ELSEVIER

Journal of Chromatography A, 917 (2001) 361–366

JOURNAL OF CHROMATOGRAPHY A

www.elsevier.com/locate/chroma

Short communication

Evaluation of the lipophilicity of bile acids and their derivatives by thin-layer chromatography and principal component analysis

Costel Sârbu[a,*], Ksenija Kuhajda[b], Slavko Kevresan[c]

[a]*Babeş-Bolyai'' University, Faculty of Chemistry and Chemical Engineering, Arany Janos 11, 3400 Cluj-Napoca, Romania*
[b]*University of Novi Sad, Faculty of Sciences, Institute of Chemistry, Trg D. Obradovica 3, 21000 Novi Sad, Yugoslavia*
[c]*University of Novi Sad, Faculty of Agriculture, Trg D. Obradovica 3, 21000 Novi Sad, Yugoslavia*

Received 29 September 2000; received in revised form 28 February 2001; accepted 28 February 2001

Abstract

The lipophilic character of bile acids and their glyco- and tauro-conjugates was studied. The classical R_{Mo} values were measured by means of reversed thin-layer chromatography using a mixture of methanol–water as the solvent system and compared with the factors scores obtained by principal component analysis based also onto the TLC-retention data. The reliability of the factor scores values as lipophilic indices are shown by their high correlation with the classical R_{Mo} values. In addition, a better correlation was observed between scores corresponding to the first principal components and the partition coefficients (log P) of bile acids. Finally, the ''lipophilicity chart'' described by the first two components has the effect of separating compounds from each other most effectively from the congeneric aspect point of view.

Keywords: Lipophilicity; Principal component analysis; Bile acids

1. Introduction

The interest concerning the role and the exact mechanism of action of bile acids and their conjugates in living organisms has increased considerably in the recent decades. Now they are widely used for example as drugs to dissolve cholesterol gallstone [1] or as promoters of drug absorption by nonparenteral routes; in each of these case their interaction with biological membranes or the lipid environment is of crucial importance [2].

Modeling of the permeation of drugs or other bioactive substances through biological membranes is typically based on the standard measure of lipophilicity (hydrophobicity), namely the logarithm of the *n*-octanol–water partition coefficient, log P [3–8]. The log P scale is a valuable reference scale of lipohilicity. It has been chosen rather arbitrarily, however, and there is evidence that partition systems other than *n*-octanol–water may offer better models of the biological activities of chemical compounds [9–11].

It is considered that the dynamic, fast equilibrium processes of bioactive compound action have much in common with the processes that are the basis of chromatographic separations. The same basic intermolecular actions determine the behavior of chemical compounds in both biological and chromatographic environments. As a consequence, the chro-

*Corresponding author.
E-mail address: csarbu@chem.ubbcluj.ro (C. Sârbu).

0021-9673/01/$ – see front matter
PII: S0021-9673(01)00726-9

matographic approach has been quite successful in duplicating log *P* data derived by traditional ''shake-flask'' technique or other procedures [9–13]. Moreover, chromatographic techniques can readily produce large amounts of precise, reproducible, numerically expressed property parameters for large series of analytes.

The R_M values obtained from various types reversed-phase thin-layer chromatography are the most widely used chromatographic alternatives to the shake flask method. The use of RPTLC is based on the assumed linear relationship between the molecular parameter (1) and log *P*.

$$R_M = \log(1/R_F - 1) \quad (1)$$

The R_M value (related to the molecular lipophilicity), determined by using of RPTLC, generally, depends linearly on the concentration of the organic component of the mobile phase:

$$R_M = R_{Mo} + bC \quad (2)$$

where R_M values were calculated using Eq. (1) and *C* is the concentration of organic modifier.

At the same time, many researchers [8–13] have also studied the relationship between the intercept R_{Mo} and the slope *b*, in the TLC [Eq. (2)]. In each case a high correlation between the two regression parameters was observed. The most important aspect of the high linear correlation was considered as a possibility of finding congeneric classes within large groups of compounds. Recently, we addressed the question if a high linear correlation between the intercept and slope can be an objective criterion to find the real (natural) congeneric classes within the chemical compounds series. This was because we demonstrated using fuzzy regression [14] that always there is a high correlation between slope and intercept. This very interesting and useful conclusion was strongly supported in a recent paper [15]. Taking into account these results and other disadvantages of the regression method applied in this scope (logarithmic scale, extrapolation) we proposed the use of scores obtained applying principal component analysis as a support for the lipophilicity scale and the ''lipophilicity chart'' as a new concept [16–20].

The purpose of this paper is to investigate the feasibility of the scores, obtained by principal component analysis using RPTLC retention data, as a measure of lipophilicity in correlation with partition coefficient (log *P*) of bile acids and some of their glyco- and tauro-conjugates. In addition, the scatterplots of the scores onto plane described by the first two components appear to be very useful having the effect of separating compounds one from each other most effectively, obtaining in this way the ''congeneric lipophilicity chart'' of the series.

2. Experimental

The chromatographic behavior of the bile acids and their derivatives presented in Table 1 was studied on the C_{18} silica gel bonded plates. RPTLC plates (20×20 cm) were obtained as a gift from Macherey-Nagel (Düren, Germany). Methanol for chromatography was supplied from Reactivul (Bucharest, Romania). Three μl of each solution in methanol (1 mg/ml) was spotted to origin of the plate by hand. Chromatography was performed in a normal developing chamber at room temperature (~20°C), the developing distance being 8 cm. Methanol was used as the organic modifier of the mobile phase in the concentration range 50–75% (v/v) in steps of 5%, as the studied compounds differed considerably in their retention.

After being developed, the dried plates were sprayed with manganese chloride and heated to 100–120°C for 10–15 min. Colored zones appeared on a colorless background and fluorescent blue–green under UV lamp ($\lambda = 365$ nm).

3. Principal component analysis

Principal component analysis [16–21] has been performed on the retention data matrix by the use of a computer program discussed in [22]. It displays objects (bile acids and their derivatives) in a reduced space by finding a direction (first principal component) that best preserves the scatter of the observations (R_F values) in the multidimensional space described by the solvent systems. As usual, PCA gives both coordinates (scores) of the studied compounds and the loading of variables (solvents) on the principal components.

Table 1
Abbreviations and some characteristics of the bile acids and their conjugates

Compound	Abbreviation	Position and orientation of hydroxyls	Origin and some characteristics
Lithocholic acid	LC	3α	Human stool, toxic
Deoxycholic acid	DC	3α, 12α	Human bile
Chenodeoxycholic acid	CDC	3α, 7α	Human bile
Cholic acid	C	3α, 7α, 12α	Human bile
Ursodeoxycholic acid	UDC	3α, 7β	Polar bear bile
Hyocholic acid	HC	3α, 6α, 7α	Pig bile
Hyodeoxycholic acid	HDC	3α, 6α	Pig bile
Glycochenodeoxycholic acid sodium salt	GCDC	Glyco conjugate of CDC	Human bile
Taurodeoxycholic acid sodium salt	TDC	Tauro conjugate of DC	Human bile
Glycocholic acid sodium salt	GC	Glyco conjugate of C	Human bile
Glycodeoxycholic acid sodium salt	GDC	Glyco conjugate of DC	Human bile
Taurolithocholic acid sodium salt	TLC	Tauro conjugate of LC	Synthetic product
Taurochenodeoxycholic acid sodium salt	TCDC	Tauro conjugate of C DC	Human bile
Glycolithocholic acid	GLC	Glyco conjugate of LC	Synthetic product
Taurocholic acid sodium salt	TC	Tauro conjugate of C	Human bile

The results obtained from the initial chromatographic data using covariance matrix (without autoscaling) can be presented as usually in three panels, although typically there are only two. The first panel shows the table of data statistics; the second is the table of components and the third panel displays the eigenvectors associated with each of the components. Table 2 lists the eigenvalues of the covariance matrix, ordered from largest to smallest, the third column of this table shows the difference between each eigenvalues and the next smaller eigenvalue and the fourth column shows the proportion. These results suggest a significant two component model, which explained 99.75% of the total variance (information), considering only the eigenvalues higher than one. The first component explains 98.73% of the total variance, the second 1.02% and the third only 0.17%; the subsequent eigenvalues are just sampling noise.

It is interesting also to mention that when the significance of the component model retained was tested applying the Bartlett's statistics [22], testing the hypothesis that $(p-k)$ eigenvalues in variance–covariance matrix are equal, a model with three components was selected.

4. Results and discussion

The results of regression analysis using Eq. (2) are compiled in Table 3. The statistics obtained (see also Table 3) illustrate that the linear equation fits in a very good way the experimental data, the linear model explaining approximately 95–99% of the total variance (see R^2 values) in the majority of cases. Although TC-acid appears as an outlier in the regression analysis [Eq. (3)] however, a good correlation has been found as usually between the R_{Mo} and b values of Eq. (2):

Table 2
The eigenvalues and the ratios of the variance explained by the six components using covariance matrix

Component	Eigenvalue	Difference	Proportion, %	Cumulative, %
1	0.13731	0.13590	98.73	98.73
2	0.00141	0.00118	1.02	99.75
3	0.00023	0.00017	0.17	99.92
4	0.00006	0.00001	0.04	99.96
5	0.00005	0.00004	0.03	99.99
6	0.00001		0.01	100.00

Table 3
Regression data and scores on the first three principal components for the bile acids and their conjugates studied in this paper[a]

Compound	R_{Mo}	b	R	R^2	PC1	PC2	PC3	log P^1	log P^2
LC	5.53	−6.63	0.9987	0.9974	0.195	0.122	0.080		
DC	4.78	−6.01	0.9922	0.9845	0.346	0.179	0.111	3.50	2.65
CDC	5.01	−6.29	0.9959	0.9918	0.340	0.177	0.082	3.28	2.25
C	3.84	−4.99	0.9938	0.9876	0.533	0.217	0.105	2.02	1.10
UDC	3.65	−4.73	0.9972	0.9944	0.545	0.226	0.114	3.00	2.20
HC	3.66	−4.77	0.9957	0.9914	0.559	0.229	0.107	2.80	1.84
HDC	3.88	−4.98	0.9968	0.9936	0.488	0.208	0.110	3.08	2.28
GCDC	2.87	−3.96	0.9818	0.9639	0.819	0.237	0.088	2.12	0.45
TDC	2.30	−3.54	0.9765	0.9536	1.165	0.191	0.090		
GC	2.30	−3.45	0.9778	0.9561	1.102	0.186	0.090	1.65	−0.40
GDC	2.90	−4.05	0.9766	0.9537	0.850	0.247	0.087	2.25	0.80
TLC	2.12	−3.13	0.9182	0.8331	1.000	0.249	0.077		
TCDC	2.27	−3.59	0.9671	0.9353	1.249	0.194	0.082		
GLC	3.39	−4.55	0.9956	0.9912	0.674	0.246	0.119		
TC	2.29	−3.99	0.9798	0.9600	1.524	0.137	0.125		

[a] The log P values of protonated (log P_{HA})^1 and ionized (log P_A)^2 bile acids were taken from [23].

$$R_{Mo} = -1.234 - 1.009b;\ r = 0.9896 \quad (3)$$

This finding indicates that the intercept, R_{Mo}, (lipophilicity) and slope, b, for the majority of these compounds are highly correlated and, in that case, they might form a homologous series of compounds as has been suggested by some authors [8–13]. Moreover, a high correlation was obtained between R_{Mo} values and the scores of the same bile acids and their derivatives on the first principal component as it is described by the linear Eq. (4):

$$R_{Mo} = 5.379–2.624PC1;\ r = -0.9203 \quad (4)$$

It is very interesting to stress also that the correlation between scores and the log P values of protonated (log P_{HA}) and ionized (log P_A) bile acids respectively [23] is higher than the correlation between R_{Mo} values and the same log P values as it is indicated by the following linear regression in Eqs. (5)–(8):

$$\log P_{HA} = 0.366 + 0.620R_{Mo};\ r = 0.8559 \quad (5)$$

$$\log P_A = -2.255 + 1.017R_{Mo};\ r = 0.8752 \quad (6)$$

$$\log P_{HA} = 3.995 - 2.195PC1;\ r = 0.8675 \quad (7)$$

$$\log P_A = 3.824 - 3.807PC1;\ r = 0.9379 \quad (8)$$

Much more, considering the scores corresponding to the first two components the correlation can be improved slightly by using multiple regression [see Eqs. (9) and (10)]:

$$\log P_{HA} = 4.407 - 2.122PC1 - 2.168PC2;\ r = 0.8715 \quad (9)$$

$$\log P_A = 3.387 - 3.885PC1 + 2.290PC2;\ r = 0.9395 \quad (10)$$

On the basis of these findings and as can been seen from data provided in Table 3, the scores on the first principal component can replace efficiently the R_{Mo} values in the estimation experiments of the lipophilicity of these compounds directly from RPTLC data or via log *P*. In addition, as it is shown in Fig. 1, scores plots are very useful as a display tool for examining the relationships between compounds, looking for trends, groupings or outliers. Hence, graphing scores onto the plane described by PC1 and PC2 we obtain ''the congeneric lipophilicity chart''. It appears clearly that the compounds studied in this paper form practically three different linear congeneric classes, in a very good agreement with the selected eigenvectors when the Bartlett test was applied and with their chemical structure. However, considering only the eigenvalues higher than one a model with only two components might be retained. The strongest acid character of the conjugates of bile acids could be the most probable explanation for

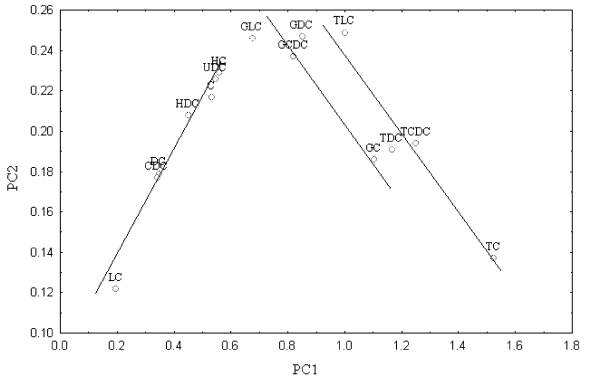

Fig. 1. Congeneric lipophilicity chart obtained by plotting scores corresponding to PC1 and PC2.

their negative slope, i.e. the bile acids have a positive increase of the R_F values with the concentration of methanol and their conjugates a negative one. The three dimensional graphically display of scores described by PC1, PC2, and PC3, respectively as is shown in Fig. 2 also supports the same conclusions. The position of each compound within the graphs is also in a good agreement with the position and orientation of hydroxyls and the presence of polar group –COOH and $-SO_3^-$ respectively. It is easy to observe that if the glyco-conjugates follow the order

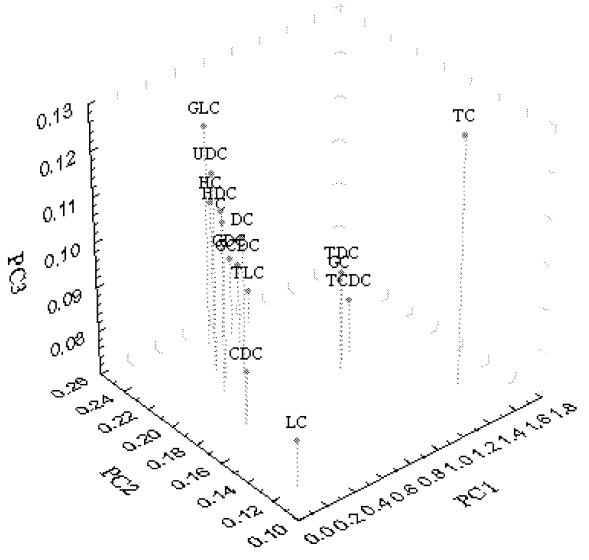

Fig. 2. 3-D scatterplot of scores corresponding to the first three principal components.

of the free bile acids the tauro-conjugates shows an opposite order.

5. Conclusions

The lipophilicity of some bile acids and their glyco- and tauro-conjugates has been determined by means of RPTLC on commercial RP-18 silica gel using a mixture of methanol–water as the solvent system. The factor scores values obtained from PCA using the initial TLC retention data enable more rational and objective estimation and comparison of lipophilicity. Scatterplot of the scores on to the plane described by the first two principal components will have the effect of separating compounds from each other more effectively into a ''congeneric lipophilicity chart'' and the scores described by the first principal component can be assessed as a relative scale of lipophilicity.

References

[1] J.E.F. Reynolds, Martindale — The Extra Pharmacopoeia, McMillan, 1996.
[2] G.S. Gordon, A.C. Moses, R.D. Silver, J.S. Flier, M.C. Carey, Proc. Natl. Acad. Sci. USA 82 (1985) 7419.
[3] A. Leo, C. Hansch, D. Elkins, Chem. Rev. 71 (1971) 525.
[4] C. Hansch, J.M. Clayton, J. Pharm. Sci. 62 (1973) 1.
[5] W.J. Lyman, W.F. Reehl, D.H. Rosenblatt, Handbook of Chemical Property Estmation Methods, McGraw-Hill, New York, 1982.
[6] R.F. Rekker, R. Mannhold, Calculation of Drug Lipophilicity, Weinheim, 1992.
[7] B. Wagner, H. Fischer, Lipophilicity Measurements by RPLC in Lipophilicity in drug Action and Toxicology, V. Pliska, B. Testa, H. Van de Waterbeemd, VCH, Basel, 1996.
[8] L.-G. Danielsson, Y.-H. Zhang, Trends Anal. Chem. 15 (1996) 188.
[9] R. Kaliszan, Trends Anal. Chem. 18 (1999) 401.
[10] M.C. Garcia Alvarez-Coque, J.R. Torres Lapasió, Trends Anal. Chem. 18 (1999) 533.
[11] T. Braumann, J. Chromatogr. 373 (1986) 191.
[12] T. Cserháti, J. Biochem. Biophys. Methods 27 (1993) 133.
[13] G.L. Biagi, A.M. Barbaro, A. Sapone, M. Recanatini, J. Chromatogr. A 662 (1994) 341.
[14] H. Pop, C. Sârbu, Anal. Chem. 68 (1996) 771.
[15] A.G. Gonzáles, M.A. Herrador, A.G. Asuero, Talanta 48 (1999) 729.
[16] S. Todor, C. Sârbu, Rev. Chim. (Bucharest) 49 (1998) 177.
[17] C. Sârbu, S. Todor, J. Planar Chromatogr. 11 (1998) 123.

[18] C. Sârbu, S. Todor, J. Chromatogr. 822 (1998) 263.
[19] M. Rusu, D. Rusu, C. Sârbu, Anal. Lett. 32 (1999) 2999.
[20] C. Sârbu, B. Malawska, J. Liq. Chrom. Relat. Technol. 23 (2000) 2143.
[21] S. Wold, K. Esbensen, P. Geladi, Chemom. Intell. Lab. Syst. 2 (1987) 37.
[22] A. Mackiewicz, W. Ratajczak, Comput. Geosci. 19 (1993) 303.
[23] A. Roda, A. Minutello, M.A. Angelloti, A. Fini, J. Lipid Res. 31 (1990) 1433.