
	E2 in class 2	−0.790 <i>0.088</i>	0.859 <i>0.060</i>	−0.215 <i>0.056</i>	0.401 <i>0.127</i>	0.463 <i>0.070</i>				40	0.912	0.149	102

Therefore, meaningful groups of solutes that behave in a comparable manner must be established.

Using smaller data sets, it was possible to establish some meaningful relationships, but only within compound families as it will be discussed in the following section.

3.4. Factorial discriminant analysis based on three classes: definition of different classes of solutes

As pointed out in Section 3.3, different reasons seem to be responsible for the enantiorecognition of different racemates. To define the groups of solutes sharing a common enantiorecognition mechanism, we adopted a procedure first suggested by Mitchell et al. [12], who investigated chiral recognition mechanisms on macrocyclic glycopeptide CSPs using HPLC. Our study used a much larger number of racemates so that the conclusion based on a better data set would be more substantiated.

Based on the multiple linear regression analyses established using Eq. (1), for achiral solutes in Part I of this series of papers and reported here in Table 2, a predicted retention factor was calculated for each solute in Table 1, both on ADMPC and on CDMPC. This retention factor is identical on one column for both enantiomers because all molecular descriptors are identical.

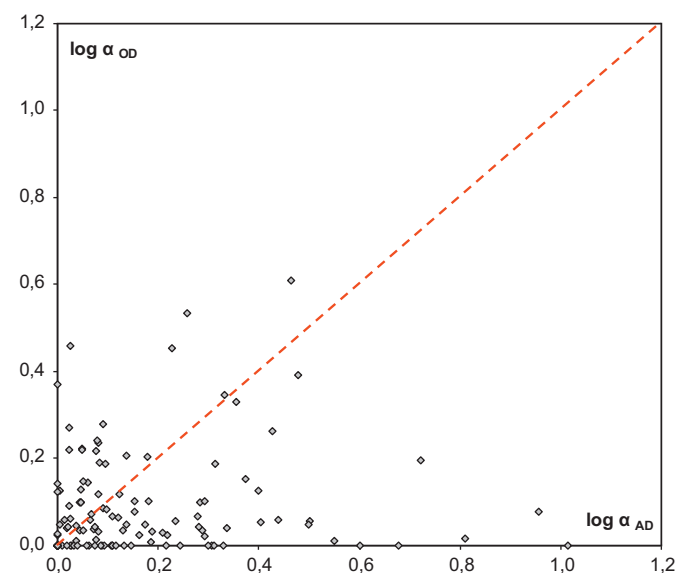


Fig. 4. Relationship between separation on the ADMPC phase ($\log \alpha_{\text{AD}}$) and separation on the CDMPC phase ($\log \alpha_{\text{OD}}$) for the racemates in Table 1. The interrupted red line is the first bisector, indicating identical separation factors on both stationary phases. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

We compared the experimental retention factors with the predicted ones, and then divided the racemates into three classes as following (Fig. 5):

- (i) Class 1 was constituted of racemates where the first eluted enantiomer was eluted before the prediction, and the second eluted enantiomer was closer to the prediction. In this case, the first eluted enantiomer was less favourable to close interaction with the stationary phase, which could result from some steric repulsion, or repulsive stereo-induced interactions. Also included in this class were racemates where both enantiomers eluted before the prediction, because we believed that inaccuracy of the descriptors calculated with Absolv software might be partially responsible for the imperfect concordance between theoretical and experimental retention factors.
- (ii) Class 2 was constituted of racemates whose second enantiomer was eluted later than the prediction, while the first eluted enantiomer was eluted closer to the prediction. In this case, attractive stereo-induced interactions may be stronger for the second eluted enantiomer. Also included in this class were the racemates with both enantiomers eluted later than the prediction.
- (iii) Class 3 was constituted on the co-eluted racemates.

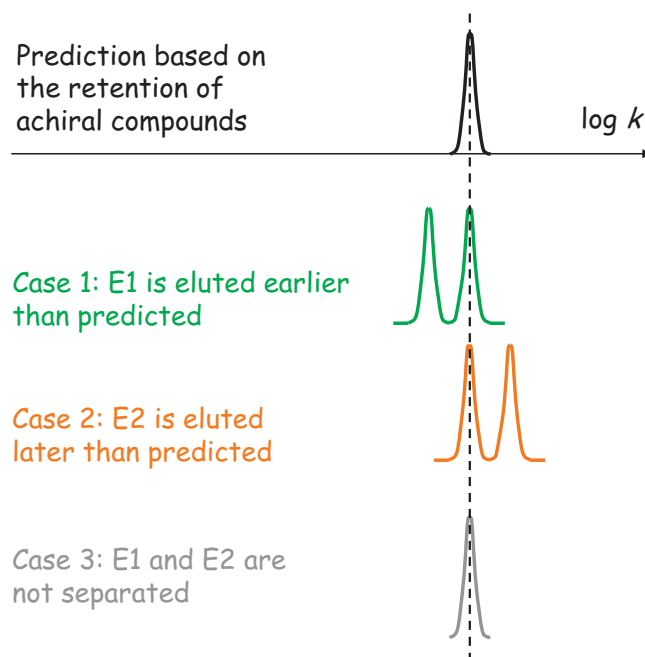


Fig. 5. Method used to determine the belonging to a class for the calculation of FDA based on three classes. See text for details.

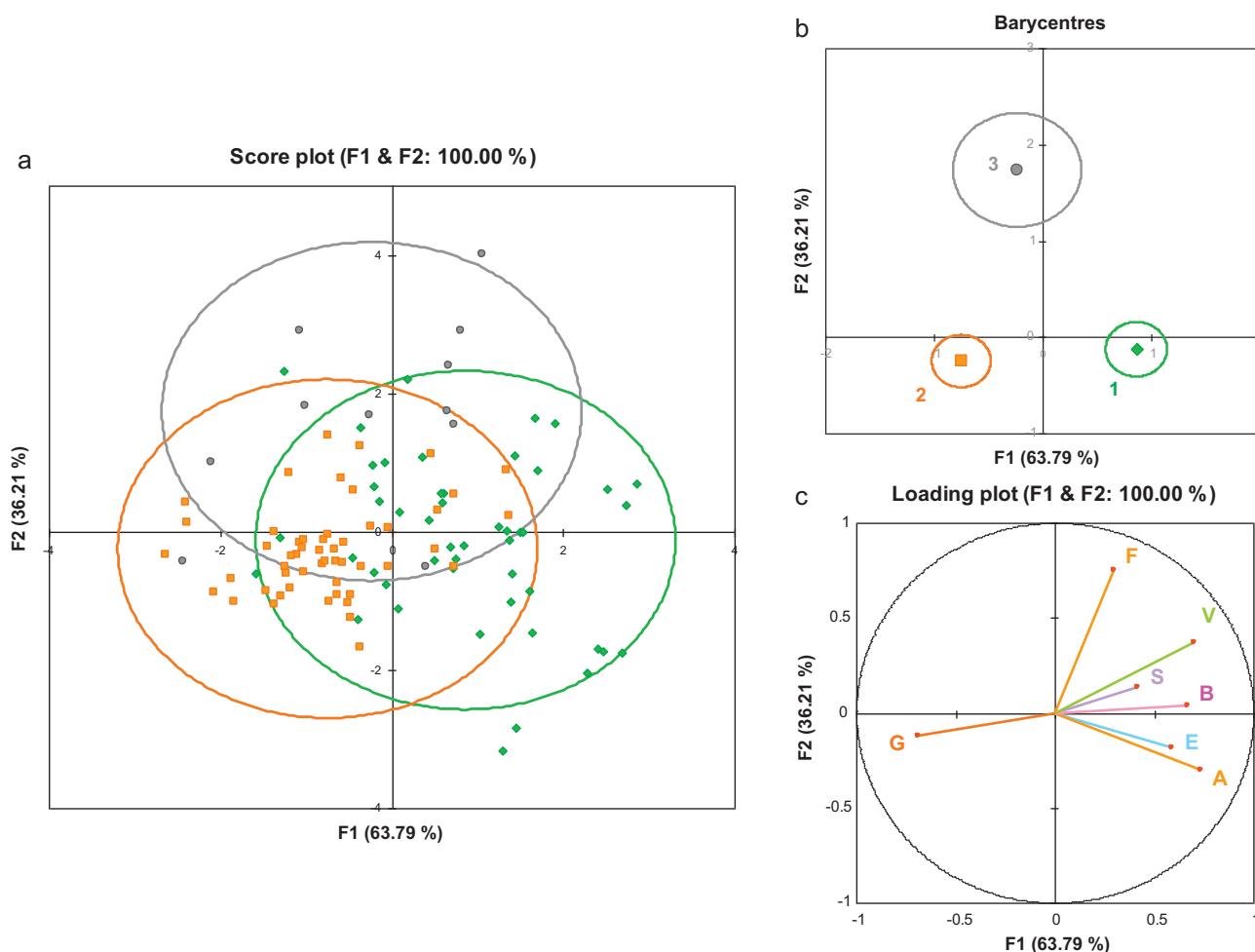


Fig. 6. Factorial discriminant analysis based on three classes of racemates (as defined in Fig. 5) on ADMPC, using the seven molecular descriptors of the modified solvation parameter model as variables. (a) Score plot, (b) barycentre plot and (c) loading plot. Green diamonds are racemates in class 1; orange squares are racemates in class 2; black circles are racemates in class 3. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

It is also possible that several enantioselective sites coexist on the CSP. It could thus happen that one enantiomer interacts with one kind of sites, while the other enantiomer interacts with another kind of sites [25].

The results are presented in Figs. 6 and 7 for ADMPC and CDMPC, respectively.

Again, the statistics associated to the FDAs must be discussed first, to ensure possible interpretation of the results. A statistical evaluation of the results can be found in the [supplementary material section](#). The statistics are reasonably good thus the results are amenable to interpretation.

It can be concluded from the number of solutes in each class that the two CSPs again behave very differently towards the enantioseparation of racemates of our test-set. Indeed, compounds belonging to one class on one phase can belong to another class on the other phase. This provides us confidence in the validity of this approach, as the scatter of points in Figs. 6 and 7a was clearly dependent on different chromatographic behaviours of the two CSPs.

As only three classes were considered, the first two discriminant axes carried 100% of the variance, indicating that the figures presented were suitable for correct interpretation of the data. Indeed, an advantage of FDA compared to PCA is that the number of discriminant axes always equals to the smaller value between V and $(C - 1)$, where V is the number of variables and C is the number of classes. Therefore, for three classes, only two discriminant functions are necessary to reach 100% variance explained. With PCA,

on the contrary, the first two principal components often carry a smaller proportion of the total information, rendering interpretation of the graphs less certain.

For the ADMPC phase, as shown in the score plot of Fig. 6a, class 3 was not clearly distinguished from the other two. Again the very small number of compounds in class 3 on ADMPC was certainly responsible for the difficulty in determining common features to the solutes in this class. However, the barycentre plot (Fig. 6b) indicated that classes 1 and 2 were clearly discriminated. Based on the Fischer weights (F) and the corresponding probabilities (p), all seven parameters were significant to explain the repartition in classes. The following conclusions can be drawn from the observation of Fig. 6c:

- Flexibility (F) is not favourable to separation, as was already observed above.
- Molecular volume and globularity both affect retention but in opposite ways. Large solutes are preferentially eluted earlier than predicted based on the retention of achiral solutes (V is pointed to the right-hand side), while small solutes are eluted later. A possible explanation is that large volume is not favourable to insertion in the chiral grooves, thus the first eluted enantiomer for large racemates would tend to be sterically excluded. For compounds of smaller volume, on the contrary, prior insertion in the chiral cavities favours additional stereo-induced interactions, causing increased retention of the second enantiomer.

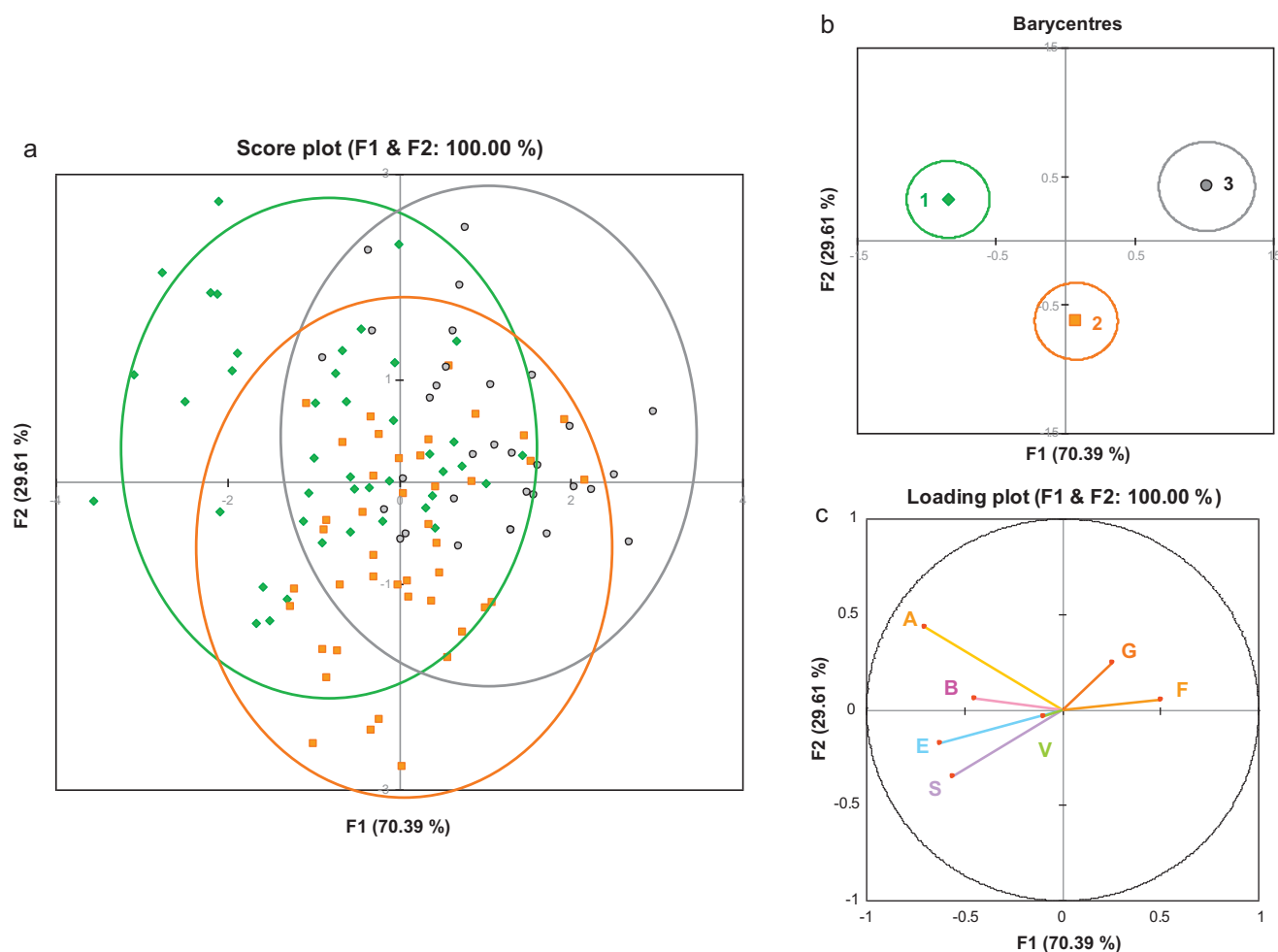


Fig. 7. Factorial discriminant analysis based on four classes of racemates (as defined in Fig. 5) on CDMPC, using the seven molecular descriptors of the modified solvation parameter model as variables. (a) Score plot, (b) barycentre plot and (c) loading plot. Green diamonds are racemates in class 1; orange squares are racemates in class 2; black circles are racemates in class 3. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

This is particularly true to globular species (*G* is pointed to the left-hand side). As mentioned above, small globular molecules may have more possibilities for simultaneous close interactions with the stationary phase, favouring a longer retention than predicted.

- All polar characteristics (*E*, *S*, *A*, *B*) are favourable for separation of solutes in class 1, but not so much for solutes in class 2. Possible interpretations are as follows: when steric fit is not favoured due to large volume, the interactions of the dipole–dipole, π – π and hydrogen-bonding type play a major role for chiral recognition. On the contrary, when steric fit is favourable and the enantiomers can enter the chiral cavities thanks to their small volume, steric adaptation and hydrogen-bonding acidity become the driving force for chiral recognition while other factors have smaller impact.

Our findings of the predominant contributions of volume, globularity and hydrogen-bonding to the separation mechanism are consistent with the literature. Booth and Wainer [5] suggested a “conformationally driven” chiral recognition process: they indicated that QSERR and molecular modelling suggested that the chiral recognition process on ADMPC involved insertion of the solute into a ravine on the surface of the CSP as well as stabilization of the solute–CSP complex by formation of a hydrogen bond within the ravine.

For the CDMPC phase, the discrimination of the three classes may be less clear when looking at the confidence ellipses in the score plot (Fig. 7a), but the barycentre figure (Fig. 7b) clearly indicates significant discrimination of the three classes of solutes. Based on the Fischer weights (*F*) and the corresponding probabilities (*p*), molecular volume was not significant to explain the class repartition; globularity was of little significance, while the other five parameters were highly significant. The following conclusions can be drawn from Fig. 7c:

Flexibility (*F*) and globularity (*G*) are not favourable features to chiral separation on CDMPC.

Polar interactions have the same effect in distinguishing class 1 from class 2 solutes, similar to ADMPC: polar solutes were preferably eluted earlier than predicted (class 1), while less polar solutes were eluted later (class 2). Hydrogen-bond acidity of the solute is again the most significant contributor to class 1 solutes, indicating the pre-dominance of hydrogen-bonding interactions in the separation process for solutes in this class.

Thus on both columns, two major classes of solutes (class 1 and class 2) were defined, in which different factors affected the enantioseparations. This conclusion is in accordance with past studies, indicating that the versatility of polysaccharide CSPs for successful chiral separation is related to the different enantioselective mechanisms occurring depending on the compounds [26].

The versatility of the polysaccharide based columns also comes with the fact that the chiral separations are hard to predict based on the analyte structures. Nevertheless, during pharmaceutical drug discovery, we often find that one particular chiral column works well for a series of molecules within one particular chemotype. Several examples in the literature indicate that molecules sharing similar structural properties also interact with the CSPs through a common mode. Our results essentially confirm this point. Indeed, when compound families are observed, they generally behave homogeneously towards one chromatographic system. For instance, on the ADMPC phase, benzodiazepines belong to class 1 while barbiturates belong to class 2. On the CDMPC phase, benzodiazepines also belong to class 1, while barbiturates belong to class 1.

3.5. Multiple linear regression analyses based on two classes of solutes

To further explain the difference in interactions between the separated enantiomers, multiple linear regression analyses were carried out again, but this time considering one group of solutes at a time, based on the classes of solutes established above.

For solutes in class 1 (as defined in the preceding paragraph), since the second eluted enantiomer was close to the prediction based on achiral solute retention, its retention behaviour can be described by the equations in Table 2 (also represented in the middle part of Fig. 8). As for the first eluted enantiomers for racemates in class 1, they must have weaker interactions with the stationary phase than the predicted based on achiral solute retention. Thus

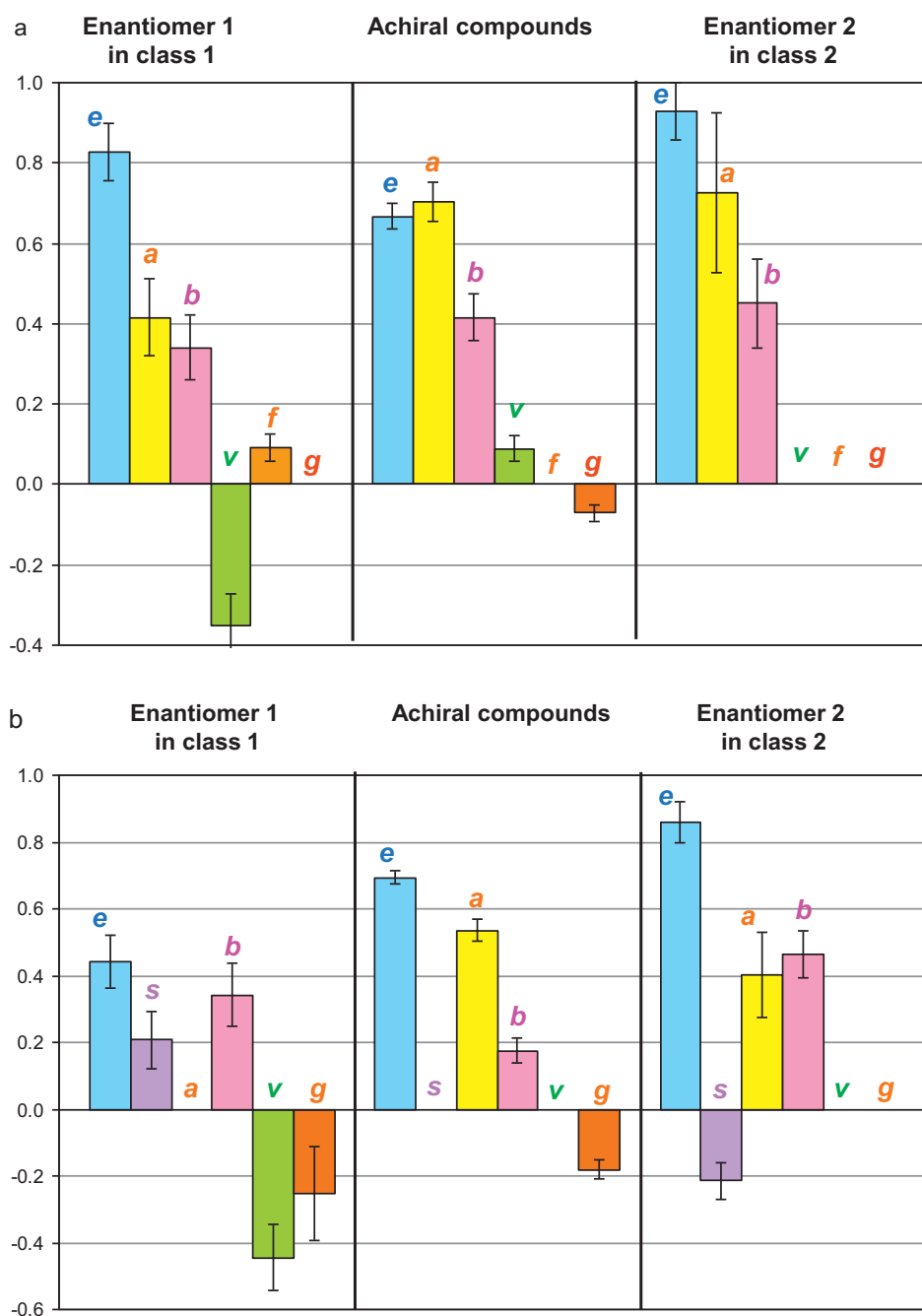


Fig. 8. Multiple linear regression analyses on (a) ADMPC and (b) CDMPC. See text for details.

considering the retention factors of the first eluted enantiomers of all racemates in class 1, new multiple linear regression analyses were calculated, using Eq. (1), for each CSP. The results appear in Table 2 and on the left hand-side of Fig. 8a for ADMPC, Fig. 8b for CDMPC.

The comparison of the models on the left and middle of Fig. 8 are thus indicative of the variation in the interactions faced by the two enantiomers of the racemates in class 1.

For solutes in class 2 (as defined in the preceding section), since the first eluted enantiomer is close to the prediction based on non-enantiospecific interactions, its retention behaviour can be described by the equations in Table 2 (also represented in the middle part of Fig. 8). As for the second eluted enantiomer for racemates in class 2, it must have stronger interactions with the stationary phase than predicted based on achiral solute retention. Thus considering the retention factors of the second eluted enantiomers of all racemates in class 2, new multiple linear regression analyses can be calculated, using Eq. (1), for each CSP. The results appear in Table 2 and on the right hand-side of Fig. 8a for ADMPC, Fig. 8b for CDMPC.

The comparison of the models on the middle and right of Fig. 8 are thus indicative of the variation of interactions faced by the two enantiomers for racemates in class 2.

The sufficient number of racemates in both classes made these calculations possible for both columns. The statistics associated with each equation can be observed in Table 2. A sufficiently large number of solutes was always retained in the final equations (between 36 and 48), although some outliers were eliminated (presumably because of the lower quality of their descriptors resulted from Absolv calculations). This was in accordance with the basic requirements of QSRRs, with more than four solutes per descriptor remaining in the test sets. Nevertheless, since the number of solutes is far less than that used to establish the equations for achiral retention, only the most significant variations will be pointed out in the following. Besides, the adjusted correlation coefficients are quite high (ranging from 0.798 to 0.912), which is comparable to the ones obtained for achiral retention.

Fig. 8 clearly displays the differences in the variation of coefficients between class 1 and class 2 solutes, and between the two CSPs. From different approaches, we arrived at the same conclusion that different enantioseparation mechanisms applied for different types of solutes on a CSP. It explains why our initial attempt in establishing a multilinear regression analysis based on the complete data set was not successful.

On ADMPC, considering class 1, the most significant variation is the large increase in the ν coefficient, followed by the increase in the a coefficient. The former indicates that the first eluted enantiomer establishes less dispersive interactions with the stationary phase, while the latter indicates that the first eluted enantiomer also established less hydrogen-bonding interactions comparing to the second enantiomer. This is consistent with the above observation that the acidic character is a strong feature of solutes in class 1.

Besides dispersive interactions, the ν coefficient is also related to the cavity formation process. In SFC, cavity formation in the mobile phase is negligible because the bulk mobile phase is not highly cohesive, contrary to what occurs with aqueous RP-HPLC mobile phases. However, the cohesiveness of the stationary phase like the macromolecular derivatized polysaccharides is not negligible because neighbouring bonded ligands may interact through dispersive interactions, π - π interactions, or hydrogen bonding. As a result, for interactions between the solute and the stationary phase to occur, the solute may have to overcome the intra-stationary phase interactions first to enter the stationary phase. Moreover, the methanol modifier may also be associated with the carbamate groups through hydrogen bonding. Thus the solutes would

also need to displace adsorbed solvent molecules in a competition process. Both mechanisms require more energy for the larger solutes. Consequently, increase of the molecular volume has a negative effect on retention. Since the stationary phases with polar ligands are highly cohesive, the ν coefficient is often negative for such phases in SFC [27].

Consequently, the negative contribution of volume to retention of the first eluted enantiomer in class 1 is likely due to steric resistance to insertion and/or competition with adsorbed solvent molecules, which can be interpreted in two manners:

- Either it means that for these enantiomers, steric constraints make it difficult to enter the chiral grooves. This is consistent with the above observation that large solutes preferably belong to class 1.
- Or it means that the free energy associated to the positive interactions occurring between this enantiomer and the stationary phase are not sufficient to counterbalance the energy required to create a cavity to accommodate this enantiomer in the stationary phase (or pseudo-stationary phase, if one considers the presence of the methanol modifier adsorbed).

In both cases, the result is the same: the enantiomer was prevented from hydrogen bonding interaction with the carbamate groups, which resulted in reduced a -type interactions. This interpretation is in accordance with previous works [28], indicating that retention on ADMPC is a two-step process, involving a first hydrogen-bonding interaction with functional groups at the outer edge of the CSP helical cavity, while conformational adjustment would permit extra- or stronger-hydrogen bonding.

The significantly larger e coefficient measured in this case is also reasonable, if one considers the fact that, even to solutes facing steric repulsion from the stationary phase, the aromatic end-groups of the bonded ligands located at the surface of the polymer remain accessible, thus become the major contributors to retention for such solutes [19]. In this respect, flexibility is also an advantage for retention of these compounds, because flexible molecules can better adapt to the surface of the CSP, possibly explaining the significant (although rather small) positive value of the f coefficient.

Considering class 2, the most significant difference resided in the increase in e coefficient, indicating that the conformation of the second eluted enantiomer was more favourable to π - π interactions than the conformation of the first enantiomer. In particular, hydrogen bonding interactions do not vary significantly in this case, indicating that both enantiomers are able to establish equally strong hydrogen-bonding interactions with the CSP. However, judging from the large standard deviations associated to the a and b terms, we would remain very cautious on this conclusion.

It seems that, for the racemates in class 2, steric effects have less part in the chiral recognition mechanism, as the ν , f and g coefficients do not vary much between the two enantiomers.

On a practical note, these results indicate how separation factors would vary within a compound family. For instance, on ADMPC, most propionic acids belonged to class 1. The enantioseparation for these compounds should thus be essentially related to the ν and a terms. In other words, those compounds exhibiting the largest molecular volume (V) or largest acidic character (A) should exhibit largest separation factors. A simple example is presented in Fig. 9a, where the separation of 2-phenylpropionic acid (solute no. 3 in Table 1) and ibuprofen (solute no. 55 in Table 1) can be compared. Judging from their descriptor values, ibuprofen displays slightly larger A and B values, but essentially a much larger V value, due to the *para*-substituted alkyl chain. This is clearly favourable for enantioseparation, as the separation factor increases from $\alpha(2\text{-phenylpropionic acid}) = 1.09$ to $\alpha(\text{ibuprofen}) = 1.22$.

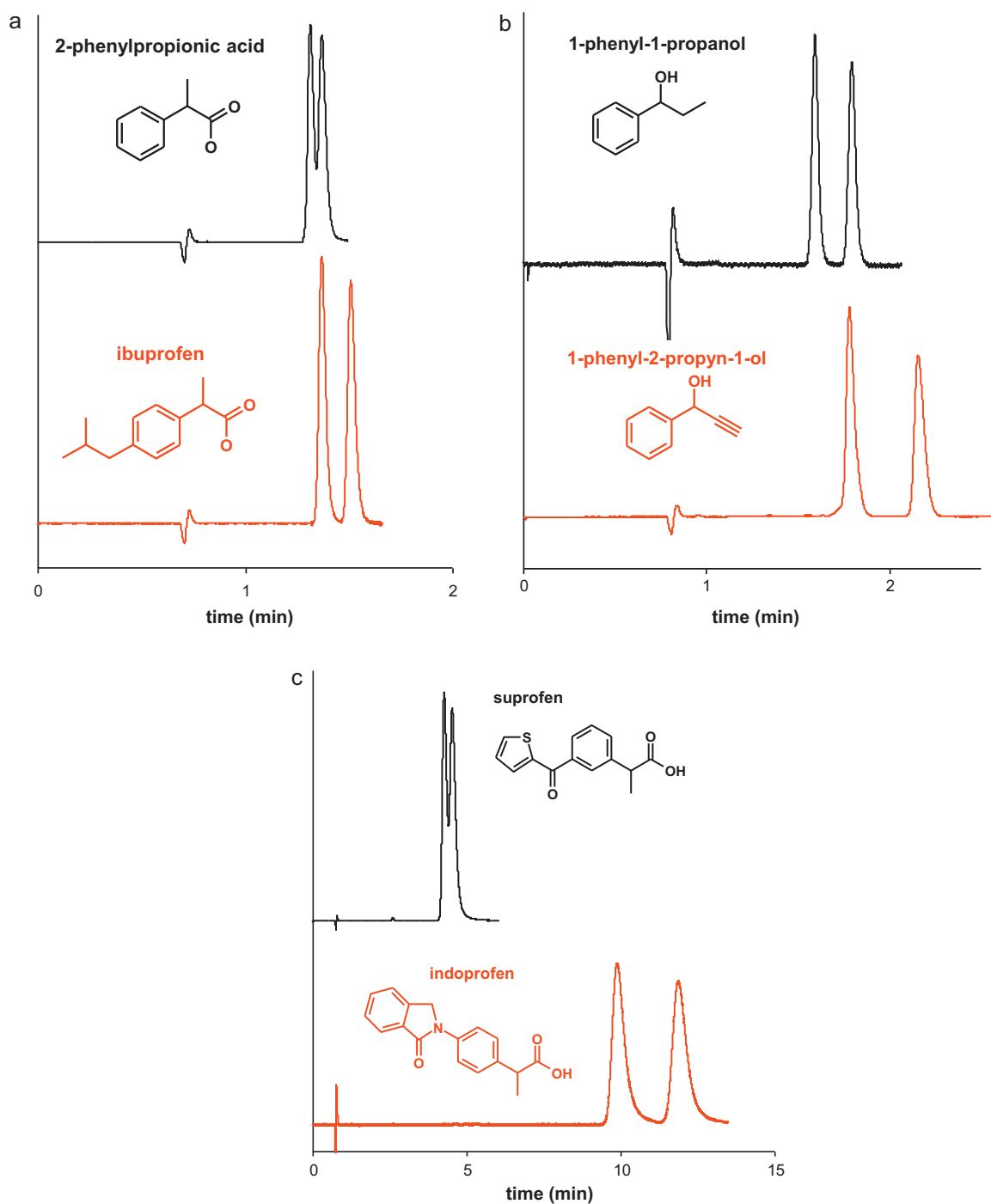


Fig. 9. Chromatographic separations obtained for (a) 2-phenylpropionic acid and ibuprofen on ADMPC, (b) 1-phenyl-1-propanol and 1-phenyl-2-propyn-1-ol on CDMPC and (c) suprofen and indoprofen on CDMPC. Conditions: CO₂-MeOH 90:10 (v/v), 25 °C, 15 MPa, 3 mL/min.

On CDMPC, for class 1 solutes, the difference between the first and the second eluted enantiomer was principally, again in the ν term (dispersive interactions and cavity effect), in the a -type hydrogen-bonding interactions, followed by the e π - π interactions. This indicates that the first eluted enantiomer, being somewhat excluded from the chiral grooves, was less able to interact with the phenylcarbamate moieties.

For class 2, the b and e coefficients vary most significantly. Hydrogen-bonding interactions of the a type (acidic solute interacting with the carbonyl group of the CSP) do not vary significantly, while hydrogen-bonding interactions of the b type (basic solutes interacting with the -NH- group of the CSP) are stronger for the sec-

ond enantiomer. Additionally, the diastereomeric complex formed by the second enantiomer with the CSP is stabilized through π - π interactions.

A significant variation of the g coefficient could also indicate that a favourable steric fit of the second enantiomer participates to the separation process for racemates in class 2.

Similarly to the above remarks, these observations can be used to deduce how well a given couple of enantiomers can be separated, based on their interaction capabilities. For instance, on CDMPC, 1-phenyl-1-propanol (solute no. 6 in Table 1) and 1-phenyl-2-propyn-1-ol (solute no. 8 in Table 1) both belong to class 1. Enantioseparation should thus be essentially related to the e , a

and ν terms. There is no significant difference between the molecular volumes of these compounds, but the capacity for e -type interactions is clearly increased for 1-phenyl-2-propyn-1-ol with its triple-bond contributing to an increased E value. This results in an increased enantioseparation as indicated by separation factor values: $\alpha(1\text{-phenyl-1-propanol}) = 1.26$ and $\alpha(1\text{-phenyl-2-propyn-1-ol}) = 1.38$ (Fig. 9b).

Another example is that of propionic acids. On CDMPC, suprofen (solute no. 85 in Table 1) and indoprofen (solute no. 57 in Table 1) both belong to class 2. Thus, the enantioseparation for these compounds should be essentially related to the e and b terms. Indoprofen exhibits significantly larger values of the E and B terms as compared to suprofen, and thus exhibits a larger separation factor, as appears in Fig. 9c: $\alpha(\text{indoprofen}) = 1.22$ while $\alpha(\text{suprofen}) = 1.07$.

Finally, it is important to note that our interpretation of the results is based on the experimental data acquired under the above-mentioned operating conditions. It remains to be proved whether they can be extrapolated to describe separation mechanisms by varying experimental conditions. This point is currently under investigation and will be discussed in subsequent papers.

4. Conclusions

In this second part of our work on characterization of enantioselective stationary phases in SFC, we have shown that a simple multiple linear regression analysis relating separation factors to molecular descriptors using a modified version of the solvation parameter model yielded no result when using the large data sets inclusively. However, multiple linear regression analyses were effective when using smaller data sets, that were classified based on the retention factor predicted from achiral solutes. It has illustrated that the mechanisms for enantioselectivity are different for different groups of solutes. This innovative approach has proven to be useful in unravelling the interactions taking part in the enantioselectivity process on the two polysaccharide CSPs, *tris*-(3,5-dimethylphenylcarbamate) of amylose and cellulose.

Using factorial discriminant analyses, we have demonstrated that the reasons for successful enantioseparation are clearly different on the two CSPs, with a clear advantage for the ADMPC phase based on our test solutes. Indeed, steric fit associated to hydrogen bonding seems to be the most important features for good enantioselectivity on ADMPC, while not on CDMPC. Enantioselectivity on CDMPC requires other interactions: not only hydrogen bonding, but also dipole–dipole and π – π interactions are necessary.

These results are important as they show that the interactions providing the principal contribution to retention (such as π – π interactions) are not necessarily the major contributor to enantioselectivity separation. Two additional descriptors (flexibility and globularity) we have introduced in the first part of this work have proven to be highly relevant in the description of the enantioselectivity process, as the statistical quality of the calculated models was significantly deteriorated when these two parameters were not considered.

FDA coupled to the modified solvation parameter model has demonstrated to be a powerful tool to discriminate separable and not-separable racemates, as the accuracy for re-classification of the test racemates reached 93% and 83% for ADMPC and CDMPC,

respectively. Good statistics and cross-validation experiments indicate that the predictive ability of the models should be quite high. Since the data set contained molecules with a wide diversity of structures, the methodology should display satisfying applicability to new molecular structures, as will be investigated in a subsequent paper. So far, no test set was considered, because the primary goal of this work was to get an insight in chiral recognition mechanisms, prior to any prediction purposes.

Clearly, this approach should not be used for small data sets comprising structurally related compounds, as they may not statistically well represent the interaction capabilities of a given chromatographic system.

Finally, it must be pointed out that this study relies upon experimental data performed under strictly identical operating conditions (methanol modifier), thus the role of mobile phase components and other operating parameters, such as temperature and pressure, in the chiral recognition process remains to be investigated.

Acknowledgment

Nathalie Percina is acknowledged for highly appreciated technical assistance.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.chroma.2010.11.085.

References

- [1] D. Mangelings, Y. Vander Heyden, *J. Sep. Sci.* 31 (2008) 1252.
- [2] A. Del Rio, *J. Sep. Sci.* 32 (2009) 1566.
- [3] R.B. Kasat, N.-H.L. Wang, E.I. Franses, *J. Chromatogr. A* 1190 (2008) 110.
- [4] R.B. Kasat, E.I. Franses, N.-H.L. Wang, *Chirality* 22 (2010) 565.
- [5] T.D. Booth, I.W. Wainer, *J. Chromatogr. A* 737 (1996) 157.
- [6] T.D. Booth, K. Azzouli, I.W. Wainer, *Anal. Chem.* 69 (1997) 3879.
- [7] T.D. Booth, W.J. Lough, M. Saeed, T.A.G. Noctor, I.W. Wainer, *Chirality* 9 (1997) 173.
- [8] M. Dervarics, R. Ötvös, T.A. Martinek, *J. Chem. Inf. Model.* 46 (2006) 1431.
- [9] C. Yhang, C. Zhong, *QSAR Comb. Sci.* 24 (2005) 1047.
- [10] A. Kovatecheva, A. Golbraikh, S. Oloff, J. Feng, W. Zheng, A. Tropsha, *SAR QSAR Environ. Res.* 16 (2005) 93.
- [11] A. Del Rio, J. Gasteiger, *J. Chromatogr. A* 1185 (2008) 49.
- [12] C.R. Mitchell, D.W. Armstrong, A. Berthod, *J. Chromatogr. A* 1166 (2007) 70.
- [13] C.R. Mitchell, N.J. Benz, S. Zhang, *J. Chromatogr. B* 875 (2008) 65.
- [14] M.F. Vitha, P.W. Carr, *J. Chromatogr. A* 1126 (2006) 143.
- [15] C. West, Y. Zhang, L. Morin-Allory, *J. Chromatogr. A* 1218 (2011) 2019.
- [16] A. Berthod, C.R. Mitchell, D.W. Armstrong, *J. Chromatogr. A* 1166 (2007) 61.
- [17] P. Borman, B. Boughtflower, K. Cattanach, K. Crane, K. Freebairn, G. Jonas, I. Mutton, A. Patel, M. Sanders, D. Thompson, *Chirality* 15 (2003) S1.
- [18] S. Svensson, A. Karlsson, O. Gyllenhaal, J. Vessman, *Chromatographia* 51 (2000) 283.
- [19] X. Chen, C. Yamamoto, Y. Okamoto, *Pure Appl. Chem.* 79 (9) (2007) 1561.
- [20] J.H. Kennedy, *J. Chromatogr. A* 725 (1996) 219.
- [21] R. Kafri, D. Lancet, *Chirality* 16 (2004) 369.
- [22] Y. Okamoto, Y. Kaida, *J. Chromatogr. A* 666 (1994) 403.
- [23] A. Medvedovici, P. Sandra, L. Toribio, F. David, *J. Chromatogr. A* 785 (1997) 159.
- [24] Y. Okamoto, T. Ikai, *Chem. Soc. Rev.* 37 (2008) 2593.
- [25] T. Fornstedt, P. Sajonz, G. Guiochon, *Chirality* 10 (1998) 375.
- [26] A. del Rio, P. Piras, C. Roussel, *Chirality* 17 (2005) S74.
- [27] C. West, E. Lesellier, *J. Chromatogr. A* 1110 (2006) 200.
- [28] T.D. Booth, I.W. Wainer, *J. Chromatogr. A* 741 (1996) 205.