ELSEVIER

Journal of Chromatography A, 922 (2001) 13–23

JOURNAL OF CHROMATOGRAPHY A

www.elsevier.com/locate/chroma

Quantitative structure–enantioselective retention relationships for chromatographic separation of arylalkylcarbinols on Pirkle type chiral stationary phases

Takahiro Suzuki[a], Simona Timofei[b], Bianca E. Iuoras[c], Georg Uray[d], Petra Verdino[d], Walter M.F. Fabian[d,*]

[a]*Chemical Resources Laboratory, Tokyo Institute of Technology, 4259 Nagatsuta, Midori-ku, Yokohama 226-8503, Japan*
[b]*Institute of Chemistry, Romanian Academy, Bul. Mihai Viteazul 24, 1900 Timisoara, Romania*
[c]*Department of Organic Chemistry, Faculty of Chemistry and Chemical Engineering, Babes–Bolyai University, Arany Janos 11, 3400 Cluj-Napoca, Romania*
[d]*Institut für Chemie, Karl-Franzens Universität Graz, Heinrichstrasse 28, A-8010 Graz, Austria*

Received 9 March 2001; received in revised form 25 April 2001; accepted 26 April 2001

Abstract

Quantitative structure–retention (QSRR, retention factors log k_1 and log k_2 for the first and second eluted enantiomer) as well as enantioselective retention relationships (QSERR, separation factor log α) for a series of 42 chiral arylalkylcarbinols on four brush-type chiral stationary phases are derived by multiple linear regression analyses and artificial neuronal network calculations using 2D and 3D molecular descriptors including those obtained by quantum chemical calculations. Separation factors are in addition modeled by the 3D-QSAR method of comparative molecular field analysis (CoMFA). For the retention factors the LUMO energy turns out to be the most important descriptor, whereas for log α it is the hydrophobicity of the analytes. With CoMFA both the steric and electrostatic field are found to be of almost comparable significance.

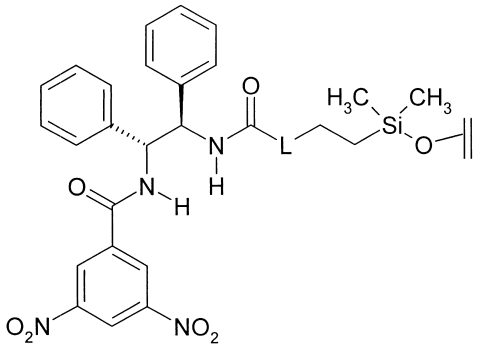
Keywords: Artificial neural network; Retention factors; Chiral selectors; Comparative molecular field analysis; Enantioselectivity; Structure–retention relationships; Arylalkylcarbinols

1. Introduction

1,2-Diamine derivatives have broad applicability as chiral selectors of chiral stationary phases (CSPs) in liquid chromatographic enantioseparation [1–4]. Recently, it has been shown that derivatives of (*R*,*R*)-*N*-3,5-dinitrobenzoyl-1,2-diphenyl-1,2-diamine (CSP I–IV, Scheme 1) make possible the enantioseparation of underivatized arylalkylcarbinols with considerable levels of enantioselectivity, good band shapes as well as short elution times [5,6]. Furthermore, separation of multifunctionalized and/or sterically demanding arylalkylcarbinols could be achieved demonstrating the versatility of these chiral stationary phases. However, on the other hand there is no reliable basis to decide which CSP should be used for a particular separation of interest. Conse-

*Corresponding author. Tel.: +43-316-380-8636; fax: +43-316-380-9840.
E-mail addresses: walter.fabian@uni-graz.at (W.M.F. Fabian), http://www-ang.uni-graz.at/~fabian (W.M.F. Fabian).

0021-9673/01/$ – see front matter
PII: S0021-9673(01)00921-9



L

CSP I	-$(CH_2)_8$-
CSP II	-$(CH_2)_2$-
CSP III	-NH-$(CH_2)_8$-
CSP IV	-NH-CH_2-

Scheme 1.

quently, to facilitate the selection of the most appropriate CSP, numerous attempts to correlate experimental retention data (log k) with a variety of molecular descriptors to obtain quantitative structure–retention relationships (QSRR) and — since the intermolecular interactions responsible for retention need not necessarily be the same as for enantioseparation [7] — quantitative structure–enantioselective retention relationships (QSERR) have been undertaken [8–17]. In addition to such multivariate regressions recently neural networks, which already have been successfully applied to retention prediction in achiral chromatography [18–31], were extended to the enantioselective chromatographic separation behavior [13]. Especially for brush-type CSPs — to which CSPs I–IV belong — quite extensive atomistic modeling of the intermolecular interaction energies responsible for enantioseparation also has been done [32–45]. As a complement and extension of classical QS(E)RR in analogy to 3D-QSAR studies of biological activity, the method of comparative molecular field analysis (CoMFA) [46–50] was also applied to retention (log k) [51,52] and enantioseparation (log α) data [7,53]. The purpose of this paper is to report a combined multiple linear regression (MLR)–artificial neural network (ANN)–comparative molecular field analysis (CoMFA) study on the enantioselectivity of the chromatographic retention behavior of a series of arylalkylcarbinols (**1**–**42**, Scheme 2) on four different stationary phases, CSP I–IV containing the same chiral backbone but different linker groups.

2. Methods

2.1. Molecular structures

The molecular structures of analytes **1**–**42** were generated by the SYBYL molecular modeling package [54], energy minimized using the Tripos force field [55] followed by a conformational analysis using the random search procedure [56] as implemented in SYBYL. Each one of these conformations was then subjected to energy minimization by the semiempirical AM1 [57] method using the VAMP program package [58]. The minimum energy structures found in this way for each compound was then used for further analyses.

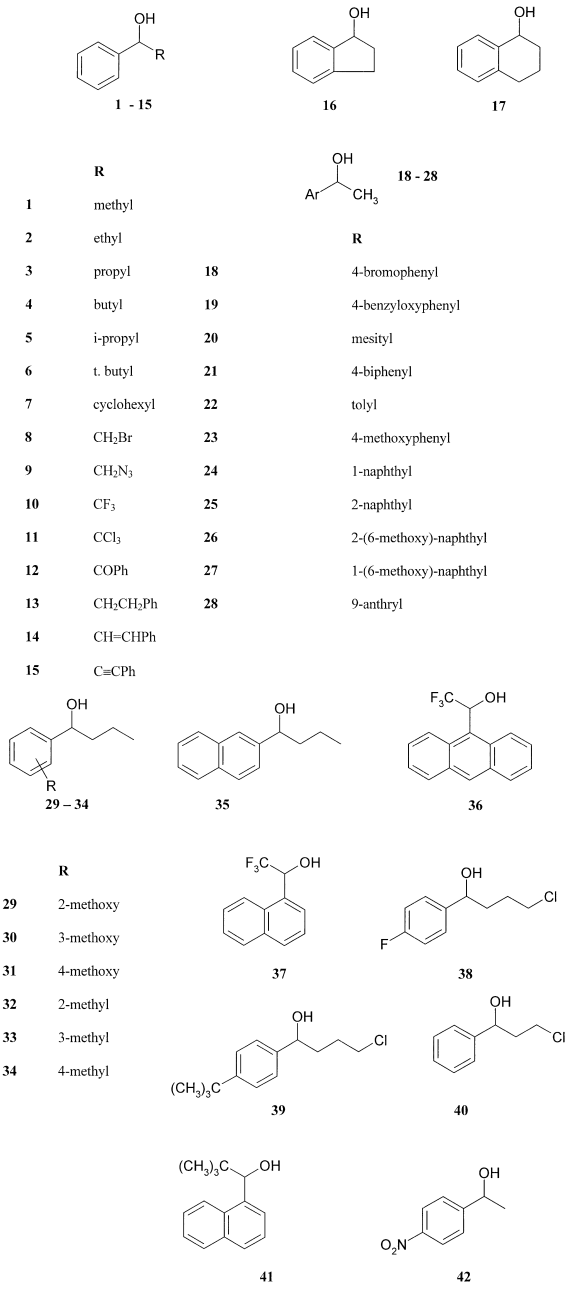
2.2. Multiple linear regression (MLR) analysis

The multiple linear regression (MLR) analysis [59] has been performed by the SYSTAT program, from SYSTAT, Evanston, IL, USA. The leave-one-out [60] and leave-half-out [61] cross-validation procedures were applied in order to verify the reliability of the results. MLR calculations were employed to model the retention capacity and enantioselectivity of the above mentioned arylalkylcarbinols on the chiral selectors CSP I–IV.

2.3. Definition of target properties and molecular descriptors

As target properties or dependent variables the individual retention factors of the respective enantiomers (log k_1, log k_2($k_2 = k_1 \times \alpha$)) and the separation factor (log α) were used. Experimental values of these variables are provided in Table 1.

Several steric descriptors derived from the optimized 3D molecular structures of analytes **1**–**42** were considered, e.g., bond and torsional angles of the atoms connected to the chiral center of the



Scheme 2.

analytes, as well as the distance from the chiral carbon atom to the last non-hydrogen atom of the alkylcarbinol moiety (DIST). The van der Waals volumes (V_W) were calculated by the additivity of van der Waals volume increments [62]. The molecular weight (MW) was used as a measure of the bulkiness of the arylalkylcarbinols. The hydrophobicity of the arylalkylcarbinols was evaluated by the logarithm of the octanol/water partition coefficient (log P) calculated by the ClogP software [63].

In addition, various quantum chemical descriptors calculated by the semiempirical AM1 method for the minimum energy structures of each one of the arylalkylcarbinols were applied: polarizability parameter (α_{mop}), dipole moment, charges on carbon atoms of the phenyl nucleus (carbon atom linked to the alkylcarbinol moiety and the carbon atoms from its neighborhood), charge on the chiral carbon atom, charge on the carbon atom attached to the chiral atom, charge on the oxygen atom attached to the chiral molecular center, charge on the hydroxy hydrogen atom, charge on the hydrogen atom bonded to the chiral center, the sum of charges of the carbon atoms of the phenyl moiety attached to the chiral center (Σq_{Car}), the maximum and minimum atomic charge of the molecule, HOMO (ϵ_{HOMO}) and LUMO (ϵ_{LUMO}) molecular orbital energies, the sum of electrophilic superdelocalizabilities of the carbon atoms of the phenyl moiety attached to the chiral center, and the difference between the maximum and minimum atomic charges. Starting from this entire data set of 27 structural descriptors, intercorrelations between these descriptors have been inspected for the set of 42 compounds. Variable selection by a stepwise regression procedure based on the Fischer test was performed. Because of the statistical quality of the obtained models, outliers have been tested by the externally Studentized residuals [64] and have not been included in the final proposed models from Table 2. All the statistical tests were done at a significance level of 5% or less.

2.4. Nonlinear modeling by ANN

A possibly nonlinear form of the QSRR and QSERR was examined with the aid of artificial neural networks (ANN) [65]. A three-layer neural network with back-propagation of errors was chosen to develop nonlinear QSRR and QSERR models. The network consists of fully connected three layers: an input, a hidden and an output layer. The input and hidden layers have an additional node, the bias neuron. The input and output (response) sets to the

Table 1
Experimental retention factors k_1 of the first eluted enantiomer and enantioselectivity α for CSP I–IV [6]

Compound	CSP I		CSP II		CSP III		CSP IV	
	k_1	α	k_1	α	k_1	α	k_1	α
1	3.85	1.12	4.74	1.11	4.00	1.13	4.62	1.16
2	2.79	1.20	3.51	1.19	3.13	1.19	3.65	1.23
3	2.45	1.22	3.15	1.20	2.84	1.21	3.41	1.24
4	2.33	1.24	3.00	1.20	2.73	1.22	3.32	1.24
5	1.93	1.30	2.43	1.27	2.25	1.27	2.66	1.32
6	1.39	1.47	1.61	1.46	1.66	1.35	2.03	1.37
7	2.08	1.33	2.49	1.29	2.47	1.28	2.88	1.34
8	2.56	1.12	3.26	1.13	3.10	1.10	3.47	1.14
9	4.16	1.11	5.49	1.11	4.88	1.09	5.76	1.11
10	3.76	1.16	4.47	1.15	5.96	1.13	6.84	1.18
11	2.91	1.30	3.23	1.31	4.70	1.19	4.85	1.30
12	4.30	1.18	4.32	1.17	4.89	1.17	5.24	1.18
13	3.70	1.20	4.38	1.17	4.72	1.17	4.95	1.17
14	6.74	n.r.	7.80	n.r.	9.35	n.r.	9.67	n.r.
15	7.59	1.02	8.59	1.07	10.73	1.03	10.97	1.06
16	5.99	1.14	7.58	1.14	6.93	1.09	7.16	1.15
17	4.29	1.16	5.14	1.15	4.95	1.09	5.50	1.16
18	4.95	1.09	5.82	1.10	5.05	1.11	5.37	1.14
19	7.47	1.11	9.86	1.12	8.90	1.13	9.58	1.17
20	2.58	1.12	3.25	1.09	3.04	1.13	3.20	1.18
21	5.97	1.11	7.31	1.12	7.11	1.13	7.02	1.17
22	3.66	1.13	4.96	1.13	3.95	1.15	4.80	1.18
23	3.16	1.21	3.26	1.27	2.73	1.19	2.67	1.25
24	7.23	1.19	8.96	1.15	8.92	1.14	9.54	1.19
25	7.62	1.29	9.76	1.29	9.14	1.29	10.63	1.36
26	6.34	1.32	7.13	1.30	8.51	1.31	7.75	1.37
27	5.35	1.43	6.14	1.39	5.64	1.50	7.97	1.60
28	7.88	1.83	9.53	1.94	11.59	1.71	12.39	1.89
29	3.49	1.25	4.38	1.19	3.53	1.20	4.50	1.26
30	4.21	1.26	5.67	1.25	4.59	1.28	5.69	1.33
31	4.67	1.21	5.85	1.21	4.80	1.22	5.81	1.28
32	1.60	1.26	1.91	1.21	1.99	1.21	2.16	1.27
33	1.61	1.27	1.97	1.25	2.02	1.25	2.18	1.29
34	1.70	1.22	2.09	1.22	2.13	1.21	2.32	1.26
35	5.08	1.49	6.42	1.52	6.49	1.47	7.25	1.57
36	4.35	2.04	4.02	2.27	8.81	1.84	6.69	2.17
37	6.95	1.48	4.02	1.58	11.51	1.40	12.17	1.55
38	2.87	1.11	3.01	1.10	3.64	1.11	3.39	1.13
39	1.88	1.15	2.10	1.19	2.49	1.16	2.48	1.21
40	2.55	1.11	2.89	1.10	3.27	1.10	3.11	1.12
41	0.66	1.85	–[a]	–[a]	–[a]	–[a]	–[a]	–[a]
42	18.00	1.00	16.20	1.03	15.70	1.06	15.20	1.07

n.r., not resolved.
[a] No experimental data available.

network were just the same as the sets of descriptors and target variables which were used in the MLR modeling. A sigmoid function was selected as the transfer function for each neuron. The connection weights of the network were adjusted iteratively by the back-propagation algorithm with the generalized delta rule to minimize the mean square error between desired and actual outputs. All input and output data were scaled between 0.05 and 0.95. The quality of the fitting of the modeling results for the training set

Table 2
Results of the multiple linear regression analysis for retention factors (log k_1, log k_2) and separation factor (log α) found for analytes **1–42** on chiral selectors CSP I–IV[a]

		Descriptor[b]			r	s	F	r^2_{LHO}	r^2_{LOO}
CSP I	log k_1	α_{mop}(0.376)	ϵ_{LUMO}(0.456)	DIST(0.168)	0.699	0.159	10.2	0.344	0.351
	log k_2	α_{mop}(0.374)	ϵ_{LUMO}(0.430)	DIST(0.195)	0.755	0.140	14.1	0.449	0.467
	log α	ClogP(0.431)	Σq_{Car}(0.373)	DIST(0.196)	0.804	0.041	23.1	0.586	0.527
CSP II	log k_1	α_{mop}(0.481)	ϵ_{LUMO}(0.323)	DIST(0.196)	0.689	0.166	8.7	0.357	0.337
	log k_2	α_{mop}(0.469)	ϵ_{LUMO}(0.314)	DIST(0.217)	0.734	0.149	11.3	0.437	0.427
	log α	ClogP(0.389)	Σq_{Car}(0.439)	DIST(0.171)	0.803	0.042	22.4	0.555	0.507
CSP III	log k_1	α_{mop}(0.327)	ϵ_{LUMO}(0.609)	DIST(0.063)	0.765	0.149	14.6	0.392	0.443
	log k_2	α_{mop}(0.336)	ϵ_{LUMO}(0.571)	DIST(0.093)	0.821	0.130	21.4	0.530	0.574
	log α	ClogP(0.401)	Σq_{Car}(0.409)	DIST(0.189)	0.787	0.035	20.0	0.550	0.501
CSP IV	log k_1	α_{mop}(0.376)	ϵ_{LUMO}(0.549)	DIST(0.074)	0.753	0.149	12.7	0. 448	0. 434
	log k_2	α_{mop}(0.376)	ϵ_{LUMO}(0.514)	DIST(0.109)	0.787	0.136	15.8	0.543	0.515
	log α	ClogP(0.376)	Σq_{Car}(0.429)	DIST(0.195)	0.828	0.037	26.9	0.607	0.577

[a] Correlation coefficient r, standard error s, F-test and cross-validated correlation coefficients r^2_{LHO} and r^2_{LOO}.
[b] Relative contributions in parentheses.

was evaluated based on the correlation coefficient r and the root-mean-square error (RMSE). The number of hidden-layer neurons was kept variable in the range mentioned to test its influence on the predictive quality of the ANN model [66]. The best architectures were determined to be (3+1):(3+1):1 (three input neurons for the three descriptors plus a bias, three hidden-layer neurons plus a bias, and one output layer neuron for a total of 16 adjustable parameters) for log k_1 and log k_2, and (3+1):(4+1):1 (21 adjustable parameters) for log α. As pointed out by Andrea and Kalayeh [67], the important quantity in ANN modeling is not the overall number of connections but the ratio of the number of data points (training compounds) to the number of model parameters or connections. According to the recommended guidelines for determining the number of hidden neurons to be employed [67], the values for the ratio of the number of training compounds to the number of model parameters should be around 2 (1.95–2.25 in this study).

2.5. Comparative molecular field analysis (CoMFA)

For alignment cyclohexyl-phenylcarbinol **7** was used as the template because of its rather large value of α. The choice of this compound as the template was guided by the results of the MLR analysis. So, for instance, compounds **28** and **41**, which show even better enantioselectivity, were rejected as template for alignment since most experimental data are based on substituted phenyl rather than polycyclic aryl carbinols. Moreover, the anthryl derivative **28** was indicated to be an outlier on the basis of MLR analyses (see below). An analogous reasoning also indicated that compound **6** is a less suitable template for alignment. The remaining analytes were then fitted onto this compound using the chiral carbon and the four atoms attached to it for the fitting procedure. Since in all known cases on (*R*,*R*)-CSP I–IV exclusively the (*R*)-enantiomer or, more generally stated, the homochiral structure, is the most retained one [6], this isomer was used for alignment. Only in cases of an apparent (*R*)/(*S*) retention inversion as in trifluoromethyl derivatives, arising solely from the Cahn–Ingold–Prelog naming convention, the alternative structure was employed [13]. For the CoMFA analysis default settings (sp^3 carbon carrying a charge of +1 as probe atom, 2.0 Å grid spacing, 30 kcal mol^{-1} cut-off, standard CoMFA scaling) with a grid size of 22×22×17 Å and AM1 atomic charges were used. Both fields (steric and electrostatic) were included. Statistical analysis was done by the PLS procedure [68] using cross-validation (leave-one-out (LOO) [60], leave-half-out (LHO) [61] and leave-seven-out (L7O) [69]) to determine the optimal number of components to be used in the final PLS analysis without cross-validation. As dependent variable the enantioselectivity (log α) was used. Gener-

ally, as found previously [50], the leave-one-out cross-validation procedure yields too optimistic results although cross-validated r^2-values obtained by the leave-seven-out method are close to r^2_{LOO}. In some cases, e.g., the leave-one-out cross-validation for CSP I, a decrease of r^2_{LOO} with an increasing number of components indicating overprediction, is evident. Furthermore, frequently increasing the number of components beyond a certain value only leads to an insignificant improvement of r^2_{CV}. Thus, in such cases fewer than the suggested optimum number of components was used [50] in the final PLS analysis without cross-validation. For a smaller subset of compounds (20 molecules) the influence of the grid resolution (1 Å vs. 2 Å) as well as of the weighting procedure (CoMFA standard vs. NONE) was tested. Generally, cross-validated r^2_{CV} are only slightly larger when the much closer grid spacing is used. Thus, using the default value of 2 Å appears to be fully sufficient. Standard scaling of the CoMFA fields yields somewhat larger values of r^2_{CV} than does scaling ''NONE''. The CoMFA analysis for the data obtained on CSP I was, in addition, performed separately in terms of the electrostatic and steric fields.

3. Results and discussion

In the following modeling results obtained for the three kinds of experimental data (log k_1, log k_2, and log α) by MLR and nonlinear modeling via ANN as well as 3D QSAR (CoMFA) analyses of enantioselectivity (logα) for analytes **1**–**42** on CSPs I–IV will be presented.

3.1. Retention behavior prediction by MLR

Starting from the set of descriptors and the variable selection procedure described above, the MLR analysis has been applied to model the retention factors (log k_1 and log k_2), i.e. retention behavior, as well as enantioselectivity (log α) of the arylalkylcarbinols for the selectors CSP I–IV, and to identify the significant factors contributing to the separation for ANN modeling.

The final regression models for the target properties obtained by the MLR analysis are collected in Table 2.

Several outliers were detected for *k*-values and hence retention, during the MLR analysis. Thus, for the retention behavior on all four stationary phases CSP I–IV compounds **16**,**28**,**36**,**37**, and **42** are common outliers. Additional outliers were detected for each individual CSP selector: compound **41** for CSP I; compounds **10**, **27**, and **39** for CSP II; compound **12** for CSP III and compounds **12** and **39** for CSP IV. Their retention behavior was different in comparison to the other compounds included in the models. This fact can be explained by the presence of some non-polar or bulky substituents attached to the chiral center or to the presence of several condensed nuclei in the aryl moiety of analytes. Adopting a conformation within the diastereomeric CSP — analyte complex differing from the lowest energy one is also a possibility. The statistical results obtained for the retention model of the first eluted enantiomer were generally worse in comparison to those of the second eluted enantiomer (especially in case of the CSP II selector — see Table 2). Except for CSP III, analogous results were also obtained by the ANN modeling (see below). As one referee pointed out the observation that the first eluting enantiomer is more poorly modeled might point to some mechanistic information hidden within. In contrast to the retention factors, no outliers were detected in the MLR analysis performed for the separation factor of the analytes. Although only four MLR models have acceptable r (>0.8) [70] and r^2_{LOO} (>0.5) [71] values from a statistical point of view, some tentative conclusions concerning the factors responsible for the retention behavior and enantioselectivity can be drawn.

For the retention behavior (log k_1 and log k_2), two descriptors, the molecular polarizability (α_{mop}) and the energy of the lowest unoccupied molecular orbital (ϵ_{LUMO}) have an important contribution. An analogous observation was reported for the enantioselective separation of a series of aromatic acids and amides on amylose-based CSPs [13]. The ϵ_{LUMO} descriptor indicates charge transfer interactions between the CSP and the analyte — exactly what is expected as an important factor for CSPs containing the 3,5-dinitrobenzoyl group. Alternatively, ϵ_{LUMO} can also be interpreted to reflect the hydrogen

bonding capability [72], which clearly also will be of considerable importance for the retention of analytes **1**–**42** on CSP I–IV. The increasing length of the alkyl chain attached to the chiral center (DIST) which expresses the steric demand of the respective substituent decreases the retention and enantioselectivity behavior of analytes, in accordance with previous studies [6]. However, compared to ϵ_{LUMO} this descriptor plays a less significant role according to its relative contribution (see Table 2). Furthermore, the present results reinforce previous findings [7,53], that the factors responsible for retention need not necessarily be the same as those responsible for separation (see Table 2). Here a hydrophobic term in the MLR models proposed for separation has an important contribution which can be explained by the face to face $\pi-\pi$ stacking of the aromatic moiety of the analytes and of the DNB function of CSP selectors, respectively [6]. The sum of the charges on the carbon atoms of the phenyl moiety (Σq_{Car}) turns out to contribute to the separation too, accounting for the basicity of the analytes in their interaction with the π-acidic function of the selector as stated previously [6]. Electrostatic interactions between the aryl moieties of the analytes with the 3,5-dinitrobenzoyl group of the CSP could be reflected by this term. The same influence on the enantioselectivity can be noticed by the bulkiness of the alkyl substituents attached to the analyte chiral center. The presence of a hydrophobic term in the MLR models proposed for separation can be explained by the face to face $\pi-\pi$ stacking of the aromatic moiety of the analytes and of the DNB function of CSP selectors, respectively [6]. The repulsive steric interactions between the alkylcarbinol moiety of the analytes and the phenyl ring at the selector's stereogenic center bearing the DNB group [6] is reflected by the influence of the DIST descriptor on the MLR model for the separation of the analytes. It is interesting to note that the most suitable descriptors are quite similar to those used by others in, e.g., modeling of modifier and/or solvent effects in chromatography [73–77].

3.2. Modeling by ANN

Since there may be a strong nonlinear component to the relationship between the chemical structures of analytes and their retention behavior, QSRR and QSERR modeling was also carried out using ANNs. The same two sets of descriptors which were found by the above MLR analysis were used as inputs for the description of log k and log α, respectively.

The statistical results obtained thereby for the retention factors (log k_1, log k_2) as well as separation coefficients (log α) are collected in Table 3.

As can be seen from Table 3, on the whole, the statistical quality of fitting generally is significantly

Table 3
Results of the ANN calculations for retention factors (log k_1, log k_2) and enantioselectivities (log α) found for analytes **1**–**42** on chiral selectors CSP I–IV[a]

CSP	Output	Inputs	n	r	RMS	r^2_{LOO}
I	log k_1	α_{mop}, ϵ_{LUMO}, DIST	36	0.828	0.117	0.316
	log k_2	α_{mop}, ϵ_{LUMO}, DIST	36	0.834	0.111	0.322
	log α	ClogP, Σq_{Car}, DIST	41	0.918	0.026	0.564
II	log k_1	α_{mop}, ϵ_{LUMO}, DIST	33	0.814	0.125	0.331
	log k_2	α_{mop}, ϵ_{LUMO}, DIST	33	0.848	0.109	0.583
	log α	ClogP, Σq_{Car}, DIST	41	0.917	0.026	0.437
III	log k_1	α_{mop}, ϵ_{LUMO}, DIST	35	0.881	0.103	0.544
	log k_2	α_{mop}, ϵ_{LUMO}, DIST	35	0.895	0.096	0.429
	log α	ClogP, Σq_{Car}, DIST	41	0.882	0.025	0.369
IV	log k_1	α_{mop}, ϵ_{LUMO}, DIST	33	0.818	0.122	0.417
	log k_2	α_{mop}, ϵ_{LUMO}, DIST	33	0.845	0.111	0.584
	log α	ClogP, Σq_{Car}, DIST	41	0.915	0.026	0.535

[a] Number of compounds used (n), correlation coefficient r, root-mean-square error RMS, and cross-validated correlation coefficients r^2_{LOO} resulting from the leave-one-out procedures, respectively. A network architecture with three input units plus a bias, three hidden-layer neurons plus a bias, and one output layer neuron ((3+1):(3+1):1) for retention factors was used. For log α a (3+1):(4+1):1 architecture was used.

better than that by MLR. These result show the intrinsically nonlinear dependence between the target properties and molecular descriptors employed. Except for CSP III, the ANN modeling performance is also best for log α.

3.3. Comparative molecular field analysis

The results of the CoMFA analysis are summarized in Table 4.

The best internal predictability is found for log α of CSP I. Considerably lower values for r^2_{CV} — especially when using the leave-half-out procedure — are found for CSP II–IV. As an example, a plot of experimental enantioselectivities (log α) found for the chiral stationary phase CSP I vs. those obtained by the CoMFA analysis (three components) is given in Fig. 1. The CoMFA results for the enantioselective separation (log α) of arylalkylcarbinols **1**–**42** on CSP I–IV indicate an almost equal contribution of the steric (≈45%) and the electrostatic (≈55%) fields.

To further assess the relative importance of the steric vs. the electrostatic fields for the enantioselective separation on CSP I, additional CoMFA calculations employing these fields separately were performed. The results obtained thereby are summarized in Table 5. As indicated by the data obtained when using both the steric and the electrostatic field in the CoMFA analysis, each one of the two fields leads to quite comparable statistical results.

In addition to traditional QSRRs and QSERRs CoMFA models offer the advantage of a graphical representation of the steric and electrostatic forces involved in enantioseparation. Representative CoMFA plots are given in Fig. 2 (steric field) and Fig. 3 (electrostatic field). The contour lines in Fig. 2 represent areas of unfavorable (top) and favorable (bottom) steric interactions. As revealed by CoMFA plots of the steric field (see Fig. 2, top), an increase of steric bulk almost exclusively has an adverse effect on the separation factor. There is only a rather small region (bottom of Fig. 2) where such an

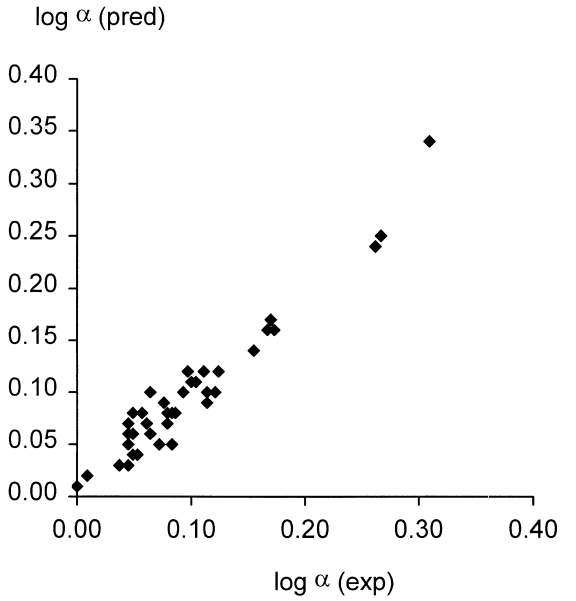


Fig. 1. Plot of experimental vs. calculated (CoMFA model with both fields, three components) log α-values for chiral selector CSP I.

Table 4
CoMFA results for the separation factors (log α) of analytes **1**–**42** on CSPs I–IV[a]

CSP	r^2	SEE	F	r^2_{LOO}	r^2_{LHO}	r^2_{L7O}	Steric	Electrostatic
I	0.944 (3)	0.016	209.4	0.700	0.587	0.605	1.652 (0.455)	1.976 (0.545)
II	0.966 (4)	0.013	245.9	0.434	0.263	0.423	1.849 (0.445)	2.309 (0.555)
III	0.920 (3)	0.016	138.7	0.389	0.173	0.303	1.600 (0.433)	2.100 (0.567)
IV	0.932 (3)	0.017	164.4	0.439	0.249	0.357	1.549 (0.426)	2.086 (0.574)

[a] Conventional squared correlation coefficients r^2, standard error of estimate SEE, F-test F, cross-validated squared correlation coefficients obtained by the leave-one-out (r^2_{LOO}), leave-half-out (r^2_{LHO}), and leave-seven-out (r^2_{L7O}) procedure, respectively (number of components are given in parentheses), and normalized coefficients (fractions in parentheses) of steric and electrostatic CoMFA fields.

Table 5
Statistical results[a] for log α of analytes **1**–**42** on CSP I obtained with steric and electrostatic CoMFA fields separately

Field	*n*	r^2	SEE	*F*
Steric	3	0.929	0.018	160.5
	4	0.950	0.016	170.0
	5	0.981	0.010	361.4
Electrostatic	3	0.931	0.018	166.7
	4	0.960	0.014	215.6
	5	0.973	0.012	252.0

[a] Number of components (*n*), conventional correlation coefficient r^2, standard error of estimate (SEE) and *F*-test.

increase should lead to higher enantioselectivity. This region matches the space occupied by the *tert*-butyl group of 1-naphthyl derivative **41** which indeed shows one of the highest α-values found for the investigated analytes on CSP I.

Similarly, in the CoMFA plot of the electrostatic field (Fig. 3), the contour lines represent regions where an increase (top) or a decrease (bottom) of positive charge should lead to an enhanced selectivity of the chiral resolution. With respect to the electrostatic field an increase of positive charge (see Fig. 3, top) generally is predicted to be favorable for high separation factors. There are only rather small

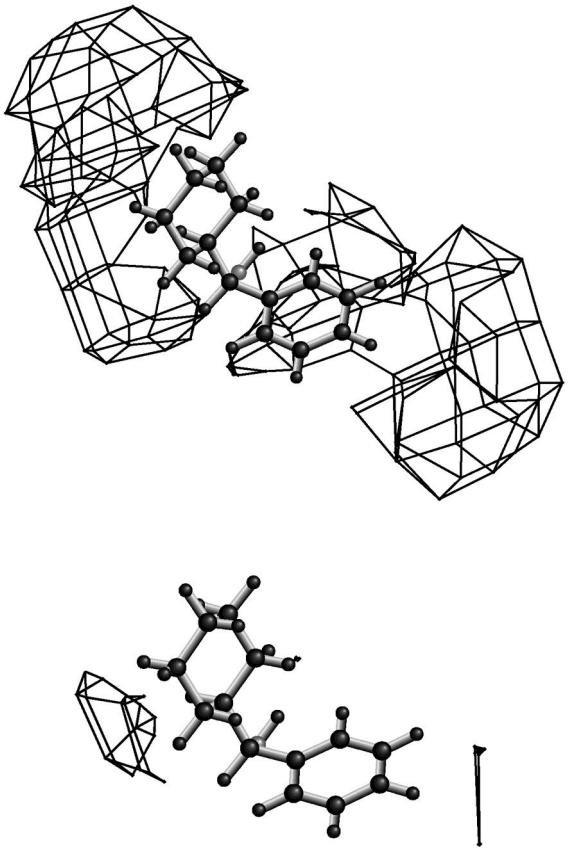


Fig. 2. Plot of steric CoMFA fields (SD* coefficients, contoured by contribution) for unfavorable (top) and favorable (bottom) interactions.

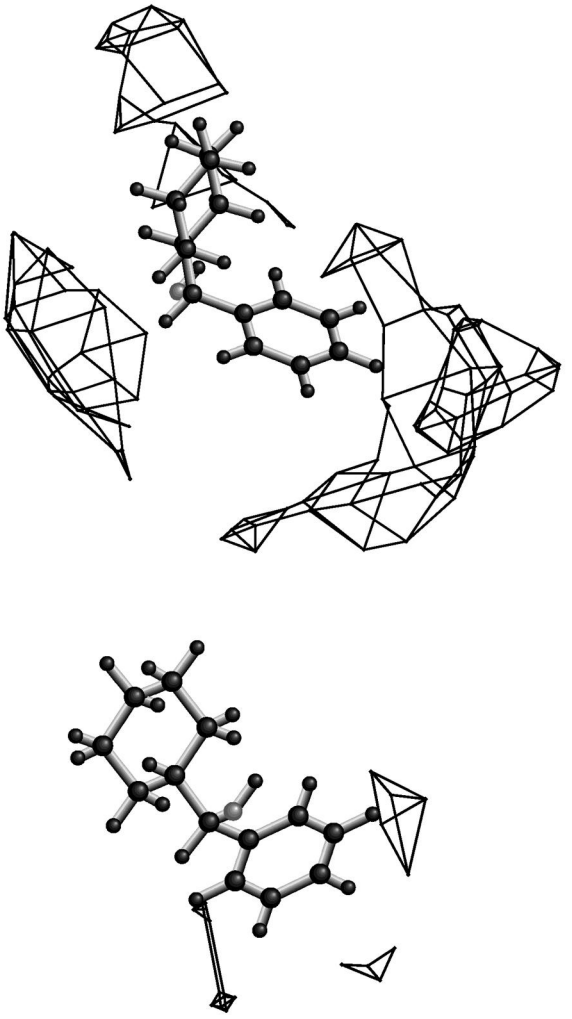


Fig. 3. Plot of electrostatic CoMFA fields (SD* coefficients, contoured by contribution) for regions where an increase (top) and decrease (bottom) of positive charge, respectively, leads to favorable interactions.

areas (Fig. 3, bottom) where a decrease of positive (or increase of negative) charge is predicted to lead to higher log α-values.

Finally, it should be noted that — from a statistical point of view — also with the CoMFA method the best results were obtained for CSP I. Since the other CSPs only differ by the length and/or nature of the linker group, this finding clearly points to the importance of the spacer on the performance of chiral stationary phases [78,79].

4. Conclusions

The retention behavior of a series of chiral arylalkylcarbinols **1**–**42** on four brush-type chiral stationary phases CSP I–IV was modeled by classical QSRRs using multiple linear regression analysis as well as with the aid of artificial neuronal network calculations. Both structural descriptors as well as electronic indices obtained by semiempirical quantum chemical calculations were used in these correlations. The best model for retention factors (log k_1, log k_2) could be obtained with the LUMO energy ϵ_{LUMO} and molecular polarizability α_{mop} as descriptors characterizing the electronic properties and — less important — the distance from the chiral carbon atom to the last non-hydrogen atom of the alkylcarbinol moiety (DIST) as a structural descriptor. The LUMO energy can be interpreted as a measure of charge transfer interactions and/or of hydrogen bonding effects [13,72].

The enantioselectivity (log α) was also treated by classical QSERRs and, in addition, by 3D-QSAR (CoMFA) methods. For this property in QSERR models bulk (or steric) as well as polar or electrostatic properties of analytes are important. This finding is in line with the CoMFA results where steric and electrostatic fields are found to contribute almost equally with a slight dominance of the electrostatic interactions.

References

[1] G. Uray, W. Lindner, Chromatographia 30 (1990) 323.
[2] N.M. Maier, G. Uray, O.P. Kleidernigg, W. Lindner, Chirality 6 (1994) 116.
[3] F. Gasparrini, D. Misiti, C. Villani, F. La Torre, J. Chromatogr. A 539 (1991) 25.
[4] N.M. Maier, G. Uray, J. Chromatogr. A 740 (1996) 11.
[5] G. Uray, K.S. Niederreiter, N.M. Maier, M.M. Spitaler, Chirality 11 (1999) 404.
[6] N.M. Maier, G. Uray, J. Chromatogr. A 732 (1996) 215.
[7] C. Altomare, S. Cellamare, A. Carotti, M.L. Barreca, A. Chimirri, A.M. Monforte, F. Gasparrini, C. Villani, M. Cirilli, F. Mazza, Chirality 8 (1996) 556.
[8] R. Kaliszan, Quantitative Structure–Chromatographic Retention Relationships, Wiley, Chichester, New York, 1987.
[9] R. Kaliszan, J. Chromatogr. B 715 (1998) 229.
[10] T.D. Booth, I.W. Wainer, J. Chromatogr. A 737 (1996) 157.
[11] R. Kaliszan, T.A.G. Noctor, I.W. Wainer, Chromatographia 33 (1992) 546.
[12] R. Kaliszan, A. Kaliszan, T.A.G. Noctor, W.P. Purcell, I.W. Wainer, J. Chromatogr. 609 (1992) 69.
[13] T.D. Booth, K. Azzaoui, I.W. Wainer, Anal. Chem. 69 (1997) 3879.
[14] V. Andrisano, C. Bertucci, V. Cavrini, M. Recanatini, A. Cavalli, L. Varoli, G. Felix, I.W. Wainer, J. Chromatogr. A 876 (2000) 75.
[15] G. Hansson, M. Ahnoff, J. Chromatogr. A 666 (1994) 505.
[16] G. Hansson, M. Ahnoff, Chemometr. Intell. Lab. Syst. 17 (1992) 223.
[17] C.A. Montanari, Q.B. Cass, M.E. Tiritan, A.L.S. De Souza, Anal. Chim. Acta 419 (2000) 93.
[18] S.D. Brown, S.T. Sum, F. Despagne, B.K. Lavine, Anal. Chem. 68 (1996) R21.
[19] K.L. Peterson, Anal. Chem. 64 (1992) 379.
[20] R.H. Zhao, B.F. Yue, J.Y. Ni, H.F. Zhou, Y.K. Zhang, Chemometr. Intell. Lab. Syst. 45 (1999) 163.
[21] G. Robertsson, G. Andersson, P. Kaufmann, Chromatographia 47 (1998) 643.
[22] B. Lucic, N. Trinajstic, J. Chromatogr. A 39 (1999) 610.
[23] A.Z. Yan, R.S. Zhang, M.C. Liu, Z.D. Hu, M.A. Hooper, Z.F. Zhao, Comput. Chem. 22 (1998) 405.
[24] R. Zhang, A. Yan, M. Liu, H. Liu, Z. Hu, Chemometr. Intell. Lab. Syst. 45 (1999) 113.
[25] G. Sacchero, M.C. Bruzzoniti, C. Sarzanini, E. Mentasti, H.J. Metting, P.M.J. Coenegracht, J. Chromatogr. A 799 (1998) 35.
[26] O. Jimenez, I. Benito, M.L. Marina, Anal. Chim. Acta 353 (1997) 367.
[27] M. Pompe, M. Razinger, M. Novic, M. Veber, Anal. Chim. Acta 348 (1997) 215.
[28] J.M. Sutter, T.A. Peterson, P.C. Jurs, Anal. Chim. Acta 342 (1997) 113.
[29] C.G. Georgakopoulos, J.C. Kiburis, J. Chromatogr. B 687 (1996) 151.
[30] Y.L. Xie, J.J. Baeza-Baeza, J.R. Torres-Lapasio, M.C. Garcia-Alvarez-Coque, G. Ramis-Ramos, Chromatographia 41 (1995) 435.
[31] J. Grunenberg, R. Herges, J. Chem. Inf. Comput. Sci. 35 (1995) 905.
[32] K.B. Lipkowitz, J. Chromatogr. A 666 (1994) 493.
[33] K.B. Lipkowitz, D.J. Malik, T. Darden, Tetrahedron Lett. 27 (1986) 1759.

[34] K.B. Lipkowitz, D.A. Demeter, R. Zegarra, R. Larter, T. Darden, J. Am. Chem. Soc. 110 (1988) 3446.
[35] K.B. Lipkowitz, B. Baker, R. Larter, J. Am. Chem. Soc. 111 (1989) 7750.
[36] K.B. Lipkowitz, S. Antell, B. Baker, J. Org. Chem. 54 (1989) 5449.
[37] K.B. Lipkowitz, D.A. Demeter, C.A. Parish, Anal. Chem. 59 (1987) 1731.
[38] R. Däppen, H.R. Karfunkel, F.J.J. Leusen, J. Comput. Chem. 11 (1990) 181.
[39] R. Däppen, H.R. Karfunkel, F.J.J. Leusen, J. Chromatogr. 469 (1989) 101.
[40] S. Topiol, M. Sabio, J. Moroz, W.B. Caldwell, J. Am. Chem. Soc. 110 (1988) 8367.
[41] M. Sabio, S. Topiol, Int. J. Quantum Chem. 36 (1989) 313.
[42] S. Topiol, M. Sabio, J. Am. Chem. Soc. 111 (1989) 4109.
[43] S. Topiol, M. Sabio, J. Chromatogr. 461 (1989) 129.
[44] S. Topiol, Chirality 1 (1989) 69.
[45] U. Norinder, E.G. Sundholm, J. Liq. Chromatogr. 10 (1987) 2825.
[46] C. Hansch, D. Hoekman, H. Gao, Chem. Rev. 96 (1996) 1045.
[47] R.D. Cramer III, D.E. Patterson, J.D. Bunce, J. Am. Chem. Soc. 110 (1988) 5959.
[48] G.R. Marshall, R.D. Cramer III, Trend Pharmacol. Sci. 9 (1988) 285.
[49] K.H. Kim, G. Greco, E. Novellino, in: H. Kubinyi, G. Folkers, Y.C. Martin (Eds.), 3D QSAR in Drug Design: Recent Advances, Vol. 3, Kluwer/Escom, Leiden, The Netherlands, 1998, p. 257.
[50] H. Kubinyi, P.v.R. Schleyer (Eds.), Encyclopedia of Computational Chemistry, Wiley, Chichester, New York, 1998, p. 448.
[51] E.R. Collantes, W. Tong, W.J. Welsh, W.L. Zielinski, Anal. Chem. 68 (1996) 2038.
[52] M.C. Nicklaus, H. Ford, L. Hegedus, W.A. Milne, J.A. Kelley, Quant. Struct.-Act. Relatsh. 14 (1995) 335.
[53] C. Altomare, A. Carotti, S. Cellamare, F. Fanelli, F. Gasparrini, C. Villani, P.A. Carrupt, B. Testa, Chirality 5 (1993) 527.
[54] SYBYL Molecular Modeling Software, Version 6.5, Tripos Associates, St Louis, MO, 1999.
[55] M. Clark, R.D. Cramer III, N. Van Opdenbosch, J. Comput. Chem. 10 (1989) 982.
[56] M. Saunders, K.N. Houk, Y.D. Wu, W.C. Still, M. Lipton, G. Chang, W.C. Guida, J. Am. Chem. Soc. 112 (1990) 1419.
[57] M.J.S. Dewar, E.G. Zoebisch, E.F. Healy, J.J.P. Stewart, J. Am. Chem. Soc. 107 (1985) 3902.
[58] T. Clark, VAMP 4.4, Erlangen Vectorized Molecular Orbital Package, Computer-Chemie-Centrum, University Erlangen-Nürnberg, Germany, 1992.
[59] S. Wold, W.J. Dunn, J. Chem. Inf. Comput. Sci. 23 (1983) 6.
[60] R.D. Cramer, J.D. Bunce, D.E. Patterson, I.E. Frank, Quant. Struct.-Act. Relatsh. 7 (1988) 18.
[61] D. Ciubotariu, E. Deretey, T.I. Oprea, T. Sulea, Z. Simon, L. Kurunczi, A. Chiriac, Quant. Struct.-Act. Relatsh. 12 (1993) 367.
[62] A. Bondi, J. Phys. Chem. 68 (1964) 441.
[63] ClogP, Version 1.0.0, BioByte, Claremont, CA, USA.
[64] P.P. Mager, Med. Res. Rev. 14 (1994) 533.
[65] J. Zupan, J. Gasteiger, Neural Networks for Chemists: An Introduction, Verlag Chemie, Weinheim, 1993.
[66] D.L. Livingstone, D.W. Salt, Biorg. Med. Chem. Lett. 3 (1993) 645.
[67] T.A. Andrea, H. Kalayeh, J. Med. Chem. 34 (1991) 2824.
[68] S. Wold, W.J. Dunn, A. Ruhe, H. Wold, Siam. J. Sci. Stat. Comput. 5 (1984) 735.
[69] G. Cruciani, M. Baroni, S. Clementi, G. Constantino, D. Riganelli, B. Skagerberg, J. Chemometrics 6 (1992) 335.
[70] G.R. Famini, L.Y. Wilson, Theor. Comput. Chem. 1 (1994) 213.
[71] T.I. Oprea, C.M.W. Ho, G.R. Marshall, Computer-Aided Molecular Design. Applications in Agrochemicals, Materials, and Pharmaceuticals, ACS Symposium Series, Vol. 589, American Chemical Society, Washington, DC, 1995.
[72] K. Azzaoui, L. Morin-Allory, Chromatographia 42 (1996) 389.
[73] C.M. Du, K. Valko, C. Bevan, D. Reynolds, M.H. Abraham, J. Chromatogr. Sci. 38 (2000) 503.
[74] F.Z. Oumada, M. Roses, E. Bosch, M.H. Abraham, Anal. Chim. Acta 382 (1999) 301.
[75] C.M. Du, K. Valko, C. Bevan, D. Reynolds, M.H. Abraham, Anal. Chem. 70 (1998) 4228.
[76] K. Valko, M. Plass, C. Bevan, D. Reynolds, M.H. Abraham, J. Chromatogr. A 797 (1998) 41.
[77] L.C. Tan, P.W. Carr, M.H. Abraham, J. Chromatogr. A 752 (1996) 1.
[78] W.H. Pirkle, M.H. Hyun, B. Bank, J. Chromatogr. 316 (1984) 585.
[79] W.H. Pirkle, C.J. Welch, J. Chromatogr. 589 (1992) 45.