Structure–retention relationships of steroid hormones in reversed-phase liquid chromatography and micellar electrokinetic capillary chromatography

M. Salo[a,*], H. Sirén[b], P. Volin[b], S. Wiedmer[b], H. Vuorela[a]

[a]*Department of Pharmacy, P.O. Box 56, FIN-00014 University of Helsinki, Helsinki, Finland*
[b]*Department of Chemistry, P.O. Box 55, FIN-00014 University of Helsinki, Helsinki, Finland*

Abstract

Quantitative structure–retention relationship (QSRR) studies are useful in retention prediction, finding the relevant structural descriptors for analytes and estimating the relative biological activities of a series of analytes. Most of the studies have been conducted by RP-HPLC and a few by micellar electrokinetic capillary chromatography (MECC). The aim of this work was to find structural parameters and characteristics related to the RP retention and electrophoretic migration of steroids in order to predict the retention/migration of steroid hormones on the basis of their molecular structure. Retention data were obtained with an ODS column using as mobile phases aqueous methanol–acetonitrile (mobile phase A) and methanol–tetrahydrofuran (mobile phase B). MECC was conducted with a sodium dodecyl sulfate (SDS)–borate system and with a mixed micellar solution of SDS and sodium cholate. Several topological indices, such as the connectivity indices, χ, were used as structural parameters. Steric factors seem to have a great effect on the retention of steroid hormones, especially with the MeCN-containing mobile phase. Retention in mobile phase B could be predicted more accurately. Topological indices can be used in the modelling and prediction of the retention/migration of steroid hormones when the solutes form a congeneric series and stereochemical properties do not govern the separation process.

Keywords: Micellar electrokinetic chromatography; Quantitative structure–retention relationships; Connectivity indices; Topological indices; Molecular descriptors; Factor analysis; Cluster analysis; Regression analysis; Steroids; Hormones

1. Introduction

Quantitative structure–retention relationship (QSRR) studies aim at understanding the connections between molecular structure and the chromatographic retention. They assist in the understanding of the separation mechanism and

* Corresponding author.

the optimization of the separation. QSRR studies can be useful in retention prediction, finding the relevant structural descriptors for analytes, evaluating the physico-chemical properties of the analytes and estimating the relative biological activities of a series of analytes. Most of the studies have been conducted by RP-HPLC, mainly because the structural properties of the analytes affecting the retention are relatively well

known. The retention is primarily affected by the size and polarity of the analyte molecule; steric effects have also been found to influence the retention. Linear solvatochromic relationships (LSER), which describe the analyte–solvent interactions by terms related to analyte size, hydrogen-bonding characteristics and dipolarity/polarizability, have been successfully used in describing the retention [1,2]. The main factors affecting the retention were the analyte size and its hydrogen-bonding basicity. RP-HPLC has been found useful in evaluating the hydrophobicity of the analytes [3] and also the pharmacological classification of drugs [4].

In micellar electrokinetic capillary chromatography (MECC), the migration behaviour of the compounds depends on the charge, size and orientation of the analytes, concentrations of all the components in the electrolyte solutions, partition of the compounds into the micelles and the instrumental parameters. LSER studies have shown that with anionic surfactants the migration is mainly governed by the size and hydrogen-bond basicity of the analytes, as in RP-HPLC [5].

Since there are no solvatochromic parameters available for complex molecules, other empirical or theoretical structural parameters must be used for the quantitative expression of the molecular structure. The hydrophobic nature of compounds often correlates with their biological and pharmacological effects. Hydrophobicity is most often described by the partition coefficient, log P, or fragmental values, such as the hydrophobic parameter, π [6]. Topological indices are based on the two-dimensional molecular structure and graph theory. Following the work of Wiener [7], who introduced an index based on the distance matrix of the molecule, the Wiener number (W), many other similar parameters have been developed. Connectivity indices were introduced by Randic [8] and further developed by Kier and Hall [9] into a system of descriptors. The kappa indices, κ, are related to the molecular shape [10]. The topological indices have been claimed to contain no size-independent information [11]. However, they performed as well or better than electronic and geometric descriptors in a thorough study by Katritzky and Gordeeva [12].

The chromatographic separation and behaviour of steroid hormones have been studied by isocratic [13,14] and gradient RP-HPLC [15]. Steroid hormones have been used in the characterization of normal-phase columns [16]. The use of HPLC in the determination of steroids has been reviewed by Makin and Heftmann [17]. Recently, the separation of ten estrogens by MECC was reported [18]. Some QSRR studies of steroids have also been conducted, relating the retention to functional groups of the steroid molecule [19–21]. Although QSRR studies by MECC are not as common as by RP-HPLC, some satisfactory correlations with topological indices have been reported [22].

Advanced statistical methods, such as factor and cluster analyses, have been used in classifying octadecyl stationary phases for RP-HPLC [23,24] and drug compounds into pharmacological groups [25,26]. These methods have previously been used in the evaluation of different molecular descriptors, mainly topological indices, and in the variable selection for quantitative structure–property studies of retinoids [27].

The aim of this work was to find structural parameters and characteristics related to the RP retention and electrophoretic migration of steroids in order to predict the retention/migration of steroid hormones on the basis of their molecular structure.

2. Experimental

The experimental data for congeneric steroid hormones (Fig. 1) are presented in Table 1. Retention data were obtained on an octadecyl stationary phase with a methanol–acetonitrile–58 mM NaH$_2$PO$_4$ monohydrate mobile phase containing 6 mM heptanesulfonic acid (4:22:74, v/v/v) (mobile phase A) for compounds **1–11**, **13–17**, **30** and **31** and with methanol–tetrahydrofuran–water (25.5:9.0:65.5, v/v/v) (mobile phase B) for compounds **3–15** and **18–32** [28]. MECC was conducted with 40 mM tetraborate–32 mM sodium dodecyl sulfate (SDS) at pH 9.3 for compounds **12–15**, **18–20** and **22** [29] and with a mixed-micellar system consisting of 49 mM 3-
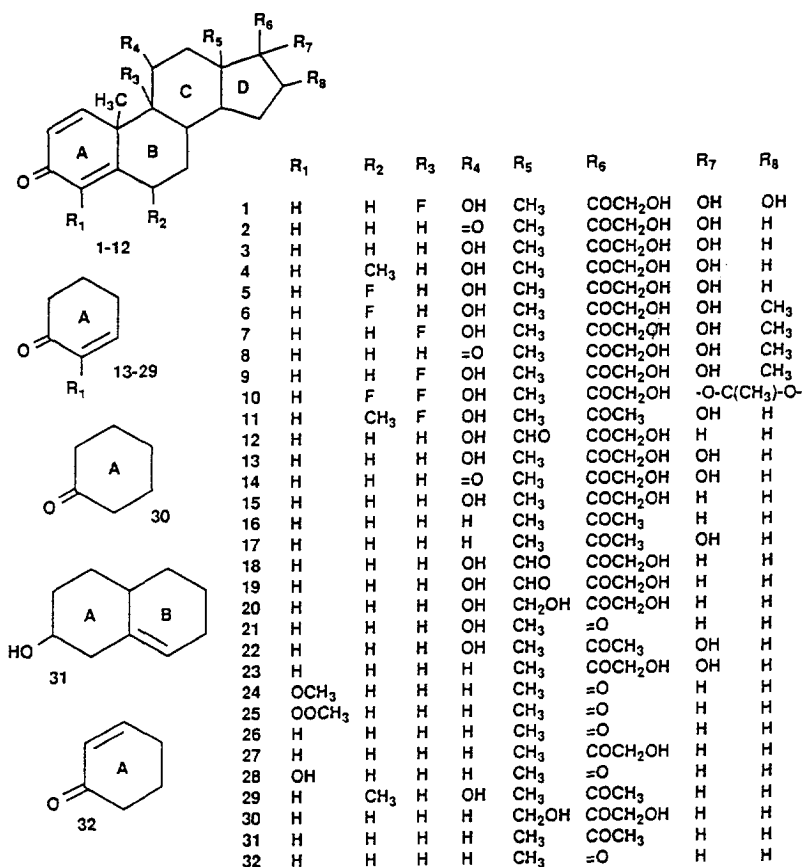
	R₁	R₂	R₃	R₄	R₅	R₆	R₇	R₈
1	H	H	F	OH	CH$_3$	COCH$_2$OH	OH	OH
2	H	H	H	=O	CH$_3$	COCH$_2$OH	OH	H
3	H	H	CH$_3$	OH	CH$_3$	COCH$_2$OH	OH	H
4	H	CH$_3$	H	OH	CH$_3$	COCH$_2$OH	OH	H
5	H	F	H	OH	CH$_3$	COCH$_2$OH	OH	H
6	H	F	H	OH	CH$_3$	COCH$_2$OH	OH	CH$_3$
7	H	H	F	OH	CH$_3$	COCH$_2$OH	OH	CH$_3$
8	H	H	H	=O	CH$_3$	COCH$_2$OH	OH	CH$_3$
9	H	H	F	OH	CH$_3$	COCH$_2$OH	OH	CH$_3$
10	H	F	F	OH	CH$_3$	COCH$_2$OH	-O-C(CH$_3$)-O-	
11	H	CH$_3$	F	OH	CH$_3$	COCH$_3$	OH	H
12	H	H	H	OH	CHO	COCH$_2$OH	H	H
13	H	H	H	OH	CH$_3$	COCH$_2$OH	OH	H
14	H	H	H	=O	CH$_3$	COCH$_2$OH	OH	H
15	H	H	H	OH	CH$_3$	COCH$_2$OH	H	H
16	H	H	H	H	CH$_3$	COCH$_3$	H	H
17	H	H	H	H	CH$_3$	COCH$_3$	OH	H
18	H	H	H	OH	CHO	COCH$_2$OH	H	H
19	H	H	H	OH	CHO	COCH$_2$OH	H	H
20	H	H	H	OH	CH$_2$OH	COCH$_2$OH	H	H
21	H	H	H	OH	CH$_3$	=O	H	H
22	H	H	H	OH	CH$_3$	COCH$_3$	OH	H
23	H	H	H	H	CH$_3$	COCH$_2$OH	OH	H
24	OCH$_3$	H	H	H	CH$_3$	=O	H	H
25	OOCH$_3$	H	H	H	CH$_3$	=O	H	H
26	H	H	H	H	CH$_3$	=O	H	H
27	H	H	H	H	CH$_3$	COCH$_2$OH	H	H
28	OH	H	H	H	CH$_3$	=O	H	H
29	H	CH$_3$	H	OH	CH$_3$	COCH$_3$	H	H
30	H	H	H	H	CH$_2$OH	COCH$_2$OH	H	H
31	H	H	H	H	CH$_3$	COCH$_3$	H	H
32	H	H	H	H	CH$_3$	=O	H	H

Fig. 1. Structures of the steroid hormones included in the study. Numbers **7** (betamethasone) and **9** (dexamethasone) are stereoisomers at R₈ and numbers **18** (17-isoaldosterone) and **19** (aldosterone) at R₆.

[(1,1 - dimethyl - 2 - hydroxyethyl)amino] - 2 - hydroxypropanesulfonic acid (AMPSO), 55 mM sodium cholate (SC) and 18 mM SDS at pH 9.0 for compounds **12–15**, **18**, **19** and **22** [30]. Retention/migration time was used as the retention parameter, because no void volume value was available. Detailed descriptions of the experimental conditions are given in the literature [28–30].

The values for the molecular descriptors were calculated using MOLCONN-X v1.0 (L.H. Hall, Quincy, MA, USA) and their statistical analysis was conducted using SYSTAT v5.1 (Systat Intelligent Software, USA). The programs were run on a Macintosh LC microcomputer. The following molecular descriptors were studied: molecular mass (M_r), connectivity indices ($^{1-10}\chi$, $^{1-10}\chi^\nu$, $^{3,4}\chi_{c/pc}$, $^{3,4}\chi^\nu_{c/pc}$) [9], graph complexity (*GrComp*), kappa indices ($^{1-3}\kappa$, $^{0-3}\kappa_\alpha$) [10], topological equivalence index (T) [31] and total topological equivalence indices (*TTS*, simple index; and *TTD*, valence index), sum of the intrinsic state values I (*sI*) [32], sum of delta-I values (*sdI*) [32], electrotopological and total electrotopological state indices (*S*, *TETS*) [32], Shannon information index (*Sh*) [33], normalized Bonchev–Trinajstic information indices [*nI(G)*, *nIW(G)*] [34], Platt's F number (*PF*) [35,36], Wiener number (W) and Wiener's P number [7]. The topological and electrotopological state values for atoms in the steroid skeleton were included in the study.

The statistical analyses included correlation, factor, cluster, stepwise regression and linear

Table 1
Retention/migration times used in the study (compounds numbered as in Fig. 1)

Compound	t_r (min)		t_M (min)	
	Mobile phase A	Mobile phase A	SDS	Mixed system
1	9.59	–	–	–
2	13.63	–	–	–
3	13.76	13.85	–	–
4	18.70	26.73	–	–
5	19.17	17.75	–	–
6	19.19	39.02	–	–
7	19.45	26.14	–	–
8	19.51	19.51	–	–
9	19.99	28.87	–	–
10	24.75	>58.0	–	–
11	26.15	47.41	–	–
12	–	4.88	11.3	11.09
13	14.37	14.05	13.2	13.79
14	14.73	10.34	12.7	11.78
15	20.80	22.63	15.2	16.26
16	25.66	–	–	–
17	29.26	–	–	–
18	–	5.46	11.7	11.21
19	–	6.24	–	12.11
20	–	7.80	8.5	–
21	–	14.44	–	–
22	–	25.17	13.9	14.66
23	–	26.53	–	–
24	–	29.26	–	–
25	–	29.66	–	–
26	–	32.97	–	–
27	–	45.85	–	–
28	–	50.53	–	–
29	–	>58.0	–	–
30	15.85	15.22	–	–
31	> 36.0	>58.0	–	–
32	–	56.58	–	–

For experimental details, see text.

regression analyses. The common factor model was used in the factor analysis. Cluster analysis was conducted using complete linkage and the Pearson's correlation coefficient as the distance metric. In multiple linear regression, intercorrelated variables were excluded from the calculations ($r > 0.800$). There were at least five data points for each dependent variable included in the equations.

3. Results and discussion

The mobile phase influenced the retention order of some of the steroids. The substitution had less effect on the shape in steroid molecules than on their hydrogen-bonding characteristics and interactions with the mobile phase [13].

In the factor analysis, in which the data included the RP-HPLC retention data, two factors obtained without rotation explained 86.6% of the variance (69.4% and 17.2%, respectively), with Varimax rotation 83.2% (66.2% and 17.1%, respectively). The factor pattern was similar in both analyses. Most of the descriptors and the retention in mobile phase B had loading values greater than 0.7 for the first factor. Exceptions include indices, such as the higher order valence connectivity indices, χ^ν, $^2\kappa_\alpha$ and the retention in mobile phase A. They had relatively small values for the first factor and values greater or about 0.7 for the second factor.

The cluster analysis classified the indices differently than did the factor analysis. The valence level connectivity indices were clearly divided into two separate groups; the first including the lower order indices, $^{0-3}\chi^\nu$, cluster and path/cluster indices, $^{3,4}\chi^\nu_{c/pc}$, and the molecular level connectivity indices, $^{0-10}\chi$. The higher order valence level indices formed a group distinctly separated from the other connectivity indices. The kappa indices were also divided into two groups. The first included $^{0-3}\kappa_\alpha$ and $^3\kappa$. Other kappa indices were like the molecular level connectivity indices. The retention data formed a third, separate group in the analysis.

The stepwise regression procedure produced no correlation for the RP-HPLC data, when all compounds were included ($R^2 = 0.06$ in mobile phase A and 0.21 in mobile phase B). For MECC the regression was reasonable (SDS, $r^2 = 0.764$, $n = 7$; mixed, $r^2 = 0.830$, $n = 7$). The retention behaviour could not be explained by the topological or electronic properties of single atoms. However, this was not surprising because of the high numbers of dependent factors in MECC compared with RP-HPLC.

The fragmental approach is based on an assumption of the additive nature of the retention

Table 2
Regression equations for the retention/migration data

Method	Conditions	Equation	r^2/R^2	n	F	s	r_{ij} [a]
MECC	SDS	$-17.54 \cdot {}^6\chi + 114.65$	0.76	7	16.19	1.14	–
	Mixed	$-16.70 \cdot {}^6\chi + 110.36$	0.83	7	24.38	0.89	–
RPLC, mobile phase A	Ring A:						
	2 C=C bonds	$-11.63 \cdot {}^3\kappa + 16.13x^4\chi^v - 21.60$	0.71	11	9.96	2.86	0.53
	1 C=C bond	$-11.53 \cdot {}^4\chi + 127.52$	0.89	5	24.63	2.50	–
	Ring C:						
	R_4 = H	$-43.88 \cdot {}^4\chi^v + 343.00$	1.00	3	1296.65	0.27	–
	R_4 = OH	$10.26 \cdot {}^2\chi^v - 55.55$	0.40	11	6.02	3.87	–
	Ring D:						
	R_6 = COCH$_2$OH	$14.65 \cdot {}^4\chi^v - 10.57x^7\chi^v - 57.69$	0.63	14	9.37	2.57	0.72
	R_6 = COCH$_2$OH, R_7 = OH	$8.96 \cdot {}^2\chi^v - 147.95x^4\chi_c^v - 34.27$	0.57	11	5.39	2.53	0.05
RPLC, mobile phase B	Ring A:						
	2 C=C bonds	$31.79 \cdot {}^3\chi_c - 82.17$	0.74	9	19.41	7.10	–
	1 C=C bond	$92.45 \cdot {}^{10}\chi - 41.35x^5\chi + 194.87$	0.81	14	23.20	6.78	0.79
	Ring C:						
	R_4 = H	$-20.40 \cdot {}^5\chi + 179.65$	0.57	8	7.97	9.82	–
	R_4 = OH	$51.06 \cdot {}^3\chi_c^v - 91.15$	0.77	15	42.57	6.32	–
	Ring D:						
	R_6 = COCH$_2$OH	$-103.18 \cdot {}^{10}\chi + 19.01x^2\chi^v + 3.54$	0.53	17	7.99	8.63	-0.32
	R_6 = COCH$_2$OH, R_7 = OH	$40.20 \cdot {}^3\chi^v - 615.71x^4\chi_c^v - 178.57$	0.81	10	15.26	4.27	0.46

[a] r_{ij} = Correlation between dependent variables.

process and has been successfully used by Davydov et al. [20] and Sadlej-Sosnowska [21] for steroid molecules. The fragment values obtained from previous studies produced poor correlations with our retention/migration parameters. When formulating a fragment model based on our data, the fragments having enough variance for a statistically sound regression analysis were too few to explain the dependence between the molecular structure and retention/migration.

Retention in RP-HPLC often depends on the hydrophobic nature of the analytes. Log P values for nine steroids measured in n-octanol–water [37] and diethyl ether–water [38] systems exhibited acceptable or good correlations with the retention/migration data ($r^2 = 0.750$–0.937).

The compounds were arranged into groups according to the substitution of ring A, C or D. The variance in the substitution of ring B was too small for statistical analysis. The stepwise regres-

sion procedure was conducted for each group separately, but not for factor or cluster analyses because of the small number of compounds remaining for each analysis. The regressions are presented in Table 2.

Steric factors have a great effect on the retention of steroid hormones, especially with the acetonitrile containing mobile phase. Retention in the tetrahydrofuran-containing mobile phase (B) was more accurately predicted by topological indices. Topological indices are insensitive to stereochemical differences between the molecules, unlike the experimental log P values, which were different, e.g., for betamethasone and dexamethasone. The predictive ability of regression equations was tested with compound **10**, for which no exact retention times were available (experimental value >58 min). The retention in mobile phase B could be predicted for compound **10** with compounds having the same substitution

in ring A $[t_R(\text{calc.}) = 71.6$ min] and ring C $[t_R(\text{calc.}) = 56.5$ min].

In conclusion, although the loading values of the indices and the retention data exhibited similar factor patterns, there was no close relationship or dependence between them according to the cluster and regression analyses. Topological indices can be used in the modelling and prediction of the retention/migration of steroid hormones when the solutes form a congeneric series and stereochemical properties do not govern the separation process.

References

[1] P.C. Sadek, P.W. Carr, R.M. Doherty, M.J. Kamlet, R.W. Taft and M.H. Abraham, Anal. Chem., 57 (1985) 2971.
[2] P.W. Carr, R.M. Doherty, M.J. Kamlet, R.W. Taft, W. Melander and C. Horváth, Anal. Chem., 58 (1986) 2674.
[3] T. Braumann, J. Chromatogr., 373 (1986) 191.
[4] R. Kaliszan, Adv. Chromatogr., 33 (1993) 147.
[5] S. Yang and M.G. Khaledi, Anal. Chem., 67 (1995) 499.
[6] C. Hansch, P.P. Maloney, T. Fujita and R.M. Muir, Nature, 194 (1962) 178.
[7] H. Wiener, J. Am. Chem. Soc., 69 (1947) 17.
[8] M. Randic, J. Am. Chem. Soc., 97 (1975) 6609.
[9] L.B. Kier and L.H. Hall, Molecular Connectivity Indices in Structure–Activity Analysis, Research Studies Press, Letchworth, 1986.
[10] L.B. Kier, Acta Pharm. Jugosl., 36 (1986) 171.
[11] K. Osmialovski and R. Kaliszan, Quant. Struct.–Act. Relat., 10 (1991) 125.
[12] A.R. Katritzky and E.V. Gordeeva, J. Chem. Inf. Comput. Sci., 33 (1993) 835.
[13] M. Olsson, L.C. Sander and S.A. Wise, J. Chromatogr., 537 (1991) 73.
[14] M. Amin, K. Harrington and R. von Wandruszka, Anal. Chem., 65 (1993) 2346.
[15] J.O. De Beer, J. Chromatogr., 489 (1989) 135.
[16] S. Hara and S. Oshishi, J. Liq. Chromatogr., 7 (1984) 69.
[17] H.C.J. Makin and E. Heftmann, Endocrinology, 30 (1988) 183.
[18] K.C. Chan, G.M. Muschik, H.I. Issaq and P.K. Siiteri, J. Chromatogr. A, 690 (1995) 149.
[19] J.M. DiBussolo and W.R. Nes, J. Chromatogr. Sci., 20 (1982) 193.
[20] V.Ya. Davydov, G.N. Filatova, E. Smolkova-Keulemansova and Yu. Zima, Chromatographia, 25 (1988) 1059.
[21] N. Sadlej-Sosnowska, J. Liq. Chromatogr., 17 (1994) 2319.
[22] P. Lukkari, H. Vuorela and M.-L. Riekkola, J. Chromatogr. A, 652 (1993) 451.
[23] M.F. Delaney, A.N. Papas and M.J. Walters, J. Chromatogr., 410 (1987) 31.
[24] B.A. Olsen and G.R. Sullivan, J. Chromatogr. A., 692 (1995) 147.
[25] M. Randic and P.C. Jurs, Quant. Struct.–Act. Relat., 8 (1989) 2249.
[26] R. Gami Yilinkou and R. Kaliszan, J. Chromatogr. 550 (1991) 573.
[27] M. Salo, S. Sarna and H. Vuorela, J. Pharm. Biomed. Anal., 12 (1994) 867.
[28] P. Volin, J. Chromatogr., 584 (1992) 147.
[29] H. Sirén, H. Haavio, J.H. Jumppanen, K.-M. Manninen, S. Wiedmer, P. Volin and M.-L. Riekkola, presented at the 6th International Symposium on HPCE, 31 January–4 February 1994, San Diego, CA; manuscript in preparation.
[30] S. Wiedmer, J.H. Jumppanen, H. Haavio and M.-L. Riekkola, Microcol. Sep., in press.
[31] L.H. Hall and L.B. Kier, Quant. Struct.–Act. Relat., 9 (1990) 115.
[32] L.B. Kier, L.H. Hall and J.W. Frazer, J. Math. Chem., 7 (1991) 229.
[33] C. Shannon and W. Weaver, Mathematical Theory of Communication, University of Illinois Press, Urbana, IL, 1949.
[34] D. Bonchev and N. Trinajstic, J. Chem. Phys., 67 (1977) 4517.
[35] J. Platt, J. Chem. Phys., 15 (1947) 419.
[36] J. Platt, J. Phys. Chem., 56 (1952) 328.
[37] H. Tomida, T. Yotsuynagi and K. Ikeda, Chem. Pharm. Bull., 26 (1978) 2832.
[38] G.L. Flynn, J. Pharm. Sci., 60 (1971) 345.