ELSEVIER

Journal of Chromatography A, 719 (1996) 37–44

JOURNAL OF CHROMATOGRAPHY A

Use of topological descriptors in chromatographic chiral separations

J.V. de Julián-Ortiz*, R. García-Domenech, J. Gálvez Alvarez, R. Soler Roca, F.J. García-March, G.M. Antón-Fos

Unit of Drug Design and Molecular Connectivity Investigation, Department of Physical Chemistry, Faculty of Pharmacy, University of Valencia, 46100 Burjassot, Valencia, Spain

Abstract

Studies of enantiomeric separations are reported, specifically the direct chromatographic separation of enantiomers using a chiral stationary phase by molecular topology. The results obtained show good correlation equations for the capacity factor, k', and the separation factor, α, for different set of compounds (hydantoins, aromatic α-amino acids and arylamides). Such equations may be useful for the selection of the optimum stationary and mobile phases for the separation of enantiomers. Further, the correlation between topological descriptors and performance in chiral separations opens up a new approach to the design of chiral stationary phases.

1. Introduction

Molecular isomerism is a subject of fundamental interest to organic chemists. A molecule has the property of chirality when its mirror image is non-superimposable on itself. The non-superimposable mirror image isomers are called enantiomers, but may also be referred to as enantiomorphs, optical isomers or optical antipodes.

Enantiomers of a given compound have identical chemical properties with regard to their reactions with non-chiral reagents, although they will give products with different configurations. In addition, they may show differences in behaviour (both in reaction rates and in product stereochemistries) in their interactions with one enantiomer of a chiral reagent.

Since the discovery of the optical isomerism of tartaric acid by Pasteur in 1848, the significance of stereoisomerism in relation to biological activity has been recognized. The case of thalidomide is an example of a problem that was, at least, complicated by the ignorance of stereochemical effects [1]. Thus, whenever a drug is to be obtained in a variety of chemically equivalent forms (such as enantiomers), it is both good science and good sense to explore the potential for in vivo differences between these forms.

The successful development of chiral stationary phases for high-performance liquid chromatography (HPLC) in the 1970s was particularly helpful to the study of direct enantiomeric separations without the need to obtain a compound by stereospecific synthesis.

A close relationship often exists between the molecular structures of organic compounds and many of their physical, chemical and biological properties [2–7]. Many of these relationships have been investigated using the topological

* Corresponding author.

0021-9673/96/$15.00
SSDI 0021-9673(95)00401-7

descriptors of molecular structures. These descriptors are used to characterize the constitution and configuration of a molecule by a single number. Topological parameters such as the connectivity [8], χ_i, charge [9], G_i, J_i, and geometrical [10] indices can be used to quantify these properties. In this work, we applied molecular connectivity to study the separation of optical isomers, particularly direct chromatographic separations of enantiomers on a chiral stationary phase (hydantoins, aromatic α-amino acids and arylamides)[11–14]. Since the choice of the stationary and mobile phases to be used in such resolutions is conditioned by the value of the column capacity factor, k', and the separation factor, α, both parameters were correlated in this study.

2. Experimental

2.1. Method of calculation

The topological descriptors used in this work were as follows.

Connectivity indices. These were calculated from a hydrogen-suppressed equation or graph of the molecule, following the method of Kier and Hall [8]. They were utilized up to the fourth order and combinations (such as differences or quotients) were also used.

Charge indices. The charge indices (C.I.) [9] G_k and J_k were defined as

$$G_k = \sum_{i=1,j=i+1}^{i=N-1,j=N} |CT_{ij}|\delta(k, d_{ij}) \quad (1)$$

and

$$J_k = \frac{G_k}{N-1} \quad (2)$$

where N = number of vertices (atoms different to hydrogen), $CT_{ij} = m_{ij} - m_{ji}$, m is the elements of the M matrix:

$$M = A \cdot D^* \quad (3)$$

A = adjacency $(N \cdot N)$ matrix in which the number of sustituents of each atom appears in the main diagonal, D^* = inverse square distance matrix, in which their diagonal entries are assigned as 0, and δ = Kronecker's delta. Thus, G_k represents the sum of all the CT_{ij} terms, with $d_{ij} = k$, d_{ij} being the energies of the topological distance matrix.

In valence C.I. terms, the presence of heteroatoms is taken into account by introducing their electronegativity values (according to Pauling's scale taking chlorine as a standard value of 2) in the corresponding entry of the main diagonal of the adjacency matrix.

Geometrical indices. Other geometrical descriptors taking into account molecular characteristics, such as shape and volume, were also used. First, the shape index E is defined, for aliphatic compounds, as

$$E = \frac{\Sigma n_i(d_i + 1)}{L} \quad (4)$$

where n_i is the number of vertices placed at a topological distance d_i from the main path, the last being the set of edges joining the two most separated vertices in the graph by the shortest path, and L represents the graph length, i.e., the number of edges constituting the main path. It is clear that the lower the E value, the more elongated the graph is. If the graph is assimilated to an ellipse, E would represent the eccentricity.

In the case of cyclic compounds, E is calculated by:

$$E = \frac{S}{L^2} \quad (5)$$

where S is the surface parameter, calculated as the sum of the S values for all the ring and aliphatic fragments in the molecule, which, in turn, are obtained as the products [10]

$$S = E \cdot L^2 \quad (6)$$

The other geometrical indices are as follows:

R = number of ramifications (i.e., number of vertices connected with three different ones);

T_{nr} = number of non-ramified terminal vertices (i.e., number of terminal vertices showing valence 1 linked to vertices with valence 2);

Pr_0 = number of vertices showing topological valence = 4;
Pr_1 = number of pairs of adjacent (separated by one edge) ramifications;
Pr_2 = number of pairs of ramifications separated by two edges;
Pr_3 = number of pairs of ramifications separated by three edges.

In addition, also included are the vertices number (N) and the Wiener index (W). Therefore, each compound is characterized by a set of 62 indices.

Once the indices have been calculated for each molecule, single and multiple regression analyses were used to find the relationship between the k' and α values and the topological indices, calculated using

$$P = A_0 + \sum_{m,t} A^m_{m,t}\chi_t \quad (7)$$

where P is a property and A_0 and $A_{m,t}$ represent the regression coefficients of the equation obtained. To test the quality of the regression equations, the following statistical parameters were used: multiple correlation coefficient (r), standard error of estimates (s), F-Snedecor function values (F) and Student's t-test (statistical significance).

2.2. Instrumentation and columns

The capacity factors of the racemic hydantoins (first eluted) used in this study were reported by Maguire [11]. Hydantoins (Table 1) were subjected to gradient chromatography using two Model 110A pumps (Altex, Berkeley, CA, USA) controlled by an Axxiom Model 710 microprocessor, an ISCO (Lincoln, NE, USA) Model V4 detector set at 254 nm, a Rheodyne (Cotati, CA, USA) Model 7125 injector with a 20-μl sample loop and a Hewlett-Packard (Avondale, PA, USA) Model 3390 integrator. A β-cyclodextrin column (5 mm; 250 × 4.6 mm I.D.) from Astec was used, with a 0.5-mm in-line filter. The column temperature was controlled at 22°C with the use of a water jacket and a thermostated circulating water-bath. A flow-rate of 0.8 ml/min and 10% methanol as the mobile phase were chosen.

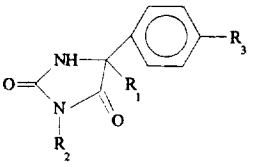
The capacity factors of racemic α-amino acids (first eluted) used in this study were reported by Armstrong et al. [12]. Amino acids (Table 2) were studied with a Shimadzu (Columbia, MD, USA) Model LC-6A liquid chromatograph with a variable-wavelength detector in the constant-flow mode; the sample injector was a Rheodyne Model 7125 with a 20-μl sample loop; a pre-saturator containing 40-mm silica gel was placed before the injection port. The column was a 25-cm α cyclodextrin bonded phase, made as reported by Ward and Armstrong [13], expect that the slurry reaction temperature was lowered to 130°C and the reaction time was increased. The column was packed at Advanced Separation Technologies (Whippany, NJ, USA) and commercially available. As the support, 5 mm diameter silica gel was used. All separations were carried out at room temperature. The flow-rate was 0.5 ml/min and the mobile phase was 1% triethylamine acetate (pH 5.1). When not in use, the column was stored in methanol.

The selectivities and capacity factors (first eluted) of the racemic arylamides used were reported by Däppen et al. [14]. Aryl derivatives (Table 3) were resolved using a Model 110 solvent metering pump (Altex, Berkeley, CA, USA). The detector was a Uvikon LCD 725 UV (Kontron, Zürich, Switzerland), set at 254 nm. A Rheodyne Model 7120 sample injector with a 20-μl sample loop was used. The column was a stainless-steel tube (250 × 3.2 mm I.D.) packed with 1-(α-naphthyl)ethylamine stationary phase prepared according to Däppen et al. [14]. A flow-rate of 1 ml/min and n-hexane–tetrahydrofuran (75:25) as the mobile phase were chosen.

2.3. Chemicals

In the experiments with hydantoins, HPLC-grade methanol from Fischer Scientific (Pittsburg, PA, USA) was used. HPLC-grade water, produced from distilled, deionized water previously treated with a Millipore Norganic filter system from Waters (Milford, MA, USA) was employed throughout. 5-Alkyl-5-

Table 1
Correlation of k' for a set of 5-(p-substituted phenyl)-hydantoins with connectivity and charge indices



Compound			$^4\chi_p^v$	G_3^v	k'(obs.) [11]	k'(calc.)[a]	Residual
R_1	R_2	R_3					
H	H	H	1.30	1.96	0.67	0.96	−0.29
Me	H	H	1.52	3.55	1.17	0.93	0.24
Et	H	H	1.87	4.00	2.30	2.18	0.12
Pr	H	H	2.11	4.06	3.30	4.35	−1.05
iPr	H	H	2.11	4.45	4.96	3.67	1.29
Bu	H	H	2.18	4.06	4.40	5.30	−0.90
Et	Me	H	2.06	5.52	2.15	2.00	0.15
Et	Et	H	2.28	5.92	2.18	3.16	−0.98
Et	Pr	H	2.41	6.13	3.13	4.36	−1.23
Et	iPr	H	2.43	6.33	2.88	4.14	−1.26
Et	Bu	H	2.51	6.19	5.28	5.60	−0.32
Et	iBu	H	2.52	6.34	6.38	5.42	0.96
Et	Pent	H	2.71	6.19	14.30	10.21	4.09
Et	Bz	H	2.90	6.05	17.80	18.74	−0.94
Et	Me	OH	2.08	5.70	2.48	1.98	0.50
Et	H	OH	1.89	4.19	2.87	2.15	0.72
Pr	H	OH	2.14	4.24	5.98	4.30	1.68

Stationary phase, β-cyclodextrin; mobile phase, 10% methanol.
[a] Selected equation:

$$\log k' = -1.314 + 1.280\,{}^4\chi_p^v - 0.185\,G_3^v \quad (10)$$
$$n = 17; \quad r = 0.9426; \quad F = 55.76; \quad p < 0.001$$

phenylhydantions were prepared via the method of Sobotka et al. [15], 3-alkyl-5-ethyl-5-phenylhydantoins were prepared as described by Maguire et al. [16] and 5-(4-hydroxyphenyl)hydantoins were available from a previous study by Butler [17]. All hydantoins were prepared as methanolic solutions at a concentration of 1.0 g/l.

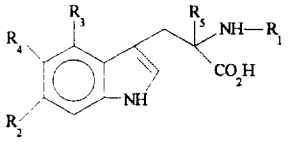
In the chromatography of amino acids, HPLC-grade methanol and water from Fischer Scientific (Plano, TX, USA) were used. The mobile phase [1% triethylamine acetate (pH 5.1)] was filtered through a 4.5-mm frit before use. Tryptophan, phenylalanine, tyrosine and their analogues were obtained from Sigma (St. Louis, MO, USA) and used without purification. All samples were dissolved in methanol or water prior to injection.

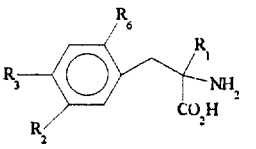
In the resolution of arylamides, LiChrosorb Si 100 (Merck, Darmstadt, Germany) and Spherisorb S5W (Phase Separations, Queensferry, UK) with a particle size of 5 mm were used. All other chemicals were purchased from Fluka (Buchs, Switzerland) or Merck.

3. Results and discussion

The effectivity of a chromatographic column with respect to enantiomeric resolution depends largely on the relative elution velocities of the

Table 2
Correlation of k' for a set of aromatic α-amino acids with connectivity and charge indices





	Structure N° 1					Structure N° 2			
Structure	R_1	R_2	R_3	R_4	R_5	R_6	k'(obs.) [12]	k'(calc.)[a]	Residual
1	H	H	H	H	H	H	2.70	3.20	−0.50
	H	H	H	H	Me	H	2.60	3.29	−0.69
	H	H	H	Me	H	H	4.90	4.80	0.10
	H	Me	H	H	H	H	6.40	6.65	−0.25
	CHO	H	H	H	H	H	8.00	7.48	0.52
	H	H	F	H	H	H	2.40	1.64	0.76
	H	H	H	F	H	H	2.70	3.12	−0.42
	H	F	H	H	H	H	3.70	3.44	0.26
2	H	H	H	–	–	H	1.10	0.47	0.63
	Me	H	OMe	–	–	H	0.70	0.88	−0.18
	Me	H	H	–	–	H	0.90	1.13	−0.23
	H	H	F	–	–	H	1.50	0.57	0.93
	H	F	H	–	–	H	1.00	1.62	−0.62
	H	H	OH	–	–	H	0.10	0.76	−0.66
	H	H	OH	–	–	Me	4.40	4.06	0.34

Stationary phase, α-cyclodextrin; mobile phase, 1% triethylamine acetate (pH = 5.1).
[a] Selected equation:

$$k' = -8.262 + 0.2935G_2 - 12.187G_4 + 13.701G_4^{\vee}$$
$$n = 15; \quad r = 0.9702; \quad s = 0.62; \quad F = 58.70; \quad p < 0.001 \qquad (11)$$

two species. Such effectivity may be quantified through the capacity factor, k' (which describes the analytes' migration velocities), and the selectivity factor, α (which distinguishes between the analytes' velocities). For a given column and phases, the optimum working conditions usually imply the values of $1 < k' < 5$ and $\alpha > 1$. The existence of a theoretical procedure, easily ready to use, able to predict the α and k' values for a given compound, would allow the a priori estimation of such parameters, and therefore the efficiency of resolution of the two enantiomers.

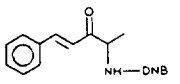
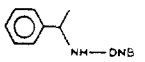
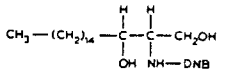
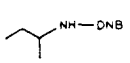
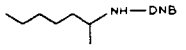
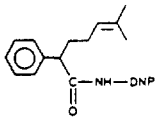
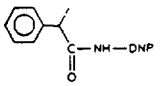
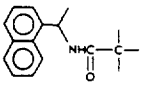
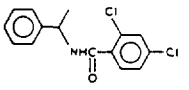
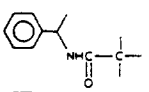
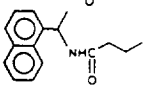
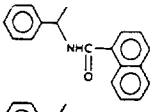
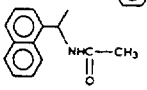
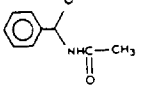
A possible approach is the use of molecular topology as a way of establishing the correlation equation between the experimental values of k' and α and suitable topological descriptors. The results obtained are given in Table 1–3. The regression equations for k' and α for the arylamides were

$$\log \alpha = 0.585 + 1.846J_3^{\vee} + 0.354({}^0\chi - {}^0\chi^{\vee}) - 0.773{}^4\chi_{px}/{}^4\chi_{pc}^{\vee}$$
$$n = 14; \quad r = 0.9795; \quad s = 0.036;$$
$$F = 78.7; \quad p < 0.001 \qquad (8)$$

$$\log k' = 3.264 - 0.209Pr_0 + 1.289({}^1\chi - {}^1\chi^{\vee}) - 1.943{}^4\chi_{pc}/{}^4\chi_{pc}^{\vee}$$
$$n = 14; \quad r = 0.9554; \quad s = 0.12;$$
$$F = 38.3; \quad p < 0.001 \qquad (9)$$

As can be observed, the equations lead to good adjustments even with a small number of

Table 3
Correlation of k' and α for a set of chiral 3,5-dintrobenzoyl, 3,5-dinitrophenyl and other arylamides with different topological indices

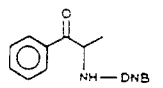
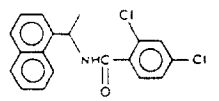
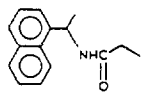
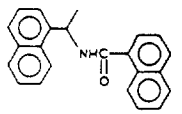
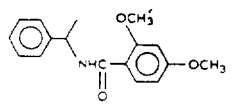
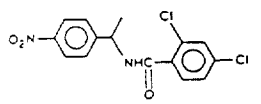
Compound	Structure	k'(obs.) [14]	α(obs.) [14]	k'(calc.)[a]	α(calc.)[a]
1		5.20	2.43	4.68	2.24
2		6.50	2.14	8.11	1.97
3		14.50	1.64	12.04	1.71
4		4.33	1.23	4.63	1.37
5		3.79	1.21	3.49	1.16
6		8.25	2.94	11.50	2.96
7		11.80	2.27	9.48	2.33
8		1.02	1.16	0.82	1.17
9		1.46	1.14	1.67	1.03
10		0.83	1.12	1.21	1.12
11		2.14	1.12	2.98	1.25
12		2.38	1.07	1.93	1.11
13		5.83	1.07	3.61	1.15
14		4.79	1.05	4.62	0.95

Stationary phase, optically active 1-(α-naphthyl)etylamine; mobile phase, n-hexane–tetrahydrofuran (75:25).
[a] From Eqs. 8 and 9.

Table 4
Randomness and stability analysis for the equations obtained

Property	Eq.	Deleted observations	$r_{av.}$	$r_{min.}$	$r_{max.}$	$s_{av.}$	$s_{min.}$	$s_{max.}$
Randomness study from random numbers assigned to dependent variable (run = 50)								
Log α	8	–	0.4313	0.1613	0.7351	0.17	0.12	0.23
Log k'	9	–	0.3849	0.2081	0.5685	0.32	0.23	0.42
Log k'	10	–	0.3378	0.0866	0.6829	0.32	0.22	0.42
k'	11	–	0.4033	0.2240	0.6410	2.17	1.56	3.12
Stability study from number of deleted observations (run > 10)								
Log α	8	1	0.9759	0.9699	0.9834	0.037	0.034	0.038
		2	0.9785	0.9621	0.9892	0.036	0.028	0.038
Log k'	9	1	0.9548	0.9478	0.9620	0.122	0.114	0.127
		2	0.9529	0.9442	0.9708	0.124	0.108	0.133
Log k'	10	1	0.9417	0.9210	0.9488	0.123	0.116	0.129
		2	0.9385	0.9182	0.9543	0.123	0.115	0.130
k'	11	1	0.9705	0.9568	0.9769	0.612	0.561	0.645

Table 5
Cross-validation results for the values of k' and α for different compounds with optical activity using the regression Eqs. 8 and 9

Compound	Structure	k'(obs.) [14]	k'(calc.)	α(obs.) [14]	α(calc.)
a		4.21	4.54	3.10	2.29
b		1.67	0.75	1.10	1.03
c		2.65	2.67	1.09	1.19
d		2.50	0.99	1.08	1.19
e		2.92	9.14	1.07	1.34
f		3.79	3.89	1.04	1.71

The equations were applied to chromatographic separation with optically active 1-(α-naphthyl)ethylamine as stationary phase and *n*-hexane–tetrahydrofuran (75:25) as mobile phase.

topological indices ($r > 0.94$), and all of them are statistically significant (tail probability over 99.9%, i.e., $p < 0.001$).

The stability analysis for these properties is illustrated in Table 4. The random value for k' and α lead, in all cases, to r values lower than 0.7351, which are far from the values of the selected equations. Moreover, the elimination of one or two molecules in the series does not appreciably change the correlations. These results clearly demonstrate that the selected equations are not random, as well as stable with respect to the elimination of one or two data.

The use, in all cases, of polar mobile phases for the enantiomeric separation conditions the presence of specific topological indices. Thus, the presence of indices such as ${}^4\chi_p^v$ and $({}^1\chi - {}^1\chi^v)$, linearly increasing the value of k', may be related to the polar character of the molecule, as shown with other types of compounds [18]. The appearance of change indices such as J_3, G_2, G_3^v, G_4 and G_4^v is representative of charge-transfer processes and clearly conditions the polar character of the correlated compounds [9].

Finally, a cross-validation study on the prediction of the α and k' values for a set of different compounds with optical activity, of course, not used in obtaining the regression equations (Eqs. 8 and 9), was carried out. The results obtained are given in Table 5. As can be seen, the predicted values are close to the experimental values which demonstrates the ability of molecular topology for the prediction, as a first approach, of an enantiomeric separation under given chromatographic conditions. Further, given a reference racemic mixture and fixed experimental conditions, and varying the chiral stationary phase, it would be possible to obtain suitable equations to optimize the molecular structure of the stationary phase.

Acknowledgement

The authors thank CICYT, SAF92-0684 (Spanish Ministry of Science and Education), for financial support of this work.

References

[1] H. Schumacher, D.A. Blake, J.M. Gurian and J.R. Gillette, J. Pharmacol. Exp. Ther., 160 (1968) 189–200.
[2] L.B. Kier and L.H. Hall, Molecular Connectivity in Structure–Activity Analysis, Research Studies Press, Letchworth, 1986.
[3] R.M. Soler, F.J. García, G.M. Antón, R. García, F. Perez and J. Gálvez, J. Chromatogr., 607 (1992) 91–95.
[4] J. Gálvez and R. García, in A. Mosqueira (Editor), Diseño de Medicamentos, Farmaindustria, Madrid, 1994, pp. 366–370.
[5] L. Strekowski and M. Mokrosz, Anti-Cancer Drug Des., 3 (1988) 79–89.
[6] G.M. Antón-Fos, F.J. García-March, F. Perez-Gimenez, M.T. Salabert-Salvador and R.A. Cercós-del-Pozo, J. Chromatogr., 672 (1994) 203–211.
[7] G.M. Antón-Fos, R. García-Domenech, F. Pérez-Gimenez, J.E. Peris-Ribera, F.J. García-March and M.T. Salabert-Salvador, Arzneim,-Forsch./Drug Res., 44 (1994) 821–826.
[8] L.B. Kier and L.H. Hall, Molecular Connectivity in Chemistry and Drug Research, Academic Press, New York, 1976.
[9] J. Gálvez, R. García, M.T. Salabert and R. Soler, J. Chem. Inf. Comput. Sci., 34 (1994) 520–525.
[10] R. García, J. Gálvez, R. Moliner and F. García, Drug Invest., 3 (1991) 344–350.
[11] J.H. Maguire, J. Chromatogr., 387 (1987) 453–458.
[12] D.W. Armstrong, X. Yang, S.M. Han and R.A. Menges, Anal. Chem., 59 (1987) 2594–2596.
[13] T.J. Ward and D.W. Armstrong, in L.J. Crane and M. Zief (Editors), Chromatographic Chiral Separations, Marcel Dekker, New York, 1987, p. 131.
[14] R. Däppen, V.R. Meyer and H. Arm, J. Chromatogr., 361 (1986) 93–105.
[15] H. Sobotka, M.F. Holzman and J. Kahn, J. Am. Chem. Soc., 54 (1932) 4697.
[16] J.H. Maguire, A.R. Murthy and I.H. Hall, Eur. J. Pharmacol., 117 (1985) 135.
[17] T.C. Butler, J. Pharmacol. Exp. Ther., 117 (1956) 160.
[18] R. Soler, M.C. Sipan, R. García and J. Gálvez, Indian J. Chem., 33B (1994) 209–215.