ELSEVIER

Journal of Chromatography A, 914 (2001) 299–314

JOURNAL OF CHROMATOGRAPHY A

www.elsevier.com/locate/chroma

Approaches to quantitative structure–enantioselectivity relationship modeling of chiral separations using capillary electrophoresis

Jeffrey P. Wolbach[a], David K. Lloyd[a,*], Irving W. Wainer[b]

[a]*Analytical Research and Development, DuPont Pharmaceuticals, Experimental Station, E353, P.O. Box 80353, Wilmington, DE 19880-0353, USA*

[b]*Department of Pharmacology, Georgetown University School of Medicine, Washington, DC 20007, USA*

Abstract

Quantitative structure–enantioselectivity relationships (QSERs) have been developed to describe the resolution of a series of chiral arylpropionic acids using capillary electrophoresis. Native β-cyclodextrin and two derivatized forms are used as the chiral resolving agents. The QSER models are developed using the results of molecular mechanics calculations as input to multivariate linear regression and also to neural networks. Single models are developed to predict the optimum cyclodextrin to resolve a given analyte, the migration order, and the magnitude of the separation. Models are also developed to predict only the optimum cyclodextrin.

Keywords: Structure–enantioselectivity relationships; Enantiomer separations; Mathematical modeling; Neural networks; Regression analysis; Molecular descriptors; Cyclodextrins; Arylpropionic acids; Propionic acids

1. Introduction

Enantiomers cannot be separated on the basis of differences in the usual physical properties, such as freezing or boiling point, or solubility. Instead many separation techniques rely upon the formation of differing transient or non-covalent diastereoisomeric complexes of the enantiomer with another chiral entity, for example in liquid chromatography with a chiral stationary phase [1], or capillary electrophoresis (CE) with a chiral additive to the background electrolyte [2,3]. Separation of the diastereoisomeric complexes can involve low selectivity, thus these types of chiral separations are an attractive application of CE, as the high efficiency of CE helps to generate adequate resolution [4].

Chiral separations using CE typically involve adding a chiral selector molecule to the background electrolyte [3]. The selector molecule preferentially complexes with one of the enantiomers, giving the two enantiomers differing effective electrophoretic mobilities. At present, choosing the optimum selector molecule for a particular separation is an inexact science. Proposed approaches for choice of selector in CE generally involve screening of a variety of potential selectors [5], followed by various approaches to optimization [6–8]. In pharmaceutical development, it is common to have to analyze many compounds within a particular therapeutic area, which typically possess some common structural motifs and significant structural diversity elsewhere.

*Corresponding author. Tel.: +1-302-695-7523; fax: +1-302-695-3705.

E-mail address: david.k.lloyd@dupontpharma.com (D.K. Lloyd).

0021-9673/01/$ – see front matter
PII: S0021-9673(01)00580-5

In such an environment, rapid prediction of critical CE method parameters, e.g. choice of chiral selector, would clearly be of value. In this work, we evaluate different strategies for prediction of the optimum selector in CE, using a model set of compounds.

One strategy for the development of predictive models of the enantioselectivity is the construction of quantitative structure–enantioselectivity relationships (QSERs). These relationships attempt to isolate and identify the structural characteristics of a series of racemic analytes deemed responsible for the observed enantioselectivity. These characteristics are used as the independent variables in multivariate regression analysis and correlated against the experimental data. Development of statistically significant equations allows for the possibility of extracting physically meaningful information relating to the selection process. This approach has been successfully applied to a number of chiral stationary phases for liquid chromatography [9–13].

The important structural characteristics, or descriptors, may also be used as inputs to neural networks. Neural networks consist of multiple layers of interconnected neurons, with the most typical networks having three layers. The first layer has one neuron for each input variable, and the second or hidden layer contains a number of neurons (the specific number depends on the network design). The final, output layer has one neuron for each parameter (e.g., retention, selectivity) to be predicted or correlated for a given analyte. The input to each neuron in a given layer is the output from every neuron in the previous layer. For example, each neuron in the hidden layer is fed by, or connected to, the output of every neuron in the first layer. The inputs are summed, and an activation function is applied to generate the output of each neuron. If the activation function is non-linear (e.g., e^{-x^2}, where x is the sum of the inputs to the neuron), a high degree of non-linearity may be introduced into the network. Typically, the activation function has the same mathematical form in every neuron of a given layer. Neural networks have been successfully applied to various areas of chromatography and electrophoresis [14–16], and have demonstrated considerable potential for retention prediction in achiral systems [17,18].

In this article, the results of experiments performed on a series of commercially available arylpropionic acids are reported. Resolution of the analytes was attempted using CE with three cyclodextrins as chiral selectors: β-cyclodextrin (β-CD), hydroxypropyl-β-cyclodextrin (HP-β-CD, avg. degree of substitution DS = 0.8), and tri-*O*-methyl-β-cyclodextrin (tri-OMe-β-CD). Analyte descriptors calculated by means of molecular modeling software were applied to multivariate regression analysis, in addition to multi-layer neural networks, to model the enantioselective behavior.

2. Experimental

2.1. Chemicals and sample preparation

Analyte structures are given in Fig. 1. All the compounds possess a cyclic, aromatic backbone, with pendant groups or side chains capable of hydrogen bonding. Analytes 4, 12, 13, 15, 18, and 22 were obtained from Sigma (St. Louis, MO, USA); all others from Aldrich (Milwaukee, WI, USA). The three types of cyclodextrin used in this study were obtained from Sigma, as were disodium hydrogenphosphate and citric acid. HPLC-grade hexanes and isopropanol were obtained from EM Scientific (Gibbstown, NJ, USA) and trifluoroacetic acid was purchased from Sigma.

The CE background electrolyte (BGE) comprised a pH 5 buffer prepared by mixing a 20 m*M* aqueous solution of Na_2HPO_4 and a 10 m*M* aqueous solution of citric acid. An appropriate amount of each cyclodextrin was dissolved in the BGE to yield concentrations of 12 m*M* (β-CD) and 50 m*M* (HP-β-CD and tri-OMe-β-CD).

All of the analytes, with the exception of suprofen (analyte 4), were dissolved in distilled water (Millipore, Milford, MA, USA) at concentrations between 0.2 and 1.0 mg/ml. Suprofen was not sufficiently soluble in water, and instead was dissolved at a concentration of ~0.4 mg/ml in the BGE with a small amount of the cyclodextrin of interest added (3 m*M* for β-CD, 5 m*M* for HP-β-CD and tri-OMe-β-CD). A neutral marker solution was prepared by dissolving 4% (v/v) acetone in the BGE with no cyclodextrin added. All solutions were passed through a 0.45 μm polyvinylidene fluoride mem-

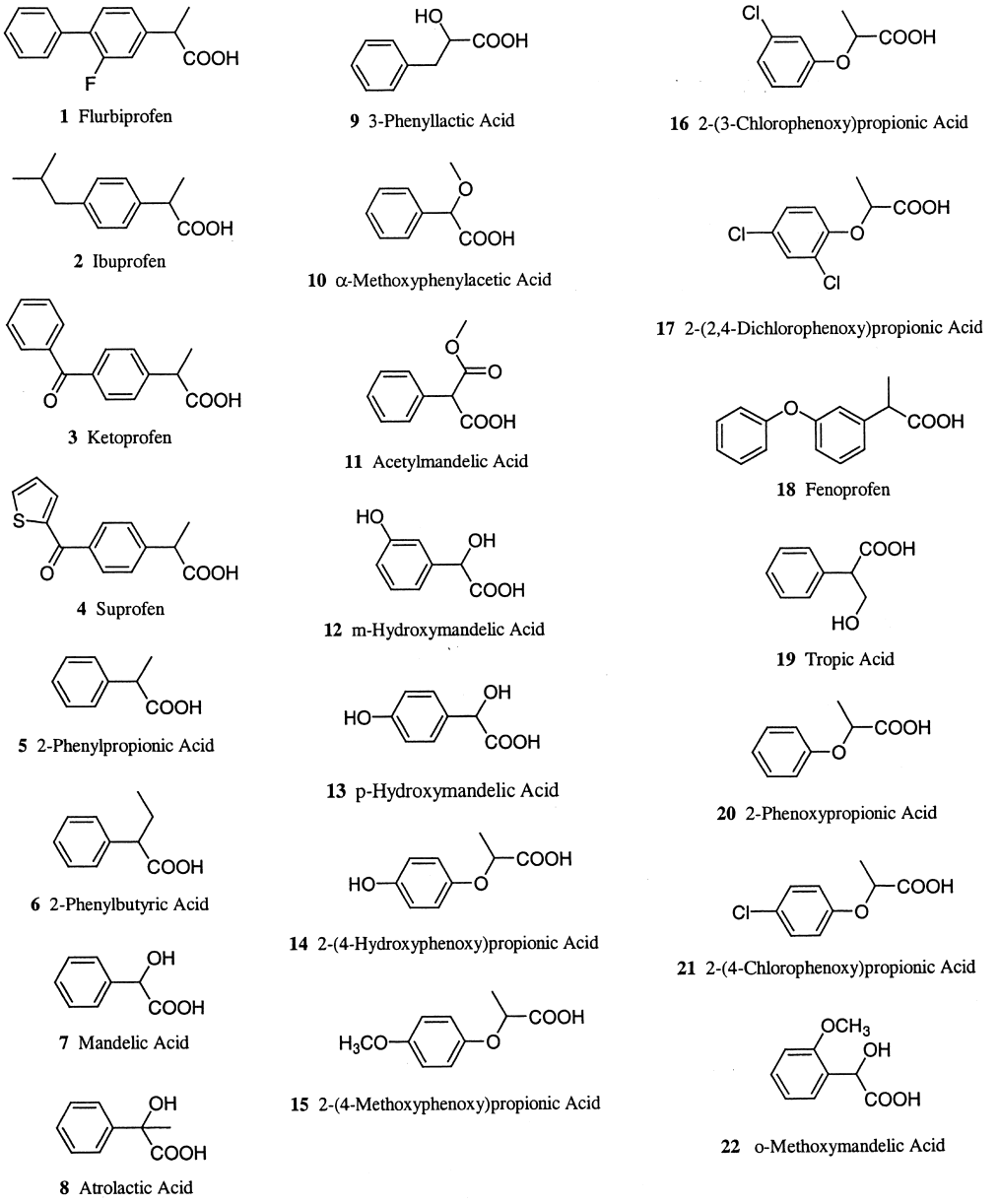

Fig. 1. Analytes considered in this work.

brane filter (Whatman, Clifton, NJ, USA) prior to use.

2.2. Capillary electrophoresis

All separations were performed on a Hewlett-Packard 3D/CE capillary electrophoresis instrument. Uncoated, fused-silica capillaries were cut from a spool of stock obtained from Polymicro Technologies (Phoenix, AZ, USA). The stock had an internal diameter of 50 μm and an external diameter of 365 μm. The capillaries were cut to a length of 64.5 cm, with an effective length of 56 cm from the injection point to the detector. The temperature of

the cartridge containing the capillary was thermostated at 35°C. All separations were performed using an applied potential of 20 kV, resulting in a current of approximately 12 μA.

Before each trial, the capillary was flushed for 3 min with 0.1 *M* NaOH, 8 min with the BGE, and then a 20 kV potential was applied for 6 min. The neutral marker and analyte solutions were injected back-to-back, each for 3 s using 30 mbar of pressure. The separations were monitored on-column at 200 and 254 nm. The migration times of the analytes were taken from the 200 nm measurement, and the migration time of the neutral marker from the 254 nm measurement. Data were collected and analyzed using the Hewlett-Packard ChemStation software (Rev. A.06.03) running on an IBM-compatible personal computer. All separations were performed in duplicate.

2.3. Chromatography

The migration order for each resolved analyte was determined by spiking the samples with one of the pure enantiomers. The spiking created enantiomeric ratios of ~3:1 in the samples. Enantiomerically pure forms of analytes 1–3, 5, 6, 11, and 14 were purchased, while for analytes 4 and 18 the migration order was taken from the literature [19,20]. The other analytes were chromatographically resolved following a published method of Booth and Wainer [13].

Separations were performed on a Hewlett-Packard 1090 chromatograph, using 250 mm×4.6 mm I.D. columns packed with ChiralPak AD (Chiral Technologies, Exton, PA, USA). Column elution order was determined using an Autopol III polarimeter (Randolph Research, Flanders, NJ, USA) operating at 365 nm. Mobile phase compositions, flow-rates, and sample preparations are listed in Table 1. All mobile phases also included 0.5% (v/v) of trifluoroacetic acid, and the mobile phase flow-rate was varied to hold the column backpressure under 430 p.s.i. (1 p.s.i. = 6894.76 Pa). Multiple 200 μl sample injections were performed to allow manual collection of ~5 mg of the first eluting enantiomer. All separations were performed at ambient temperature (~25°C).

2.4. Determination of pK_a

If the pK_a of an analyte could not be found in the literature, it was determined experimentally by potentiometric titration using a GLpKa Titrator (Sirius Analytical Instruments, Forest Row, East Sussex, UK). If the analyte was not sufficiently soluble in water, a mixed solvent of water and methanol was used. When using the co-solvent procedure a series of titrations were performed at three different weight percentages of methanol. Co-solvent pK_a values are extracted from the titration data using a least squares refinement procedure and extrapolated to 100% aqueous conditions using the Yasuda–Shedlovsky plot.

2.5. Computational chemistry

Neutral and anionic structures for each analyte were created in Cerius[2] (Molecular Simulations, San Diego, CA, USA). Complete conformational search-

Table 1
Liquid chromatographic conditions for manual fraction collection[a]

Compound	Sample conc. (mg/ml)	Sample solvent hexane–isopropanol	Mobile phase hexane–isopropanol	Flow-rate (ml/min)	Elution order
8	15	80:20	95:5	1.00	+ before −
12	10	80:20	85:15	0.75	+ before −
13	5	85:15	85:15	0.75	− before +
15	10	95:5	95:5	1.00	− before +
16	10	95:5	95:5	1.00	− before +
17	10	95:5	95:5	1.00	+ before −
19	10	80:20	95:5	1.00	+ before −
20	6	95:5	95:5	1.00	− before +
21	10	95:5	95:5	1.00	− before +

[a] In each case, the mobile phase also contained 0.5% (v/v) of trifluoroacetic acid.

es were performed using the Conformers module of Cerius[2], and the molecular geometries were optimized using the consistent-valence force field [21–28]. Molecular descriptors were calculated for the minimum energy conformation of the neutral and anionic form of each analyte using the QSAR+ and Descriptor+ modules of Cerius[2]. All calculations were performed on a Silicon Graphics O2 workstation (Silicon Graphics, Sunnyvale, CA, USA).

2.6. Statistical analysis and neural networks

The first statistical technique used was multivariate linear regression directed by a genetic function approximation (GFA) as implemented in the GFA module of Cerius[2] [29]. In the GFA, a series of initial equations are formed using random combinations of the molecular descriptors. The randomly generated equations are subjected to multivariate linear regression, and the best-fitting equations are allowed to ''breed'' or exchange terms to form new equations. As many generations are bred, the approximation rapidly converges to sets of descriptors that best describe the experimental data. Unlike other techniques, the GFA generates multiple equations with similar qualities of fit to the data. Although a single equation may then be chosen based on a single criterion, e.g. a correlation coefficient or the lack of fit parameter, further insight is often gained by examination of the population of well-fitting equations [29].

Two types of neural networks were also used: a general regression neural network (GRNN) [30] to build a continuous-valued model of the data, and a probabilistic neural network (PNN) [31,32] to model the data as a categorization problem. These networks have three-layers, with the hidden layer containing a neuron for each record in the training set. The proper descriptors to use as inputs to the networks were determined using a genetic algorithm. The neural network models were developed using the program NeuroShell 2, V. 4.0 (Ward Systems Group, Frederick, MD, USA) running on a Pentium III personal computer.

3. Results and discussion

The experimentally determined effective mobilities are presented in Table 2. The data in this and subsequent tables are separated into the training set (1–17) and the test set (18–22) used in the multivariate linear regression QSER analysis. The reported mobilities are the average of two analyses. If an analyte is resolved, the effective mobility of each enantiomer is listed.

Selectivity may be defined as:

$$\xi_{+/-} = \frac{\mu_{\text{eff}}(+\text{ enantiomer})}{\mu_{\text{eff}}(-\text{ enantiomer})} \tag{1}$$

The selectivities as defined in Eq. (1) are presented in Table 3. With this definition, the magnitude of the separation and the migration order are contained in a single number. If $\xi_{+/-}$ is greater than 1, the (−)-enantiomer migrates first. If it is less than 1, the (+)-enantiomer migrates first. Representative chromatograms of the separations are presented in Fig. 2.

Enantiomeric resolution may occur due to differences in the binding strength of either the neutral species or the anionic species. Since both mechanisms are possible, the descriptor values used are the mole-fraction weighted averages of the descriptor values for the neutral and anionic forms of each analyte. The fraction of each analyte that was ionized (X) at pH 5 was determined by re-arrangement of the Henderson–Hasselbalch equation:

$$K_{\text{eq}} = 10^{-(\text{p}K_{\text{a}}-\text{pH})} = \frac{X}{1-X}$$
$$X = \frac{10^{-(\text{p}K_{\text{a}}-\text{pH})}}{1+10^{-(\text{p}K_{\text{a}}-\text{pH})}} = \frac{1}{1+10^{(\text{p}K_{\text{a}}-\text{pH})}} \tag{2}$$

The pK_{a} values and fraction ionized are presented in Table 4.

There are four categories of analyte properties that could potentially affect inclusion behavior. The size, shape, and hydrophobicity of the analyte are all important for interaction with the internal cavity, while successful interaction with the groups around the rim of the cyclodextrin requires the analyte to have the capacity to form hydrogen bonds. Twenty-four descriptors were calculated for each neutral and anionic species reflecting these and other properties. The descriptors are listed in Table 5. The descriptors were not enantiospecific, as none would yield a different value for the (+)- and (−)-enantiomers of the same analyte.

Table 2
Effective mobilities of the analytes ($\mu_{eff} \times 10^4$) ($cm^2\ V^{-1}\ s^{-1}$)[a]

Compound	β-CD	HP--β-CD	Tri-OMe--β-CD	Compound	β-CD	HP--β-CD	Tri-OMe--β-CD
1	−1.06	−0.68	−0.58 −0.64	13	−2.57	−2.03 −2.06	−2.51
2	−0.87	−0.59	−0.48 −0.56	14	−2.12 −2.21	−1.60	−2.27
3	−1.16	−0.71	−0.87 −0.94	15	−1.93 −1.96	−1.42	−2.14
4	−1.11	−0.76	−0.66 −0.71	16	−1.89 −1.95	−1.23 −1.26	−1.91 −1.93
5	−1.41	−0.82 −0.91	−2.06 −2.14	17	−1.74	−1.09 −1.12	−1.50 −1.55
6	−1.16	−0.51 −0.62	−1.61 −1.70				
7	−3.07	−2.40	−2.98	18	−0.99	−0.60	−0.36
8	−2.49 −2.53	−1.92 −1.96	−2.70	19	−2.20	−1.40	−0.44 −2.20
9	−2.27	−1.57	−2.55 −2.56	20	−2.41	−1.53 −1.88	−2.26 −2.48
10	−2.81	−2.25	−2.66		−2.48	−1.92	
11	−2.60	−2.03 −2.11	−2.48	21	−1.86 −1.90	−1.27 −1.29	−1.60 −1.65
12	−2.82	−2.29 −2.31	−2.57	22	−2.93	−2.46	−2.75

[a] Experimental conditions: 35°C, 20 kV applied potential, ~15 m*M* pH 5 buffer with 12 m*M* β-CD, 50 m*M* HP-β-CD, or 50 m*M* Tri-OMe-β-CD. For analytes that are resolved, mobilities for both enantiomers are listed.

Table 3
Selectivities as a function of cyclodextrin: $\xi_{+/-}$ defined by Eq. (1)

Compound	β-CD	HP--β-CD	Tri-OMe--β-CD	Compound	β-CD	HP--β-CD	Tri-OMe--β-CD
1	1.000	1.000	1.098	13	1.000	0.984	1.000
2	1.000	1.000	1.175	14	1.040	1.000	1.000
3	1.000	1.000	1.074	15	1.016	1.000	1.000
4	1.000	1.000	1.084	16	1.032	1.028	1.010
5	1.000	0.917	0.963	17	1.000	0.976	0.968
6	1.000	0.830	0.949				
7	1.000	1.000	1.000				
8	1.017	1.018	1.000	18	1.000	1.000	1.233
9	1.000	1.000	0.994	19	1.000	1.098	1.027
10	1.000	1.000	1.000	20	1.027	1.019	1.000
11	1.000	0.975	1.000	21	1.018	1.014	1.031
12	1.000	0.992	1.000	22	1.000	1.000	1.000

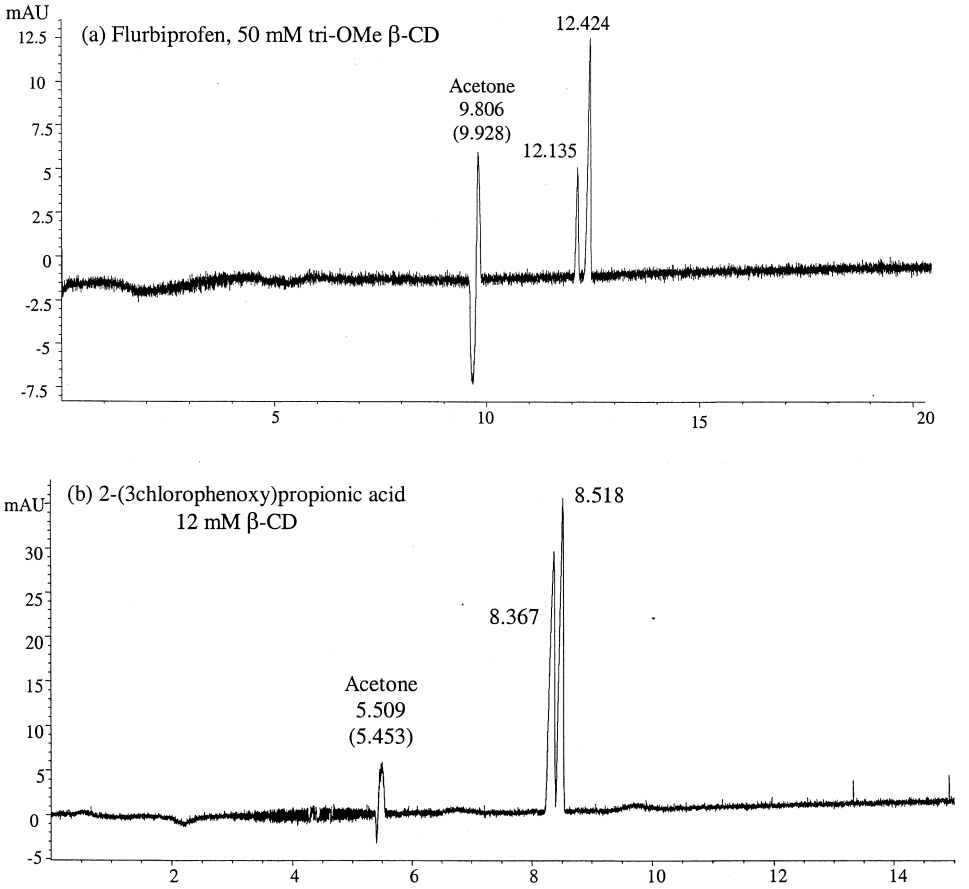

Fig. 2. Representative chromatograms of the separations. Chromatograms measured at 200 nm. Migration times in min. The times in parentheses are the migration times of acetone taken from measurements at 254 nm.

Table 4
pK_a and fraction of each analyte ionized at pH 5

Compound	pK_a	Fraction ionized	Compound	pK_a	Fraction ionized
1	4.340	0.820	13	3.271	0.982
2	4.520	0.751	14	3.061	0.989
3	4.360	0.814	15	3.150	0.986
4	3.910	0.925	16	3.150	0.986
5	4.640	0.696	17	2.927	0.992
6	4.230	0.862			
7	3.410	0.975			
8	3.465	0.972	18	4.660	0.686
9	3.586	0.963	19	4.120	0.884
10	2.968	0.991	20	3.200	0.984
11	2.550	0.996	21	3.110	0.987
12	3.243	0.983	22	3.590	0.963

Table 5
List of calculated descriptors

Size/shape descriptors	Jurs descriptors [33,34]
V_m	Total apolar surface area
PMI	Total polar surface area
Rad. of gyration	Fractional neg. surface area
Surface area	Partial neg. surface area
Shadow-ν	Relative apolar surface area
Density	Relative polar surface area
No. of rotational bonds	
Solvent acc. surface area	
Electronic descriptors	**Thermodynamic descriptors**
LUMO	*A* Log *P* (atom-based)
Superdelocalizability	Log *P*
Polarizability (atom-based)	Molar refractivity
HOMO	No. of H-bond acceptors
Dipole moment	No. of H-bond donors

3.1. Model 1

The initial modeling technique used was multivariate linear regression. In contrast to more mathematically intensive techniques (such as neural networks), a model built using multivariate linear regression that accurately describes the experimental data more easily lends itself to physical interpretation. The desire for physical interpretability is also reflected in the selection of calculated descriptors, as many other easily calculated alternatives may be chosen. The quantity of greatest interest was the identity of the optimum cyclodextrin to resolve a given analyte. The optimum cyclodextrin is defined as the cyclodextrin that maximizes the absolute value of the quantity $(\xi_{+/-} - 1)$. It was also desired to correctly predict or correlate the migration order of the resolved analytes.

The first model developed answered the relevant question indirectly by using multivariate linear regression to describe the selectivities obtained using each cyclodextrin. The model consisted of three separate equations (one for each cyclodextrin), and the optimum cyclodextrin was determined by comparing the magnitude of the response of each equation for each analyte. The quantity $(\xi_{+/-} - 1)$ was modeled rather than $\xi_{+/-}$, so the equations would not require a leading constant. Analytes 1–17 were used as the training set to develop each equation, and the analytes 18–22 were reserved as a test set. The test set was chosen to be representative of the whole data set, and its composition was not an adjustable parameter in this study.

The modeled data is discontinuous, as not every analyte is separated by every cyclodextrin. To account for this discontinuity, spline functions in the descriptors were allowed in addition to normal linear terms. The derivation of the equations was directed by the GFA. In the GFA, the knot of a spline in a given descriptor may only occur at a value of the descriptor present in the training set. For example, the knot of a spline in molar volume may only occur at the calculated molar volume of one of the compounds in the training set. Each location of the spline knot is treated as a separate descriptor for breeding purposes (e.g., $\langle V_m - 180 \rangle$ and $\langle V_m - 210 \rangle$ are separate, unique descriptors). The initial population was 300 equations and breeding continued for 5000 generations. The equations generated by the GFA were restricted to a fixed length of three terms, with no leading constant.

The goal of this work was to build a model that properly identified the optimum cyclodextrin to resolve a given analyte. In this first model, the given analyte is presented to an equation for each cyclodextrin, and the equation with the response largest in magnitude represents the optimum cyclodextrin. Initially, the model was built by selecting from the output of the GFA the single equation that best fit (largest R^2) the training set data for each cyclodextrin. Surprisingly, this model did not perform very well in determining the optimum cyclodextrin for resolving a given analyte. This may have occurred because it was not possible to generate highly accurate correlations of the data for the individual cyclodextrins ($R^2 > 0.9$). For a correlation to be highly accurate, all of the non-zero terms (resolved analytes) would have be nearly correct both in sign and magnitude. Comparison of the responses of such equations to determine the optimum cyclodextrin would have a high potential for accurately reflecting the experimental reality. In this work, the quality of fit achieved by the multivariate linear regression indicated errors occurring in some of the non-zero points. The errors could be of differing severity: omitting the separation of a compound by a non-optimum cyclodextrin would likely be benign, significantly overestimating the magnitude of the separation by a non-optimum cyclodextrin less so.

Table 6
Equations comprising model 1 (multivariate linear regression)

Cyclodextrin	Equation
β-CD	$2.17\times10^{-4}\langle 349.3-\text{TASA}\rangle+7.0\times10^{-3}\langle -14.811-\text{HOMO}\rangle-0.37\langle 0.512-\text{RASA}\rangle$ R^2: 0.696 $R^2_{\text{C.V.}}$: 0.563
HP-β-CD	$-1.37\langle 1.527-\nu\rangle-3.35\times10^{-4}\langle 195.8-\text{TASA}\rangle+0.75\langle -10.933-\text{LUMO}\rangle$ R^2:0.870 $R^2_{\text{C.V.}}$: 0.685
Tri-OMe-β-CD	$-0.12\langle\nu-1.998\rangle+2.1\times10^{-3}\langle\text{TASA}-274.7\rangle-0.04\langle\text{A log P}-5.628\rangle$ R^2: 0.748 $R^2_{\text{C.V.}}$: 0.548

These types of errors were present in all of the equations generated by the GFA.

However, the GFA generates a large number of equations for a given cyclodextrin [29], and for each cyclodextrin there were many equations of nearly equal quality of fit. It was thus possible to build many unique models combining high-scoring equations for each cyclodextrin, and compare their performance on identifying the optimum cyclodextrins. This was done, and the combination that best correlated the training set data for optimum cyclodextrin was retained in the model. The data in the test set were not used to judge the prospective models. The equations composing the model are presented in Table 6, with the R^2 value and the cross-validated R^2 value. The notation $\langle\ \rangle$ denotes a spline function.

The retained equations for the three cyclodextrins are comprised of seven unique descriptors. The values of these descriptors for each analyte are listed in Table 7. Shadow-ν is the ratio of the longest primary axis of the analyte to the smallest, and is a measure of the shape of the analyte [33]. The

Table 7
Descriptor values for the analytes in this study

Compound	Shadow-ν	*A* Log *P*	HOMO (eV)	LUMO (eV)	SASA (cm^2/mol)	TASA (cm^2/mol)	RASA
1	1.976	6.757	−14.618	−7.645	461.4	327.2	0.709
2	1.998	6.296	−15.589	−6.561	455.4	349.3	0.767
3	1.865	6.132	−9.476	−4.564	477.9	355.9	0.745
4	1.818	5.968	−10.957	−5.824	460.3	313.7	0.682
5	1.449	4.525	−16.708	−6.010	345.2	250.1	0.725
6	1.427	5.467	−17.656	−8.022	367.9	274.7	0.747
7	1.764	4.448	−19.021	−10.773	324.2	195.8	0.604
8	1.504	4.640	−18.460	−10.747	348.4	234.9	0.674
9	2.092	4.660	−17.278	−9.909	354.4	226.9	0.640
10	1.527	4.779	−18.381	−9.950	365.3	267.2	0.732
11	1.500	4.774	−18.041	−10.155	393.0	265.4	0.675
12	1.617	4.190	−17.283	−10.122	343.2	148.2	0.432
13	1.861	4.186	−17.226	−10.614	337.1	145.5	0.432
14	1.955	4.834	−11.061	−9.579	375.0	209.9	0.560
15	2.210	4.857	−9.476	−8.439	415.2	287.3	0.692
16	1.919	5.628	−15.309	−10.970	383.1	196.2	0.512
17	2.069	6.164	−14.811	−10.933	405.8	165.7	0.408
18	1.975	5.922	−8.921	−2.797	473.0	358.9	0.759
19	1.436	4.138	−17.449	−8.594	347.9	232.3	0.668
20	1.819	5.104	−15.487	−10.873	365.6	263.0	0.719
21	2.071	5.632	−15.397	−10.940	391.1	211.5	0.541
22	1.679	4.155	−17.737	−10.292	359.2	219.3	0.611

hydrophobicity of an analyte is reflected in *A* log *P*, the sum of the atomic contributions to the octanol–water partition coefficient. The highest-occupied and lowest-unoccupied molecular orbital (HOMO, LUMO) are related to the hydrogen bond donating and accepting capacity of the analyte, respectively. SASA is the solvent-accessible surface area, a measure of the size of the analyte, while TASA [34] is the total apolar surface area, the surface area of an analyte due to atoms with a formal charge smaller in magnitude than 0.2. Larger values of TASA indicate either larger or more hydrophobic analytes. RASA [34], the relative apolar surface area, is the ratio of TASA to SASA.

The optimum cyclodextrins determined using this model are displayed in Table 8. For a given analyte, if none of the equations produce a response greater in magnitude than 0.010, the analyte is assumed not be separable and the optimum cyclodextrin is ''None''. Although the individual equations do not produce strong fits to the data in the training set ($R^2 < 0.9$), the composite model performs well in correlating the optimum cyclodextrin for resolving a particular analyte. The only error in the training set is the optimum separation of atrolactic acid (compound 8) having the incorrect migration order, albeit with the proper cyclodextrin (HP-β-CD).

There are two errors in the model predictions for the test set of compounds. Compound 19 (tropic acid) is predicted to optimally separate with the proper cyclodextrin, but with the incorrect migration order. The errors in migration order for compounds 8 and 19 may be due to the similarity of these compounds to compounds 5 and 6 (2-phenylpropionic and 2-phenylbutyric acids). Compounds 5 and 6 separate strongly using HP-β-CD and migrate (−) before (+), while compounds 8 and 19 migrate (+) before (−).

The second erroneous result in the test set is compound 21 (2-(4-chlorophenoxy) propionic acid)

Table 8
Selection of optimum cyclodextrin by models using multivariate linear regression[a]

Compound Training set	Experimental Optimum			Model 1, Optimum		Model 2, Optimum	
	CD	$(\xi_{+/-}-1)$	$(\xi-1)$	CD	$(\xi_{+/-}-1)$	CD	$(\xi-1)$
1	Tri-OMe	0.098	0.098	Tri-OMe	0.063	Tri-OMe	0.099
2	Tri-OMe	0.175	0.175	Tri-OMe	0.128	Tri-OMe	0.092
3	Tri-OMe	0.074	0.074	Tri-OMe	0.148	Tri-OMe	0.064
4	Tri-OMe	0.084	0.084	Tri-OMe	0.067	Tri-OMe	0.087
5	HP	−0.083	0.083	HP	−0.106	HP	0.078
6	HP	−0.170	0.170	HP	−0.137	HP	0.100
7	None	–	–	None	–	None	–
8	HP/β	0.018	0.018	*HP*	*−0.032*	HP	0.023
9	Tri-OMe	−0.006	0.006	Tri-OMe	−0.011	Tri-OMe	0.014
10	None	–	–	None	–	None	–
11	HP	−0.025	0.025	HP	−0.037	HP	0.027
12	HP	−0.008	0.008	HP	−0.016	HP	0.014
13	HP	−0.016	0.016	HP	−0.017	HP	0.016
14	β	0.040	0.040	β	0.030	β	0.030
15	β	0.016	0.016	β	0.013	β	0.013
16	β	0.032	0.032	β	0.030	β	0.030
17	Tri-OMe	−0.032	0.032	Tri-OMe	−0.030	Tri-OMe	0.030
Test set							
18	Tri-OMe	0.233	0.233	Tri-OMe	0.163	Tri-OMe	0.046
19	HP	0.098	0.098	*HP*	*−0.124*	HP	0.091
20	β	0.026	0.026	β	0.014	β	0.014
21	Tri-OMe	0.031	0.031	*β*	*0.026*	*β*	*0.026*
22	None	–	–	None	–	None	–

[a] Erroneous results are italicized. If an analyte cannot be resolved, the optimum cyclodextrin is ''None''.

being predicted to behave in the same manner as compound 16 (2-(3-chlorophenoxy)propionic acid). The only difference between these two analytes is the position of the –Cl substitution on the phenyl ring. However, these two compounds are optimally resolved using different cyclodextrins. This leads to the incorrect cyclodextrin being predicted for optimum resolution of compound 21.

3.2. Model 2

The first model was unable to predict the proper migration order for a pair of analytes, and this may be related to the descriptors used not being enantiospecific. To test this hypothesis, a second model was derived using linear regression, ignoring the migration order. The selectivity was thus defined as:

$$\xi = \frac{\mu_{\text{eff}}(\text{2nd migrating enantiomer})}{\mu_{\text{eff}}(\text{1st migrating enantiomer})} \quad (3)$$

The GFA was used to obtain an equation for each cyclodextrin. The quantity $(\xi - 1)$ was modeled, to avoid the need for a leading constant in the equations. The initial population was 300 equations, and breeding continued for 5000 generations. The equations were restricted to three terms, with spline functions allowed. The same training and test sets were used as were used for the first model. As was the case for model 1, the best model was found by trying various combinations of high-scoring equations, rather than simply combining the single highest-scoring equation found by the GFA for each cyclodextrin.

The retained equations for model 2 are presented in Table 9, and the performance of model 2 is displayed in Table 8. As would be expected, this second model removed the errors in migration order in both the training set and the test set. However, the model still incorrectly predicts compound 21 to optimally resolve using β-CD.

Despite the improved success of the second model, it is not possible to extract meaningful physical information from the component equations. This may indicate the data is not amenable to linear regression allowing physical insight. There are at least nine compounds that are unresolved using each cyclodextrin, and as a result the selectivity data are rather discontinuous. Mathematically, QSAR modeling is an attempt to perform a coordinate transformation that maps the activity (retention, selectivity) data onto a line, a plane, or a hyperplane, depending on the dimensionality of the equation, with the calculated descriptors as the available coordinate choices. The discontinuity of the selectivity data required the use of spline functions to exclude the non-active (unresolved) portions of the coordinate space.

It should be noted that the discontinuity does not automatically imply unsuitability for linear regression, as an ideal case for discontinuous data is presented in Fig. 3. When a hierarchy of spline functions can be established, it may be possible to assign real physical meaning to the generated QSER equation. In the case of Fig. 3A, the physical situation could be interpreted as: if an analyte is sufficiently hydrophobic (A log P), and is the right size to fit into the cyclodextrin cavity (SASA), and has the right degree of hydrogen-bonding (HOMO), then it may be resolved.

Instead, the typical case in this work is represented by Fig. 3B. For either model, in the regressed equation for a given cyclodextrin each spline function included only a few compounds. There was little overlap between the included compounds of each spline function, and it was this lack of overlap that made it difficult to assign physical meaning to any of

Table 9
Equations comprising model 2 (multivariate linear regression)

Cyclodextrin	Equation
β-CD	$2.17 \times 10^{-4}\langle 349.3 - \text{TASA}\rangle + 7.0 \times 10^{-3}\langle - 14.811 - \text{HOMO}\rangle - 0.37\langle 0.512 - \text{RASA}\rangle$ R^2: 0.696 $R^2_{\text{C.V.}}$: 0.563
HPβ-CD	$1.07\langle 1.527 - \nu\rangle + 7.9 \times 10^{-4}\langle 165.7 - \text{TASA}\rangle + 0.14\langle - 10.773 - \text{LUMO}\rangle$ R^2: 0.934 $R^2_{\text{C.V.}}$: 0.930
Tri-OMe-β-CD	$0.06\langle \nu - 1.861\rangle + 1.35 \times 10^{-3}\langle \text{SASA} - 393.0\rangle - 3.21 \times 10^{-2}\langle \text{HOMO} + 11.061\rangle$ R^2: 0.778 $R^2_{\text{C.V.}}$: 0.740

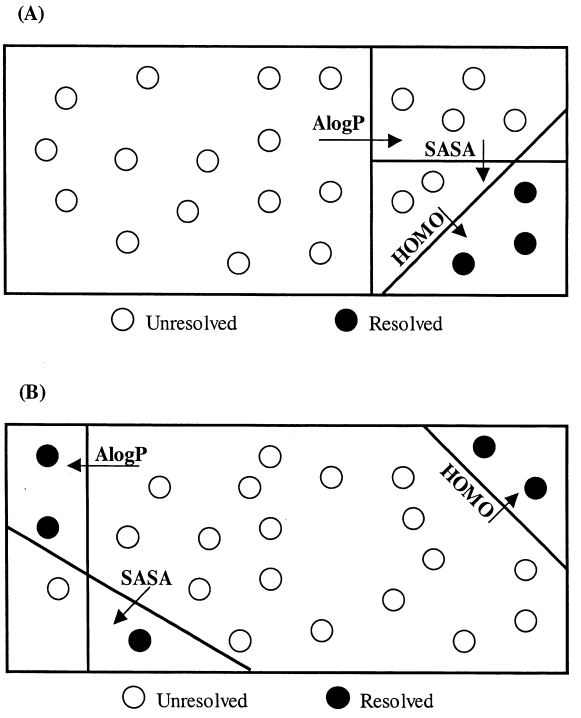

Fig. 3. (A) Ideal case when using spline functions. (B) Non-ideal case when using spline functions. Arrows indicate direction of included points.

the equations. It was not apparent whether the spline functions were representative of underlying physical phenomena, or were simply mathematically convenient.

A second possible difficulty is that while the descriptors used in this study are well suited to describe the binding strength between an analyte and a given cyclodextrin, the enantioselectivity is not necessarily related to the binding strength. Instead, the selectivity is dependent on the difference in binding strength between enantiomers, and it was not possible to calculate meaningful enantiospecific descriptors.

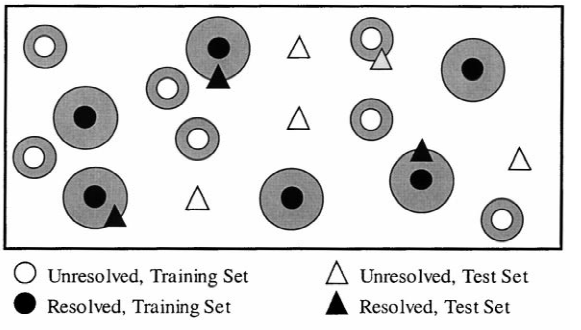

Fig. 4. Representation of Gaussian Functions. Gray areas surrounding points in training set indicate the domain of each Gaussian.

3.3. Model 3

Since it was not possible to assign strong physical meaning to the models developed using simple linear regression, a more sophisticated modeling approach was chosen whereby better predictive ability was expected, albeit at the expense of losing any ability to directly interpret the model. The third model was developed using a general regression neural network (GRNN). A GRNN is a three-layer network, with one hidden layer. The hidden layer contains a neuron for each record in the training set, and the hidden-layer activation function is a Gaussian. The form of the activation function for neuron j is given below:

$$A_j = \exp\left(-\frac{\sum_i \sigma_i (v_i - V_{ij})^2}{2\sigma_u^2} \right) \quad (4)$$

In Eq. (4), v_i is the value of descriptor i being input to neuron j, and V_{ij} is the value of descriptor i for record j in the training set. Thus, the center of the Gaussian activation function for neuron j is located at the point in descriptor space corresponding to record j. As the value of any of the descriptors deviates from that of record j, the response of the activation function (A_j) decreases in magnitude. The standard deviation of the Gaussian (σ_u) is referred to as the universal smoothing factor. The larger this number, the less sensitive is the response of each neuron to deviations from the center-point.

The Gaussian activation functions of the GRNN allowed formation of a highly discontinuous representation of the descriptor space, and thus should be appropriate to model the discontinuous selectivity data. Each Gaussian is centered at a point in descriptor space corresponding to a compound in the training set, and possesses a local neighborhood of influence. The GRNN takes a new compound and generates a response if the new compound lies in one of the local neighborhoods. This is represented pictorially in Fig. 4. To lie in the neighborhood of

influence of one the compounds in the training set, the new compound must have descriptor values very similar to those of the training compound. The GRNN may thus be thought of as performing pattern matching on the test set data.

The model was built by training a network on the experimental results for each cyclodextrin. The optimum cyclodextrin for resolving a particular analyte was determined by comparing the responses of the three networks. The quantity $(\xi_{+/-} - 1)$ was modeled, with $\xi_{+/-}$ defined by Eq. (1). The model was thus expected to predict optimum cyclodextrin, magnitude of separation, and migration order. The model was structured in this fashion for comparison with model 1, where the same quantity was modeled using multivariate linear regression.

A genetic algorithm (GA) directed descriptor selection for the networks. Starting with a number of possible descriptors, the GA generates a population of sets of individual smoothing factors (σ_i in Eq. (4)), one for each descriptor. A network is developed for each set of individual smoothing factors by optimizing the value of the universal smoothing factor to best fit the test set data. The sets of individual smoothing factors that produce the best fits to the test set data have the greatest chance of having their elements propagated into future generations. This is equivalent to using the GFA to direct multivariate linear regression, with the algorithm selecting individual smoothing factors instead of descriptor identities. Network training is stopped when 20 successive generations have not reduced the overall mean-squared error fit to the test set data by more than 1%. This requirement prevented overtraining of the network. Larger individual smoothing parameters yield a narrower Gaussian, and indicate the activation function is more sensitive to deviations in a given descriptor. Very small (~0) smoothing parameters indicate descriptors that have little effect on network performance. The number of possible descriptors was reduced by successively eliminating those with very small individual smoothing parameters, until the best set of three descriptors was obtained. The possible descriptors were restricted to those presented in Table 7.

Each GRNN had three adjustable parameters (σ_i) varied to fit the data in the test set. To maintain a proper ratio of data points to adjustable parameters (>2:1), analytes 3, 12, and 15 were moved from the training set to the test set. The training set was then comprised of fourteen analytes, and the test set of eight analytes. The descriptors used in each network, and their associated smoothing parameters, are displayed in Table 10. As evidenced by the values of the individual smoothing parameters, in each case all three descriptors made a significant contribution to the network. The overall performance of this model in selecting the optimum cyclodextrin for separation is presented in Table 11.

The identity of the optimum cyclodextrin and the magnitude of the optimum separations for compounds in the training set are correlated almost exactly, as with this type of network, deviations in the fit of the training data occur only when two of the Gaussians have overlapping regions of influence. The model correctly predicts the optimum cyclodextrin and migration order for seven of the eight analytes in the test set. However, as was the case for model 1, this model predicts compound 21 to behave in the same way as compound 16, resulting in an incorrect optimum cyclodextrin.

3.4. Model 4

The final model built treated the data as a categorization problem, directly answering the relevant question. This model ignored the migration order and the magnitude of the separation, but allowed the optimum cyclodextrin to be determined with a single neural network.

Table 10
Included descriptors and smoothing factors for model 3 (GRNNs)

Cyclodextrin	Smoothing factor	
β-CD	Universal	0.134
	HOMO	2.165
	TASA	2.200
	Shadow-ν	1.976
HP-β-CD	Universal	0.068
	LUMO	1.365
	TASA	2.765
	Shadow-ν	2.529
Tri-OMe-β-CD	Universal	0.057
	LUMO	1.294
	TASA	2.247
	Shadow-ν	2.612

Table 11
Selection of optimum cyclodextrin by models using neural networks[a]

Compound Training set	Experimental		GRNN		PNN	
	CD	$(\xi_{+/-}-1)$	CD	$(\xi_{+/-}-1)$	CD	Confidence
1	Tri-OMe	0.098	Tri-OMe	0.098	Tri-OMe	0.999
2	Tri-OMe	0.175	Tri-OMe	0.175	Tri-OMe	1.000
4	Tri-OMe	0.084	Tri-OMe	0.084	Tri-OMe	0.998
5	HP	−0.083	HP	−0.083	HP	0.993
6	HP	−0.170	HP	−0.170	HP	0.845
7	None	–	None	–	None	1.000
8	HP/β	0.018	β	0.018	HP	0.343
9	Tri-OMe	−0.006	Tri-OMe	−0.006	Tri-OMe	1.000
10	None	–	None	–	None	0.503
11	HP	−0.025	HP	−0.016	HP	0.410
13	HP	−0.016	HP	−0.016	HP	0.999
14	β	0.038	β	0.040	β	0.999
16	β	0.031	β	0.032	β	0.994
17	Tri-OMe	−0.033	Tri-OMe	−0.032	Tri-OMe	0.985
					Conf.	**86.2**
Test set						
3	Tri-OMe	0.074	Tri-OMe	0.084	Tri-OMe	0.999
12	HP	−0.008	HP	−0.016	HP	0.387
15	β	0.016	β	0.017	β	1.000
18	Tri-OMe	0.233	Tri-OMe	0.175	Tri-OMe	0.940
19	HP	0.089	HP	0.018	HP	0.844
20	β	0.026	β	0.030	β	0.981
21	Tri-OMe	0.030	*β*	*0.018*	Tri-OMe	0.910
22	None	–	None	–	None	0.858
					Conf.	**86.5**

[a] Erroneous results are italicized. If an analyte cannot be resolved, the optimum cyclodextrin is ''None''.

The single network used was a probabilistic neural network (PNN). The PNN was similar to the GRNN, as it had three layers with one node in the hidden layer for each record in the training set, and the activation function in the hidden layer was a Gaussian (Eq. (4)). There were four output nodes: one for each cyclodextrin and ''None'' in case an analyte could not be resolved. Each output neuron (or cyclodextrin) is connected only to those neurons in the hidden layer that represent cases in the training set that are optimally resolved using the particular cyclodextrin. For example, the output neuron for β-CD is connected to the hidden neurons corresponding to analytes 8, 14, and 16. When a test analyte is presented to the network, the output neuron that fires the strongest (largest magnitude) is designated as optimum.

The training and test sets are the same as those used for model 3. The data used to train the network is a four-element array for each analyte, with a one representing the optimum cyclodextrin to perform the separation and zeroes for the other elements. Atrolactic acid (analyte 11) was assigned a one for both β-CD and HP-β-CD, as the selectivity using these cyclodextrins was identical within experimental error. The descriptors to include in the network were chosen using the GA, in the same fashion as for model 3. Network training is again stopped when 20 successive generations have not reduced the overall mean-squared error fit to the test set data by more than 1%, to prevent over-training of the network.

The descriptors retained in model 4 were HOMO, TASA, and Shadow-ν. The individual smoothing factors were 1.682, 0.482, and 1.282, respectively. The universal smoothing factor was 0.088. The performance of model 4 is presented in Table 11. The confidence of the model prediction is defined as:

$$\text{Confidence} = \frac{\text{Max(actual node values)}}{\sum \text{actual node values}} \quad (5)$$

Maximum confidence in a result is obtained when only the hidden neurons connected to a single output neuron fire. Confidences less than 1 for the training set occur when hidden neurons corresponding to different optimum cyclodextrins have overlapping regions of influence.

Model 4 performs well using a single network, correctly predicting the optimum cyclodextrin for each analyte in the test set. However, this model does not give migration order or an estimate of the separation magnitude, and thus may be less useful than the other models, depending on the needs of the user.

3.5. Use of R/S vs. $\pm$ designation

The data can also be modeled using the R/S designation to identify the order of migration. The direction of optical rotation was chosen because it was known absolutely for the first eluting enantiomer of each analyte. For five of the analytes, the R/S designation was not known absolutely, and would have had to be inferred from the relationship between R/S and direction of optical rotation of similar analytes in the data set.

Table 12 shows the migration order for our data set using the R/S designation, including the best guesses at the unknown designations. Models 1 and 3 have been re-derived with the selectivity defined as:

$$\xi_{S/R} = \frac{\mu_{\text{eff}}(S \text{ enantiomer})}{\mu_{\text{eff}}(R \text{ enantiomer})} \quad (6)$$

The models thus developed perform identically to the models developed with the selectivity defined by Eq. (1). The same errors are committed on the same compounds. Models 2 and 4 do not depend on the migration order.

3.6. Comparison of models

The models developed to simultaneously predict the optimum cyclodextrin for resolution, migration order, and magnitude of separation performed fairly well. With either multivariate regression or neural networks, in only one case was the incorrect optimum cyclodextrin identified. However, this error and errors that occurred in prediction of the migration order may indicate that neither mathematical approach was capable of recognizing the more subtle distinctions between the compounds present in the data set.

If the only quantity of interest is the identity of the optimum cyclodextrin for resolution, the modeling approach using a PNN worked excellently. Model 4, developed with a PNN, was able to accurately predict the optimum cyclodextrin for resolving a test set of analytes. Model 2, developed using multivariate linear regression, retained an error in predict-

Table 12
First eluting enantiomerof analytes: (R/S) designation

Compound	β-CD	HP--β-CD	Tri-OMe--β-CD	Compound	β-CD	HP--β-CD	Tri-OMe--β-CD
1	–	–	*R*	13*	–	*R*	–
2	–	–	*R*	14	*S*	–	–
3	–	–	*R*	15*	*S*	–	–
4	–	–	*R*	16*	*S*	*S*	*S*
5	–	*S*	*S*	17*	–	*R*	*R*
6	–	*S*	*S*				
7	–	–	–				
8	*R*	*R*	–	18	–	–	*R*
9	–	–	*S*	19	–	*S*	*S*
10	–	–	–	20	*S*	*S*	–
11	–	*R*	–	21	*S*	*S*	*S*
12*	–	*R*	–	22	–	–	–

* Relationship between R/S and optical rotation inferred from similar compounds.

ing the optimum cyclodextrin to resolve one of the analytes in the test set.

The inability to assign physical significance to the equations in the linear regression models affects the range where the models may be confidently applied. For compounds that represent interpolation from the training set compounds, the models should produce results similar to those obtained for the test and training sets. However, because in this case the spline functions appear to be mathematical conveniences rather than representative of underlying physical phenomena, the models may behave poorly when presented with data requiring extrapolation from the training set. The distinction between interpolation and extrapolation would be based on the values of the calculated descriptors for the new compounds.

By the nature of their activation functions, both GRNN and PNN may be thought of as local networks, as each Gaussian has only a limited neighborhood of influence. The neural network models should be successfully applied to new compounds that represent interpolation of the test and training sets. However, the limited range of influence of each Gaussian would make it difficult to apply the neural network models to compounds whose descriptor values fall significantly outside the range of the training set compounds.

The modeling approach described in this work has been shown to yield useful answers for a series of compounds with similar structural features (such as those from a discovery effort in a single therapeutic area). It has been shown that a PNN may be successfully applied to identify the optimum chiral selector to resolve a given analyte. The PNN may be confidently applied to new compounds that represent interpolation from the data used to develop the network, and once the network is trained, it is simple to identify those compounds that require extrapolation from the data used to develop the network. The GRNN and multivariate linear regression models described here can also provide useful information, albeit with errors not present in the model based on a PNN. Because of the inability to extract physical information from the model equations, it is not clear how successfully this approach could be applied to a set of compounds with a high degree of structural diversity.

References

[1] I.W. Wainer, in: I.W. Wainer (Ed.), Drug Stereochemistry, 2nd ed., Dekker, New York, 1987, p. 139.
[2] R. Vespalec, P. Bocek, Electrophoresis 20 (1999) 2579.
[3] B. Chankvetadze, G. Blaschke, Electrophoresis 20 (1999) 2592.
[4] A. Shibukawa, D.K. Lloyd, I.W. Wainer, Chromatographia 35 (1993) 419.
[5] M. Fillet, P. Hubert, J. Crommen, Electrophoresis 19 (1998) 2834.
[6] S. Fanali, S. Furlanetto, Z. Aturki, S. Pinzauti, Chromatographia 48 (1998) 395.
[7] V. Dohnal, M. Farkova, J. Havel, Chirality 11 (1999) 616.
[8] X.-F. Zhu, B.-C. Lin, A. Jakob, U. Epperlein, B. Koppenhoefer, J. High Resolut. Chromatogr. 22 (1999) 449.
[9] R.M. Wolfe, E. Francotte, D. Lohman, J. Chem. Soc., Perkin Trans. II. (1988) 893.
[10] U. Norinder, J. Hermansson, Chirality 3 (1991) 422.
[11] R. Kaliszan, T.A.G. Noctor, I.W. Wainer, Chromatographia 33 (1992) 546.
[12] C. Roussel, C. Popescu, Chirality 6 (1994) 251.
[13] T.D. Booth, I.W. Wainer, J. Chromatogr. A 737 (1996) 157.
[14] M. Farkova, E.-M. Pena-Mendez, J. Havel, J. Chromatogr. A 848 (1999) 365.
[15] R.-H. Zhao, B.F. Yue, J.-Y. Ni, H.-F. Zhou, Y.-K. Zhang, Chemometr. Intell. Lab. Syst. 45 (1999) 163.
[16] T.D. Booth, K. Azzaoui, I.W. Wainer, Anal. Chem. 69 (1997) 3879.
[17] R. Kaliszan, M.A. van Straten, M. Markuszewski, C.A. Cramers, H.A. Claessens, J. Chromatogr. A 855 (1999) 455.
[18] J.Y. Dai, S.C. Yao, Y.B. Ding, L.S. Wang, J. Liq. Chromatogr. Relat. Technol. 22 (1999) 2271.
[19] M. Blanco, J. Coello, H. Iturriaga, C. Perez-Maseda, J. Chromatogr. A 793 (1998) 165.
[20] S. Fanali, Z. Aturki, J. Chromatogr. A 694 (1995) 297.
[21] A.T. Hagler, E. Huler, S. Lifson, J. Am. Chem. Soc. 96 (1974) 5319.
[22] A.T. Hagler, S. Lifson, J. Am. Chem. Soc. 96 (1974) 5327.
[23] S. Lifson, A.T. Hagler, P. Dauber, J. Am. Chem. Soc. 101 (1979) 5111.
[24] A.T. Hagler, S. Lifson, P. Dauber, J. Am. Chem. Soc. 101 (1979) 5122.
[25] A.T. Hagler, P. Dauber, S. Lifson, J. Am. Chem. Soc. 101 (1979) 5131.
[26] D.H. Kitson, A.T. Hagler, Biochemistry 27 (1988) 5246.
[27] D.H. Kitson, A.T. Hagler, Biochemistry 27 (1988) 7176.
[28] P. Dauber-Osguthorpe, V.A. Roberts, D.J. Osguthorpe, J. Wolff, M. Genest, A.T. Hagler, Proteins Struct. Funct. Genet. 4 (1988) 31.
[29] D. Rogers, A.J. Hopfinger, J. Chem. Inf. Comput. Sci. 34 (1994) 854.
[30] D. Specht, IEEE Trans. Neural Networks 2 (6) (1991) 568.
[31] D. Specht, P. Shapiro, Proc. Int. Joint Conf. On Neural Networks 1 (1991) 887.
[32] D. Specht, Neural Networks 3 (1990) 109.
[33] R.H. Rohrbaugh, P.C. Jurs, Anal. Chim. Acta 199 (1987) 99.
[34] D.T. Stanton, P.C. Jurs, Anal. Chem. 62 (1990) 2323.