AKADÉMIAI KIADÓ

Acta Chromatographica

DOI:
10.1556/1326.2022.01096


ORIGINAL RESEARCH PAPER

# High-performance liquid chromatography evaluation of lipophilicity and QSRR modeling of a series of dual DNA gyrase and topoisomerase IV inhibitors

VLADIMIR DOBRIČIĆ[1]*, JELENA SAVIĆ[1], TIHOMIR TOMAŠIČ[2], MARTINA DURCIK[2], NACE ZIDAR[2], LUCIJA PETERLIN MAŠIČ[2], JANEZ ILAŠ[2], DANIJEL KIKELJ[2] and OLIVERA ČUDINA[1]

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

[2] Department of Pharmaceutical Chemistry, Faculty of Pharmacy, University of Ljubljana, Aškerčeva cesta 7, 1000 Ljubljana, Slovenia

Received: September 6, 2022 • Accepted: November 7, 2022

**ABSTRACT**

Bacterial DNA gyrase and topoisomerase IV control the topological state of DNA during replication and represent important antibacterial drug targets. To be successful as drug candidates, newly synthesized compounds must possess optimal lipophilicity, which enables efficient delivery to the site of action. In this study, retention behavior of twenty-three previously synthesized dual DNA gyrase and topoisomerase IV inhibitors was tested in RP-HPLC system, consisting of C8 column and acetonitrile/phosphate buffer (pH 5.5 and pH 7.4) mobile phase. logD was calculated at both pH values and the best correlation with logD was obtained for retention parameter *φ0*, indicating that this RP-HPLC system could be used as an alternative to the shake-flask determination of lipophilicity. Subsequent QSRR analysis revealed that intrinsic lipophilicity (*logP*) and molecular weight (*bcutm13*) have a positive, while solubility (*bcutp3*) has a negative influence on this retention parameter.

**KEYWORDS**

RP-HPLC, DNA gyrase and topoisomerase IV inhibitors, lipophilicity prediction, PLS, MLR and SVM models

## 1. INTRODUCTION

The increasing emergence of pathogenic bacteria resistant to antibacterial drugs is a serious threat to global health because commonly accessible antibiotics will no longer be effective in treating these infections. To overcome the problem of bacterial resistance, many studies continue to investigate this field using modern approaches, particularly through multi-targeting as a promising tactic [1].

DNA gyrase and topoisomerase IV catalyze changes in DNA topology by breaking and rejoining double-stranded DNA [2]. Both enzymes modify the topological state of DNA, which is vital to DNA replication, repair, and decatenation and this is essential for cell viability. DNA gyrase is involved in the negative supercoiling of DNA during replication, whereas topoisomerase IV is involved in the decatenation during DNA replication. DNA gyrase is a heterotetrameric protein consisting of two GyrA and two GyrB subunits ($A_2B_2$), while topoisomerase IV is composed of two ParC and two ParE subunits ($C_2E_2$) that are homologous to GyrA and GyrB, respectively. The GyrA and ParC subunits are involved in

*Corresponding author.
E-mail: vladimir@pharmacy.bg.ac.rs

DNA transit, while the GyrB and ParE subunits contain ATPase domains [3]. Because of the structural similarities between DNA gyrase and topoisomerase IV, dual targeting is possible in most bacteria, which prolongs the onset of resistance development and makes these two enzymes attractive targets for discovering novel antibacterial drugs [4]. Novel synthetic classes of GyrB and ParE inhibitors have been reported in recent years [5], which have included derivatives of benzimidazoles [6], pyrrolamides [7–9], pyrrolopyrimidines [10], pyridylureas [11] and pyrazolopyridones [12, 13]. Despite all the efforts to discover dual-targeting GyrB/ParE inhibitors, none have advanced into the clinic.

Careful attention to physicochemical properties of newly synthesized compounds must be paid because it can improve their delivery to the site of action and their biological activity [14]. Lipophilicity has long been considered a predictor of a drug's successful passage through preclinical and clinical development. It contributes to the ADMET (absorption, distribution, metabolism, excretion, and toxicity) properties and biological activity of drug candidates. Increasing evidence suggests that monitoring lipophilicity may contribute significantly to the overall quality of drug candidates at different stages of drug discovery [15]. Shake-flask is the traditional method for the determination of lipophilicity, but it is time-consuming, requires large amounts of tested substances and cannot provide very reliable values when logP > 3. Reversed-phase high-performance liquid chromatography (RP-HPLC) represents a reliable alternative to the shake-flask method for the lipophilicity evaluation. In contrast to the shake-flask method, it requires low amounts of tested compounds and enables the determination of a wide range of logP values ($-3 < \log P < 8$) [16]. According to the literature overview, RP-HPLC retention data were used for the lipophilicity evaluation of various classes of pharmacologically active compounds in the early phases of drug discovery [17–19].

The aim of this study was to investigate the retention properties of a selected group of twenty-three dual DNA gyrase and topoisomerase IV inhibitors using RP-HPLC, to select the most reliable RP-HPLC retention parameter for the lipophilicity evaluation, and to identify the structural properties that most influence their retention.

## 2. EXPERIMENTAL

### 2.1. Materials and reagents

Acetonitrile HPLC purity (JT Baker, Deventer, Netherlands), sodium hydrogenphosphate (Sigma Aldrich, Steinheim, Germany), phosphoric acid (Sigma Aldrich, Steinheim, Germany) and deionized water (TKA water purification system, Niederelbert, Germany) were used for the mobile phase preparation. Dimethyl sulfoxide, used for the preparation of stock solutions, was purchased from Fisher (Loughborough, UK).

The design, synthesis and biological evaluation of dual DNA gyrase and topoisomerase IV inhibitors (Fig. 1) tested in this study were previously published [8, 20, 21]. Tested compounds were dissolved in dimethyl sulfoxide to prepare

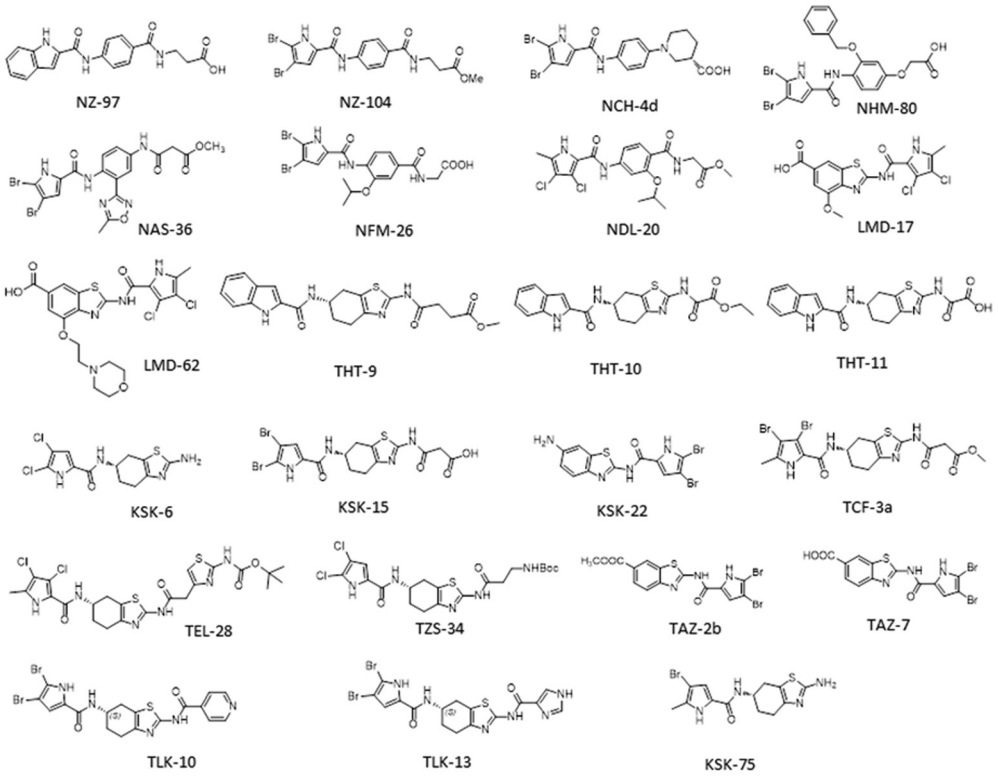


*Fig. 1.* Chemical structures of tested compounds

stock solutions (1 mg $mL^{-1}$), which were then diluted with the mobile phase (see below) to obtain working solutions (0.01 mg $mL^{-1}$).

## 2.2. RP-HPLC analysis

Retention behavior was tested on HP 1100 HPLC chromatograph, using the Zorbax Eclipse Plus C8 column (150 × 4.6 mm, 5 µm particle size; Agilent technologies, USA). Phosphate buffer was prepared by dissolving sodium hydrogenphosphate in deionized water and pH was adjusted to 5.5 or 7.4 by adding phosphoric acid. Mobile phase consisted of acetonitrile and phosphate buffer (pH 5.5 or 7.4). Retention factor ($k$) of each compound was determined with the mobile phase containing 50% of acetonitrile (v/v). According to these results, four different ratios of acetonitrile and phosphate buffer were then selected for each compound, making sure that $k$ is higher than 0.3 (content of acetonitrile ranged from 20 to 65%, in 5% increments, v/v). Column temperature was set to 25 °C, flow rate 1 mL $min^{-1}$, injection volume 20 µL and detection was performed at 254 nm. For each compound, retention factor was calculated and *logk* values were plotted against percentage of acetonitrile ($\varphi$). Following retention parameters were then calculated: $logk_w$ (y-axis intercept), $S$ (slope) and $\varphi 0$.

$$\boldsymbol{logk = logkw + S\varphi} \quad (1)$$

$$\boldsymbol{\varphi 0 = -logkw/S} \quad (2)$$

## 2.3. Calculation of molecular descriptors and logD

Molecular descriptors were calculated using the web-based platform ChemDes [22], which allows the calculation of more than 3,000 molecular descriptors. In this study, 3D Chemopy descriptors were calculated and after deletion of those without variance, 479 molecular descriptors were retained for modeling. ChemDes uses the MOPAC software as default to optimize each molecule. Prior to the descriptor calculation, the dominant form of each molecule at pH 5.5 was determined using MarvinSketch 21.4.0. logD was calculated in MarvinSketch 21.4.0 at two pH values (5.5 and 7.4) by the consensus method, which utilizes ChemAxon and Klopman's models and the PhysProp database. For these calculations, electrolyte concentrations ($Na^+$, $K^+$ and $Cl^-$) were set to the default values [23].

## 2.4. QSRR modelling

Descriptor selection as well as multiple linear regression (MLR), partial least squares (PLS) and support vector machine (SVM) modelling were performed in Statistica 13.3 [24].

Prior to modelling, the number of descriptors has to be reduced and only the most relevant ones should be retained. Molecular descriptors can be selected using various approaches, such as genetic algorithm [25], principal component analysis [26] or stepwise MLR [27, 28]. In this study, forward stepwise MLR was used for the selection of descriptors prior to model creation. In forward stepwise MLR, descriptors are included in the model sequentially and evaluated at each step. Descriptors are retained or removed according to specified criteria (F to enter and F to remove). In this study, F to enter was 8, whereas F to remove was 3. The following descriptors were selected for SVM and MLR modelling: *logP*, *bcutp3* and *bcutm13*. For the PLS modelling, nine most influential descriptors were selected following the same procedure (*MoRSEU14*, *RDFC20*, *DPSA1*, *E2e*, *bcutv1*, *MATSv5*, *logP*, *bcutp3* and *bcutm13*).

The quantitative structure-retention relationship (QSRR) studies were performed to investigate the relationships between $\varphi_0$ (dependent variable) of tested compounds and their calculated molecular descriptors (independent variables). In order to perform a relevant comparison between different methodologies used to build QSRR models, the same training and test sets were prepared. For MLR($\varphi 0$), PLS($\varphi 0$) and SVM ($\varphi 0$), the test set consisted of six compounds (NDL-20, NHM-80, TAZ-7, TCF-3a, THT-9 and THT-11), while remaining derivatives formed the training set. The test set was formed with the aim to cover structural and physico-chemical diversity of all tested compounds and so the $\varphi_0$ of these compounds were evenly distributed over the entire range of $\varphi_0$ values.

MLR was applied to assess the linear relationship between selected molecular descriptors and $\varphi 0$. In this study, the standard MLR model building method was applied.

PLS modeling is useful when analyzing data with collinear, noisy and numerous descriptors. The optimal number of PLS components was selected by analyzing the $R^2(Y)$ value and the cumulative $R^2(Y)$ value of each component. Scaled regression coefficients were used to evaluate the influence of the descriptors on the created model and for their ranking. The final PLS ($\varphi 0$) model consisted of three components.

SVM was developed as a binary classification tool [29]. However, in recent years, it has also been used as a nonlinear method in QSAR and QSPR modelling [30, 31]. In this study, the optimal SVM($\varphi 0$) model was obtained using radial basis function (RBF) Kernel type and regression type 1. Gamma value was set to 0.25, while optimal capacity (C) and epsilon (ε) values were 8 and 0.1, respectively. The model consisted of 11 support vectors (6 bounded).

***2.4.1. Evaluation of quality of created QSRR models.*** Following statistical parameters were calculated and used for the evaluation of the quality of created QSRR models: RMSEE, RMSEP, the F ratio, the $P$ value, $r$, $Q^2$ (Eq. (3)) and $R^2_{pred}$ (Eq. (4)).

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

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

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

RMSEE value represents root mean squared error of estimation, while RMSEP is defined as the root mean squared error of prediction. These errors were calculated for the

training and the test set, respectively. $Q^2$ is an internal validation parameter calculated for the training set according to Eq. (3) and leave-one-out (LOO) procedure [32, 33]. In this equation, $\overline{Y}_{training}$ is the average value, whereas $Y_{obs(training)}$ is an observed $\varphi 0$ value of the training set compounds. PRESS was calculated after the completion of the LOO procedure, according to Eq. (5). In this equation $e_{(i)}$ represents the difference between the observed and predicted $\varphi 0$ values. $R^2_{pred}$ is an external validation parameter calculated for the test set according to Eq. (4), which is used to assess predictive potential of a model for compounds that are structurally different from the training set [34]. In this equation $Y_{obs(test)}$ is an observed value of $\varphi 0$ of a test set compound, whereas $\overline{Y}_{training}$ is the average $\varphi 0$ value of the training set compounds. Values of $Q^2$ and $R^2_{pred}$ higher than 0.5 indicate good predictive potential of the model [32, 35, 36]. The *F*-test is based on the ratio MS Regression/MS Residual and evaluates the significance of the model. The *P*-value indicates the probability level where a model with this F-value may be the result of just chance. The model is considered statistically significant if the *P*-value is lower than 0.05 [36].

## 3. RESULTS AND DISCUSSION

### 3.1. RP-HPLC analysis

The tested compounds contain ionizable groups and therefore, effective lipophilicity (logD) is a more appropriate lipophilicity parameter than logP (logP represents intrinsic lipophilicity, i.e. lipophilicity of unionized form of a compound). The pH values that were chosen in this study were 5.5 and 7.4 because they are the most relevant for the pharmacokinetic (gastrointestinal absorption, distribution and elimination) and pharmacological behavior of drugs. For the initial lipophilicity estimation, logD was calculated at both pH values using MarvinSketch. There are several highly lipophilic compounds with logD > 3 (TEL-28, TAZ-2b, NAS-36, TLK-10, KSK-22 and TCF-3a). Therefore, the shake-flask method might give unreliable results, which justifies the development of HPLC method for the lipophilicity evaluation of these compounds.

Retention parameters ($logk_w$, $S$ and $\varphi 0$), as well as calculated logD values of the tested compounds are presented in Table 1. Due to high correlations between $logk$ and $\varphi$ at both pH values (calculated correlation coefficients were from 0.9837 to 0.9999), $logk_w$ could be calculated by the y-axis extrapolation of these curves.

For the development of HPLC method, C8 column was chosen due to the similarity of its hydrophobicity with the hydrophobicity of octanol alkyl chain (octanol is the most frequently used organic solvent in the shake-flask method). It was also expected that the analysis would be faster than with the C18 column. Phosphate buffer was used for the adjustment of pH values (pH 5.5 and pH 7.4) of the mobile phases because it is present in biological fluids.

Of the three retention parameters presented in Table 1 ($logk_w$, $S$ and $\varphi_0$), $\varphi 0$ showed the highest correlation with

*Table 1.* RP-HPLC retention parameters and calculated logD values

| Compound | pH 5.5 | | | | pH 7.4 | | | |
|---|---|---|---|---|---|---|---|---|
| | $logk_w$[a] | $S$ | $\varphi 0$ | logD | $logk_w$ | $S$ | $\varphi 0$ | logD |
| TEL-28 | 2.96 ± 0.10 | −0.044 ± 0.002 | 67.48 | 5.07 | 2.98 ± 0.12 | −0.044 ± 0.002 | 67.89 | 5.07 |
| TAZ-2b | 3.15 ± 0.11 | −0.050 ± 0.002 | 63.55 | 4.13 | 3.61 ± 0.23 | −0.064 ± 0.005 | 56.55 | 4.13 |
| KSK-75 | 1.95 ± 0.10 | −0.041 ± 0.002 | 47.26 | 2.05 | 3.07 ± 0.27 | −0.064 ± 0.007 | 47.88 | 2.09 |
| NFM-26 | 3.88 ± 0.16 | −0.096 ± 0.005 | 40.26 | −0.36 | 3.88 ± 0.14 | −0.103 ± 0.004 | 37.68 | −1.04 |
| NAS-36 | 3.05 ± 0.13 | −0.050 ± 0.003 | 61.34 | 3.24 | 3.39 ± 0.28 | −0.054 ± 0.005 | 62.39 | 3.24 |
| LMD-17 | 2.19 ± 0.11 | −0.053 ± 0.003 | 41.28 | 1.90 | 3.09 ± 0.10 | −0.090 ± 0.003 | 34.37 | 0.45 |
| TLK-10 | 2.74 ± 0.15 | −0.053 ± 0.003 | 51.84 | 3.07 | 3.79 ± 0.41 | −0.075 ± 0.010 | 50.61 | 3.06 |
| LMD-62 | 2.43 ± 0.13 | −0.067 ± 0.003 | 36.28 | 0.62 | 3.19 ± 0.09 | −0.096 ± 0.003 | 33.39 | 0.21 |
| KSK-22 | 2.47 ± 0.10 | −0.045 ± 0.002 | 54.52 | 3.30 | 3.35 ± 0.35 | −0.065 ± 0.008 | 51.44 | 3.29 |
| THT-11 | 2.71 ± 0.08 | −0.081 ± 0.003 | 33.57 | −1.02 | 2.77 ± 0.06 | −0.086 ± 0.002 | 32.04 | −1.35 |
| NZ-104 | 2.71 ± 0.11 | −0.052 ± 0.003 | 52.01 | 2.24 | 2.84 ± 0.08 | −0.058 ± 0.002 | 49.38 | 2.24 |
| THT-9 | 2.99 ± 0.13 | −0.059 ± 0.003 | 50.44 | 2.18 | 2.57 ± 0.12 | −0.050 ± 0.003 | 51.19 | 2.18 |
| TZS-34 | 2.88 ± 0.15 | −0.049 ± 0.003 | 58.47 | 2.92 | 2.94 ± 0.14 | −0.052 ± 0.003 | 56.83 | 2.92 |
| NHM-80 | 3.21 ± 0.20 | −0.072 ± 0.005 | 44.80 | 1.82 | 3.99 ± 0.17 | −0.097 ± 0.005 | 41.24 | 0.73 |
| NZ-97 | 2.51 ± 0.07 | −0.076 ± 0.002 | 33.13 | 0.19 | 2.55 ± 0.05 | −0.084 ± 0.002 | 30.52 | −1.38 |
| KSK-6 | 2.37 ± 0.12 | −0.052 ± 0.003 | 45.16 | 2.01 | 2.04 ± 0.10 | −0.044 ± 0.002 | 46.56 | 2.04 |
| NCH-4d | 2.85 ± 0.13 | −0.070 ± 0.004 | 40.91 | 2.44 | 2.99 ± 0.09 | −0.087 ± 0.003 | 34.20 | 0.78 |
| NDL-20 | 2.95 ± 0.11 | −0.045 ± 0.002 | 66.05 | 2.79 | 2.66 ± 0.10 | −0.039 ± 0.002 | 67.62 | 2.79 |
| TCF-3a | 2.97 ± 0.12 | −0.054 ± 0.002 | 54.96 | 3.20 | 3.21 ± 0.17 | −0.059 ± 0.004 | 53.97 | 3.20 |
| THT-10 | 2.87 ± 0.13 | −0.058 ± 0.003 | 49.16 | 2.91 | 3.06 ± 0.13 | −0.065 ± 0.003 | 46.74 | 2.90 |
| TLK-13 | 2.65 ± 0.15 | −0.056 ± 0.004 | 46.83 | 2.54 | 2.63 ± 0.14 | −0.056 ± 0.003 | 46.85 | 2.51 |
| TAZ-7 | 2.69 ± 0.10 | −0.071 ± 0.003 | 37.98 | 2.14 | 2.76 ± 0.04 | −0.086 ± 0.001 | 31.97 | 0.55 |
| KSK-15 | 2.86 ± 0.10 | −0.081 ± 0.003 | 35.27 | 0.62 | 3.02 ± 0.03 | −0.090 ± 0.001 | 33.57 | −0.94 |

[a] four different ratios of acetonitrile and phosphate buffer were used for the construction of the curves; all $logk_w$ values were calculated by the extrapolation method (correlation coefficients between $logk$ and $\varphi$ were from 0.9837 to 0.9999)

logD at both pH values ($r$ = 0.85 and 0.89 with logD calculated at pH 5.5 and pH 7.4, respectively). Correlation coefficients between $logk_w$ and logD were 0.01 (pH 5.5) and 0.07 (pH 7.4), while correlation coefficients between $S$ and logD were 0.80 (pH 5.5) and 0.79 (pH 7.4). Therefore, *φ0* could be considered the most reliable RP-HPLC parameter for logD prediction at both pH values. This retention parameter was introduced by Valko and Slegel [37] and it represents the volume fraction of organic modifier in the mobile phase at which equal partitioning of the solute between the mobile and stationary phases is obtained ($k$ = 1, $logk$ = 0). It relies on the pH, temperature and organic modifier, while it does not rely on column type and length, flow rate and mobile phase composition. It also represents concentration of the organic modifier in the mobile phase resulting in the retention time ($t_R$) that is double the dead time ($t_0$). Therefore, it can be accurately estimated which is considered a distinguishable advantage over extrapolating of logk values to pure water mobile phase [38]. This parameter was also proved to be more suitable for the estimation of lipophilicity of sets of structurally unrelated compounds [39, 40]. Although introduced almost thirty years ago, this parameter is still used for the estimation of lipophilicity of various classes of bioactive compounds [38, 41]. Correlations between *φ0* and logD at both pH values were presented in Fig. 2.

Compound NDL-20, with an unexpectedly high *φ0* at both pH values, was an outlayer. At both pH values, compounds with the highest *φ0* values were TEL-28, TAZ-2b and NAS-36, and their calculated logD values were higher than 3.2. TEL-28, TAZ-2b and NAS-36 contain ester groups, halogen atoms (chlorine or bromine) and three or four rings. These moieties increase lipohilicity and consequently increase *φ0* values. At pH 5.5, compounds with the lowest *φ0* values were NZ97 and THT11, and their logD values were lower than 0.19. At pH 7.4, compounds with the lowest *φ0* values were also NZ-97 and THT-11, as well as TAZ-7 and KSK-15, for which calculated logD values were lower than 0.55. NZ-97, THT-11, TAZ-7 and KSK-15 contain carboxylic acid moiety, and as pH increases, the proportion of the ionized form of these compounds also increases, which could explain their low *φ0* values determined at pH 7.4.

## 3.2. QSRR modelling

Due to high correlation between *φ0* determined at pH = 5.5 and pH = 7.4 ($r$ = 0.98), only *φ0* determined at pH = 5.5 was used for further QSRR modelling.

Results of the validation of created models are presented in Table 2.

PLS (*φ0*) model cannot be considered reliable due to the high discrepancy between RMSEE and RMSEP. This means that using this model, *φ0* values of compounds that are structurally different from those in the training set cannot be reliably predicted. The remaining QSRR models (MLR (*φ0*) and SVM (*φ0*)) pass all validation tests. Although RMSEE was higher (2.75 *vs* 2.48), due to lower RMSEP, lower difference between RMSEE and RMSEP, higher $Q^2$, $r$ and $R^2_{pred}$, MLR (*φ0*) could be considered more reliable model for the logD prediction. Descriptors which form both MLR (*φ0*) and SVM (*φ0*) models were *logP*, *bcutp3* and *bcutm13*. Cross-correlation between these decscriptors was not higher than 0.41 (*logP* and *bcutp3*: $r$ = 0.32; *logP* and *bcutm13*: $r$ = 0.07; *bcutp3* and *bcutm13*: $r$ = 0.41). A rule of thumb in regression analysis is that at least 5 observations per descriptor are needed and the more acceptable ratio would be 10:1 [42]. Therefore, the dataset size in this study (twenty-three compounds) allows the use of three descriptors for reliable QSRR modelling. In addition, satisfactory values of validation parameters for the test set (Table 2) show that the model was not overfitted. Statistical data calculated for the MLR (*φ0*) model are presented in Table 3.

Intrinsic lipophilicity, *logP*, was calculated by the Crippen method. The relationship between this descriptor and *φ0* is linear and positive (Fig. 3) and the increase in lipophilicity results in the increase in *φ0*. This supports our previous observation that applied RP-HPLC systems can be used for reliable prediction of the lipohilicity of dual DNA gyrase and topoisomerase IV inhibitors.

Descriptors *bcutp3* and *bcutm13* belong to the Burden descriptors and are based on polarizability and atomic masses, respectively. Descriptors *bcut* are defined as 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. These descriptors are based on a weighted version of the Burden matrix [43, 44] and the weights are a variety of

*Table 2.* Validation parameters of created models

| | RMSEE | $Q^2$ | RMSEP | $r$ | $R^2_{pred}$ |
|---|---|---|---|---|---|
| SVM (*φ0*) | 2.48 | 0.78 | 4.19 | 0.92 | 0.80 |
| MLR (*φ0*) | 2.75 | 0.86 | 3.77 | 0.94 | 0.88 |
| PLS (*φ0*) | 1.41 | 0.93 | 5.60 | 0.88 | 0.73 |

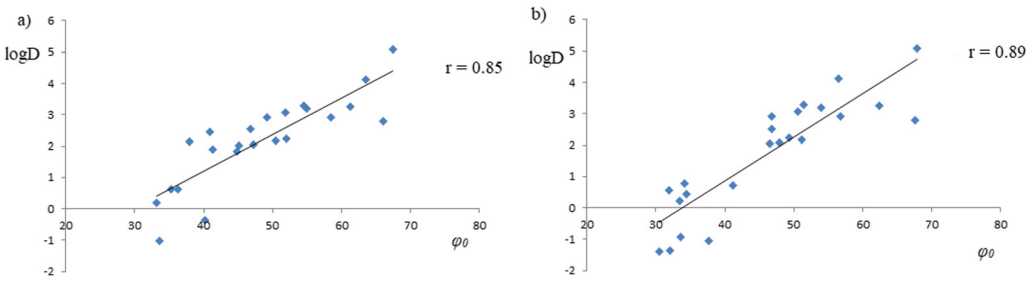


*Fig. 2.* a) Correlation between *φ0* and logD (pH 5.5) and b) correlation between *φ0* and logD (pH 7.4)

*Table 3.* Statistical data (coefficients, standard errors and *P*-values) of the MLR (*φ0*) model

| | Coefficient (B) | Standard error of B | *P*-value |
|---|---|---|---|
| Intercept | 44.351 | 18.519 | 0.032 |
| logP | 8.618 | 0.727 | 0.000 |
| bcutp3 | −14.819 | 7.089 | 0.043 |
| bcutm13 | 26.946 | 9.316 | 0.013 |

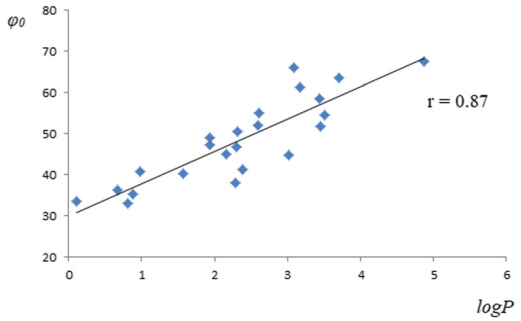


*Fig. 3.* Correlation between *logP* and *φ0*

atomic properties placed along the diagonal of the Burden matrix (atomic weight, partial charge (Gasteiger–Marsili) and polarizability weighting schemes are used) [45].

Descriptor *bcutp3* represents the highest eigenvalue number 3 of the Burden matrix/weighted by atomic polarizabilities which encode information about the polarizability of the molecules. It was shown that the increase in polarizability increases solubility, i.e. that molecular polarizability, both of the solute and solvent, is a major factor that dictates solubility [46–48]. According to the MLR model, this variable has negative impact on *φ0* and the increase in *bcutp3* results in the decrease in *φ0,* which indicates that the water solubility of the molecules is inversely related to *φ0.* Derivatives with the highest values of this descriptor are NHM-80 and TEL-28, while those with the lowest values are KSK-22 and NDL-20.

Descriptor *bcutm13* belongs to the Burden descriptors based on atomic mass. The relationship between *bcutm13* and *φ0* is positive, according to the positive value of the coefficient in the MLR model. This also means that the molecular weight of tested compounds positively affects their retention in the applied RP-HPLC systems. Compounds with the highest value of this descriptor are LMD-62 and TEL-28 (molecular weights of these compounds are higher than 530), while those with the lowest values are TAZ-7 and KSK-22 (molecular weights of these two compounds are lower than 450). It is difficult to estimate *φ0* only on the basis of descriptor values. For example, LMD-62 has low *φ0* despite the high value of *bcutm13.* This could be explained by the high value of *bcutp3*, which has a negative impact on *φ0.* TEL-28 also has high *bcutm13* value and despite high value of *bcutp3* (which has a negative effect on *φ0*), this compound has high value of *φ0.* Compound KSK-22 has low values of both *bcutm13 and bcutp3* and a high value of *logP*. Its high *φ0* could be explained by the greater influence of *bcutp3* and *logP* on *φ0* compared to the influence of *bcutm13.* Therefore, all selected descriptors have to be taken into account, which underlines the importance of use of the created MLR (*φ0*) model for valid prediction of *φ0.*

## 4. CONCLUSION

The retention behavior of a group of twenty-three dual DNA gyrase and topoisomerase IV inhibitors was tested in an RP-HPLC system, employing a C8 column and a mobile phase consisiting of acetonitrile/phosphate buffer (pH was adjusted to 5.5 or 7.4). The HPLC parameter *φ0* had the highest correlation with logD values calculated at pH 5.5 and 7.4, which indicates that the applied HPLC system could be used as an alternative to the shake-flask method for the evaluation of logD. QSRR analysis showed that lipophilicity and molecular weight have a positive effect, while solubility has a negative effect on retention parameter *φ0.* It was also shown that prediction of *φ0* is possible by the use of developed MLR (*φ0*) model. These results could facilitate the design of new dual DNA gyrase and topoisomerase IV inhibitors with more optimal lipophilicity and biological properties.

## ACKNOWLEDGMENT

This research was funded by the Ministry of Education, Science and Technological Development, Republic of Serbia through the Grant Agreement with University of Belgrade-Faculty of Pharmacy No: 451-03-68/2022-14/200161, Slovenian Research Agency (Grant No. P1-0208) and the project of bilateral cooperation between Republic of Slovenia and Republic of Serbia (BI-RS/18-19-034).

## REFERENCES

1. East, S. P.; Silver, L. L. *Expert Opin. Drug Discov.* **2013**, *8*, 143.
2. Sissi, C.; Palumbo, M. *Cell. Mol. Life Sci.* **2010**, *67*, 2001.
3. Champoux, J. J. *Annu. Rev. Biochem.* **2001**, *70*, 369.
4. Sanyal, G.; Doig, P. *Expert Opin. Drug Dis.* **2012**, *7*, 327.
5. Durcik, M.; Tomašič, T.; Zidar, N.; Zega, A.; Kikelj, D.; Peterlin Mašič, L.; Ilaš J. *Expert Opin. Therap. Pat.* **2019**, *29*, 171.
6. Charifson, P. S.; Grillot, A. L.; Grossman, T. H.; Parsons, D. J.; Badia, M.; Bellon, S.; Deininger, D. D.; Drumm E. J.; Gross, C. H.; LeTiran, A.; Liao, Y.; Mani, N.; Nicolau, D. P.; Perola, E.; Ronkin, S.; Shannon, D.; Swenson, L. L.; Tang, Q.; Tessier, P. R.; Tian, S.; Trudeau, M.; Wang, T.; Wei, Y; Hong Zhang, H.; Stamos, D. *J. Med. Chem.* **2008**, *51*, 5243.
7. Zidar, N.; Tomašič, T.; Macut, H.; Sirc, A.; Brvar, M.; Montalvao, S.; Tammela, P.; Ilaš, J.; Kikelj, D. *Eur. J. Med. Chem.* **2016**, *117*, 197.
8. Zidar, N.; Macut, H.; Tomašič, T.; Brvar, M.; Montalvao, S.; Tammela, P.; Solmajer, T.; Peterlin-Mšič, L.; Ilaš, J.; Kikelj, D. *J. Med. Chem.* **2015**, *58*, 6179.

9. Tomašič, T.; Mirt, M.; Barančokova, M.; Ilaš, J.; Zidar, N.; Tammela, P.; Kikelj, D. *Bioorg. Med. Chem.* **2017**, *25*, 338.
10. Tari, L. W.; Trzoss, M.; Benssen, D. C.; Li, X.; Chena, Z.; Lama, T.; Zhang, J.; Creightona, C. J.; Cunningham, M. L.; Kwana, B.; Stidhama, M.; Shawa, K. J.; Lightstone, F. C.; Wong, S. E.; Nguyen, T. B.; Nix, J.; Finna, J. *Bioorg. Med. Chem. Lett.* **2013**, *23*, 1529.
11. Basarab, G. S.; Manchester, J. I.; Bist, S.; Boriack-Sjodin, A.; Dangel, B.; Illingworth, R.; Sherer, B.; Sriram, S.; Uria-Nickelsen, M.; Eakin, A. *J. Med. Chem.* **2013**, *56*, 8712.
12. Mesleh, M. F.; Cross, J. B.; Zhang, J.; Kahmann, J.; Andersen, O. A.; Barker, J.; Cheng, R. K.; Felicetti, B.; Wood, M.; Hadfield, A. T.; Scheich, C.; Moy, T. I.; Yang, Q.; Shotwell, J.; Nguyen, K.; Lippa, B.; Dolle, R.; Ryan, D. M. *Med. Chem. Lett.* **2016**, *26*, 1314.
13. Cross, J. B.; Zhang, J.; Yang, Q.; Mesleh, M. F.; Romero, J. C.; Wang, B.; Bevan, D.; Poutsiaka, K. M.; Epie, F.; Moy, T.; Daniel, A.; Shotwell, J.; Chamberlain, B.; Carter, N.; Andersen, O.; Barker, J.; Ryan, M. D.; Metcalf, C. A.; Silverman, J.; Nguyen, K.; Blaise Lippa, B.; Dolle, R. E. *ACS Med. Chem. Lett.* **2016**, *7*, 374.
14. Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J. *Adv. Drug Deliv. Rev.* **2001**, *46*, 3.
15. Lobo, S. *Expert Opin. Drug Discov.* **2020**, *15*, 261.
16. Kempinska, D.; Chmiel, T.; Kot-Wasik, A.; Mroz, A.; Mazerska, Z.; Namiesnik, J. *TrAC Trends Anal. Chem.* **2019**, *113*, 54.
17. Mrkvičkova, Z.; Kovarikova, P.; Balikova, S.; Klimes, J. J. *J. Pharm. Biomed. Anal.* **2008**, *48*, 310.
18. Szatkowska-Wandas, P.; Koba, M.; Smolinski, G.; Wandas, J. *J. Med. Chem.* **2016**, *12*, 592.
19. Jevric, L. R.; Karadzic, M. Z.; Mandic, A. I.; Podunavac Kuzmanovic, S. O.; Kovacevic, S. Z.; Nikolic, A. R.; Okljesa, A. M.; Sakac, M. N.; Penov Gasi, K. M.; Stojanovic, S. Z. *J. Pharm. Biomed. Anal.* **2017**, *134*, 27.
20. Tomašič, T.; Katsamakas, S.; Hodnik, Ž.; Ilaš, J.; Brvar, M.; Solmajer, T.; Sofia Montalvão, S.; Tammela, P.; Banjanac, M.; Ergović, G.; Anderluh, M.; Peterlin Mašič, L.; Kikelj, D. *J. Med. Chem.* **2015**, *58*, 5501.
21. Durcik, M.; Nyerges, Á.; Skok, Ž.; Skledar, D. G.; Trontelj, J.; Zidar, N.; Ilaš, J.; Zega, A.; Cruz, C. D.; Tammela, P.; Welin, M.; Kimbung, Y. R.; Focht, D.; Benek, O.; Révész, T.; Draskovits, G.; Szili, P. E.; Daruka, L.; Pál, C.; Kikelj, D.; Peterlin Mašič, L.; Tomašič, T. *Eur. J. Med. Chem.* **2021**, *213*, 113200.
22. Dong, J.; Cao, D.; Miao, H.; Liu, S.; Deng, B.; Yun, Y.; Wang, N.; Lu, A.; Zeng, W.; Chen, A. *J. Cheminformatics* **2015**, *7*, 60.
23. MarvinSketch 15.1.26. ChemAxon: Budapest, **2015**, (http://www.chemaxon.com).
24. Statistica, version 13; TIBCO Software Inc.; Palo Alto, USA, **2018**, (http://statistica.io).
25. Gupta, V. K.; Khani, H.; Ahmadi-Roudi, B.; Mirakhorli, S.; Fereyduni, E.; Agarwal, S. *Talanta* **2011**, *83*, 1014.
26. Zhang, Y. X. *Talanta* **2007**, *73*, 68.
27. Gonzalez-Arjona, D.; Lopez-Perez, G.; Gustavo-Gonzalez, A. *Talanta* **2002**, *56*, 79.
28. Dobričić, V.; Marković, B.; Nikolic, K.; Savić, V.; Vladimirov, S.; Čudina, O. *Eur. J. Pharm. Sci.* **2014**, *52*, 95.
29. Cortes, C.; Vapnik, V. *Mach. Learn.* **1995**, *20*, 273.
30. Heitkamp, K.; Bajorath, J. *Expert Opin. Drug. Discov.* **2014**, *9*, 93.
31. Vucicevic, J.; Popovic, M.; Nikolic, K.; Filipic, S.; Obradovic, D; Agbaba, D. *Environ. Res.* **2017**, *28*, 235.
32. Ojha, P. K.; Roy, K. *Chemometr. Intell. Lab.* **2011**, *109*, 146.
33. Snedecor, G. W.; Cochran, W. G. Statistical Methods. Oxford & IBH: New Delhi, **1967**.
34. Marshall, G. R. In 3D QSAR in Drug Design – Theory, Methods and Applications; Kubinyi, H., Ed. ESCOM: Leiden, **1994**; pp. 80–116.
35. Tropsha, A. *Mol. Inf.* **2010**, *29*, 476.
36. Eriksson, L.; Johansson, E.; Kettaneh-Wold, N.; Trygg, J.; Wikstrom, C.; Wold, S. Multi-and Megavariate Data Analysis. Basic Principles and Applications I, 2nd ed.; Umetrics Academy: Umea, **2001**.
37. Valko, K.; Slegel, P. *J. Chromatogr.* **1993**, *631*, 49.
38. Elmansi, H.; Nasr, J. J.; Rageh, A. H.; El-Awady, M. I.; Hassan, G. S.; Abdel-Aziz, H. A.; Belal, F. *BMC Chem.*, **2019**, *13*, 1.
39. Wiczling, P.; Kawczak, P.; Nasal, A.; Kaliszan, R. *Anal. Chem.*, **2006**, *78*, 239.
40. Valko, K. *Trends Anal. Chem.* **1987**, *6*, 214–6.
41. Klose, M. H.; Theiner, S.; Varbanov, H. P.; Hoefer, D.; Pichler, V.; Galanski, M.; Meier-Menches, S. M.; Keppler, B. K. *Inorganics*, **2018**, *6*, 130.
42. Hair, J. F.; Black, W. C.; Babin, B. J.; Anderson, R. E. Multivariate Data Analysis, 7th ed.; Upper Saddle River: Prentice Hall, **2009**.
43. Burden, F. R. *J. Chem. Inf. Comput. Sci.* **1989**, *29*, 225.
44. Burden, F. R. *Quant. Struct-Act. Relat.* **1997**, *16*, 309.
45. Kang, Y. K.; Jhon M. S. *Theor. Chim. Acta.* **1982**, *61*, 41.
46. Abraham, M. H.; Le, J. *J. Pharm. Sci.* **1999**, *88*(9), 868.
47. Cowan, M. J.; Higaki, T.; Jin, R., Mpourmpakis, G. *J. Phys. Chem. C.* **2019**, *123*(32), 20006.
48. David, M. R.; Subbiah, P.; Selvaraj, S. *Curr. Comput. Aided Drug Des.* **2018**, *15*(4), 294.