JPC – Journal of Planar Chromatography – Modern TLC (2020) 33:245–253
https://doi.org/10.1007/s00764-020-00029-w

ORIGINAL RESEARCH PAPER

Evaluation of the lipophilicity of chalcones by RP-TLC and computational methods

Vladimir Dobričić [1] · Nemanja Turković [1,2] · Branka Ivković [1] · Oszkár Csuvik [1,3] · Zorica Vujić [1]

Received: 16 February 2020 / Accepted: 10 May 2020 / Published online: 8 June 2020

Abstract

Retention behaviour of twenty-one chalcones synthesized in our laboratory was tested in three thin-layer chromatography (RP-TLC) systems (acetonitrile–water, ethanol–water and acetone–water) and chromatography parameters R_M^0, S and C_0 were calculated. The most suitable RP-TLC system (acetonitrile–water) and chromatography parameter (C_0) for lipophilicity prediction of tested compounds were selected on the basis of the highest correlations with calculated log$P$ values. In selected system, compound *12* had the highest, whereas *47* had the lowest C_0 value. QSRR analysis was performed and three models representing relationships between C_0 and selected molecular descriptors were created—MLR(C_0), PLS(C_0) and SVM(C_0). Interpretation of molecular descriptors which form statistically the most reliable SVM(C_0) model identified the most important structural and physico-chemical properties that influence retention behaviour of tested compounds. In addition, descriptors with the highest influence on R_M^0 as well as on C_0 calculated in the remaining two RP-TLC systems were identified and interpreted.

Keywords RP-TLC · Chalcones · log$P$ · QSRR models

1 Introduction

Lipophilicity is recognized as one of the most important physico-chemical properties, due to its influence on biological activity and pharmacokinetics (absorption, distribution, metabolism and excretion) of drug molecules. It is expressed by the octanol/water partition coefficient (P), but virtually the logarithm of P is used to describe it (log$P$). The traditional approach to determine log$P$ is the shake-flask method, but it is time-consuming, requires large amounts of tested substances and cannot provide very reliable values when log$P > 3$. Reversed-phase thin-layer chromatography (RP-TLC) can be used as an alternative method and can yield quantitatively comparable and precise data [1–4].

✉ Vladimir Dobričić
vladimir@pharmacy.bg.ac.rs

[1] Department of Pharmaceutical Chemistry, Faculty of Pharmacy, University of Belgrade, Vojvode Stepe 450, Belgrade 11000, Serbia

[2] Agency for Medicines and Medical Devices of Montenegro, 64A Bulevar Ivana Crnojevića, 81000 Podgorica, Montenegro

[3] Institute of Pharmaceutical Chemistry, University of Szeged, Eötvös u. 6, Szeged 6720, Hungary

Chalcones (1,3-diphenyl-2-propen-1-ones) are precursors for biosynthesis of flavonoids and isoflavonoids, with wide-range of biological activities, such as antiproliferative, cytotoxic, anti-inflammatory, antiviral, antimalarial, antibacterial, etc. Additionally, two chalcone derivatives have been approved for clinical treatments, metochalcone as a choleretic drug, and sofalcone as anti-ulcer drug. The general structure of chalcones consists of two aromatic rings (A and B) spaced by three-carbon α, β-unsaturated carbonyl bridge, preferably with E isomerism [5–7]. Chalcones have a simple chemistry that enables to synthesize largely variable analogues. The most common method is Claisen–Schmidt condensation, in which the starting materials are benzaldehyde and acetophenone derivatives. The reaction is usually carried out in the presence of alkaline or acid catalysts, and the required temperature of the reaction can be achieved by both classical and microwave-assisted heating [7].

Chalcones tested in this study (Fig. 1) showed satisfactory antimicrobial activity against clinical isolates of dermatophytes, redox activity as well as anti-HIV-1 protease activity. Compounds *18*, *22* and *52* showed similar anti-HIV-1 protease activity in comparison with lopinavir [8].

The aim of this study was evaluation of the lipophilicity of a series of chalcone derivatives, synthesized in our laboratory, by use of RP-TLC and computational methods as well as

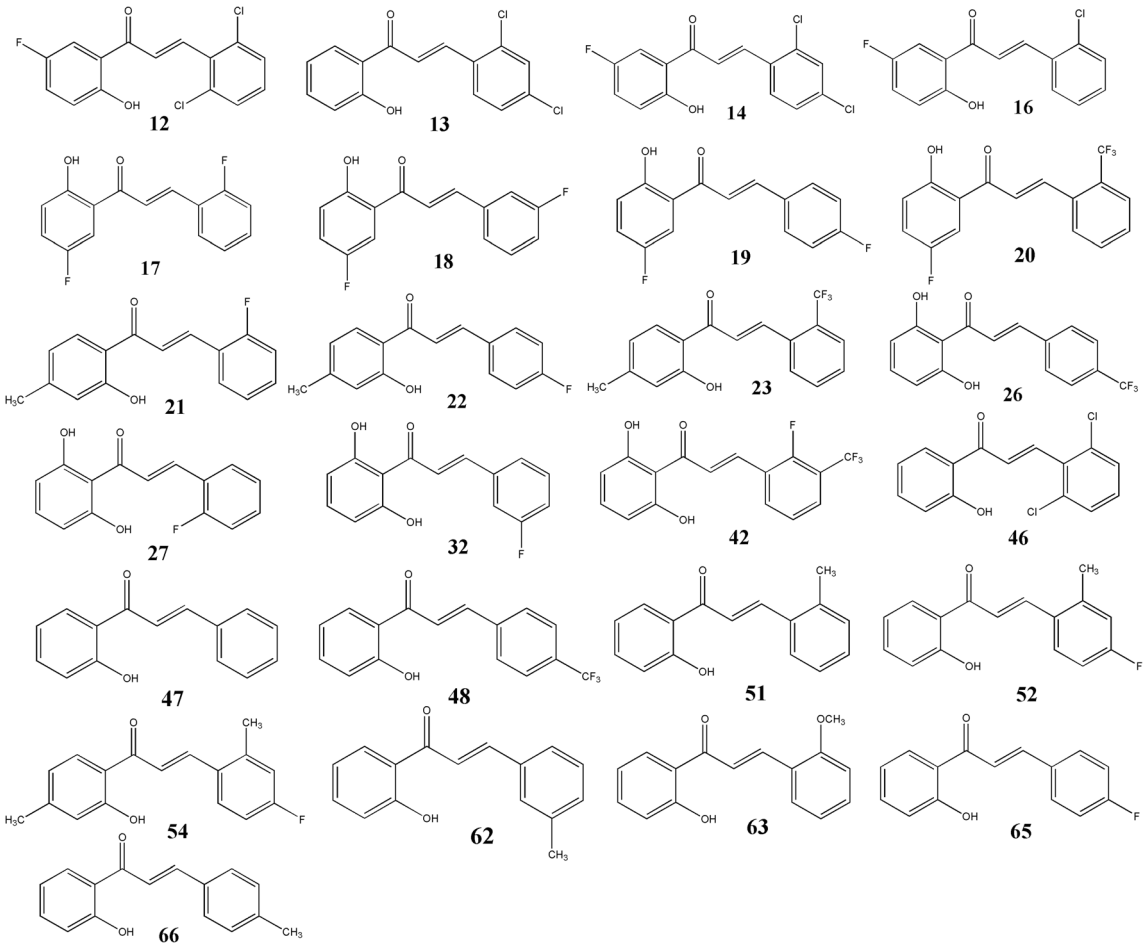

Fig. 1 Chemical structures of the tested compounds

identification of major structural and physico-chemical properties that influence their retention in selected RP-TLC system.

2 Experimental

2.1 Materials and reagents

Chalcones tested in this study were *12* ((E)-3-(2,6-dichlorophenyl)-1-(5-fluoro-2-hydroxyphenyl)prop-2-en-1-one), *13* ((E)-3-(2,4-dichlorophenyl)-1-(2-hydroxyphenyl)prop-2-en-1-one), *14* ((E)-3-(2,4-dichlorophenyl)-1-(5-fluoro-2-hydroxyphenyl)prop-2-en-1-one), *16* ((E)-3-(2-chlorophenyl)-1-(5-fluoro-2-hydroxyphenyl)prop-2-en-1-one), *17* ((E)-1-(5-fluoro-2-hydroxyphenyl)-3-(2-fluorophenyl)prop-2-en-1-one), *18* ((E)-1-(5-fluoro-2-hydroxyphenyl)-3-(3-fluorophenyl)prop-2-en-1-one), *19* ((E)-1-(5-fluoro-2-hydroxyphenyl)-3-(4-fluorophenyl)prop-2-en-1-one), *20* ((E)-1-(5-fluoro-2-hydroxyphenyl)-3-(2-(trifluoromethyl)phenyl)prop-2-en-1-one), *21* ((E)-3-(2-fluorophenyl)-1-(2-hydroxy-4-methylphenyl)prop-2-en-1-one), *22* ((E)-3-(4-fluorophenyl)-1-(2-hydroxy-4-methylphenyl)prop-2-en-1-one), *23* ((E)-1-(2-hydroxy-4-methylphenyl)-3-(2-(trifluoromethyl)phenyl)prop-2-en-1-one), *46* ((E)-3-(2,6-dichlorophenyl)-1-(2-hydroxyphenyl)prop-2-en-1-one), *47* ((E)-1-(2-hydroxyphenyl)-3-phenylprop-2-en-1-one), *48* ((E)-1-(2-hydroxyphenyl)-3-(4-(trifluoromethyl)phenyl)prop-2-en-1-one), *51* ((E)-1-(2-hydroxyphenyl)-3-(o-tolyl)prop-2-en-1-one), *52* ((E)-3-(4-fluoro-2-methylphenyl)-1-(2-hydroxyphenyl)prop-2-en-1-one), *54* ((E)-3-(4-fluoro-2-methylphenyl)-1-(2-hydroxy-4-methylphenyl)prop-2-en-1-one), *62* ((E)-1-(2-hydroxyphenyl)-3-(m-tolyl)prop-2-en-1-one), *63* ((E)-1-(2-hydroxyphenyl)-3-(2-methoxyphenyl)prop-2-en-1-one), *65* ((E)-3-(4-fluorophenyl)-1-(2-hydroxyphenyl)prop-2-en-1-one), and *66* ((E)-1-(2-hydroxyphenyl)-3-(p-tolyl)prop-2-en-1-one) (Fig. 1).

The following solvents were used for the mobile phase preparation: acetonitrile (LC/MS grade, Fisher Scientific, Loughborough, UK), ethanol (HPLC grade, Fisher Scientific, Loughborough, UK), acetone (Avantor Performance Materials, Gliwice, Poland) and distilled water (TKA water purification system, Niederelbert, Germany).

Dimethyl sulfoxide (Fisher Scientific, Loughborough, UK) and methanol (Fisher Scientific, Loughborough, UK) were used for the preparation of solutions of tested compounds.

2.2 Chromatography analysis

Thin-layer chromatography analysis was performed using 10 × 10 cm reversed-phase silica 60 RP-18 F_{254}s plates (Merck, Darmstadt, Germany). Chalcones were dissolved in DMSO (2 mg mL^{-1}) and diluted with methanol to obtain final solutions (0.1 mg mL^{-1}). The plates were spotted with 5 μL of each solution, and three binary combinations of organic solvent and water were used as mobile phases (Table 1). The developed chromatograms were observed under UV light (λ = 254 nm). Finally, R_f and R_M values were calculated for each tested compound.

The relation between R_M and organic modifier percentage (φ) is expressed by Soczewinski–Wachtmeister eq. (1) [9].

$$R_M = R_M^0 + S\varphi \quad (1)$$

R_M^0 is determined by extrapolation technique and represents R_M value in the hypothetical system that contains 0% of organic modifier. Parameter S is the slope of the regression curve (eq. (1)), whereas C_0 is calculated by dividing R_M^0 by S (eq. (2)).

$$C_0 = -\frac{R_M^0}{S} \quad (2)$$

2.3 Calculation of log*P*

For the calculation of log*P* values of tested compounds, ChemDraw 8 [10] and MarvinSketch [11] were used. Four log*P* values were calculated in ChemDraw 8—Viswanadhan logP, Brotto logP, Crippen logP and CLogP [12–14].

2.4 Calculation of molecular descriptors

ChemoPy molecular descriptors (1135 1D, 2D and 3D molecular descriptors) were calculated using freely available web-based platform ChemDes, which utilizes MOPAC software for the molecular geometry optimization [15]. After the elimination of those without variance, 873 descriptors were retained for further QSRR modelling.

2.5 QSRR analysis

Statistica 13 software [16] was used for molecular descriptor selection, multiple linear regression (MLR), partial least squares (PLS) and support vector machine (SVM) modelling.

The quantitative structure-retention relationships (QSRR) studies were performed to investigate the relationships between C_0 (dependent variable) of the tested compounds and their calculated molecular descriptors (independent variables). For MLR(C_0), PLS(C_0) and SVM (C_0), test set consisted of 5 compounds (*13*, *18*, *20*, *23* and *63*), while other compounds were used as training set. Test set was formed so C_0 of these compounds were homogenously distributed in the whole range of C_0 values.

2.5.1 Descriptor selection

Prior to QSRR modelling, molecular descriptor selection has to be performed. For this purpose, several methodologies have been applied, such as genetic algorithm [17], principal component analysis [18] and stepwise MLR [19, 20]. In this study, Statistica's algorithm feature selection and variable screening (FSVS) was applied. FSVS can be applied on extremely large sets of descriptors and enables evaluation of both linear and nonlinear relationships between dependent variable and descriptors. The range of values of each descriptor is separated into k intervals. If $k = 2$, only monotonous relationships between descriptors and dependent variable are investigated. In this study, k was set to 5 and for MLR and SVM modelling following descriptors were selected: *bcutp8*, *ATSp4* and *kappam2*. For PLS modelling, 30 most significant descriptors were selected using the same criterion ($k = 5$).

2.5.2 QSPR and QSRR model building

MLR was applied to assess linear relationship between calculated molecular descriptors and C_0. In this study, the MLR(C_0) model was created using "all effects" method, which means that all descriptors are simultaneously added into the model.

PLS modelling is useful when analysing data with collinear, noisy and numerous descriptors. Optimal number of PLS components was determined on the basis of each component's

Table 1 RP-TLC systems applied in this study

RP-TLC system	Mobile phase composition
I	Acetonitrile–water (50:50, 60:40, 70:30 and 80:20, *v/v*)
II	Ethanol–water (50:50, 60:40, 70:30 and 80:20, *v/v*)
III	Acetone–water (50:50, 60:40, 70:30 and 80:20, *v/v*)

R2(Y) value and cumulative R2(Y) value. The influence of descriptors on created model was evaluated on the basis of their scaled regression coefficient values. In this study, optimal PLS(C_0) model consisted of three components.

Although SVM was initially developed as a binary classification tool [21], it can also be used for the development of nonlinear QSAR and QSRR models [22, 23]. In this study, SVM(C_0) model was created using radial basis function (RBF) Kernel type and regression type 1, while optimal gamma value was 0.333. Subsequently, capacity (C) and epsilon ($\mathcal{E}$) values were automatically optimized by the software and optimal values were $C = 59$ and $\mathcal{E} = 0.3$. Finally, the SVM(C_0) model consisted of 4 supported vectors (0 bounded).

Evaluation of the reliability of created QSRR models was performed on the basis of following statistical parameters: RMSEE (root mean squared error of estimation), RMSEP (root mean squared error of prediction), the F ratio, the p value, r, Q^2 (eq. (3)) and R^2_{pred} (eq. (4)).

$$Q^2 = 1 - \frac{PRESS}{\sum\left(Y_{obs(training)} - \overline{Y}_{training}\right)^2} \tag{3}$$

$$R^2_{\text{pred}} = 1 - \frac{PRESS}{\sum\left(Y_{obs(test)} - \overline{Y}_{training}\right)^2} \tag{4}$$

$$PRESS = \sum_{i-1}^{n} e^2_{(i)} \tag{5}$$

RMSEE value was calculated for training, whereas RMSEP was calculated for test set and these values represent errors of prediction for training and test set compounds. Q^2 is an internal validation parameter used to assess predictive potential of a model for compounds similar to training set. This parameter was calculated according to the leave-one-out (LOO) procedure. Briefly, each training set compound was deleted once while the remaining compounds were used to create a model. The model thus created was used to predict the C_0 value of the deleted compound. The procedure was applied for all training set compounds and finally Q^2 was calculated (eq. (3)) [24, 25]. In this equation, $\overline{Y}_{\text{training}}$ is average value, whereas $Y_{\text{obs(training)}}$ is an observed C_0 value of the training set compounds. PRESS was calculated according to the eq. (5) after the LOO procedure ($e_{(i)}$ represents difference between observed and predicted C_0 values). R^2_{pred} (eq. (4)) is an external validation parameter used to assess predictive potential of a model for compounds that are structurally different than those from the training set [26]. In this equation, $Y_{\text{obs(test)}}$ is an observed value of C_0 of a test set compound, while $\overline{Y}_{\text{training}}$ is mean C_0 value of the training set compounds. PRESS value was calculated for the test set according to the eq. (5). High predictive potential is expected for models with Q^2 and R^2_{pred} higher than 0.5 [24, 27, 28].

The F test evaluates significance of the model and it is based on the ratio MS Regression/MS Residual. The p value indicates probability level where a model with this F value may be the result of just chance. A model is considered statistically significant if p value is lower than 0.05 [28].

3 Results and discussion

3.1 RP-TLC analysis and calculated log$P$ values

Chromatography parameters (R^0_M, S and C_0) in acetonitrile–water, acetone–water and ethanol–water RP-TLC systems as well as calculated log$P$ values of all tested compounds are presented in Table 2. There is a high degree of concordance between calculated log$P$ values since correlation coefficients between pairs of these values were from 0.96 to 1.00.

Parameter S is related to the specific hydrophobic surface area [29]. High correlation coefficients between R^0_M and S indicate that the tested compounds form congeneric set and similar mechanisms influence their retention behaviour. Lack of correlation between these parameters can be observed with chemically versatile compounds and also in case of the presence of ionizable groups, which modify the interactions with stationary and mobile phases [30, 31]. In all chromatography systems tested in this study, high correlation between R^0_M and S was observed ($r = 0.87$–0.97), which is due to the high degree of structural similarity among tested compounds, and it can be expected that in each applied system similar mechanisms are responsible for the interactions of compounds with stationary and mobile phases.

All chromatography parameters (R^0_M, S and C_0) can be used as lipophilicity descriptors in QSPR, QSRR and QSAR studies as well as for the estimation of lipophilicity and pharmacokinetic properties [29, 32–36]. However, in some cases, low correlations between lipophilicity and chromatography parameters were observed [37], which indicates the importance of careful evaluation of potential of RP-TLC as alternative method for lipophilicity estimations of different sets of compounds.

The correlation coefficients between calculated log$P$ and chromatography parameters were 0.80–0.84 (log$P$ vs C_0, acetonitrile–water), 0.23–0.29 (log$P$ vs C_0, acetone–water), 0.47–0.56 (log$P$ vs C_0, ethanol–water), 0.39–0.44 (log$P$ vs R^0_M, acetonitrile–water), 0.26–0.29 (log$P$ vs R^0_M, acetone–water), 0.53–0.61 (log$P$ vs R^0_M, ethanol–water), below 0.05 (log$P$ vs S, acetonitrile–water), 0.29–0.36 (log$P$ vs S, acetone–water) and 0.38–0.43 (log$P$ vs S, ethanol–water). Due to the highest correlation coefficients between C_0 determined in acetonitrile–water RP-TLC system and calculated log$P$ values (from 0.80 to 0.84), this RP-TLC system and corresponding C_0 parameter were selected as the most suitable for

Table 2 Calculated log*P* values and lipophilicity parameters determined in the applied RP-TLC systems

Compound	Calculated log*P* values					Acetonitrile–water			Acetone–water			Ethanol–water		
	Viswanadhan log*P*	Broto log*P*	Crippen log*P*	Clog*P*	Marvin log*P*	R_M^0	*S*	C_0	R_M^0	*S*	C_0	R_M^0	*S*	C_0
12	4.57	4.33	4.47	5.65	5.59	3.4536	−0.0373	92.5898	4.7533	−0.0519	91.5857	4.2983	−0.0488	88.0799
13	4.43	4.20	4.31	5.38	5.44	3.9170	−0.0426	91.9484	5.4817	−0.0620	88.4145	4.1480	−0.0471	88.0679
14	4.57	4.33	4.47	5.65	5.59	3.4576	−0.0392	88.2041	4.9182	−0.0566	86.8940	3.6590	−0.0429	85.2914
16	4.05	3.72	3.91	4.94	4.98	3.6574	−0.0408	89.6422	4.8513	−0.0552	87.8859	3.5681	−0.0412	86.6044
17	3.67	3.23	3.51	4.37	4.52	3.5691	−0.0412	86.6286	5.2654	−0.0615	85.6163	3.4150	−0.0403	84.7395
18	3.67	3.23	3.51	4.37	4.52	3.6126	−0.0422	85.6066	5.0341	−0.0589	85.4686	3.2600	−0.0381	85.5643
19	3.67	3.23	3.51	4.37	4.52	3.5763	−0.0431	82.9768	4.9077	−0.0560	87.6375	3.9698	−0.0491	80.8513
20	4.41	4.08	4.27	5.11	5.26	3.5066	−0.0392	89.4541	4.8328	−0.0552	87.5507	3.7488	−0.0449	83.4922
21	4.00	3.51	3.84	4.60	4.89	3.7371	−0.0427	87.5199	4.8531	−0.0548	88.5602	4.4154	−0.0526	83.9430
22	4.00	3.51	3.84	4.60	4.89	3.4649	−0.0394	87.9416	4.9529	−0.0578	85.6903	3.4826	−0.0415	83.9181
23	4.74	4.36	4.60	5.34	5.63	4.1238	−0.0465	88.6839	4.9493	−0.0561	88.2228	3.6793	−0.0427	86.1663
46	4.43	4.20	4.31	5.38	5.44	3.9111	−0.0434	90.1175	4.9731	−0.0558	89.1237	4.2856	−0.0492	87.1057
47	3.39	2.96	3.20	3.96	4.24	3.7433	−0.0455	82.2703	5.0706	−0.0594	85.3636	3.3953	−0.0410	82.8122
48	4.27	3.94	4.12	4.84	5.11	3.3444	−0.0389	85.9743	4.8673	−0.0567	85.8430	3.4286	−0.0413	83.0169
51	3.86	3.38	3.68	4.46	4.75	3.4652	−0.0392	88.3980	4.8821	−0.0562	86.8701	3.9320	−0.0469	83.8380
52	4.00	3.51	3.84	4.60	4.89	3.4954	−0.0397	88.0453	4.8117	−0.0549	87.6448	3.7729	−0.0450	83.8422
54	4.47	3.93	4.33	5.10	5.41	4.0665	−0.0456	89.1776	4.8475	−0.0538	90.1022	4.3749	−0.0519	84.2948
62	3.86	3.38	3.68	4.46	4.75	3.5507	−0.0408	87.0270	5.0112	−0.0584	85.8082	3.5926	−0.0415	86.5687
63	3.14	3.09	3.07	3.88	4.08	3.4131	−0.0404	84.4827	4.6424	−0.0546	85.0256	3.2562	−0.0388	83.9227
65	3.53	3.10	3.35	4.10	4.38	3.2629	−0.0392	83.2372	3.8105	−0.0405	94.0864	3.2556	−0.0388	83.9072
66	3.86	3.38	3.68	4.46	4.75	3.5521	−0.0411	86.4258	4.0757	−0.0447	91.1790	3.5402	−0.0409	86.5575

Table 3 Calculated statistical parameters

QSRR model	RMSEE	Q^2	RMSEP	r	R^2_{pred}
MLR(C_0)	1.20	0.64	2.87	0.43	0.79
PLS(C_0)	1.11	0.63	2.36	0.56	0.86
SVM(C_0)	1.14	0.68	1.57	0.83	0.94

lipophilicity prediction. In this system, compound *12* had the highest, whereas *47* had the lowest C_0 value.

3.2 QSRR analysis

Three QSRR models were created using C_0 determined in acetonitrile–water RP-TLC system as dependent variable and selected molecular descriptors as independent variables. Statistical parameters were calculated and presented in Table 3.

According to the presented statistical parameters, the most reliable QSRR model is SVM(C_0) due to the lowest values and lowest difference between RMSEE and RMSEP as well as due to the highest values of other statistical parameters. This model has good predictive potential because both Q^2 and R^2_{pred} are higher than 0.5.

Descriptors which form SVM(C_0) model are *bcutp8*, *ATSp4* and *kappam2*.

bcutp8 belongs to Burden descriptors based on polarizability. *bcut* descriptors are the eigenvalues of a connectivity matrix which takes into account both connectivity and atomic properties of a molecule, such as atomic weight, partial charge and polarizability. The descriptor is based on a weighted version of the Burden matrix [38, 39]. The weights are a variety of atom properties placed along the diagonal of the Burden matrix. The following three weighting schemes are employed: atomic weight, partial charge (Gasteiger–Marsili) and polarizability [40].

Relation between C_0 and bcutp8 is presented in Fig. 2. All compounds could be sorted into two groups: compounds with bcutp8 between 1.90 and 2.00 (first group) and compounds with bcutp8 between 2.10 and 2.45 (second group). Compounds from the second group have median values of C_0 (from 84 to 90), whereas in the first group (lowest bcutp8) those with highest and lowest values of C_0 can be found.

ATSp4 is a 2D topological descriptor and belongs to Moreau–Broto autocorrelation descriptors based on atomic polarizability (Broto–Moreau autocorrelation of a topological structure − lag 4 / weighted by atomic polarizabilities). It is related to the polarity of a molecule [41]. In 2D autocorrelation descriptors, atoms represent a set of discrete points in space and the atomic property and function are evaluated at those points. The symbol for each of the autocorrelation descriptors is followed by two indices d and w (d stands for the lag and w stands for the weight). The lag is defined as the topological distance d between pairs of atoms. The weight can be m (relative atomic mass), p (polarizability), e (Sanderson electronegativity) and v (Van der Waals volume) [42].

Relation between C_0 and ATSp4 is presented in Fig. 3**.** The highest values of this descriptor possess compounds with high C_0 and high calculated log*P* values (e.g. *12* and *46*), whereas the lowest values possess compounds with low C_0 and low calculated log*P* values (e.g., *47* and *65*), indicating the presence of relation between lipophilicity and retention properties of tested compounds.

kappam2 belongs to kappa descriptors and represents Molecular shape Kappa index for 2 bonded fragments. The

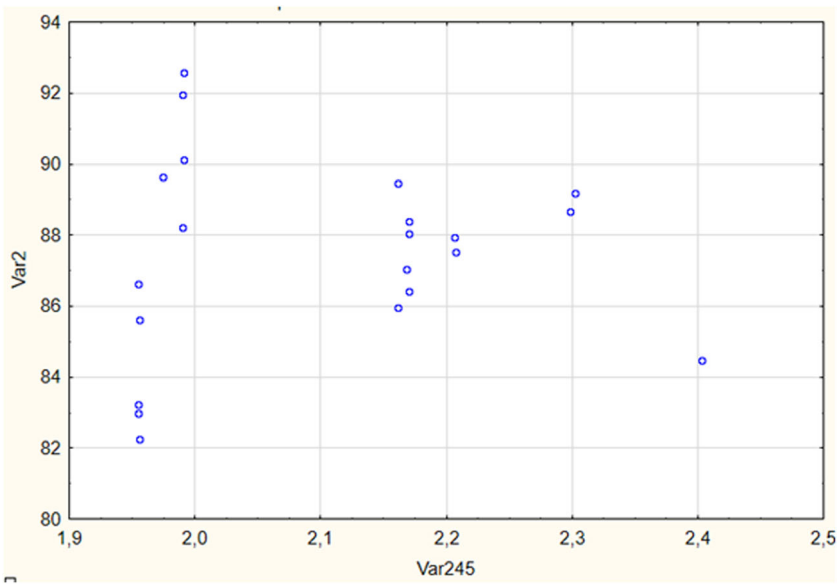

Fig. 2 C_0 (var2) vs bcutp8 (var245) plot

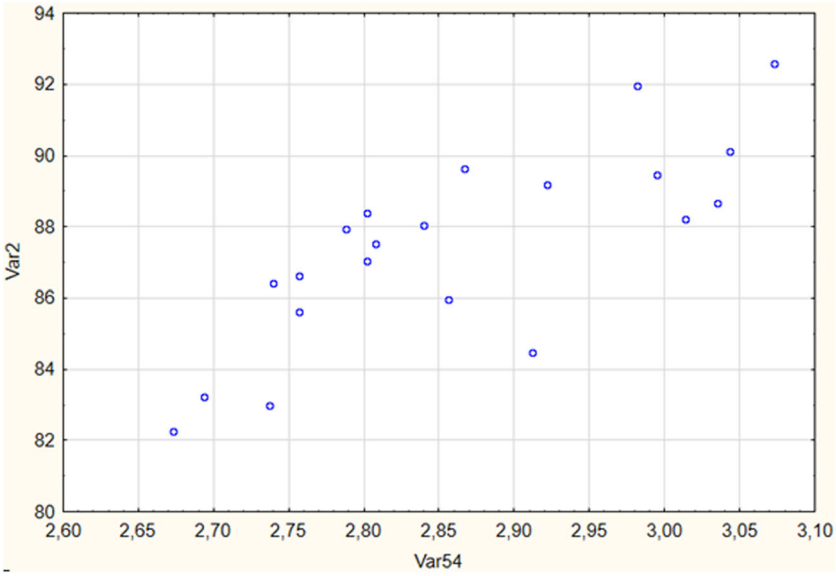

Fig. 3 C_0 (var2) vs ATSp4 (var54) plot

kappa indexes could be defined as shape parameters based on the degree of branching of the molecular graph [43].

Relation between C_0 and kappam2 is presented in Fig. 4**.** The highest values of this descriptor possess compounds with high C_0 (e.g. *12*, *13* and *46*), whereas the lowest values possess compounds with low C_0 (e.g., *47* and *65*). Compounds *12*, *13* and *46* possess higher number of substituents on benzene rings (3 and 4 substituents) and therefore higher degree of branching in comparison to *47* and *65* (1 and 2 substituents).

In order to investigate the relationships between other chromatography parameters and calculated descriptors, the FSVS variable selection procedure was performed for following dependent variables: C_0 calculated in acetone–water and ethanol–water RP-TLC systems as well as for R_M^0 calculated in acetonitrile–water RP-TLC system. The relationships were investigated using SVM. Descriptors with the highest influence on C_0 in ethanol–water system were *MR*, *logP2* and *bcutp7*. Descriptors with the highest influence on R_M^0 calculated in acetonitrile–water RP-TLC system were *ATSp7* and *bcutv10*. However, there were no statistically significant relationships between calculated descriptors and C_0 in acetone–water RP-TLC system. *bcutp7* and *bcutv10* belong to Burden descriptors based on atomic polarizability and volumes, respectively. *MR* is molar refractivity and *logP2* is square of log*P* value based on the Crippen method [12]. *ATSp7* belongs to Moreau–Broto autocorrelation descriptors based on atomic polarizability (Broto–Moreau autocorrelation of a topological structure − lag 7 / weighted by atomic polarizabilities) and it is related to the polarity of a molecule [41].

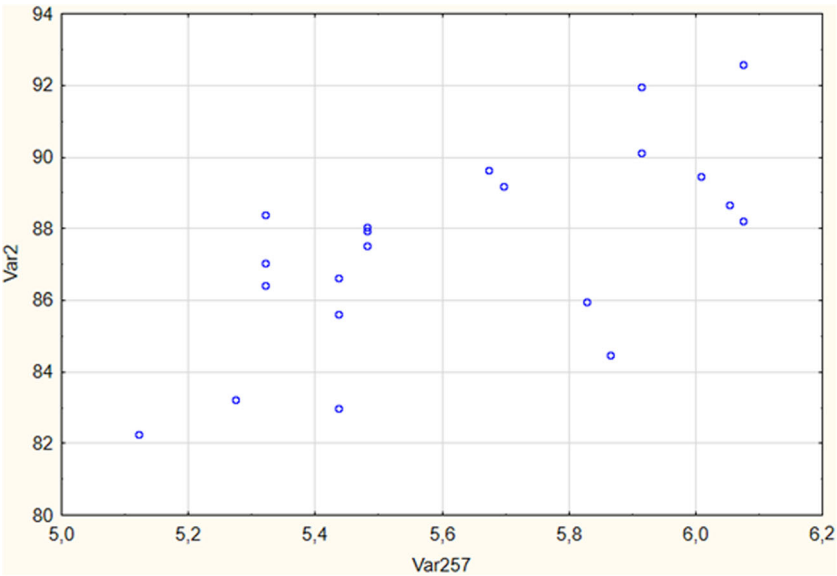

Fig. 4 C_0 (var2) vs kappam2 (var257) plot

4 Conclusion

RP-TLC with three binary combinations of water and organic solvents as the mobile phase (acetonitrile–water, ethanol–water and acetone–water) was used for the testing of retention behaviour of twenty-one chalcones synthesized in our laboratory. The most suitable RP-TLC system (acetonitrile–water) and chromatography parameter (C_0) for lipophilicity prediction were selected according to the highest correlations with calculated log*P* values. In this system, compound *12* had the highest, whereas *47* had the lowest C_0 value. QSRR analysis of obtained results resulted in creation of three models (MLR(C_0), PLS(C_0) and SVM(C_0)), and on the basis of calculated statistical parameters, SVM(C_0) was selected as the most reliable. Interpretation of descriptors that form this model identified the most important structural and physico-chemical properties that influence retention behaviour of tested compounds in the selected RP-TLC system. In addition, descriptors with the highest influence on R_M^0 as well as on C_0 calculated in the remaining two RP-TLC systems, were identified and interpreted.

Availability of data and material All raw data are available upon request.

Funding information This work was financially supported by the Ministry of Education, Science and Technological Development, Republic of Serbia, as part of Project No.172041.

Compliance with ethical standards

Conflicts of interest The authors declare that they have no conflict of interest.

References

1. Morak-Młodawska E, Nowak M, Pluta K (2007) Determination of the lipophilicity parameters RM0 and LogP of new azaphenothiazines by reversed-phase thin-layer chromatography. J Liq Chromatogr R T 30:1845–1854. https://doi.org/10.1080/10826070701360749
2. Csermely T, Kalász H, Deák K, Mohammed Y, Hasan MY, Darvas F, Petroianu G (2008) Lipophilicity determination of some ACE inhibitors by TLC. J Liq Chromatogr R T 31:2019–2034. https://doi.org/10.1080/10826070802198410
3. Odović J, Karljković-Rajić K, Trbojević-Stanković J, Stojimirović B, Vladimirov S (2012) Lipophilicity examination of some ACE inhibitors and hydrochlorothiazide on cellulose in RP thin-layer chromatography. Iran J Pharm Res 11:763–770
4. Dobričić V, Vladimirov S, Čudina O (2016) Synthesis and RP-TLC lipophilicity evaluation of a novel fluocinolon acetonide soft drug derivative. Kragujevac J Sci 38:107–114. https://doi.org/10.5937/KgJSci1638107D
5. Xia Y, Yang Z-Y, Xia P, Bastow KF, Nakanishi Y, Lee K-H (2000) Antitumor agents. Part 202: novel 2′-amino chalcones: design, synthesis and biological evaluation. Bioorg Med Chem Lett 10:699–701. https://doi.org/10.1016/s0960-894x(00)00072-x
6. Venkatesan P, Sumathi S (2010) Piperidine mediated synthesis of n-heterocyclic chalcones and their antibacterial activity. J Heterocyclic Chem 47:81–84. https://doi.org/10.1002/jhet.268
7. Gomes MN, Muratov EN, Pereira M, Peixoto JC, Rosseto LP, Cravo PVL, Andrade CH, Neves BJ (2017) Chalcone derivatives: promising starting points for drug design. Molecules 22:E1210. https://doi.org/10.3390/molecules22081210
8. Turkovic N, Ivkovic B, Kotur-Stevuljevic J, Tasic M, Marković B, Vujic Z (2020) The molecular docking, synthesis and anti-HIV-1 protease activity of novel chalcones. Curr Pharm Des 26:802–814. https://doi.org/10.2174/1381612826666200203125557
9. Soczewiński E, Wachtmeister CA (1962) The relation between the composition of certain ternary two-phase solvent systems and RM values. J Chromatogr A 7:311–320. https://doi.org/10.1016/S0021-9673(01)86422-0
10. CambridgeSoft Corporation. 2005. ChemDraw ultra version 8.0.3. Cambridge, MA, USA
11. MarvinSketch 15.1.26 (2015). ChemAxon, Budapest. http://www.chemaxon.com
12. Ghose AK, Crippen GM (1987) Atomic physicochemical parameters for three-dimensional-structure-directed quantitative structure-activity relationships. 2. Modeling dispersive and hydrophobic interactions. J Chem Inf Comput Sci 27:21–35. https://doi.org/10.1021/ci00053a005
13. Viswanadhan VN, Ghose AK, Revankar GR, Robins RK (1989) Atomic physicochemical parameters for three dimensional structure directed quantitative structure-activity relationships. 4. Additional parameters for hydrophobic and dispersive interactions and their application for an automated superposition of certain naturally occurring nucleoside antibiotics. J Chem Inf Comput Sci 29:163–172. https://doi.org/10.1021/ci00063a006
14. Broto P, Moreau G, Vandycke C (1984) Molecular structures: perception, autocorrelation descriptor and SAR studies. System of atomic contributions for the calculation of n-octane/water partition coefficients. Eur J Med Chem- Chim Theor 19:71–78
15. Dong J, Cao DS, Miao HY, Liu S, Deng B-C, Yun Y-H, Wang N-N, Lu AP, Zeng W-B, Chen AF (2015) ChemDes: an integrated web-based platform for molecular descriptor and fingerprint computation. Aust J Chem 7:60. https://doi.org/10.1186/s13321-015-0109-z
16. TIBCO Software Inc., 2017. Statistica (data analysis software system), version 13. http://statistica.io
17. Gupta VK, Khani H, Ahmadi-Roudi B, Mirakhorli S, Fereyduni E, Agarwal S (2011) Prediction of capillary gas chromatographic retention times of fatty acid methyl esters in human blood using MLR, PLS and back-propagation artificial neural networks. Talanta 83:1014–1022. https://doi.org/10.1016/j.talanta.2010.11.017
18. Zhang YX (2007) Artificial neural networks based on principal component analysis input selection for clinical pattern recognition analysis. Talanta 73:68–75. https://doi.org/10.1016/j.talanta.2007.02.030
19. Gonzalez-Arjona D, Lopez-Perez G, Gustavo-Gonzalez A (2002) Non-linear QSAR modeling by using multilayer perceptron feedforward neural networks trained by back-propagation. Talanta 56:79–90. https://doi.org/10.1016/S0039-9140(01)00537-9
20. Dobričić V, Marković B, Nikolic K, Savić V, Vladimirov S, Čudina O (2014) 17β-carboxamide steroids–in vitro prediction of human skin permeability and retention using PAMPA technique. Eur J Pharm Sci 52:95–108. https://doi.org/10.1016/j.ejps.2013.10.017
21. Cortes C, Vapnik V (1995) Support-vector networks. Mach Learn 20:273–297. https://doi.org/10.1007/BF00994018
22. Heitkamp K, Bajorath J (2014) Support vector machines for drug discovery. Expert Opin Drug Discovery 9:93–104. https://doi.org/10.1517/17460441.2014.866943

23. Vucicevic J, Popovic M, Nikolic K, Filipic S, Obradovic D, Agbaba D (2017) Use of biopartitioning micellar chromatography and RP-HPLC for the determination of blood–brain barrier penetration of α-adrenergic/imidazoline receptor ligands, and QSPR analysis. SAR QSAR Environ Res 28:235–252. https://doi.org/10.1080/1062936X.2017.1302506
24. Ojha PK, Roy K (2011) Comparative QSARs for antimalarial endochins: importance of descriptor-thinning and noise reduction prior to feature selection. Chemometr Intell Lab 109:146–161. https://doi.org/10.1016/j.chemolab.2011.08.007
25. Snedecor GW, Cochran WG (1967) Statistical methods. Oxford & IBH, New Delhi
26. Marshall GR (1994) Binding-site modeling of unknown receptors. In: Kubinyi H (ed) 3D QSAR in drug design – theory, methods and applications. ESCOM, Leiden, pp 80–116
27. Tropsha A (2010) Best practices for QSAR model development, validation and exploitation. Mol Inf 29:476–488. https://doi.org/10.1002/minf.201000061
28. Eriksson L, Johansson E, Kettaneh-Wold N, Trygg J, Wikstrom C, Wold S (2001) Multi-and Megavariate Data Analysis Basic Principles And Applications I, 2nd edn. Umetrics Academy, Umeå
29. Kuchar M, Kraus E, Jelínková M (1991) Influence of mobile phase composition on evaluation of lipophilicity by partition chromatography. J Chromatogr 557:399–411
30. Starek M, Komsta Ł, Krzek J (2013) Reversed-phase thin-layer chromatography technique for the comparison of the lipophilicity of selected non-steroidal anti-inflammatory drugs. J Pharm Biomed Anal 85:132–137. https://doi.org/10.1016/j.jpba.2013.07.017
31. Tesic ZL, Milojkovic-Opsenica DM (2013) TLC determination of drug lipophilicity. In: Komsta L, Waksmundzka-Hajnos M, Sherma J (eds) Thin Layer Chromatography in Drug Analysis. CRC Press, Taylor & Francis Group, Boca Raton, pp 225–246
32. Biagi GL, Barbaro AM, Sapone A, Recanatini M (1994) Determination of lipophilicity by means of reversed-phase thin-layer chromatography: I. Basic aspects and relationship between slope and intercept of TLC equations. J Chromatogr A 662:341–361. https://doi.org/10.1016/0021-9673(94)80521-0
33. Rageh AH, Atia NN, Abdel-Rahman HM (2017) Lipophilicity estimation of statins as a decisive physicochemical parameter for their hepato-selectivity using reversed-phase thin layer chromatography. J Pharm Biomed Anal 142:7–14. https://doi.org/10.1016/j.jpba.2017.04.037
34. Morak B, Nowak M, Pluta K (2007) Determination of the lipophilicity parameters R M0 and log P of new azaphenothiazines by reversed-phase thin-layer chromatography. J Liq Chromatogr R T 30:1845–1854
35. Ciura K, Fedorowicz J, Andrić F, Greber KE, Gurgielewicz A, Sawicki W, Sączewski J (2019) Lipophilicity determination of quaternary (fluoro) quinolones by chromatographic and theoretical approaches. Int J Mol Sci 20:E5288. https://doi.org/10.3390/ijms20215288
36. Constantinescu T, Lungu CN, Lung I (2019) Lipophilicity as a central component of drug-like properties of chalchones and flavonoid derivatives. Molecules 24:E1505. https://doi.org/10.3390/molecules24081505
37. Ciura K, Nowakowska J, Pikul P, Struck-Lewicka W, Markuszewski MJ (2015) A comparative quantitative structure-retention relationships study for lipophilicity determination of compounds with a phenanthrene skeleton on cyano-, reversed phase-, and normal phase-thin layer chromatography stationary phases. J AOAC Int 98:345–353. https://doi.org/10.5740/jaoacint.14-187
38. Burden FR (1989) Molecular identification number for substructure searches. J Chem Inf Comput Sci 29:225–227. https://doi.org/10.1021/ci00063a011
39. Burden FR (1997) A chemically intuitive molecular index based on the eigenvalues of a modified adjacency matrix. Quant Struct-Act Relat 16:309–314. https://doi.org/10.1002/qsar.19970160406
40. Kang YK, Jhon MS (1982) Additivity of atomic static polarizabilities and dispersion coefficients. Theor Chim Acta 61:41–48. https://doi.org/10.1007/BF00573863
41. Takaku T, Nagahori H, Sogame Y, Takagi T (2015) Quantitative structure–activity relationship model for the fetal–maternal blood concentration ratio of chemicals in humans. Biol Pharm Bull 38:930–934. https://doi.org/10.1248/bpb.b14-00883
42. Hajimahdi Z, Safizadeh F, Zarghi A (2016) QSAR analysis for some 1, 2-benzisothiazol-3-one derivatives as caspase-3 inhibitors by stepwise MLR method. Iran J Pharm Res IJPR 15:439–448
43. Madden JC (2011) Introduction to QSAR and other in silico methods to predict toxicity. In: Cronin MTD, Madden JC (eds) In silico toxicology: principles and applications, RSC Publishing, pp 11-30