Journal of Pharmaceutical and Biomedical Analysis 50 (2009) 613–621

Contents lists available at ScienceDirect

Journal of Pharmaceutical and Biomedical Analysis

journal homepage: www.elsevier.com/locate/jpba

Derived chromatographic indices as effective tools to study the self-aggregation process of bile acids

Benedetto Natalini*, Roccaldo Sardella, Emidio Camaioni, Antonio Macchiarulo, Antimo Gioiello, Giovanni Carbone, Roberto Pellicciari

Università degli Studi di Perugia, Dipartimento di Chimica e Tecnologia del Farmaco, Via del Liceo, 1, 06123 Perugia, Italy

ARTICLE INFO

Article history:
Received 21 July 2008
Accepted 15 October 2008
Available online 6 November 2008

Keywords:
Critical micellar concentration
Hydrophobic/hydrophilic balance
Molecular field analysis
Chromatographic indices
Multiple linear regression analysis

ABSTRACT

Reversed-phase high-performance liquid chromatography (RP-HPLC) is more and more adopted as a fruitful tool for the estimation of several physico-chemical properties of diverse classes of organic compounds. In this frame, derived chromatographic indices have been proposed as effective parameters to measure the lipophilicity (log *P* or log *D*) of compounds. Instead, a limited attention has been directed towards the chromatographic evaluation of the Critical Micellar Concentration (CMC), one of the most important parameters employed to study the bile acid physico-chemical profile. We have recently reported on the effectiveness of the derived chromatographic index φ_0 for the study of the self-aggregation process of bile acids. Here we show that this index is independent upon the adopted chromatographic environment so as to be instrumental for the evaluation of the hydrophobic/hydrophilic balance of bile acids. Molecular modelling studies have also been undertaken with the aim of rationalizing the experimental findings.

1. Introduction

Over the last three decades, high-performance liquid chromatography in the reversed-phase separation mode (RP-HPLC) has been established as a powerful tool for the estimation of several physico-chemical properties of diverse classes of organic compounds [1–6]. Due to the relevant implications of molecular lipophilicity in drug action (adsorption, blood–brain distribution, drug-receptor interaction, etc.), a particular attention has been addressed to the octanol–water partition coefficient ($\log P_{o/w}$ or $\log D_{o/w}$ in the case of ionisable compounds) determination through RP-HPLC measurements [1–6]. Indeed, partitioning between an aqueous/organic mobile phase and a reversed-phase material can be exploited in a direct quantification of lipophilicity.

Oppositely, only a scanty interest has been devoted to the chromatographic Critical Micellar Concentration (CMC) estimation [7,8]. The RP-HPLC mobility of surfactants and their propensity to associate each other into micelles should reflect the free energy of the transfer of detergent molecules from an aqueous to a hydrophobic phase [7,8]. As a general rule, the chromatographic approach provides several practical and operative advantages: speed, reproducibility, insensitivity to impurities, reduced amount of sample, easy automation, etc. [4,9,10]. According to the following Collander-type equation (Eq. (1)) [11], the retention factor (k) of a selected compound in a reversed-phase environment is linearly related to its partition coefficient (P) and hence to its own lipophilicity:

$$\log P_{o/w} = a \log k + b \quad (1)$$

Eq. (1) has also been adapted for surface active compounds considering their retention as a measure of their self-aggregation ability [7]. However, instead of the retention factor determined at a single eluent composition, the one attainable with a 'virtual' totally water-containing mobile phase (k_w) is generally accepted as more fruitful for lipophilicity assessments [1]. In a completely aqueous mobile phase, the solvophobic effect relates to the specific hydrophobicity extent of each analyte [10,12]. Retention factors from a pure water-containing mobile phase can be effortlessly derived by utilizing the 'linear solvent strength (LSS) model' proposed by Snyder in the late 70s [13]. This simple model bases on a linear relationship between the organic content in the mobile phase (solvent strength) and the retention factor (Eq. (2))

$$\log k = \log k_w + S\varphi \quad (2)$$

where S (which is dependent on the chromatographic system and solute) is the slope of the equation and φ is the fraction of the organic solvent in the eluent. In accordance with the above equation (Eq. (2)), k values are measured for each compound with eluents containing different amounts of the organic modifier. Consequently, regression analysis of the linear portion in the $\log k$ vs φ plot allows the extrapolation of the correspondent k_w parameter.

* Corresponding author. Tel.: +39 075 5855131 fax: +39 075 5855161.
E-mail address: natalini@chimfarm.unipg.it (B. Natalini).

0731-7085/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.jpba.2008.10.035

The engagement of the log k_w value limits several experimental disadvantages which are often encountered when physico-chemical properties are estimated from a single mobile phase composition. First, the unacceptable long analysis time often required for eluting particularly hydrophobic compounds within a mixture and, consequently, the possible occurrence of peak inversion when a gradient-mode elution is adopted to overcome this drawback. This could direct to uncorrected assignments in the relative hydrophobicity among the tested compounds.

In the log k_w vs φ plot, a lost of linearity is observed when $\varphi < 0.2$ and $\varphi > 0.8$ [1,2]. In these two regions, many of the properties of the mobile phase and the stationary phase which result to be significant for retention in a RP-LC system do not undergo a linear variation. It is commonly assumed that the curvature (concave) observed in the presence of low percentage of organic modifier ($\varphi < 0.2$) is deputable to some change occurring in the stationary phase solvation, thus reflecting in a different retention mechanism. In this case, the non-linear profile observed is generally accepted being associated with some pore exclusion effect within the stationary phase particles [2]. Conversely, at higher organic modifier concentrations ($\varphi > 0.8$) the concave or convex deviation from linearity can be partially attributed to silanophilic interactions [2,4,9]. Several equations taking account of the above lost of linearity have been proposed. Eq. (3), expressing the Shoenmakers solubility parameter model [14], well fits the generally observed profile when the whole organic modifier range is considered:

$$\log k = A\varphi + B\varphi^2 + E\sqrt{\varphi} + \log k_w \quad (3)$$

While the $B\varphi^2$ term accounts for the curvature in the correspondence of high organic modifier contents, $E\sqrt{\varphi}$ expresses the lack of linearity when water-rich eluents are used. However, the extrapolated log k_w value for a selected compound can result different when different organic modifiers are used [1,9]. Accordingly, it has been found that when MeCN is used instead of MeOH as the major organic component in the eluent, a most pronounced curvature in the correspondence of water-rich eluents occurs, thus amplifying the uncertainty in the log k_w value determination [1,9,14,15]. The above observation can be explained upon the fact that with MeCN, the two parameters S and log k_w vary independently from the analyte structure. Conversely, a high correlation between these results occurs when MeOH is employed [14,16,17]. The adsorption isotherms reported for RP-HPLC systems using water–methanol mixtures indicate that the compositions of alkyl-bonded stationary phases remain fairly constant in the range $0.2 < \varphi < 0.8$ [18,19]. The above pore exclusion effect, also observed in the non-linear portion, is currently attributed to the presence of not-alkylated silanols onto the packing material.

In the last years, many efforts have been spent with the aim of reducing the free silanols interference in the overall chromatographic process. Their inactivation can be achieved in different ways. The most promising results have been reached either introducing a polar group (generally ureas, carbamates, amides) within the hydrocarbon chain or through a polar-endcapping with short carbon chains (3 or 4 carbon atoms) with a polar end [2,20,21]. One of the most undertaken strategies to limit the deleterious effect from the free silanols is represented by the addition of a small amount of an amine (usually diethylamine, triethylamine or decylamine) to the mobile phase. Such minor eluent component, which acts as a masking agent, also improves the peak symmetry provoked by the otherwise occurring peak tailing phenomenon [22,23].

Several authors [1,3,5,24] have reported on the influence of 1-octanol in the mobile phase as a powerful artifice for significantly increasing the correlation between the derived log k_w parameters and the log P (either experimental or calculated) values. The benefit rising from the addition of the long-chain alcohol to the mobile phase can dwell in a most faithful reproduction of the intermolecular interactions experienced by a solute in a classical 1-octanol/water shake-flask determination [24]. A relatively new derived chromatograhic index, φ_0, is increasingly proposed as an effective parameter to measure the hydrophobicity of compounds [9,10,25,26]. Recently we have suggested this index as a valid tool for the bile acid (BA) CMC estimation [8]. To the best of our knowledge, this has been the first attempt in correlating the derived chromatographic index φ_0 with the hydrophobic/hydrophilic balance (HHB) of BAs. Moreover, owing to their peculiar 3D-structure, BAs do not easily lend themselves to classical HHB calculations as proposed by Davies [27]. In this frame and in accordance with what already stated by Valkó [9], the expression 'hydrophobicity' has been favoured to 'lipophilicity' since a molecular property rather than a strict distribution between two phases is here considered. The φ_0 value for a selected compound represents the percentage (in volume) of organic modifier which is necessary to achieve its equal distribution between the two phases of the chromatographic medium [9,25,26]. Consequently, this corresponds to a mobile phase composition providing a retention time which is twice the dead time (i.e., log $k = 0$). The φ_0 index can be calculated from Eq. (4) as

$$\varphi_0 = -\frac{\log k_w}{S} \quad (4)$$

where S has a reciprocal unit to afford a dimensionless φ_0 value. Most properly, the log k_w factor should be replaced with 'intercept' as it represents a definite point in the log k vs φ plot. For both particularly hydrophobic (when log $k > 0$ with 100% organic phase) and markedly hydrophilic (when log $k < 0$ with a totally aqueous mobile phase) compounds, the φ_0 value is practically unattainable in these extreme conditions. Analogously to the log k_w determination, the method requires retention factor measurement (log k values) with at least three mobile phase compositions (preferably as closer as the concentration region where log $k = 0$). In principle, compared to the most frequently proposed log k_w parameter, φ_0 exempts from the uncertainty characterizing the former, so that the latter has to be reasonably expected as a more representative hydrophobicity index. Being φ_0 dependent only on the distribution constant, it results in a very close correlation with the organic modifier type and percentage as well as to the system temperature. Additionally, also the eluent pH plays a crucial role for ionisable compounds [10].

Herein, we provide further evidences on the effectiveness of the φ_0 index as a tool for estimating the bile acids HHB. Indeed, its independence upon the adopted chromatographic system has been evaluated and demonstrated. Moreover, molecular modelling studies have been undertaken with the aim of rationalizing the experimental findings.

2. Experimental

2.1. Materials and reagents

All reagents were of analytical grade. Methanol (MeOH), 1-octanol, triethylamine and ethanol were purchased from Sigma–Aldrich (Milano, Italy). While CA, DCA, UDCA, HDCA, HCA, UCA, CDCA and NCA were generously gifted by Erregierre (Bergamo, Italy), conjugated bile acids such as GCA, TCA, GUDCA, TUDCA, GCDCA, TCDCA, GDCA, TDCA were purchased from Calbiochem (La Jolla, CA, USA). All the remaining unconjugated BAs were synthesized in our laboratory. NaH_2PO_4 and NaOH were purchased from Carlo Erba (Milano, Italy). HPLC-grade water was obtained from a tandem Milli-Ro/Milli-Q apparatus (Millipore, Bedford, MA, USA). The mobile phase was prepared by dissolving NaH_2PO_4 in HPLC-grade water; the latter solution was adjusted to pH 7.4 with NaOH.

The solutions were then filtered through a 0.45 μm Millipore filter (Bedford, MA, USA) and degassed with 10 min sonication; finally, the pH was adjusted again and the desired amount of organic modifier added. BAs were prepared in approximate concentrations between 1.0 and 1.5 mg mL^{-1} (always below the corresponding aqueous CMC value) in filtered mobile phase components and sonicated until completely dissolved.

2.2. Instrumentations

The HPLC analytical-scale experiments were carried out on a Shimadzu (Kyoto, Japan) LC-Workstation Class LC-10 equipped with a CBM-10A system controller, two LC-10AD high pressure binary gradient delivery systems, a SPD-10A variable-wavelength UV–vis detector and a Rheodyne 7725i injector (Rheodyne, Cotati, CA, USA) with a 20 μL stainless steel loop. The UV detection wavelength was set at 205 nm. An Ultra Aqueous C18 (Restek, Bellefonte, PA, USA) 250 mm × 4.6 mm i.d., 5 μm, 100 Å analytical column was used after previous conditioning by passing through the column the selected mobile phase for at least 30 min. All the analyses in the presence of 1-octanol and triethylamine as the minor components in the mobile phase were made on a LiChrospher 100 RP-18 (Merck, Darmstadt, Germany) 250 mm × 4.0 mm i.d., 5 μm, 100 Å analytical column. This was used after previous conditioning by passing through the column the selected mobile phase for at least 30 min.

The flow rate was 1.0 mL min^{-1} while the $NaNO_2$ injection peak was used for a completely unretained marker in all analyses. The retention factors were measured as $k = (t - t_0)/t_0$, where t_0 is the retention time of $NaNO_2$ and t is the retention time of the chromatographed bile acid. Retention factor values are the average result of three consecutive injections of each compound.

2.3. Conductimetric analyses

Conductimetric measurements were carried out as reported elsewhere [8].

2.4. Molecular modelling methods

The 20 bile acids were built using the builder module of Cerius-2 (Accelrys, San Diego, CA). The geometry of each compound was optimized using the semiempirical method AM1 as implemented in MOPAC v.6. Partial atomic charges were calculated using the ESP method. All the molecules were mutually aligned on their global minimum conformations using a consensus strategy and the subgraph matching routine as implemented in the Align module of Cerius-2. The resulting alignment was manually refined to maximize the overlapping volume of the compounds.

MFA studies were performed with the QSAR module of Cerius-2. Two molecular interaction fields were generated using the methyl probe (M) and the hydroxyl probe (O), encoding respectively for the steric and hydrogen bonding interactions. The interaction energies of the molecular fields were sampled at each point of a regularly grid built around the aligned compounds with a step size of 1 Å. The grid box was defined in order to occupy a volume of 2 Å^3 bigger then the global volume occupied by all the bile acids. The multiple linear regression analysis was carried out using the forward stepwise variable selection method. Briefly, this method is based on the F value of the independent variables. The F value allows the appraisal of the relative importance of the variable among the original pool of descriptors to explain the dependent variable (chromatographic indices). In particular, a variable is entered into the regression equation if its F value is greater than a defined entry value ($F_{min} = 0.05$), thereafter the F values of the rest variables in the model are recalculated and those with F values less than the removal value ($F_{max} = 0.10$) are removed. This procedure is continued until the F values of the remaining variables out of the regression equation are all less than the defined F_{min}. The predictive ability of the regression equation was assessed using a leave-one-out (LOO) cross-validation procedure and measured through the r^2_{xv} value.

Further regression and the statistical analyses were performed with Microsoft Excel® 2002 by using the REG.LIN function.

3. Results and discussion

As previously stated, we have recently proposed the use of the derived chromatographic index φ_0 as an effective tool for estimating the bile acid HHB [8], particularly profitable when small amount of synthetic or rare species is available. Since the attitude of a surfactant to associate into micelles strictly depends on his HHB, the availability of comparative values can be of practical advantage for the physico-chemical characterization of these compounds.

With the aim of further evaluating the validity of such an index as a measure of the above intrinsic structural property of each molecule, we tried to modify the chromatographic environment. Accordingly, when the 1-octanol is present both in the mobile phase (as the minor alcoholic component) and, consequently, hydrophobically adsorbed onto the reversed-phase carbon chains, adjunctive interactions are promoted. Triethylamine has been utilized with the aim of limiting the secondary, unproductive and perturbing interactions with underivatized silanols which are unavoidably present onto a conventional RP-18 stationary phase.

Conversely, no additive to the eluent was used when the polar Base Deactivated RP-18 column (RP-18-BD) was employed. While the chromatographic system with the RP-18-BD column will be subsequently referred to as the *System A*, the one composed of the common RP-18 stationary phase combined with an eluent carrying both 1-octanol and triethylamine will be labelled as the *System B*. The set of compounds reported in Table 1 and represented by the general structure in Fig. 1 covers a rather wide range of hydrophobicity values (3 mM < CMC < 60 mM) [8] and chemical diversity as well. Indeed, both bi- (CDCA, DCA, HDCA, UDCA, 23MUDCA [28], 6FCDCA [29], 6MCDCA, 6ECDCA [30] and tri-hydroxylated (CA, HCA, UCA) as well as glycine- (GCA, GCDCA, GCDA, GUDCA) and taurine-conjugated (TCA, TCDCA, TDCA, TUDCA) species have been included in this study. Moreover, the investigated set of compounds also contains the 3α-hydroxy-7-keto BA (NCA).

3.1. Relationship between φ_0 and pCMC

As well known, in an aqueous environment and above a specific concentration value (namely the CMC), bile salts aggregate to form micelles. From a structural point of view, bile salts are characterized by a rigid steroidal backbone bearing hydroxyl groups on the concave α-face, and methyl and methylene carbons on the convex β-face. This arrangement is responsible for their characteristic 'facial amphiphilicity' enabling them to aggregate with each other [8,31]. Primary bile salt molecule aggregation is hence mainly driven by hydrophobic β-faces association (the so called *back-to-back* or *facial* association). Secondary micelles are then formed at higher sample concentrations through hydrogen bonding interactions [31]. Table 2 shows the molar CMC values and the measured log k_w, $|S|$ and φ_0 chromatographic indices for all the investigated compounds and with both adopted systems (*System A* and *B*).

The previously discussed benefit gained from the use of MeOH instead of MeCN as the organic modifier in the eluent, prompted us to select the alcoholic component for all analyses.

Table 1
Analytes investigated in this work.

BA	Trivial name	R_1	R_2	R_3	R_4	R_5
CA	Cholic acid	H	α–OH	α–OH	H	OH
CDCA	Chenodeoxycholic acid	H	α–OH	H	H	OH
DCA	Deoxycholic acid	H	H	α–OH	H	OH
HCA	Hyocholic acid	α–OH	α–OH	H	H	OH
HDCA	Hyodeoxycholic acid	α–OH	H	H	H	OH
NCA	Nutriacholic acid	H	=O	H	H	OH
UCA	Ursocholic acid	H	β–(H	α–OH	H	OH
UDCA	Ursodeoxycholic acid	H	β–(H	H	H	OH
23MUDCA	23Me-ursodeoxycholic acid	H	β–(H	H	CH_3	OH
6FCDCA	6αF-chenodeoxycholic acid	α-F	α–OH	H	H	OH
6MCDCA	6αMe-chenodeoxycholic acid	α–$CH_($	α–OH	H	H	OH
6ECDCA	6αEt-chenodeoxycholic acid	α–$CH_($$CH_($	α–OH	H	H	OH
GCA	Glycocholic acid	H	α–OH	α–OH	H	$NHCH_2CO_2H$
GCDCA	Glycochenodeoxycholic acid	H	α–OH	H	H	$NHCH_2CO_2H$
GDCA	Glycodeoxycholic acid	H	H	α–OH	H	$NHCH_2CO_2H$
GUCDA	Glycoursodeoxycholic acid	H	β–(H	H	H	$NHCH_2CO_2H$
TCA	Taurocholic acid	H	α–OH	α–OH	H	$NH(CH_2)_2SO_3H$
TCDCA	Taurochenodeoxycholic acid	H	α–OH	H	H	$NH(CH_2)_2SO_3H$
TDCA	Taurodeoxycholic acid	H	H	α–OH	H	$NH(CH_2)_2SO_3H$
TUDCA	Tauroursodeoxycholic acid	H	β–(H	H	H	$NH(CH_2)_2SO_3H$

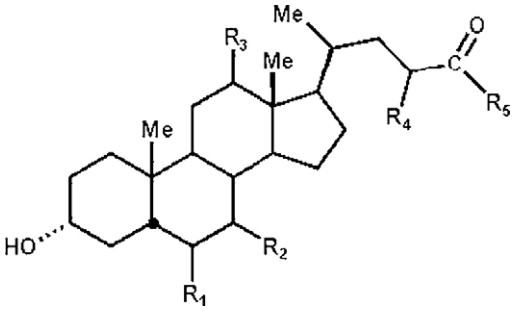

Fig. 1. General structure of the bile acids investigated in this work.

As from Eq. (4), the φ_0 values of all samples were determined by measuring their log k values in the presence of three different mixtures of methanol (50, 60 or 70%) and phosphate buffer (0.01 M NaH_2PO_4, pH 7.4). Differently from *System A*, *System B* is characterized by the presence of 1-octanol (0.25%, v/v, with respect to the volume of methanol) and triethylamine (0.15%, v/v, with respect to the total volume of mobile phase). Only after its preliminary saturation with 1-octanol, the phosphate buffer was added to the 1-octanol containing organic modifier. Triethylamine was then added to the obtained solution thus achieving the mobile phase to be employed.

Experimental CMC determinations were carried out by means of the non-invasive conductimetric technique [8]. For both studied chromatographic systems, φ_0 and pCMC values revealed to be highly correlated when considering separately the two classes of congeners (conjugated, C, and non-conjugated, N) (Table 3, Eqs. (5)–(8).

This finding holds a noticeable conceptual relevance since further interaction modes between analyte and stationary phase are furnished in the *System B* by the hydrophobically coated 1-octanol onto the RP material. Indeed, while exclusively hydrophobic interactions (through the β-face) are realized in the *System A*, some H-bond contact can plausibly take place by the BA hydroxyl groups (occupying the α-face) in the *System B*. Consequently, for bile salts, φ_0 has evidently proved to be independent on the chromatographic milieu but strictly related to structural peculiarities of each molecule (Table 3). Accordingly, the above chromatographic index reveals to be highly instrumental in estimating the HHB values for

Table 2
Experimental CMC values and derived chromatographic indices (log k_w, $|S|$ and φ_0) for both adopted systems.

BA	CMC (M)	pCMC	*System A*			*System B*		
			log k_w	$\|S\|$	φ_0	log k_w	$\|S\|$	φ_0
CA	0.012	1.92	3.69	5.15	0.72	3.45	4.79	0.72
CDCA	0.009	2.05	4.26	5.54	0.77	4.75	6.22	0.76
DCA	0.010	2.00	4.37	5.66	0.77	5.15	6.74	0.76
HCA	0.016	1.80	3.53	5.16	0.68	3.25	4.78	0.68
HDCA	0.015	1.82	3.68	5.20	0.71	3.48	5.00	0.70
NCA	0.022	1.66	3.39	4.89	0.69	3.25	4.71	0.69
UCA	0.062	1.21	2.54	4.36	0.58	2.46	4.20	0.59
UDCA	0.019	1.72	3.28	4.77	0.69	3.26	4.77	0.68
23MUDCA	0.024	1.62	3.21	4.77	0.67	3.11	4.75	0.65
6FCDCA	0.011	1.96	3.73	5.14	0.73	3.58	4.91	0.73
6MCDCA	0.007	2.15	4.72	5.89	0.80	2.71	3.10	0.87
6ECDCA	0.003	2.52	4.87	5.89	0.83	3.67	4.15	0.88
GCA	0.009	2.05	3.44	4.89	0.70	3.72	5.24	0.71
GCDCA	0.008	2.12	4.03	5.27	0.76	4.55	6.01	0.76
GDCA	0.007	2.15	4.02	5.19	0.78	4.61	5.98	0.77
GUCDA	0.010	2.00	3.09	4.66	0.66	3.56	5.32	0.67
TCA	0.009	2.05	3.51	5.05	0.69	3.63	5.12	0.71
TCDCA	0.008	2.10	4.00	5.31	0.75	4.59	6.10	0.75
TDCA	0.008	2.12	4.03	5.29	0.76	4.79	6.29	0.76
TUDCA	0.009	2.05	3.15	4.78	0.66	3.49	5.23	0.67

Table 3
pCMC/φ_0 relationships according to the general equation pCMC = $a + b\varphi_0$, and statistical parameters.

	Regression equations	r	r^2	n	F	Equation
N (*A*)	pCMC = 4.68 (±0.36) φ_0 − 1.50 (±0.26)	0.97	0.94	12	98.26	(5)
C (*A*)	pCMC = 1.06 (±0.13) φ_0 + 1.32 (±0.09)	0.95	0.90	8	65.37	(6)
N (*B*)	pCMC = 3.61 (±0.41) φ_0 − 0.7 (±0.30)	0.94	0.88	12	78.83	(7)
C (*B*)	pCMC = 1.16 (±0.17) φ_0 + 1.24 (±0.13)	0.94	0.88	8	43.79	(8)

A: *System A*; *B*: *System B*; N: unconjugated bile acids; C: conjugated bile acids; r: correlation coefficient; n: number of objects; F: Fisher test.

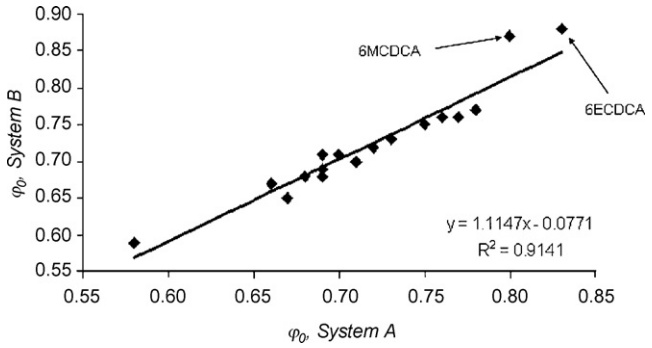

Fig. 2. Plot of the φ_0 values obtained with the *System A* vs those obtained with the *System B*.

bile acids. Moreover, its independence upon the adopted experimental methodology renders it markedly reproducible and hence particularly suitable for interlaboratory studies.

It is worth nothing that the *F* value of the Fisher test diminishes for both classes of congeners by passing from *System A* to *B*. This means that a lower significance level of the model occurs when the two additives in the mobile phase are included. A plot of φ_0 values from *System A* against those from *System B* for all the investigated analytes (Fig. 2) displays the presence of two outliers (6MCDCA and 6ECDCA). An explanation of this dissimilar behaviour is tentatively given in the following chapter. As above reported, an approximately unitary slope is clearly confirmed.

The effectiveness of the φ_0 index as a tool for the bile acid HHB estimation is furthermore emphasized when pCMC and log k_w values are put into correlation for both systems and classes of congeners (Table 4, Eqs. (9)–(12). Accordingly, appreciable correlations in both systems are maintained for conjugated (C) compounds (Table 4, Eqs. (10) and (12), while for non-conjugated (N) species the correlation resulted satisfactory in *System A* (Eq. (9) and very poor in *System B* (Table 4, Eq. (11). Conversely, in *System B*, identical and highly significant correlations are experienced by the two classes when the φ_0 index is engaged (Table 3, Eqs. (7) and (8). Despite of log k_w, φ_0 has revealed to be dependent only on the structural peculiarity of each molecule.

3.2. Relationship between S and log k_w

Only a limited attention has been directed towards the physicochemical meaning of the derived parameter *S*. From a mathematical point of view, the *S* index can be considered as the log *k* variation which occurs with a 1% modification in the eluent organic phase concentration. Formally, this can be written as follows (Eq. (13))

$$S = \log K_{x+1} - \log K_x \quad (13)$$

where *K* represents the distribution constant which is a measure of the equilibrium distribution of a specific compound between the stationary and the mobile phase, while x and $x+1$ symbolize the x% and $(x+1)$% fractions of organic modifier, respectively [10]. Moreover, it is generally accepted to be related to the specific hydrophobic surface area of each molecule and reflects its interaction with the solvent. Accordingly, this parameter is strictly related to the number of mobile phase solvent molecules in the solute solvation sphere and hence is a function of its own size as well as type, number and orientation of polar substituents on its backbone [1,3]. Consequently, a linear relationship between *S* and log k_w values has to be expected when the solute/stationary phase interaction is governed by the same mechanism as the analyte solvation. Since all the slopes are negative, the relative *S* values have been reported as absolute numbers. A |*S*| vs log k_w plot can be useful as a tool for assessing how uniform is the retention mechanism for a selected

Table 4
pCMC/log k_w relationships according to the general equation pCMC = $a + b$ log k_w and statistical parameters.

Data set	Regression equations	r	r^2	n	F	Equation
N (*A*)	pCMC = 0.46 (±0.05) log k_w + 0.15 (±0.18)	0.95	0.90	12	98.26	(9)
C (*A*)	pCMC = 0.12 (±0.02) log k_w + 1.64 (±0.07)	0.92	0.85	8	37.12	(10)
N (*B*)	pCMC = 0.19 (±0.06) log k_w + 1.20 (±0.21)	0.45	0.20	12	2.57	(11)
C (*B*)	pCMC = 0.87 (±0.02) log k_w − 1.72 (±0.07)	0.91	0.83	8	29.85	(12)

A: *System A*; *B*: *System B*; N: unconjugated bile acids; C: conjugated bile acids; r: correlation coefficient; n: number of objects; F: Fisher test.

Table 5
|*S*|/log k_w relationships according to the general equation |*S*| = $a + b$ log k_w and statistical parameters.

Data set	Regression equations	r	r^2	n	F	Equation
N + C (*A*)	|*S*| = 0.68 (±0.03) log k_w + 2.60 (±0.11)	0.98	0.96	20	553	(13)
N + C (*B*)	|*S*| = 1.07 (±0.43) log k_w + 1.14 (±0.11)	0.91	0.83	20	90	(14)
N + C (*B*/no out)	|*S*| = 0.94 (±0.04) log k_w + 1.74 (±0.16)	0.99	0.98	18	530	(15)
N (*A*)	|*S*| = 0.70 (±0.03) log k_w + 2.58 (±0.11)	0.99	0.98	12	584	(16)
N (*B*)	|*S*| = 1.08 (±0.18) log k_w + 1.04 (±0.64)	0.89	0.79	12	36	(17)
N (*B*/no out)	|*S*| = 0.96 (±0.05) log k_w + 1.64 (±0.20)	0.99	0.98	10	312	(18)
C (*A*)	|*S*| = 0.60 (±0.06) log k_w + 2.86 (±0.20)	0.98	0.96	8	118	(19)
C (*B*)	|*S*| = 0.83 (±0.07) log k_w + 2.25 (±0.29)	0.98	0.96	8	139	(20)

A: *System A*; *B*: *System B*; N: unconjugated bile acids; C: conjugated bile acids; r: correlation coefficient; n: number of objects; F: Fisher test.

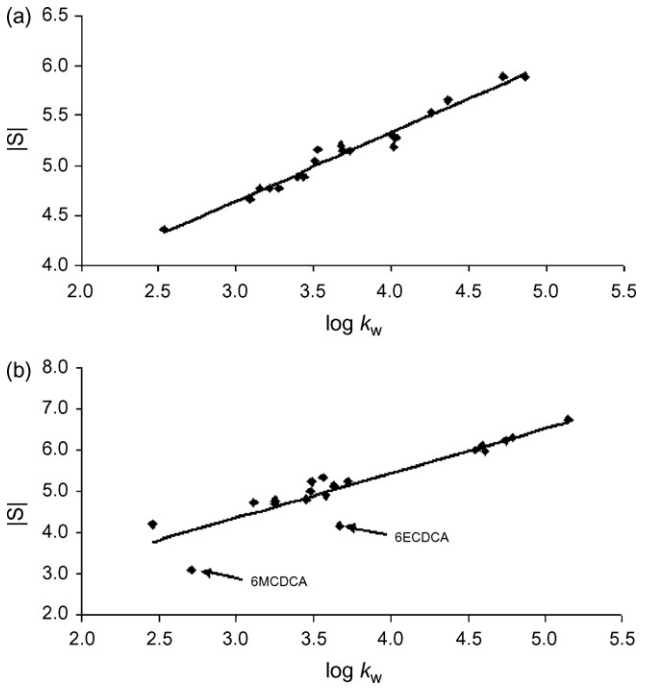

Fig. 3. |*S*| vs log k_w plot for the whole set of compounds: (a) in the absence of additives to the mobile phase and with a RP-18 Base Deactivated as the analytical column (*System A*); (b) in the presence of 1-octanol and triethylamine as the additives to the mobile phase and with a common RP-18 as the analytical column (*System B*).

set of compounds [4–6] and as the starting point for further investigations when different mechanisms are encountered [32].

Inherently, regression equations were derived with the aim of getting a deeper insight into the effect provided by the presence of 1-octanol and triethylamine in the mobile phase (*System B*).

Moreover, analogous calculations were carried out for the system lacking of such mobile phase additives (*System A*). The entire set (conjugated, C, and non-conjugated, N) as well as, separately, the two classes of congeners were studied under the two chromatographic conditions (Table 5, Eqs. (13)–(20)). A pronounced reduction in the statistical quality was observed in the case of *System B* (Table 5, Fig. 3 and Eqs. (14), (17), and (20)). However, a satisfactory degree of linearity was still kept safe even with this more complex environment. Worth mentioning is the slope in the |*S*|/log k_w plot being close to unity. In principle, a unit slop in such a plot indicates that the solute/solvent interactions and solute/stationary phase adsorption are homoenergetic, i.e. characterized by the same free energy changes. In other words, *S* and log k_w values are controlled by the same forces [5,32]. The lower correlation extent observed when 1-octanol and triethylamine are added to the mobile phase can be evidently ascribed to the presence of two outliers (viz 6MCDCA and 6ECDCA) as shown in Fig. 3. Interestingly, the other three 6-substituted, namely the more polar 6FCDCA, HDCA and HCA, did not display any controversial behaviour. Substantially, both in *System A* and *B*, their interaction with the solvent molecules is governed by the same mechanism as their adsorption onto the modified stationary phase. In *System B*, for the two most lipophilic 6MCDCA and 6ECDCA, a difference between the solute solvation and the solute/stationary phase interaction mechanisms can be proposed for tentatively explaining their incoherence with respect to the observed general trend. The previously shown difference in the φ_0 values (Table 2) for both 6MCDCA and 6ECDCA with the two adopted chromatographic systems can be explained on this basis.

It is widely known the controversial behaviour displayed by BAs when chromatographed either in normal or in reversed phase. Indeed, by changing the chromatographic mode, a complete reversion of their elution order is not always observed [33]. This evidence resides in their structural peculiarity. The individual HHB value mainly depends on the number, type and orientation of the substituents onto the steroidal backbone. Polar functionalities (and particularly, in this instance, the hydroxyl groups) can undertake H-bond interactions which result is emphasized by reducing the lipophilic character of the stationary phase (which also produces generally lower retention). Additionally, the presence of 1-octanol and triethylamine in the mobile phase plainly mutates the solvation process of the whole investigated set of compounds. As a consequence, besides the hydrophobic forces, supplementary interactions can play a crucial role in the adsorption mechanism. Moreover, due to the presence of a α-alkyl group in both 6MCDCA and 6ECDCA, a steric hindrance in approaching the 'polar' octanol-sturated stationary phase can be experienced by these two compounds.

Absolute *S* values (Table 2) always result higher in *System B* than in *System A* for all the conjugated compounds. An opposite situation has been found for the unconjugated ones. An explanation of this controversial finding can be demanded to the bimodal effect put forth by MeOH depending on the adopted system. In *System B*, the polar protic MeOH acts as a strong competitor for the intermolecular analyte-stationary phase hydrogen-bonding interactions. The produced 'shielding effect' becomes more and more significant as its content in the eluent is increased. Moreover, it is particularly felt by molecules bearing more H-bridge engaging groups, which presently are represented by the conjugated analytes. The stronger

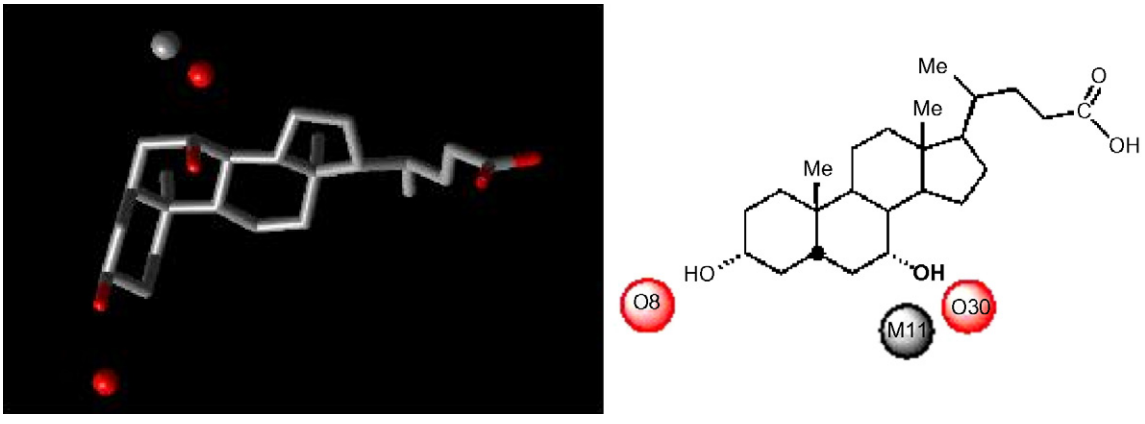

Fig. 4. Molecular interaction points affecting log k_w as observed around the molecular surface of bile acid scaffold.

Table 6
List of statistical parameters from MFA/MLRA for log k_w values.

Parameter	*System A*[a]	*System B*[b]
r^2	0.84	0.64
r^2_{xv}	0.71	0.48
F	26.90	9.44
Observations	20	20
Ratio Obs/Vi	6.70	6.70

(a) In the absence of additives to the mobile phase and with a RP-18 Base Deactivated as the analytical column; (b) in the presence of 1-octanol and triethylamine as the additives to the mobile phase and with a common RP-18 as the analytical column. r: correlation coefficient; r^2_{xv}: prediction coefficient; F: Fisher test

Table 7
Calculated and experimental log k_w values and residual analysis.

BA	Calculated log k_w values	Experimental log k_w values	Residuals
CA	3.84	3.69	0.15
CDCA	4.17	4.26	−0.09
DCA	3.91	4.37	−0.46
HCA	3.92	3.53	0.39
HDCA	3.88	3.68	0.20
NCA	3.00	3.39	−0.39
UCA	3.03	2.54	0.54
UDCA	3.03	3.28	0.25
23MUDCA	3.02	3.21	−0.19
6FCDCA	3.89	3.73	0.16
6MCDCA	4.57	4.72	0.15
6ECDCA	4.60	4.87	−0.27
GCA	3.86	3.44	0.42
GCDCA	3.88	4.03	−0.15
GDCA	3.90	4.02	−0.12
GUDCA	3.03	3.09	−0.06
TCA	3.90	3.51	0.39
TCDCA	3.91	4.00	−0.09
TDCA	4.02	4.03	−0.01
TUDCA	2.06	3.15	−1.09

eluotropic effect exerted in *System B* on this class of compounds can be cautiously explained having recourse to the reduced interaction capability suffered by not only the hydroxyl steroidal groups but even by the glycine and taurine tags. In *System A*, where retention is exclusively deputed to hydrophobic interactions, the organic modifier is added to lower the polarity of the aqueous mobile phase. The lower the polarity of the mobile phase, the greater its eluting strength. In *System A*, this typical reversed-phase mode is particularly experienced by unconjugated molecules being characterized by a lower number of polar moieties.

3.3. Molecular modelling studies

In order to investigate the quantitative relationships between some chromatographic indices and the interaction properties of the molecular shape of the analytes, we engaged in a Molecular Field Analysis (MFA) study of the previously selected set of BAs.

Two quantitative structure-property relationships (QSPR) models were developed using multiple linear regression analysis (MLRA) and the molecular field interaction points: the *Model A* which refers to *System A* and the *Model B* instead referred to *System B*. The principle of MLRA is to model a quantitative dependent variable, the log k_w of the compounds, as measured in *System A* and *B*, through a linear combination of quantitative explanatory variables, namely the interaction points [34]. The model yields an equation that is used to infer the QSPR of the compounds. While the major advantage of the method is its computational simplicity that offers the possibility to easily interpret the resulting equation, its limitation is mainly linked to the ratio of objects and independent variables (Obs/Vi) that, as a rule of thumb, leads to a large risk in chance correlation if below a value of five [35]. Conversely to the QSPR model of *Model A* that shows a good statistical significance, with the regression coefficient (r^2) being 0.84 and the predictive coefficient being $r^2_{xv} = 0.71$, *Model B* is affected by poor regression and predictive coefficients ($r^2 = 0.64$; $r^2_{xv} = 0.47$). Thus, we deemed it advisable to further analyze only the *Model A* since the calculated molecular field interaction points were not sufficient to explain a higher complexity of factors ruling the log k_w obtained in *System B*.

According to *Model A*, the MLRA gives a QSPR equation (Eq. (21)) that is composed of three independent field variables comprising one steric interaction point (M11) and two hydrogen bonding interaction points (O8, O30) (Fig. 4).

$$\log k_w = 3.135 + 0.048(M11) - 0.083(O8) + 0.032(O30) \quad (21)$$

The variables composing the equation are independent, as $r^2 = 0.31$ is the largest intercorrelation observed for M11 and O30.

The independent variables are mapped onto positions C_3, C_6 and C_7 of the steroid nucleus as shown in Fig. 4. The hydrogen bond interaction point O8 has the highest coefficient in the equation and is placed near the C_3-α position of the BA scaffold. In agreement with the negative value of its coefficient, O8 indicates that the hydroxyl group in this position may form hydrogen bonding interactions that favour the increasing of log k_w value.

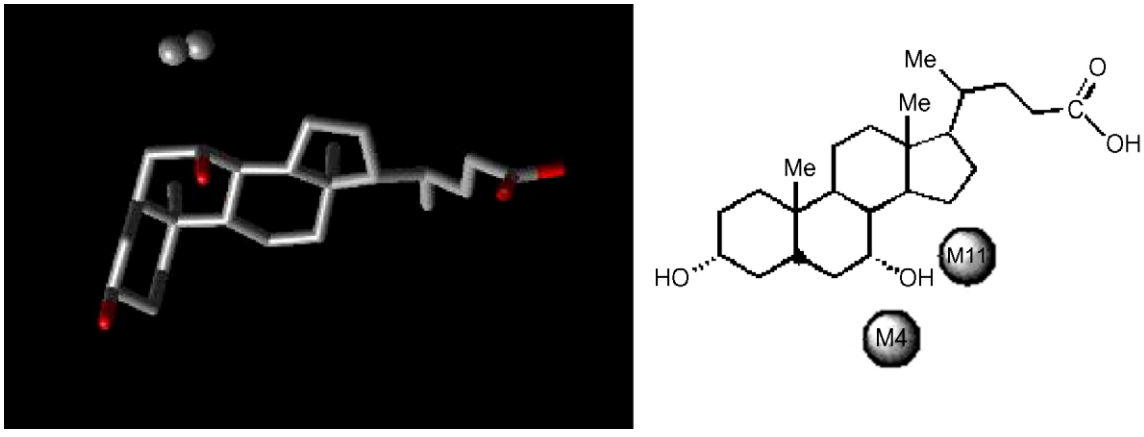

Fig. 5. Molecular interaction points affecting *S* as observed around the molecular surface of bile acid scaffold.

Table 8
List of statistical parameters from MFA/MLRA for |*S*| values in the absence of additives to the mobile phase and with a RP-18 Base Deactivated as the analytical column.

Parameter	*System A*
r^2	0.84
r^2_{xv}	0.80
F	45.97
Observations	20
Ratio Obs/Vi	10

r: correlation coefficient; r^2_{xv}: prediction coefficient; *F*: Fisher test.

Conversely, the positive value of the interaction point O30 suggests that hydrophobic interactions occur at the region of the BA lining positions C_6 and C_7 where the presence of hydroxyl groups negatively affects the log k_w. Furthermore, such hydrophobic interactions are sterically demanding since the steric interaction point located at C_7-β of the steroid nucleus has also a negative effect on the log k_w.

After the leave-one-out cross-validation, the QSPR equation is able to correctly predict the chromatographic index log k_w of the analytes with the $r_{xv}{}^2$ being 0.71 (Table 6).

The experimental log k_w vs calculated log k_w values and residual analysis for the analytes of this study are shown in Table 7.

Fig. 5.

In order to explain the effects of solvation, analogously to *Model A*, a MLRA study was carried out using the chromatographic index *S* as dependent variable. As reported in Table 8, the QSPR model that we obtained shows satisfactory statistical parameters, with the regression coefficient (r^2) being 0.84 and the predictive coefficient being $r^2_{xv} = 0.80$. Thus, according to this model, the chromatographic index *S* is related to two steric interaction points (M4 and M11, Fig. 5) which are located nearby position 6 of the steroid nucleus (Eq. (22)).

$$|S| = 4.702 - 0.045(\mathrm{M4}) + 0.044(\mathrm{M11}) \quad (22)$$

The interaction point M4, in particular, is placed on the 6-β epimer and unfavourable steric interactions at this point negatively affect the *S* value. Conversely, M11 is located at 6-α epimer and unfavourable steric interactions at this point positively affect the *S* value.

Since both M4 and M11 appear in the QSPR equation with similar loading values, the contribution of these independent variables to the chromatographic index *S* is almost similar.

Table 9
Calculated and experimental |*S*| values and residual analysis.

BA	Experimental |*S*| values	Calculated |*S*| values	Residuals
CA	5.18	5.15	0.03
CDCA	5.39	5.54	−0.15
DCA	5.32	5.66	−0.34
HCA	5.31	5.16	0.15
HDCA	5.33	5.20	0.13
NCA	4.82	4.89	−0.07
UCA	4.66	4.36	0.30
UDCA	4.66	4.77	−0.11
23MUDCA	4.66	4.77	−0.11
6FCDCA	5.27	5.14	0.13
6MCDCA	5.87	5.89	−0.02
6ECDCA	5.88	5.89	−0.01
GCA	5.17	4.89	0.28
GCDCA	5.18	5.27	−0.09
GDCA	5.28	5.19	0.09
GUDCA	4.66	4.66	0.00
TCA	5.17	5.05	0.12
TCDCA	5.11	5.31	−0.20
TDCA	5.32	5.29	0.03
TUDCA	4.66	4.78	−0.12

In Table 9 are reported the experimental and calculated values of *S* according to the QSPR Eq. (21).

4. Conclusions

RP-HPLC has proved to be a fruitful tool for the CMC estimation of bile acids. Among the practical and operative advantages arising from this approach, particularly worth noting are the reduced number of samples required and the short analysis times. Conversely to the classical log k_w index, the derived φ_0 parameter showed its independence upon the adopted chromatographic system, thus revealing itself instrumental for the bile acid HHB evaluation. Indeed, despite the severely different interactions which occur into the two investigated chromatographic environments (exclusively hydrophobic in *System A* and both polar and apolar in *System B*), the φ_0 index remained practically unaffected. Consequently, this derived parameter has been profitably engaged for evaluating the self-aggregation process of new bile acids. Moreover, molecular modelling studies cast new lights onto the differences governing the retention mechanism in the two chromatographic media. Accordingly, Molecular Field Analysis (MFA) studies have been successfully engaged in investigating the quantitative relationships between the solute/stationary phase adsorption by exploiting the different interaction properties of the molecular shape of the analytes. Analogously, solute/solvent interactions have been rationalized on this basis as well.

References

[1] D.J. Minick, J.H. Frenz, M.A. Patrick, D.A. Brent, J. Med. Chem. 31 (1988) 1923–1933.
[2] P.A. Tate, J.G. Dorsey, J. Chromatogr. A 1042 (2004) 37–48.
[3] C. Giaginis, S. Theocharis, A. Tsantili-Kakaulidou, J. Chromatogr. A 1166 (2007) 116–125.
[4] C. Giaginis, S. Theocharis, A. Tsantili-Kakoulidou, Anal. Chim. Acta 573–574 (2006) 311–318.
[5] X. Liu, H. Tanaka, A. Yamauchi, B. Testa, H. Chuman, J. Chromatogr. A 1091 (2005) 51–59.
[6] X. Liu, H. Tanaka, A. Yamauchi, B. Testa, H. Chuman, Helv. Chim. Acta 87 (2004) 2866–2876.
[7] R. Shaw, W.H. Elliott, B.G. Barisas, Mikrochim. Acta 3 (1991) 137–145.
[8] B. Natalini, R. Sardella, E. Camaioni, A. Gioiello, R. Pellicciari, Anal. Bioanal. Chem. 388 (2007) 1681–1688.
[9] K. Valkó, C. Bevan, D. Reynolds, Anal. Chem. 69 (1997) 2022–2029.
[10] K. Valkó, P. Slégel, J. Chromatogr. 631 (1993) 49–61.
[11] R. Collander, Acta Chem. Scand. 5 (1951) 774–780.
[12] Cs. Horvát, W. Melander, I. Molnár, J. Chromatogr. 125 (1976) 129–156.
[13] L.R. Snyder, J.W. Dolan, J.R. Gant, J. Chromatogr. 165 (1979) 3–30.
[14] D.J. Schoenmakers, H.A.H. Billiet, J. de Galan, J. Chromatogr. 185 (1979) 179–195.
[15] K. Valkó, J. Liq. Chromatogr. 7 (1984) 1405–1424.
[16] T. Braumann, G. Weber, L.H. Grimme, J. Chromatogr. 261 (1983) 329–343.
[17] M. Harnisch, H.J. Möckel, G. Schulze, J. Chromatogr. 282 (1983) 315–322.
[18] E.H. Slaats, W. Markovsky, J. Fekete, H. Poppe, J. Chromatogr. 125 (1981) 299–323.
[19] J. Stählberg, M. Almegren, Anal. Chem. 57 (1985) 817–821.
[20] M.R. Euerby, P. Petersson, J. Chromatogr. A 1088 (2005) 1–15.
[21] N.S. Wilson, J. Gilroy, J.W. Dolan, L.R. Snyder, J. Chromatogr. A 1026 (2004) 91–100.
[22] N.E. Basci, A. Temizer, A. Bozkurt, A. Isimer, J. Pharm. Biomed. Anal. 18 (1998) 745–750.
[23] A.M. Enlund, D. Westerlund, J. Chromatogr. A 895 (2000) 17–25.
[24] F. Lombardo, M.Y. Shalaeva, K.A. Tupper, F. Gao, M.H. Abraham, J. Med. Chem. 43 (2000) 2922–2928.
[25] C.M. Du, K. Valkó, C. Bevan, D. Reynolds, M.H. Abraham, Anal. Chem. 70 (1998) 4228–4234.
[26] K. Valkó, J. Chromatogr. A 1037 (2004) 299–310.
[27] J.T. Davies, Proceedings of the Second International Congress of Surface Activity, Butterworths Scientific Publications, London, 1957, pp. 409–421.
[28] R. Pellicciari, H. Sato, A. Gioiello, G. Costantino, A. Macchiarulo, B.M. Sadeghpour, G. Giorgi, K. Schoonjans, J. Auwerx, J. Med. Chem. 50 (2007) 4265–4268.
[29] R. Pellicciari, G. Costantino, E. Camaioni, B.M. Sadeghpour, A. Entrena, T.M. Willson, S. Fiorucci, C. Clerici, A. Gioiello, J. Med. Chem. 47 (2004) 4559–4569.

[30] R. Pellicciari, S. Fiorucci, E. Camaioni, C. Clerici, G. Costantino, P.R. Maloney, A. Morelli, D.J. Parks, T.M. Willson, J. Med. Chem. 45 (2002) 3569–3672.
[31] M.C. Carey, M.D. Small, Arch. Intern. Med. 130 (1972) 506–527.
[32] D. Vrakas, I. Panderi, D. Hadjipavlou-Litina, A. Tsantili-Kakoulidou, Quant. Struct. -Act. Relat. 24 (2005) 254–260.
[33] T. Nambara, J. Goto, in: K.D.R. Setchell, D. Kritchevsky, P.P. Nair (Eds.), The Bile Acids Vol.4 Methods and applications, Plenum Press, New York, 1988, pp. 43–63.
[34] K. Héberger, J. Chromatogr. A 1158 (2007) 273–305.
[35] J.G. Topliss, R.P. Edwards, J. Med. Chem. 22 (1979) 1238–1244.