Sandra Šegan[1]
Filip Andrić[2]
Aleksandra Radoičić[2]
Dejan Opsenica[1]
Bogdan Šolaja[2]
Mario Zlatović[2]
Dušanka Milojković-Opsenica[1]

[1]Institute of Chemistry, Technology and Metallurgy, University of Belgrade, Njegoševa, Belgrade, Serbia
[2]Faculty of Chemistry, University of Belgrade, Belgrade, Serbia

Received March 1, 2011
Revised June 21, 2011
Accepted June 22, 2011

Research Article

Correlation between structure, retention and activity of cholic acid derived *cis–trans* isomeric bis-steroidal tetraoxanes

Both quantitative structure–retention (QSRR) and quantitative structure–activity relationship (QSAR) studies have been performed to correlate the molecular characteristics of seven pairs of *cis–trans* isomeric bis-steroidal tetraoxanes with their reversed-phase thin-layer chromatography (RPTLC) retention as well as with their antiproliferative activity. 2D and 3D molecular descriptors as whole molecule representations together with retention parameters as well as with biological activity data were subjected to the multivariate statistical analysis (principal component analysis – PCA and hierarchical cluster analysis – HCA) in order to determine the most influential factors governing the retention and activity against human cervix carcinoma (HeLa) and human malignant melanoma (Fem-X) cell lines. Both QSRR and QSAR models were built by means of the partial least-squares (PLS) statistical method. It was found that hydrogen bond donating (HBD), hydrogen bond accepting (HBAcc), hydrophilic surface percentage (%HS) and hydrophilic–lipophilic balance (HLB) exhibit the strongest influence on retention. The most prominent factors affecting antiproliferative activity of the investigated substances are those relating to the size and shape of a molecule such as: connectivity indices, refractivity (Ref), surface area (SA), molecular volume and weight, polarizability (Pol) and those regarding the ability of hydrogen bonding (HB).

Keywords: Bis-steroidal tetraoxanes / Cholic acid derivatives / Quantitative structure–activity relationship (QSAR) / Quantitative structure–retention relationship (QSRR) / Reversed-phase thin-layer chromatography (RPTLC)
DOI 10.1002/jssc.201100185

1 Introduction

Reversed-phase liquid chromatography is one of the most frequently used separation methods among the other separation methods utilized in biomedical analyses. As it is known, the retention behaviour of analyte in LC is determined by three main factors: chemical structure of the analyte, physico-chemical properties of the mobile phase and physico-chemical properties of the stationary phase. The methods of relating molecular structure of solutes, expressed as descriptors, to their chromatographic (retention) behaviour are commonly denoted as quantitative structure–retention relationships (QSRR). The same assumption that structure descriptors of substances are correlated with their properties and biological activity is widely used in quantitative structure–activity relationship (QSAR) studies. Because it is considered that the same basic intermolecular interactions determine the behaviour of substance in both biological and chromatographic environment [1], a similar influence of particular structural descriptors on both the retention and biological activity might be expected.

Moreover, QSRR studies are useful in providing the insight into the following aspects: (i) revealing the molecular mechanism of the chromatographic separation, (ii) emphasis on the most influential descriptors, (iii) the quantitative comparison of different properties of individual chromatographic conditions [2].

Correspondence: Dr. Dušanka Milojković-Opsenica, Faculty of Chemistry, University of Belgrade, P. O. Box 51, 11158 Belgrade, Serbia
E-mail: dusankam@chem.bg.ac.rs
Fax: +38-11-12639357

Abbreviations: **HCA**, hierarchical cluster analysis; **HBAcc**, hydrogen bond acceptor; **HBD**, hydrogen bond donor; **HLB**, hydrophilic–lipophilic balance; **MD**, molecular depth; **MWe**, molecular weight; **PLS**, partial least squares; **Pol**, polarizability; **PCA**, principal component analysis; **QSAR**, quantitative structure–activity relationship; **QSRR**, quantitative structure–retention relationship; **Ref**, refractivity; **RPTLC**, reversed-phase thin-layer chromatography; **RMSEC**, root mean-square errors of calibration; **RMSECV**, root mean-square errors of cross-validation; **SA**, surface area; **VIP**, variable importance in the projection

The main parameters used in QSRR studies are physico-chemical parameters, non-specific parameters and topological indices. Lipophilicity is one of the most widely used among them. This parameter governs numerous processes such as transport, distribution and metabolism of biologically active molecules.

According to Soczewinski and Matysik [3] the use of reversed-phase thin-layer chromatography (RPTLC) in estimating solute lipophilicity is based on the linear relationship between the retention constant R_M and the concentration of the organic modifier in the mobile phase expressed by the equation

$$R_M = R_M^0 + bC \quad (1)$$

where C is the volume fraction of organic modifier in the mobile phase and b is the change in R_M value due to the 1% increase in the organic modifier (associated with the specific hydrophobic surface area (SA)). The advantages of this method are: applicability to compounds with higher lipophilicity (log P ranging from 0 to 7), only a small amount of the sample required, low sensitivity to impurities, rapid determination, good accuracy and excellent reproducibility [4]. The estimation of lipophilicity parameter strongly varies depending on the type of chromatographic system employed, and a few multivariate statistical techniques can be used to select the most appropriate ones [5]. Unlike this method, the traditional shake-flask method [6] faces problems such as poor reproducibility, long time of analysis and it requires a larger amount of pure compounds.

Most QSRRs reported in the chromatographic literature were derived by means of the multiple linear regression (MLR) analysis using retention data as a dependent variable and various empirical, semi-empirical and non-empirical structural parameters as the independent ones. Unfortunately, the independent variables applied are often mutually inter-correlated, which sometimes may result in statistically poorly established QSRR equations [7, 8]. Nowadays, a widely used powerful tool in the field of modelling is partial least-squares (PLS) because, unlike MLR, it can analyze strongly collinear and numerous *X*-variables, and it simultaneously models several response variables. In addition, principal component analysis (PCA) and hierarchical cluster analysis (HCA) are useful in providing data overview, and are usually carried out at the introductory level, revealing outliers or highlighting similarities among objects and variables.

Cholic acid-derived bis-steroidal tetraoxanes are recently synthesized compounds that were evaluated as antimalarials and antiproliferatives. They exist as *cis–trans* diastereoisomers, which is the consequence of the different orientations of two steroidal parts of a molecule with respect to tetraoxane ring. Accordingly, in *cis*-C(2)C(2a) series hydrophobic surfaces of a steroid molecule are placed on the opposite sides of the molecule, whereas in *trans*-C(2)C(2a) series hydrophobic areas are oriented towards the same side. It was found that potent antimalarial activity of these compounds is determined by the pharmacophore, which contains two hydrogen bond acceptors (lipid) and one aliphatic hydrophobic site (aliphatic) [9–14].

TLC behaviour of bis-steroidal tetraoxanes on different sorbents was the subject of our previous investigations [15]. At that time, highly selective separations were achieved in different normal- and reversed-phase systems.

Since the pharmacophore for potent antimalarial activity of these compounds was suggested earlier, in the present work we were interested in the selection of the most influential descriptors regarding both chromatographic behaviour in reversed-phase conditions and antiproliferative activity, as well as in the estimation of lipophilicity parameters for the studied compounds. The latter is of a special interest because the calculated log P values for *cis–trans* isomers are identical.

2 Materials and methods

2.1 Studied compounds

Studied tetraoxanes (Table 1) were synthesized according to the procedure described in the literature [11].

2.2 TLC

The chromatographic investigations were performed using the horizontal TLC on 10×10 cm plates covered with octadecyl or cyano-modified silica. An HPTLC developing chamber (Camag, Muttenz, Switzerland) in the tank configuration was used for this purpose.

The investigated substances were dissolved in dichloromethane, and the plates were spotted with 1.0 μL aliquots of freshly prepared solutions (C ~2 mg/mL). Before development, the spotted plates were equilibrated for 15 min in the chromatographic chamber with vapours of the corresponding mobile phase. All solvents used throughout the present study were of analytical-grade purity.

The following chromatographic systems were used: RP-18 W F_{254s} (Art. 13124, Merck, Darmstadt, Germany) with mobile phase 0–14 Vol% water in methanol (increment 2%), 10–30 Vol% water in acetone (increment 5%) and 10–35 Vol% water in dioxane (increment 5%); CN-silica F_{254s} (Art. 16464, Merck) with mobile phase 10–30 Vol% water in methanol (increment 5%) and 10–40 Vol% water in acetone (increment 5%).

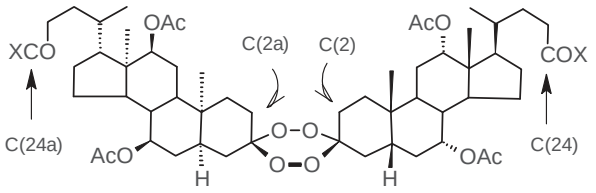
Development distance was 4.5 cm. Detection of individual zones was performed by spraying the plates with 50% sulphuric acid and heating until the spots became visible. All experiments were performed at ambient temperature ($22 \pm 2^{\circ}$C).

2.3 Structural descriptors and molecular modelling

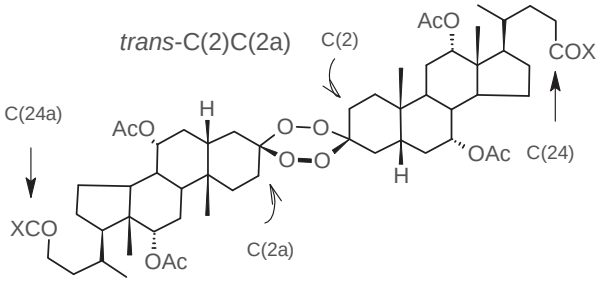
Various structural descriptors encoding significant structural information such as topology, geometry and electronic environment were calculated.

Table 1. The structures of the investigated bis-steroidal tetraoxanes

cis-C(2)C(2a)



trans-C(2)C(2a)



X	Compound	
	cis-C(2)C(2a)	*trans*-C(2)C(2a)
OCH_3	1	2
NH_2	3	4
$NHPr^n$	5	6
OH	7	8
$NHCH_2CO_2CH_3$	9	10
$N(Pr^n)_2$	11	12
	13	14

The Hyperchem Professional software (version 7.0, Hybercube, Gainesville, FL, USA) was used to calculate some quantum chemical descriptors such as: polarizability (Pol), the energy of HOMO and LUMO orbitals, dipole moment (DM) and the QSAR properties such as: SA (grid) and refractivity (Ref). Molecular Modelling Program Plus (MMP plus) software [16] (http://norgwyn.com/mmpplus.html) was used to calculate additional descriptors related to the molecular size and branching (molecular length (ML), molecular width (MWe), and molecular depth (MD), Kier-type connectivity indices 0–3 (CI 0–3), Kier-type valence indices 0–3 (VI 0–3), Kier-type kappa shape index 2 (KI 2), Wiener index (WI)), various polymer and surfactant properties (Van Krevelen and Hoftyzer's 3-D solubility parameters – molar volume (MV), SA, MWe, hydrophilic–lipophilic balance (HLB), surface polarity (SP), dispersion (D), polarity (P), hydrogen bonding (HB), percent hydrophilic surface (%HS), hydrogen bond acceptor (HBAcc), hydrogen bond donor (HBD), lipophilicity (log *P*)). They were all calculated from the hydrogen-suppressed graph of the molecule, encoding information about the degree of branching and the size of the molecules. Geometry optimization was performed using the Hyperchem software. The geometry of a given molecule was first optimized at the empirical level using an MM+ molecular mechanics force field as a rough and introductory run, followed by unrestricted geometric optimization at the semi-empirical level (AM1 Hamiltonian) using a calculation with the convergence limit set at 0.1 kcal/mol (Polak–Ribiere algorithm).

2.4 Multivariate statistical analysis and modelling

PCA, HCA and PLS have been performed using a demo version of PLS_Toolbox statistical package (Eigenvectors, v. 5.7) for MATLAB version 7.4.0.287 (R2007a) (MathWorks, Natick, MA, USA). The data were autoscaled before building the model in order to prevent variables with larger magnitude to prevail in the final models, although it does not necessarily have to be the case with chromatographic data, since they are usually of the same order of magnitude. PCA was carried out as an exploratory data analysis by using single value decomposition algorithm (SVD) and 0.95 confidence level for Q and T^2 Hotelling limits for outliers. In that way, employing only a limited number of PCs, the dimensionality of the retention data space was reduced, further analysis was simplified, and the substances were grouped according to their intrinsic ability for specific interactions. The agglomerative HCA was performed in addition to the PCA in order to group similar objects more easily and provide some help in explaining the results of PCA. The Euclidian distance and Ward's method were used to form well-shaped clusters. The PLS method was employed by means of SIMPLS algorithm without forcing orthogonality constrains to the model which are supposed to condense Y-block variance into the first few latent variables. Validation of the models was performed by the leave-one-out cross-validation procedure [17], and the following parameters were monitored: the squared values of Pearson's correlation coefficients for calibration and cross-validation, R^2 cal (cum) and R^2_{cv} (cum), respectively, being as high as possible, and RMSEC (root mean-square errors of

calibration) and RMSECV (root mean-square errors of cross-validation) being as low as possible, with the lowest difference between them. Low value of RMSEC is desirable but if the high values of RMSECV are present at the same time, it indicates the poor predictability of the calibration model.

3 Results and discussion

In all instances, it was found that the retention of the investigated compounds decreased with increasing concentration of the organic modifier in the mobile phase. The R_M values obtained by Eq. (1) depend linearly ($r > 0.99$) on the concentration of the organic modifier in the mobile phase and therefore they might be used to assess lipophilicity. The chromatographically determined parameters of lipophilicity expressed as R_M^0 values are summarized in Table S1 (Supporting Information).

The greater values of lipophilicity parameters determined by TLC are obtained on RP-18 (higher R_M^0 values) with those on CN-silica. Such chromatographic behaviour is probably the consequence of stronger hydrophobic interactions between the analytes and octadecyl alkyl-chains of RP-18 silica in comparison to less hydrophobic propyl chains of CN-silica. In addition, the specific influence of the organic modifier on retention is evident. Namely, methanol, which is the most frequently used organic modifier for the determination of lipophilicity due to its great polarity and ability for strong association with water molecules [18], contributes to higher R_M^0 values in comparison to acetone or dioxane. Although Biagi et al. [19] observed that R_M^0 values are independent of the nature of the organic modifier in the mobile phase when the organic modifier is acetone, acetonitrile or methanol, this finding cannot be accepted as general for every chromatographic process. Some other works have also pointed to significant deviations in R_M^0 values obtained when methanol and other organic modifiers are used. This is explained by strong specific solute–eluent interactions, i.e. the formation of hydrogen bonds that influence solubility of an analyte in a mobile phase [20]. Compounds **11** and **12** have the greatest values of lipophilicity in all the systems used, while primary amides (**3** and **4**) show the lowest lipophilicity in most cases.

Regardless of chromatographic system used lower lipophilicity was established for *cis*-isomers in comparison to the corresponding *trans*-isomers.

3.1 Introductory multivariate analysis based on retention data and molecular descriptors

Application of PCA to retention data can reveal some similarities among the studied compounds that are governed by both their intrinsic structural properties and specific interactions that occur in different chromatographic systems. Loading plots highlight the most influential chromatographic systems responsible for such clustering. Furthermore, a PCA carried out on the set of calculated molecular descriptors can cluster compounds based on their structural features alone. It is therefore useful to perform a PCA on both retention data and molecular descriptors separately. If a congeneric series of compounds is studied, outliers might be detected by using PCA and removed prior to the final modelling.

In order to obtain some basic insight into the chromatographic behaviour of the studied compounds, a PCA was first performed on chromatographic data (R_M values) with autoscaling resulting in two-component model that explains 95.06% of the total variance. The first principal component (PC1) accounted for 91.11% of data variance and the second one (PC2) for 8.39%. Score values, i.e. their mutual projections, for PC1 and PC2 are shown in Fig. 1A and the mutual projections of the loading vectors in Fig. 1B. All the data lay inside the Hotelling T^2 ellipse, suggesting that there were no outliers. The obtained results show that PC1 separates acids (**7**, **8**), primary amides (**3**, **4**) and secondary amides (**5**, **6**, **9**, **10**) from esters (**1**, **2**), tertiary amides (**11**, **12**) and piperidine derivatives (**13**, **14**). This reveals that the classification of bis-steroidal tetraoxanes is based on their ability to accept protons (compounds **1**, **2**, **11–14**) and donate protons (compounds **3–10**). This should be supported by the data related to the mobile-phase composition, expecting that dioxane or acetone-containing mobile phase will mostly influence the retention of proton-donating compounds, while methanol will be the cause of the decreased retention of proton-accepting compounds. Unfortunately, the loading graph does not reveal any significant influence of the mobile-phase composition along the PC1 direction. On the contrary, the chromatographic systems are much better differentiated along the PC2 direction (Fig. 1B), which shows the highest positive impact of systems with dioxane on RP-18 (compounds with PC2 scores above zero have stronger retention in the chromatographic systems containing dioxane) and the lowest negative values of systems with methanol on RP-18 (compounds with PC2 scores below zero exhibit stronger retention in the chromatographic systems using methanol as a modifier).

The PCA performed on descriptors resulted in a five-component model that explains 94.71% of total variance. It reveals a quite different classification (Fig. 1C), and shows no outliers among the studied compounds. Going along the PC1 axis from its negative end towards positive values, the primary amides (**3**, **4**) and acids (**7**, **8**) are positioned almost at the end, grouped very close to each other, and well separated from the rest of the compounds. The next group of compounds are secondary amides (**5**, **6**, **9**, **10**) located in the middle and differentiated in small clusters along the PC2 direction, which is in accordance with their ability to act as both proton donors and proton acceptors. Since the tertiary amides (**11–14**) do not possess any of proton donating abilities, they are placed at the opposite end of the PC1 axis. The only exception is a pair of methyl esters (**1**, **2**)

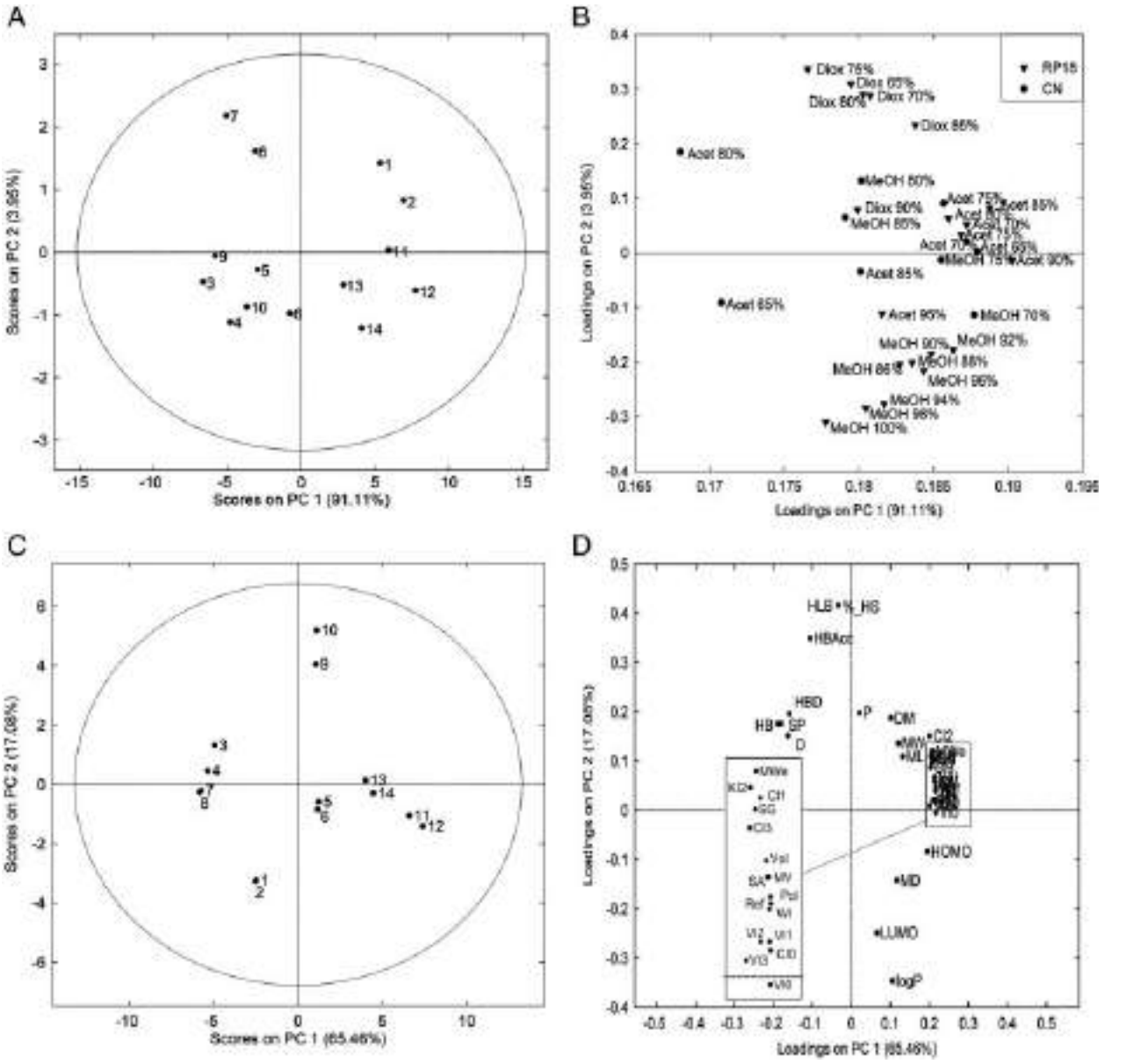


Figure 1. PC1–PC2 score plots of retention parameters and descriptors (A and C respectively) and factor loadings (B and D respectively).

which possesses only proton-accepting properties and is placed in the middle of the PC score graph. As it can be observed from Fig. 1D, the majority of descriptors have a positive impact on PC2, while only HOMO, MD, LUMO and log *P* have a negative influence. The parameters %HS, HLB, and HBAcc exhibit the highest positive impact on PC2. The log *P* has the highest negative influence on the PC2 score.

In addition to this, PCA indicates that there is no grouping of bis-steroidal tetraoxanes according to the type of isomerisation (*cis*- or *trans*-).

HCA was performed in order to confirm the grouping of compounds already obtained by the PCA. It is sometimes used to make the process of grouping similar objects in PCA score graphs easier. In this case it was performed on the basis of both retention parameters (R_M values) and previously calculated molecular descriptors. The obtained dendrogram based on retention parameters in all chromatographic systems shows two well-separated clusters (Fig. 2A). The first cluster includes the following compounds: primary amides (**3**, **4**), secondary amides (**5**, **6**, **9**, **10**) and carboxylic acids (**7**, **8**); all three sorts of compounds express major proton-donating abilities. The second cluster is made of methyl esters (**1**, **2**), and tertiary amides (**11–14**), both lacking a proton-donating ability and possessing strong proton-accepting functional groups.

The cluster analysis performed on descriptors resulted in two main clusters, with slight differences compared with the previously described results (Fig. 2B). The first cluster is

made of primary amides (**3**, **4**) and acids (**7**, **8**) while the second cluster is formed from the rest of the compounds, i.e. secondary amides (**5**, **6**, **9**, **10**), tertiary amides (**11–14**) and methyl esters (**1**, **2**).

3.2 Retention and biological activity modelling

After performing PCA and HCA in order to provide some basic insight into the factors that are responsible for retention of the studied compounds as well as a necessary step prior to the final modelling, PLS modelling was performed on both retention data and the data of antiproliferative activity (Fem-X and HeLa cell lines). The number of latent variables was selected according to the following criteria: the minimum value of RMSECV and the minimum difference between RMSEC and RMSECV. The obtained results are summarized in Table 2.

The best models are obtained with acetone–water as mobile phase on both RP-18 and CN-modified silica sorbents with the following statistical performance: R^2 cal = 91.00%, R^2_{CV} = 83.04%, RMSEC = 0.434 and RMSECV = 0.595 for RP-18, and R^2 cal = 88.56%, R^2_{CV} = 77.21%, RMSEC = 0.173 and RMSECV = 0.247 for CN-silica.

The assessment of descriptors that have the greatest influence on retention was done based on variable importance in the projection (VIP) scores. The variables with VIP scores higher than 1 were considered to be the most relevant to explaining the dependent variable *Y*, while those significantly lower than 1 (arbitrarily, the value lower than 0.5 was taken) had little or almost no influence. Descriptors with VIP > 1.2 are denoted by asterisks in the regression graphs (Fig. 3), and they are summarized in Table 2 in descending order of their coefficient values in regression graphs, as well as the statistical performance of the final model.

It is evident that, more or less, the same descriptors are included in all retention models, mostly describing the overall polarity of compounds and their proton donating or accepting abilities. Since the best models were obtained using acetone as an organic modifier, these models will be explained here in detail.

The most relevant descriptors influencing the R^0_M values of bis-steroidal tetraoxanes for the RP-18/acetone–water system are: HLB, %HS, HBAcc, HBD and log *P*, and, apart from these, in the case of the CN-silica/acetone–water system, P, MD, SP and HB were also included (Fig. 3A and B).

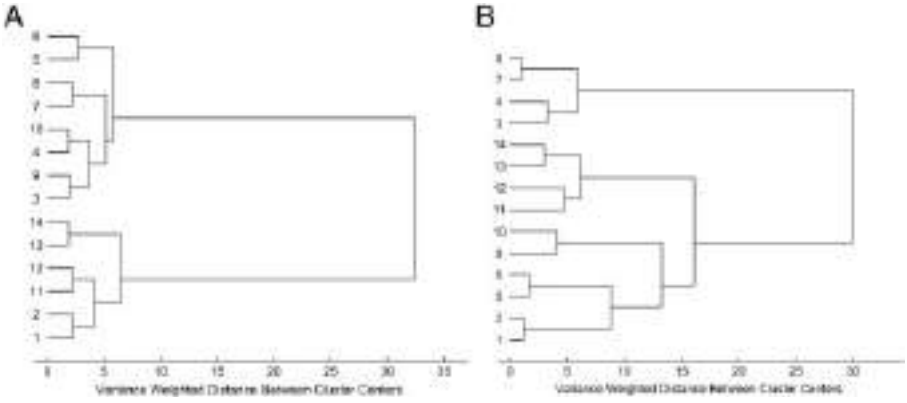


Figure 2. Dendrogram based on R_M values and descriptors, A and B, respectively.

Table 2. The results of PLS analysis

Chromatographic system used		Statistical performance of the models				Molecular descriptors included in the model
Stationary phase	Mobile phase	R^2_{CV}	R^2_{cal}	RMSEC	RMSECV	
RP-18	Methanol/water	0.784	0.886	0.960	1.325	HLB, %HS, HBAcc, logP, HBD, SP, D
	Acetone/water	0.830	0.910	0.434	0.595	HLB, %HS, HBAcc, logP, HBD
	Dioxane/water	0.667	0.798	0.736	0.948	HBAcc, logP, HLB, %HS, HBD, P
CN-silica	Methanol/water	0.469	0.700	0.531	0.730	HLB, %HS, logP, MD, D, HBAcc, SP, HBD, P
	Acetone/water	0.772	0.885	0.173	0.247	HLB, %HS, P, logP, HBAcc, HBD, MD, SP, HB
Biological activity						
Fem-X		0.478	0.636	0.335	0.410	MWe, MV, SA, SP, HB, HBD, CI 0-3, VI1-3, Pol, Ref, SA (Grid), WI, HOMO, LUMO
HeLa		0.512	0.660	0.348	0.425	MV, SA, HB, CI 0-3, VI 1-3, Pol, Ref, SA (Grid), WI, HOMO, LUMO

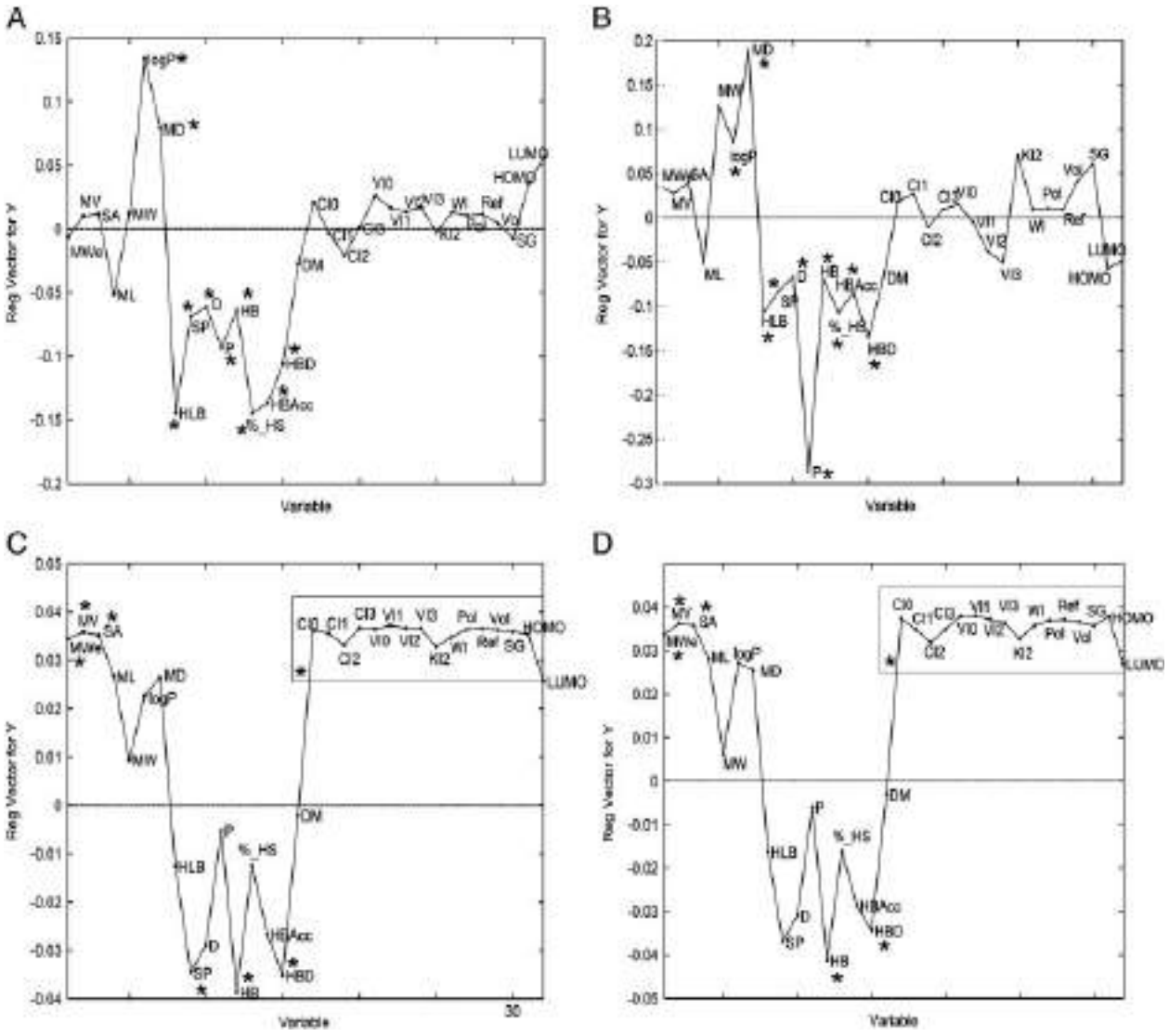


Figure 3. Plot of the regression coefficients of descriptors of PLS models for acetone/water on RP-18 and CN-silica (A and B, respectively) and for biological activity against Fem-X and HeLa (C and D, respectively).

It can be observed from the regression plots (Fig. 3A and B) that the descriptors describing polarity of the studied molecules and their abilities to form hydrogen bonds, such as: HLB, %HS, HBAcc, HBD, SP and HB make a negative contribution to the R_M^0 values. Since the polar interactions in RPTLC are mainly established between the solute and mobile-phase molecules, it is therefore logical that the higher value of the aforementioned parameters leads to lower retention. On the other hand, descriptors such as log *P* and MD, which describe nonpolar properties of the solutes and their abilities for dispersive interactions with stationary phase, positively contribute to retention (R_M^0 values) under the studied chromatographic conditions.

As it is well known, electronic, hydrophobic or steric properties can have an impact on biological activity and, additionally, lipophilicity seems to be one of the major properties that can control drug-likeness of a compound.

Since, the phamacophore of antimalarial activity of bis-steroidal tetraoxanes was already determined in the previous paper [12], the antiproliferative activity (human melanoma Fem-X and human cervix carcinoma HeLa cell lines) of the same compounds was of special interest in our present study.

The log *C* values of minimal inhibitory concentrations for the investigated bis-steroidal tetraoxanes are presented in Table S2 (Supporting Information).

The most significant value of antiproliferative activity against Fem-X and HeLa cell lines is exhibited by compound **3** (the activity of this compound against HeLa is very similar to those of *cis*-platin). By applying a paired *t*-test it was found that the type of isomerisation

(*cis*- or *trans*-) does not influence biological activity of the investigated compounds either in the case of Fem-X ($t_{paired} = 0.099$, $t_{crit} = 2.571$, $P = 0.05$, and $n = 6$) or in the case of HeLa cells ($t_{paired} = 2.485$, $t_{crit} = 3.182$, $P = 0.05$, and $n = 4$).

The relationship between the proliferative activity as the dependent variables and structural descriptors as the independent variables was analyzed by PLS (Table 2). The obtained QSAR models are not of satisfactory statistical quality to be used for accurate prediction of biological activity, but they may qualitatively indicate descriptors that play an important role in the activity exhibited by these compounds. Based on the VIP scores, the following molecular descriptors are identified as the most influential factors against both Fem-X and HeLa cell lines (Table 2): MV, HB, Ref, Pol, surface area grid (SA (Grid)), HOMO, LUMO, and those from group of connectivity indices. In addition, SP, MWe and HBD show high influence against Fem-X cells. Regression vectors of the obtained models indicate that descriptors HB, HBD and SP influence antiproliferative activity in a negative manner (positive coefficients), while the rest of them show a significant positive effect (Fig. 3C and D). The similar conclusion was reported by Cvijetić et al. [21] in the case of 33 1,2,4,5-tetraoxane derivatives who claimed that hydrophobicity and H-bond donor properties were the main parameters influencing potency of compounds towards human cervix carcinoma (HeLa) and human malignant melanoma (Fem-X) cell lines. In addition, according to Bhattacharjee et al. [8], the following functional features are required for potent anatimalarial activity of 1,2,4,5-tetraoxane derivatives: two hydrogen bonds acceptors and one aliphatic hydrophobic site, and one hydrogen bond acceptor function or at least one oxygen atom of trioxane and tetraoxane moiety.

It is usually considered that the same properties are responsible for the behaviour of an analyte in biological and reversed-phase chromatographic environment. However, in the obtained QSRR and QSAR models, different features of the studied analytes influence the retention in investigated chromatographic systems and specific antiproliferative activity. It can be noticed from Fig. 3 that regression vectors of retention and antiproliferative activity are almost opposite to each other. The most obvious difference is between regression coefficients of connectivity indices that are close to zero and have no significant VIP scores in the first case, while in the second case they are high and strongly affect antiproliferative activity on both investigated cell lines. Since these descriptors describe the shape, the branching degree and steric properties of a molecule, it can be concluded that, in the case of antiproliferative activity, the factors sensitive to molecular shape and size are of particular importance. In the case of retention under reversed-phase chromatographic conditions the factors responsible for a partition between a polar mobile phase and a nonpolar stationary phase, such as: log *P*, HLB, DM, P, and HB, both HBAcc and HBD, are dominant.

4 Concluding remarks

The present work focuses on identifying the most important descriptors affecting reversed-phase chromatographic behaviour as well as on antiproliferative activity of *cis–trans* isomeric pairs of bis-steroidal tetraoxanes on RP-18 and CN-silica thin layers. For this purpose PCA and HCA followed by PLS were performed.

PCA and HCA revealed that analytes can be classified according to their structural characteristics. The best statistical results were obtained with acetone as the organic modifier on both sorbents. The application of PLS revealed that the most dominant factors governing the retention of the investigated bis-steroidal tetraoxanes were HLB, %HS, HBAcc, log *P* and HBD, i.e. the descriptors which describe polarity of compounds and their proton-donor and proton-acceptor abilities. The PLS revealed that the biological activity of bis-steroidal tetraoxanes is influenced by MWe, MV, SP, HB, HBD, Ref, Pol, WI, HOMO, LUMO and the descriptors from the group of connectivity indices.

This work was realized within the scope of research project no. 172008 supported by the Ministry of Education and Science, the Republic of Serbia.

The authors have declared no conflict of interest.

5 References

[1] Baošić, R., Radojević, A., Radulović, M., Miletić, S., Natić, M., Tešić, Ž., *Biomed. Chromatogr.* 2008, *22*, 379–386.

[2] Baczek, T., Kaliszan, R., Novotná, K., Jandera, P., *J. Chromatogr. A* 2005, *1075*, 109–115.

[3] Soczewinski, E., Matysik, G., *J. Chromatogr.* 1968, *32*, 458–471.

[4] Sarbu, C., Daković-Sekulić, T., Perišić-Janjić, N., *J. Pharm. Biomed. Anal.* 2002, *30*, 739–745.

[5] Djaković-Sekulić, T., Smolinski, A., Perisić-Janjić, N., Janicka, M., *J. Chemom.* 2008, *22*, 195–202.

[6] Hansch, C., Fujita, T., *J. Am. Chem. Soc.* 1964, *86*, 1616–1626.

[7] Andries, J. P. M., Claessens, H. A., Vander Heyden, Y., Buydens, L. M. C., *Anal. Chim. Acta* 2009, *652*, 180–188.

[8] Koba, M., Stasiak, J., Bober, L., Bączek, T., *J. Mol. Model.* 2010, *16*, 1319–1331.

[9] Opsenica, D., Pocsfalvi, G., Juranić, Z., Tinant, B., Declercq, J.-P., Kyle, D. E., Milhous, W. K., Šolaja, B. A., *J. Med. Chem.* 2000, *43*, 3274–3282.

[10] Šolaja, B., Terzić, N., Pocsfalvi, G., Gerena, L., Tinant, B., Opsenica, D., Milhous, W. K., *J. Med. Chem.* 2002, *45*, 3331–3336.

[11] Opsenica, D., Angelovski, G., Pocsfalvi, G., Juranić, Z., ŽiŽak, Ž., Kyle, D., Milhous, W. K., Šolaja, B. A., *Bioorgan. Med. Chem.* 2003, *11*, 2761–2768.

[12] Bhattacharjee, A. K., Carvalho, K. A., Opsenica, D. M., Šolaja, B. A., *J. Serb. Chem. Soc.* 2005, *70*, 329–345.

[13] Opsenica, I., Terzić, N., Opsenica, D., Angelovski, G., Lehnig, M., Eilbracht, P., Tinant, B., Juranić, Z., Smith, K. S., Yang, Y.S., Diaz, D.S., Smith, P. L., Milhous, W. K., Doković, D., Šolaja, B. A., *J. Med. Chem.* 2006, *49*, 3790–3799.

[14] Terzić, N., Opsenica, D., Milić, D., Tinant, B., Smith, B. K., Milhous, W. K., Šolaja, B. A., *J. Med. Chem.* 2007, *50*, 5118–5127.

[15] Šegan, S. B., Opsenica, D. M., Šolaja, B. A., Milojković-Opsenica, D. M., *J. Planar Chromatogr. – Mod. TLC* 2009, *22*, 175–181.

[16] Molecular Modeling Software–NGMSI NorGwyn Montgomery Software INC, accessed 30 November 2010.

[17] Mevik, B.-H., Cederkvist, H. R., *J. Chemom.* 2004, *18*, 422–429.

[18] Natalini, B., Sardella, R., Camaioni, E., Natalini, S., Pellicciari, R., *Chromatographia* 2006, *64*, 343–349.

[19] Biagi, G., Barbaro, A., Sapone, A., *J. Chromatogr. A* 1994, *662*, 341–361.

[20] Niewiadomy, A., Żabinska, A., Matysiak, J., Różylo, J., *J. Chromatogr. A* 1997, *791*, 237–243.

[21] Cvijetić, I. N., Žižak, Ž. P., Stanojković, T. P., Juranić, Z. D., Terzić, N., Opsenica, I. M., Opsenica, D. M., Juranić, I. O., Drakulić, B. J., *Eur. J. Med. Chem.* 2010, *45*, 4570–4577.

