J Pharm Pharm Sci (www.cspsCanada.org) 16(4) 622 - 647, 2013

Partial Least Square and Hierarchical Clustering in ADMET Modeling: Prediction of Blood – Brain Barrier Permeation of α-Adrenergic and Imidazoline Receptor Ligands

Katarina Nikolic[1], Slavica Filipic[1], Adam Smoliński[2], Roman Kaliszan[3], and Danica Agbaba[1]

[1] Institute of Pharmaceutical Chemistry, Faculty of Pharmacy, University of Belgrade, Vojvode Stepe Belgrade, Serbia. [2] Central Mining Institute, Department of Energy Saving and Air Protection, Katowice, Poland. [3] Department of Biopharmaceutics and Pharmacodynamics, Medical University of Gdańsk, Gdańsk, Poland.

Received, June 26, 2013; Revised, September 18, 2013; Accepted, October 13, 2013; Published, October 30, 2013

ABSTRACT - PURPOSE. Rate of brain penetration (logPS), brain/plasma equilibration rate (logPS-brain), and extent of blood-brain barrier permeation (logBB) of 29 α-adrenergic and imidazoline-receptors ligands were examined in Quantitative-Structure-Property Relationship (QSPR) study. **METHODS.** Experimentally determined chromatographic retention data (logKw at pH 4.4, slope (*S*) at pH 4.4, logKw at pH 7.4, slope (*S*) at pH 7.4, logKw at pH 9.1, and slope (*S*) at pH 9.1) and capillary electrophoresis migration parameters (μ_{eff} at pH 4.4, μ_{eff} at pH 7.4, and μ_{eff} at pH 9.1), together with calculated molecular descriptors, were used as independent variables in the QSPR study by use of partial least square (PLS) methodology. **RESULTS.** Predictive potential of the formed QSPR models, QSPR(logPS), QSPR(logPS-brain), QSPR(logBB), was confirmed by cross- and external validation. Hydrophilicity (Hy) and H-indices (H7m) were selected as significant parameters negatively correlated with both logPS and logPS-brain, while topological polar surface area (TPSA(NO)) was chosen as molecular descriptor negatively correlated with both logPS and logBB. The principal component analysis (PCA) and hierarchical clustering analysis (HCA) were applied to cluster examined drugs based on their chromatographic, electrophoretic and molecular properties. Significant positive correlations were obtained between the slope (*S*) at pH 7.4 and logBB in A/B cluster and between the logKw at pH 9.1 and logPS in C/D cluster. **CONCLUSIONS**. Results of the QSPR, clustering and correlation studies could be used as novel tool for evaluation of blood-brain barrier permeation of related α-adrenergic/imidazoline receptor ligands.

INTRODUCTION

Earlier findings that clonidine and its analogues produce central hypotensive effect by interaction with both the α_2-adrenoreceptors (α_2-AR) and imidazoline receptors (IRs) (1, 2) influenced on development of novel guanidines/imidazolines as multipotent centrally acting antihypertensives (3-5). Until today three subtypes of imidazoline receptors, I_1-IR, I_2-IR, and I_3-IR, were experimentally characterized (6, 7). The I_1-IRs were defined as one of the binding sites involved in the control of blood pressure (8). The I_1-IRs have high affinity for clonidine, moxonidine, and rilmenidine, while the I_2-IRs has high affinity for idazoxan and its analogues (9, 10).

Selective I_1-IR ligands as rilmenidine and moxonidine induce fall in plasma catecholamines, rennin, and antidiuretic hormone (11, 12), which result in increase of renal blood flow, natriuresis, potassium excretion, and diuresis (13).

For the active sites of I_2-IRs were shown to reside on monoamine oxidase-B (MAO-B) (10), well known drug target for several neurological diseases. Therefore the I_2-IRs ligands are examined as novel therapeutic drugs in treatment of various neurological diseases (14). Research on agmatine, an endogenous ligand for all imidazoline receptor(s), has confirmed its implication in mediation of analgesia, stress responses, convulsions, and neuroprotection (15). Since the IRs are involved in many neurophysiologic and pathologic functions the potential for new IRs ligands development is very intriguing (14, 15).

Thus, the main goal of the study was to examine brain penetration of the drugs interacting with IRs/α-ARs and to develop Quantitative Structure-Property Relationship (QSPR) models capable to predict and describe rate and extent of the brain penetration for the related IRs/α-ARs ligands.

Corresponding Author: Katarina Nikolic, Institute of Pharmaceutical Chemistry, Faculty of Pharmacy, University of Belgrade, Belgrade, Serbia; E-mail: knikolic@pharmacy.bg.ac.rs

Initially, we have selected ligands of α-adrenergic and imidazoline-receptors for the study.

Because the I-IR ligands exert additional CNS and diuretic effects, we have decided to include few structurally related CNS drugs (clozapine, maprotiline, mianserin, and olanzapine) and diuretics (amiloride, clopamide, indapamide, and triamterene) in the QSPR study.

Rate of brain penetration (logPS), brain/plasma equilibration rate (logPS-brain), and extent of blood-brain barrier permeation (logBB) of the 29 α-adrenergic/imidazoline receptors ligands and related compounds were used as dependent (Y) variables in the QSPR study.

The QSPR (logPS, logPS-brain, logBB) study of the 29 drugs was performed by use of principal component analysis (PCA) and partial least square (PLS) methodologies.

Chromatographic retention parameters (logKw at pH 4.4, slope (*S*) at pH 4.4, logKw at pH 7.4, slope (*S*) at pH 7.4, logKw at pH 9.1, and slope (S) at pH 9.1), capillary electrophoresis migration parameters (μ_{eff} at pH 4.4, μ_{eff} at pH 7.4, and μ_{eff} at pH 9.1) and computed molecular parameters of the drugs were used as independent variables (X) for the QSPR modeling. The PLS methodology was applied to select molecular parameters of the drugs with the strongest impact on their brain penetration and to create QSPR (logPS, logPS-brain, logBB) models able to predict blood-brain barrier permeation properties of the related α-adrenergic/imidazoline receptors ligands.

Since the examined drugs are structural very diverse principal component analysis (PCA) and hierarchical clustering analysis (HCA) were applied to cluster examined drugs based on their chromatographic, electrophoretic and molecular properties. The PCA/HCA clusters of the drugs were further examined by the correlation study between chromatographic/electrophoretic parameters, as independent variables, and blood-brain permeation parameters, as dependant variables.

The performed QSPR, clustering and correlation studies are first reported theoretical investigation of brain penetration process of the α-adrenergic/imidazoline receptor ligands.

MATHERIALS AND METHODS

Chemicals and reagents

All reagents used were of analytical grade purity. Methanol-HPLC gradient grade, glacial acetic acid (J.T. Baker Deventer, Netherlands), ammonium acetate, sodium hydroxide, sodium dihydrogen phosphate, disodium hydrogen phosphate, (Merck, Darmstadt, Germany), boric acid (Sigma-Aldrich, St. Louis, MO, USA), ammonium hydroxide (Carlo Erba, Milan, Italy) and water-HPLC grade were used to prepare mobile phases and background electrolytes.

The following standards were used: clonidine hydrochloride, moxonidine hydrochloride, guanfacine hydrochloride, brimonidine hydrochloride, efaroxan hydrochloride, idazoxan hydrochloride, rilmenidine hemifumarate, harmane, harmine hydrochloride, tizanidine hydrochloride, triamterene, clopamide, indapamide, naphazoline hydrochloride, xylometazoline hydrochloride, tetrahydrozoline hydrochloride, oxymetazoline hydrochloride, ephedrine hydrochloride, pseudoephedrine hydrochloride, maprotiline hydrochloride, tamsulosin hydrochloride, mianserin hydrochloride, carvedilol, clozapine and olanzapine (Sigma-Aldrich, St. Louis, MO, USA); tramazoline hydrochloride and doxazosin mesilat (Zdravlje, Leskovac, Serbia); amiloride hydrochloride (Galenika, Belgrade, Serbia); phenilephrine hydrochloride (Ivančić i sinovi, Belgrade, Serbia).

SAMPLE PREPARATION

Capillary Electrophoresis

Stock solutions of harmane, harmine, clopamide, indapamide, rilmenidine, mianserin, doxazosin, carvedilol, clozapine, olanzapine and triamterene were prepared in methanol and diluted with water to a final concentration of methanol 2 % for triamterene and 1 % for remaining substances. Brimonidine was dissolved in 0.1% formic acid and the other 17 compounds were dissolved in water. Acetone 2% was used as EOF marker. Samples were prepared in different concentrations, from 5.8 to 60 μg mL^{-1}, depending on their UV response and solubility.

HPLC analysis

Working solutions were prepared by dissolving the substances into the methanol in order to obtain the concentration of 0.7 mg mL^{-1} for rilmenidine and 0.1 mg mL^{-1} for the remaining compounds.

CE equipment

All experiments were carried out on SpectraPhoresis 500-capillary electrophoretic system (Spectra Physics Analytical, USA) equipped with UV detector. Data were recorded and analyzed with ChromQuest software version 4.0 (Thermo Finnigan, USA).

Electrophoretic conditions

Separations were performed using an uncoated fused capillary (31,5 cm_50 mm id, effective length 24 cm, Polymicro Technologies, USA) at 25 ºC, 11 kV, and 200 nm. The new capillary was gradually flushed with 0.1 M NaOH (15 min), water (10 min) and running buffer (10 min). Finally, a 10 min application of the high voltage through the capillary filled with background electrolyte were applied in order to equilibrate the inner surface, stabilize electroosmotic flow and to maintain proper reproducibility of run-to-run injections. Background solutions of constant ionic strength (I=25 mmol/L) were prepared at three different pH values (4.4; 7.4 and 9.1). The samples were injected hydrodynamically three times at each pH (4.4; 7.4 and 9.1). Between runs the capillary was rinsed with the background electrolyte for 1 min.

The effective electrophoretic mobility, μ_{eff}, of the analyte at three different pH values 4.4; 7.4 and 9.1 ($\mu_{eff\ pH\ 4.4}$, $\mu_{eff\ pH\ 7.4}$ and $\mu_{eff\ pH\ 9.1}$ respectively) were calculated and used for QSPR study.

HPLC equipment

A chromatographic system Agilent Technologies 1200 (Wilmington, DE, USA) equipped with on-line degasser, binary pump, column oven and photo diode array detector was used for HPLC analysis. Sample injection was performed through Rheodyne injector valve with a 20 μL sample loop. Data were recorded and analyzed with Agilent's ChemStation software.

HPLC conditions

All chromatographic measurements were performed on XTerra® RP18 column, 4.6 x 100 mm, particle size 3.5 μm (Waters Corporation, Milford, MA, USA). The flow rate was 0.8 ml min^{-1} with UV detection in the range 200-280 nm. The temperature was set at 25°C. The retention parameters were obtained at the isocratic elution mode using at least six mobile phases methanol/buffer with concentration of methanol varying from 75-2%, depending on compounds retention properties. Buffers solutions of constant ionic strength (I=25 mmol/L) were prepared at three different pH values: 4.4, 7.4 and 9.1. Dead volume was measured with KNO_3 as a non-retained marker. The values corresponding to 100% of buffered eluent, log K_w and slope (*S*), were obtained by extrapolation, following the Snyder-Soczewinski equation (16, 17). The $logK_w$ and slope (*S*)of the analyte at three different experimental conditions, methanol/buffer pH 4.4, methanol/buffer pH 7.4 and methanol/buffer pH 9.1 ($logK_{w\ pH\ 4.4}$, slope $(S)_{pH\ 4.4}$, $logK_{w\ pH\ 7.4}$, slope $(S)_{pH\ 7.4}$, $logK_{w\ pH\ 9.1}$, and slope $(S)_{pH\ 9.1}$ respectively) were calculated for all examined compounds and used for QSPR study.

Computational Method

Among analyzed compounds (Figure 1), cyclic amidines and guanidines structures such as moxonidine, clonidine, tizanidine, rilmenidine, brimonidine and tramazoline can exist in two major tautomeric forms (amino and imino).

The dominant amino/imino tautomers were selected by use of the B3LYP/6–31G(d, p) level of the Density Functional Theory (DFT) (18, 19) incorporated in the Gaussian 98 program (20). The selected basis set was proved to be good choice for examination of the related amidines and guanidines (21). Based on the obtained Self Consistent Field Energy (SCF energy), rilmenidine amine was selected as predominant tautomeric form while moxonidine, clonidine, tizanidine, brimonidine and tramazoline exist as more stable imino tautomers.

Calculation of pKa and selection of dominant molecules/cations species at pH 4.4, 7.4, and 9.1, have been performed for 29 α-adrenergic and imidazoline-receptors ligands using the Marvin 5.5.1.0 ChemAxon program (22). The Marvin program defined nitrogens of the analyzed structures with the highest potential to attract proton at analytical pH.

The geometries of the examined ligands (selected tautomers and molecules/cations species at three different pH values) were completely optimized at B3LYP/3–21(d,p) levels of the Density Functional Theory using the Gaussian 98 program (20).

Molecular descriptors

The molecular descriptors, as numerical parameters representing the chemical structures, were calculated for the optimized molecular models. The Density-Functional-Theory/B3LYP/3-21G(d,p) basis set (18-20), the Marvin 5.5.1.0 ChemAxon (22), Chem3D Ultra 7.0.0 programs (23), and Dragon programs (24) were applied for computation of constitutional, physico-chemical, thermodynamic and electronic parameters of the optimized molecular models. Also, quantum chemically-based reactivity descriptors, such as chemical potential (μ), electronegativity (χ), hardness (η), global softness (GS), and electrophilicity index (ω), were added to the set of calculated molecular descriptors of the analyzed data set (25, 26).

QSPR modeling

Pharmacokinetic parameters of the examined 29 α-adrenergic and imidazoline-receptors ligands (logPS, logPS-brain, and logBB), calculated by use of ACD/i-Lab program (27), were used as dependant (Y) variable in the QSPR study. The HPLC retention parameters (logKw and Slope at pH 4.4, 7.4, and 9.1), CE migration parameters (μ_{eff} at pH 4.4, 7.4, and 9.1), and calculated molecular descriptors (18-20, 22-24) of the examined ligands were used as independent (X) variables in the QSPR study by use of the partial least square (PLS) Regression (28).

The Soft Independent Modeling of Class Analogy SIMCA P+ 12.0 program (29) was used for the PLS analysis and QSPR modeling.

Based on the Principal Component Analysis (PCA) plots (t1 *vs*. t2 and t1 *vs.* u1) the data set of 29 compounds is divided on Training Set (22 compounds for QSPR models building) and Verification set (7 compounds for QSPR models validation). Based on the PCA plots were not defined any outlier in the QSPR models.

Partial least square (PLS) regression, recently developed generalization of multiple linear regressions (MLR) (28, 29), has been used for calculation of VIP parameter and QSAR models building. In multilinear modeling, a summary of the importance of each variable (x_k), for both Y and X matrices, is presented as VIP_k parameter. The x-variables with VIP value larger than 1 are the most relevant for explaining the regression model, the x-variables with 1.0>VIP>0.5 are moderately influential, while x-variables with VIP value smaller than 0.5 are not relevant for the model (28, 29).

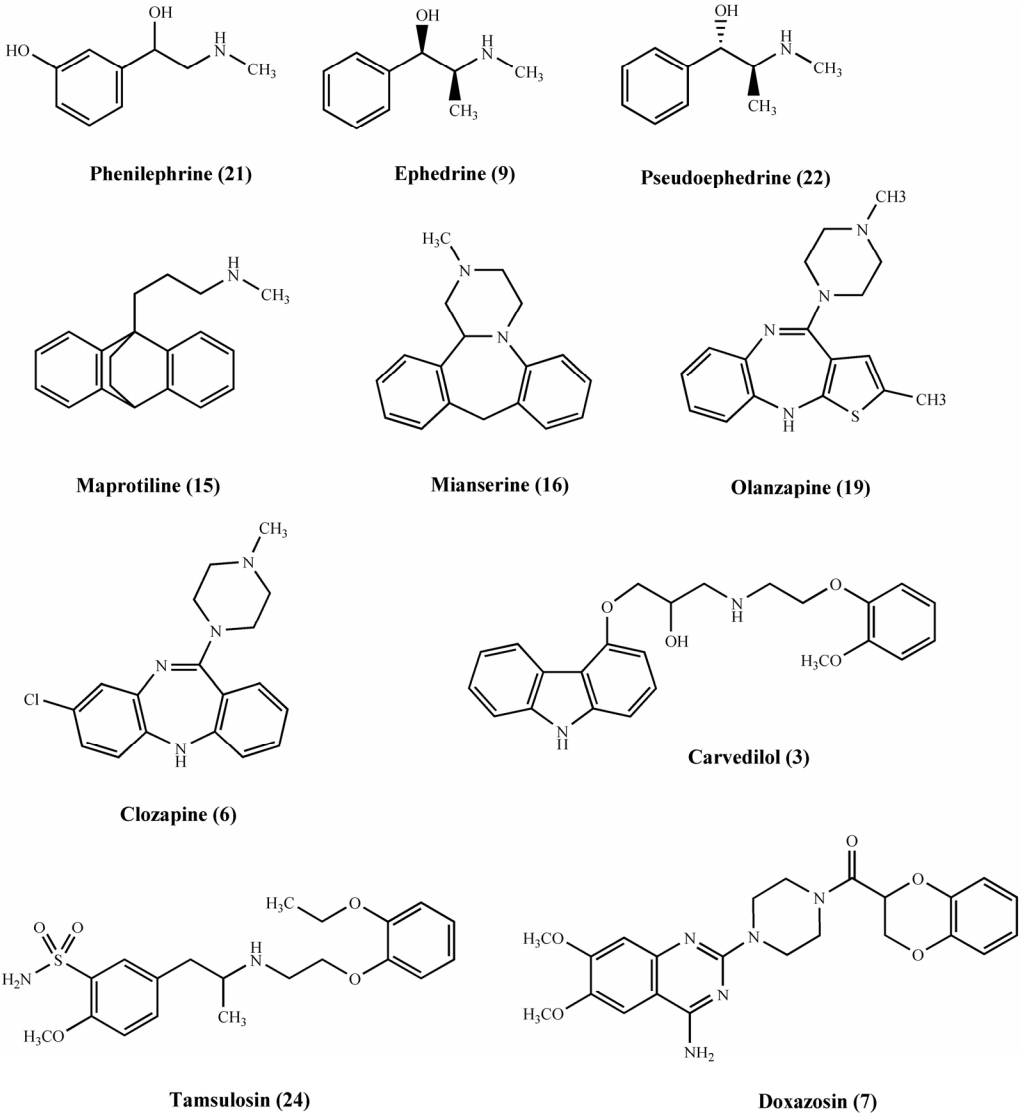

Figure 1. Structural formulas of the examined α-adrenergic/imidazoline receptor ligands and related compounds.

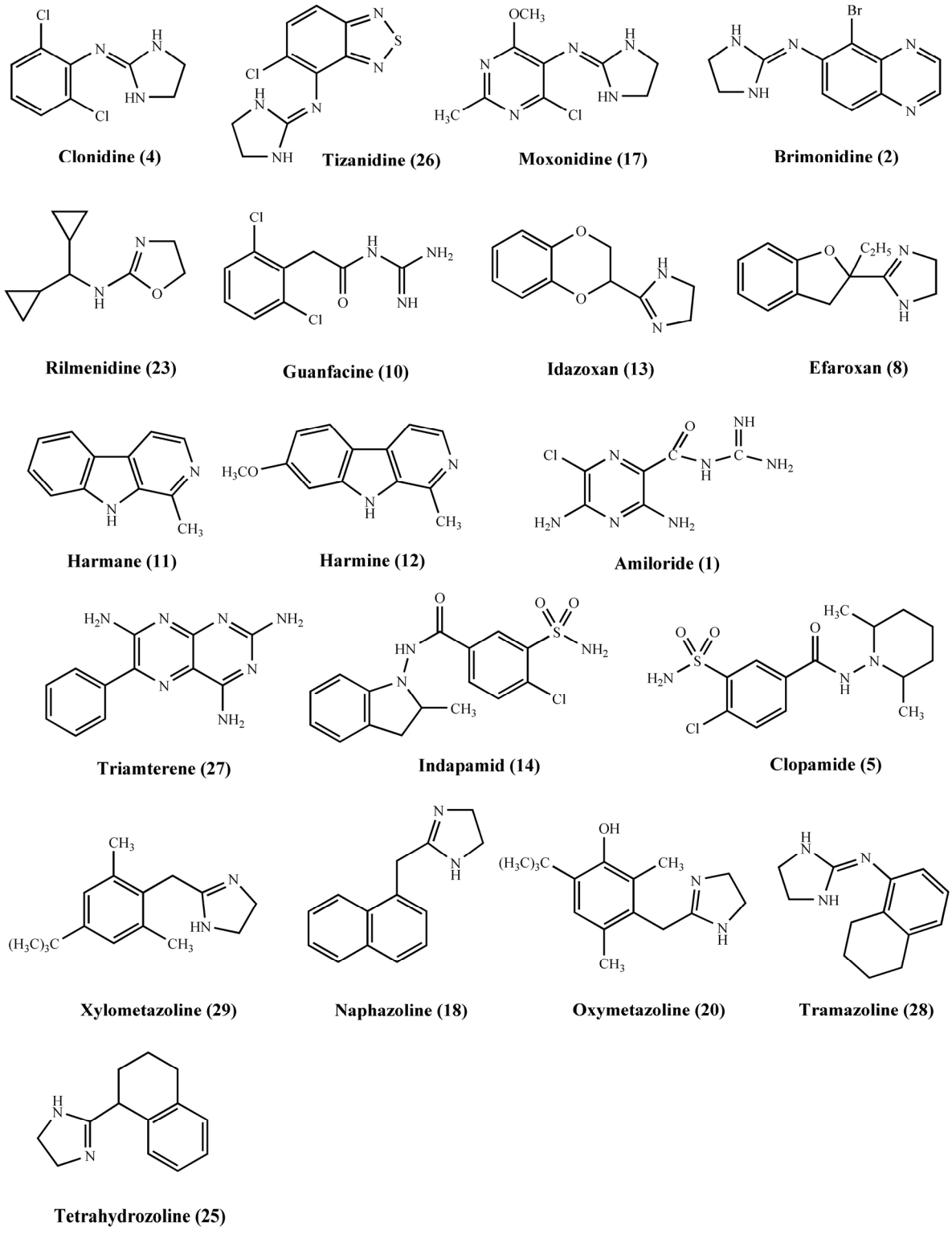

Figure 1. Structural formulas of the examined α-adrenergic/imidazoline receptor ligands and related compounds.

Descriptors with lowest VIP-value are successively removed from the PLS model and new PLS model is created. For each new model are calculated regression factors R^2, Q^2, F ratio, P-value, root mean square error of estimation (RMSEE), and compared with the previous model. The procedure is repeated until the best model is created.

Quality of the obtained QSPR models was examined by use of the R^2, Q^2 and RMSEE, CV-ANOVA analysis of variance testing of cross-validated predictive residuals (F- and P-value), and external validation (root mean square error of prediction (RMSEP), R^2_{pred}, $R^2_{Obs\ vs.\ Pred}$) (28, 30, 31).

Predictive power of the model is determined by Q^2, which is Leave-One-Out Cross-Validated (LOO-CV)/or Leave-n-Out Cross-Validated (LnO-CV) version of R^2. A model is fitted to the data leaving one/or more compounds out, computing VIP, selecting the best variables, and than predicting Y for the left–out compounds. This procedure is repeated until all compounds have been left out, which result in many parallel models. The difference between observed and the predicted Y values are calculated ($e_{(i)}$) for each model. In this setting were defined PRESS (Predicted Sum of Squares), RMSE (RMSEE and RMSEP) and Q^2 as:

$$PRESS = \sum_{i=1}^{n} e_{(i)}^2 \qquad (1)$$

Difference between observed and the predicted Y values - ($e_{(i)}$)

$$RMSE = \sqrt{\frac{PRESS}{n}} \qquad (2)$$

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

QSPR models with $Q^2 \geq 0.5$ can be considered to have good predictive capability (28, 29)

Quality of the QSPR models prediction was assessed on test set by use of the RMSEP, $R^2_{Obs\ vs.\ Pred}$, and R^2_{pred} (28, 30, 31).

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

QSPR models with $R^2_{pred} \geq 0.5$ can be considered to have good predictive capability (31).

The response permutation test (Y scrambling) examined the statistical significance of the R^2 and Q^2 and overfitting due to the chance correlation (28, 30). In this test the Y-matrix is randomly re-ordered (100 times in this project) while the X-matrix is kept intact. Model is fitted to the new Y-data and the new $R^2(Y)$, Q^2 and VIP parameters are calculated. All model selection steps are repeated on the scrambled Y-response data. Lines are fitted through the R^2-values and through the Q^2-values, yielding two separate intercepts. For a valid model, the R^2 - intercept should not exceed the 0.4 while the Q^2-intercept < 0.05 (28).

The F-test, based on the ratio MS Regression/MS Residual, formally assesses 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 common practice is to interpret a P-value lower than 0.05 as pointing to a significant model (28).

Principal component analysis and hierarchical clustering analysis

The Principal component analysis (PCA) (32) is one of the most common methods used in a two-way data analysis. In PCA a two-way matrix, **X** (m × n), is decomposed into two matrices **S** (m × fn) and **D** (n × fn):

$$\mathbf{X}\ (m \times n) = \mathbf{S}\ (m \times fn) \cdot \mathbf{D}'\ (n \times fn) + \mathbf{E}\ (m \times n)$$

where m and n denote, respectively, the number of objects and the number of variables, **S** represents the scores matrix, **D** represents the loading matrix, **E** is the residuals matrix, and fn denotes the number of significant factors. Scores and loadings matrices are orthogonal, i.e. $\mathbf{S'S} = \mathbf{D'D} = \mathbf{I}$. The columns of matrix **S** are called the Principal Components (PC), or eigenvectors. Each PC is constructed as a linear combination of original variables with weights maximizing description of the data variance (i.e. $\mathbf{S} = \mathbf{XD}$). The sum of the squared elements of each eigenvector (PC) is called an eigenvalue. The first PC describes the largest amount of the data variance, so that the associated eigenvalue also has the highest value. The sum of the eigenvalues defines the total variance of the data. Scores vectors (i.e. the columns of matrix **S**) and loading vectors (i.e. the columns of matrix **D**) are used to visualize relationships between the objects and the parameters in a matrix **X**.

Hierarchical clustering analysis (HCA) can be applied to multidimensional data sets, in order to study similarities (or dissimilarities) of objects in the variables space, or similarities of variables in the objects space (33). Final results of hierarchical clustering are presented in a form of a dendrogram. The indices of clustered objects (or variables) are displayed on axis *x* of the dendrogram, whereas axis *y* represents the corresponding linkage distances (or an adequate measure of similarity) between the two objects or clusters, which are merged.

RESULTS

Quantitative Structure-Property Relationship (QSPR) study

The pharmacokinetic properties (logPS, logPS-brain, and logBB) of the 29 ligands (Figure 1), calculated by use of ACD/i-Lab program (27), were used as dependent (Y) variables in the QSPR study. Values of the logPS, logPS-brain, and logBB parameters are reflecting rate and extent of blood-brain permeation process, where higher logPS, logPS-brain, and logBB values indicate on higher rate and extent of brain penetration of a ligand.

The HPLC retention parameters ($\log Kw_{pH\ 4.4}$, $S_{pH\ 4.4}$, $\log Kw_{pH\ 7.4}$, $S_{pH\ 7.4}$, $\log Kw_{pH\ 9.1}$, and $S_{pH\ 9.1}$) and CE migration parameters ($\mu_{effpH\ 4.4}$, $\mu_{effpH\ 7.4}$, and $\mu_{effpH\ 9.1}$), and calculated molecular parameters of the optimized molecular models were used as independent (X) variables. The Density-Functional-Theory/B3LYP/3-21G(d,p) basis set (18-20), Marvin 5.5.1.0 ChemAxon (22), Chem3D Ultra 7.0.0 (23), and Dragon (24) programs were applied for computation of constitutional, physico-chemical, thermodynamic, electronic, and quantum chemically-based reactivity descriptors (25, 26) of the optimized molecular models.

Based on the obtained Score Plots (t1 *vs*. t2 and t1 *vs*. u1) the data set of 29 α-adrenergic and imidazoline-receptors ligands is divided on Training set (22 compounds for QSPR models building) and test set (7 compounds for QSPR models validation) (28, 30, 31).

The rate of brain penetration (logPS) was spanned 3.9 log units (-5.3 – (-1.4)), the brain/plasma equilibration rate (logPS-brain) interval was 2.9 log units (-5.3 – (-2.4)), while the extent of blood-brain barrier (BBB) permeation (logBB) was spanned 2.97 log units (-2.00 – 0.97) in the Training sets. The domain of applicability for the QSPR models is defined by the rate of brain penetration (logPS: -5.3 – (-1.4)), the brain/plasma equilibration rate (logPS-brain: -5.3 – (-2.4)), and the extent of blood-brain barrier (BBB) permeation (logBB: -2.00 – 0.97) of the Training sets. Relatively wide Y intervals of the Training sets provide extensive applicability domain for the formed QSPR models.

Quality of the logBB prediction (27) was examined on three drugs from the Data set (clonidine (34), mianserine (35) and olanzapine (36)) and seven related drugs (imipramine (34), mirtazapine (35), phenytoin (37), tacrine (38), acebutolol (39), alprenolol (39), and pindolol (39). Relatively high correlation coefficient (r: 0.776) between experimental (34-39) and predicted (27) logBB values (Figure 2) indicated on good predictive potential of the ACD/i-Lab program for calculation of logBB parameter for the examined α-adrenergic and imidazoline-receptors ligands (Table 1).

The PLS methodology was applied for selection of the most relevant molecular descriptors and QSPR models building.

Descriptors with lowest Variable Importance in the Projection (VIP) value are successively removed from the PLS model, each time new PLS model is created, and the obtained R^2, Q^2, F ratio, P-value, RMSEE of new model were compared with the previous one. The procedure is repeated until the best model is created.

The applied PLS methodology has selected sets of significant descriptors during development of each QSPR model. Overfitting of the QSPR models is avoided by continual monitoring of the RMSE for training and test set during the QSPR modeling. The QSPR models were selected when RMSEP of test set begins to increase while RMSEE of training set continues do decrease. Furthermore, the QSPR models with different number of selected significant descriptors were compared and optimal one is chosen by comparing statistical parameters of the training set, such as R^2, Q^2 (Leave-One-Out Cross Validation (LOO-CV), Leave-n-Out Cross Validation (LnO-CV)), RMSEE, F ratio, and P-value, and also statistical parameters of the test set, such as $R^2_{Obs\ vs.\ Pred}$, R^2_{pred}, and RMSEP. Apart from the Leave-One-Out Cross Validation, in the QSPR study was performed Leave-2-Out Cross-Validation (L2O-CV) and Leave-3-Out Cross-Validation (L3O-CV).

Optimal QSPR models for each pharmacokinetic parameter (logPS, logPS-brain, and logBB) were selected by use of statistical parameters of the training and test set (Tables 2-4).

In order to investigate the statistical significance of the R^2 and Q^2 and to test the model for overfitting due to the chance correlation the response permutation test (Y scrambling) is applied for all QSPR models. The R^2 – intercepts were less than 0.1 (upper limit is 0.4), while the Q^2-intercepts were less than -0.1 (upper limit is 0.05) for all developed QSPR models. Thus the response permutation test results have proved good quality of the formed QSPR models.

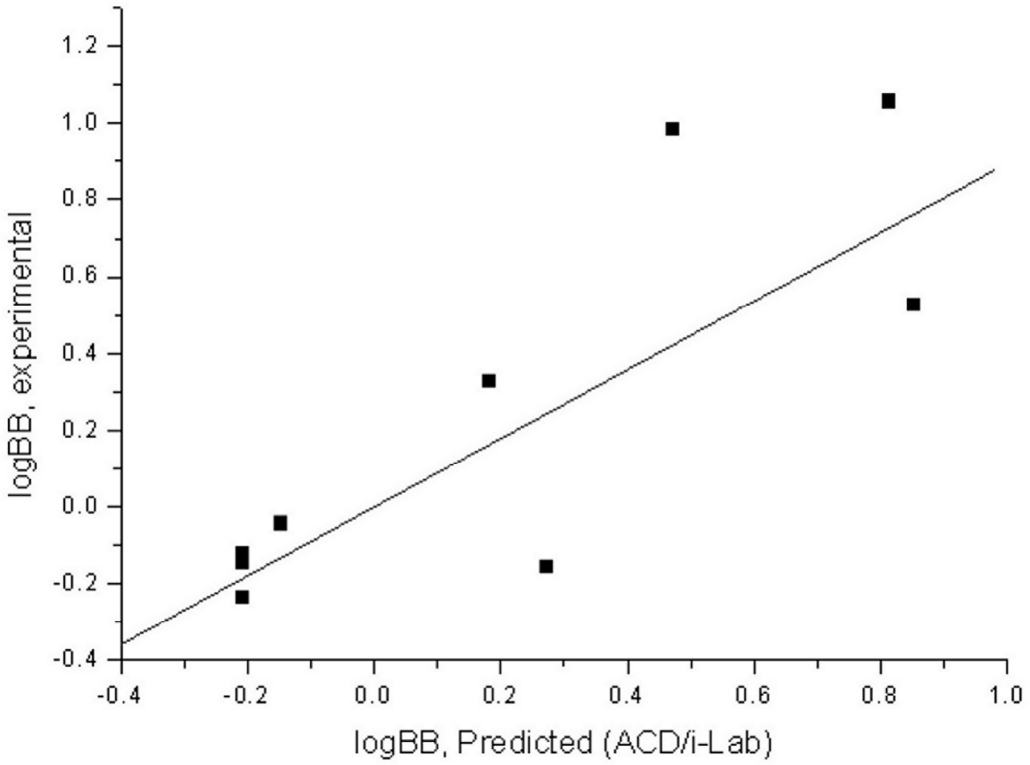

Figure 2. Observed *vs.* predicted logBB for adrenergic/imidazoline receptor ligands and related compounds.

Table 1. Experimental and predicted logBB for adrenergic/imidazoline receptor ligands and related compounds.

Name	Experimental LogBB	Predicted logBB, ACD/i-Lab (27)
Clonidine	0.11 (34)	-0.46
Mianserine	0.99 (35)	0.47
Olanzapine	0.33 (36)	0.18
Imipramine	1.06 (34)	0.81
Mirtazapine	0.53 (35)	0.85
Phenytoin	-0.04 (37)	-0.11
Tacrine	-0.12 (38)	-0.21
Acebutolol	-0.15 (39)	0.27
Alprenolol	-0.23 (39)	-0.21
Pindolol	-0.14 (39)	-0.21
	coefficient of correlation, r	**0.776**

The optimal QSPR (logPS) model with one significant components (A=1), relating five variables (Mi, H7m, R3s+, Hy, and TPSA(NO)), R^2: 0.794, LOO-CV/Q^2: 0.748, L2O-CV/Q^2: 0.740, L3O-CV/Q^2: 0.744, F-ratio: 28.149, P-value: 2.08E-06, and RMSEE: 0.409 (Table 2), was selected.

Significant descriptors of the QSPR (logPS) model are: Mi - mean first ionization potential (scaled on Carbon atom)/*Constitutional indices-Basic descriptors*; H7m - H autocorrelation of lag 7 (weighted by mass GETAWAY descriptors)/*H-indices*; R3s+ - R maximal autocorrelation of lag 3 (weighted by I-state GETAWAY descriptors)/*R-indices*, Hy - hydrophilic factor/*Molecular properties-Basic descriptors*, and TPSA(NO) - topological polar surface area using N,O polar contributions/*Molecular properties-Basic descriptors*.

Table 2. Statistical parameters and predicted brain penetration obtained by QSPR (logPS) study. SD - Standard Deviation, SE – Standard Error: $SE = SD/\sqrt{N}$.

Training Set		**Brain penetration**	
Primary ID	Name	logPS (27)	QSPR-predicted logPS
1	Amiloride	-5.3	-5.607
2	Brimonidine	-3.3	-2.805
3	Carvedilol	-2.0	-2.616
5	Clopamide	-2.3	-2.753
6	Clozapine	-3.3	-2.397
7	Doxazosin	-2.9	-2.849
8	Efaroxan	-2.1	-1.986
9	Ephedrine	-3.0	-2.512
10	Guanfacine	-2.3	-2.833
11	Harmane	-1.4	-1.773
12	Harmine	-1.4	-1.831
13	Idazoxan	-2.6	-2.139
15	Maprotiline	-1.4	-1.685
17	Moxonidine	-3.1	-3.011
18	Naphazoline	-1.6	-1.788
19	Olanzapine	-1.9	-2.043
22	Pseudoephedrine	-3.0	-2.475
23	Rilmenidine	-2.1	-2.174
24	Tamsulosin	-3.4	-3.214
26	Tizanidine	-3.2	-2.912
28	Tramazoline	-2.5	-2.203
29	Xylometazoline	-1.5	-1.992
F-Ratio	28.149	**R^2**	0.794
P-value	2.08E-06	**LOO - Q^2**	0.748
		RMSEE	0.409
PLS equation:			
Descriptor	PLS coefficient	**SE** (PLS coefficient)	PLS - Intercept
Mi	-0.208015	0.114865	-2.73863
H7m	-0.196653	0.106972	
R3s+	-0.218820	0.067949	
Hy	-0.203340	0.039503	
TPSA(NO)	-0.186372	0.099012	
Test Set			
Primary ID	Name	logPS (27)	QSPR-predicted logPS
4	Clonidine	-2.8	-2.658
14	Indapamid	-2.2	-2.601
16	Mianserine	-1.1	-1.553
20	Oxymetazoline	-2.0	-2.407
21	Phenilephrine	-3.6	-2.848
25	Tetrahydrozoline	-1.8	-1.925
27	Triamterene	-3.2	-3.211
SD (predicted logPS)=0.758443		**$R^2_{Obs\ vs.\ Pred}$**	0.823
		R^2_{pred}	0.747
		RMSEP	0.402

Since all PLS coefficients of the QSPR (logPS) model are negative numbers (Figure 3a, Table 2) was concluded that all significant descriptors (Mi, H7m, R3s+, Hy, and TPSA(NO)) are in negative correlation with rate of brain penetration - logPS. Actually, increase in hydrophilicity (Hy), topological polar surface area (TPSA(NO)), mean first ionization potential (Mi), H-indices (H7m), and R-indices (R3s+) of a ligand will lead to decrease of rate of brain penetration (logPS).

Also, decrease in hydrophilicity (Hy), topological polar surface area (TPSA(NO)), mean first ionization potential (Mi), H-indices (H7m), and R-indices (R3s+) of a ligand will lead to increase of rate of brain penetration (logPS).

Constitutional descriptors are descriptors reflecting the chemical composition of a compound without any information about its molecular geometry or atom connectivity. The mean first ionization potential (such as Mi) of a molecule is amount of energy required to remove an electron from the molecule in the gaseous state.

Generally, the GETAWAY (GEometry, Topology, and Atom-Weights AssemblY) descriptors (40, 41) are chemical structure descriptors derived from the Molecular Influence Matrix. *H-indices* (such as H7m) are based on the spatial autocorrelation formulas, weighting the molecule atoms by physico-chemical properties w together with 3D information encoded by the elements of the molecular influence matrix H. *R-indices* (such as R3s+) include some descriptors obtained by applying traditional matrix operators to the influence/distance matrix *R* and autocorrelations calculated by weighting the molecule atoms by physico-chemical properties *w* together with 3D information encoded by the elements of the influence/distance matrix. The H, R and maximal R+ indices are molecular descriptors based on spatial autocorrelation, encode information on structural fragments and, therefore, seem to be particularly suitable for describing differences in congeneric series of molecules. Also, GETAWAY descriptors are geometrical descriptors encoding information on the effective position of substituents and fragments in the molecular space and provide information about molecular size, shape, and specific atomic properties (40, 41).

The hydrophilic factor (Hy) is a hydrophilicity descriptor defined by Todeschini et al (42).

The topological polar surface area (TPSA) is based on a group contribution method, calculated according to the model proposed by Ertl (43). The TPSA(NO) is topological polar surface area derived only from polar fragments with nitrogen and oxygen. The TPSA(NO) of a molecule is determined by the summation of tabulated surface contributions of the polar atom types. The topological polar surface area descriptors were successfully used in validation studies based on sets of published absorption data, including intestinal absorption, Caco-2 monolayer penetration, and blood-brain barrier penetration (43).

Predictive potential of the QSPR (logPS) model was tested by use of leave-one-out cross validation of the training set (LOO-CV/Q^2: 0.748, L2O-CV/Q^2: 0.740, L3O-CV/Q^2: 0.744, RMSEE: 0.409) and evaluation of the verification set ($R^2_{\text{Observed } vs. \text{ Predicted}}$: 0.823, RMSEP: 0.402, and R^2_{pred}: 0.747) (Table 2). The obtained statistical parameters proved that the QSPR (logPS) model has acceptable accuracy for predicting rate of brain penetration (logPS) of the α-adrenergic and imidazoline receptor ligands.

The QSPR (logPS-brain) model with three significant components (A=3), relating five variables (GATS1e, H7m, nHDon, T(N..N), and Hy), R^2: 0.954, LOO-CV/Q^2: 0.923, L2O-CV/Q^2: 0.918, L3O-CV/Q^2: 0.926, F-ratio: 29.935, P-value: 1.57E-07, and RMSEE: 0.122 (Table 3), was derived.

Significant descriptors of the QSPR (logPS-brain) model are: GATS1e - Geary autocorrelation of lag 1 weighted by Sanderson electronegativity/*2D autocorrelations-Geary autocorrelations*; H7m - H autocorrelation of lag 7 (weighted by mass GETAWAY descriptors)/*H-indices*; nHDon - number of donor atoms for H-bonds (N and O)/*Functional group counts-Basic descriptors*, T(N..N) - sum of topological distances between N..N/*2D Atom Pairs-Weighted topological atom pairs*, and Hy - hydrophilic factor/*Molecular properties-Basic descriptors*.

Negative sign of all PLS coefficients in the QSPR (logPS-brain) model (Figure 3b, Table 3) indicated that all significant descriptors (GATS1e, H7m, nHDon, T(N..N), and Hy) are in negative correlation with the brain/plasma equilibration rate – logPS-brain. Actually, increase in Geary autocorrelations (GATS1e) parameter, H-indices (H7m), number of donor atoms for H-bonds (nHDon), sum of topological distances between N..N (T(N..N)), and stronger hydrophilicity (Hy) of a ligand will lead to decrease of logPS-brain Also, decrease in Geary autocorrelations (GATS1e) parameter, H-indices (H7m), number of donor atoms for H-bonds (nHDon), sum of topological distances between N..N (T(N..N)), and stronger hydrophilicity (Hy)of a ligand will lead to increase of brain/plasma equilibration rate – logPS-brain.

Table 3. Statistical parameters and predicted brain/plasma equilibration rates obtained by QSPR (logPS-brain) study. SD - Standard Deviation, SE – Standard Error: $SE = SD/\sqrt{N}$.

Training Set		**Brain/plasma equilibration rate**	
Primary ID	Name	logPS-brain (27)	QSPR-predicted logPS-brain
1	Amiloride	-5.3	-5.320
2	Brimonidine	-3.5	-3.429
3	Carvedilol	-3.3	-3.498
5	Clopamide	-2.7	-2.605
6	Clozapine	-3.4	-3.393
7	Doxazosin	-3.2	-3.219
8	Efaroxan	-2.8	-2.673
9	Ephedrine	-3.2	-3.166
10	Guanfacine	-2.9	-2.895
11	Harmane	-2.4	-2.721
12	Harmine	-2.5	-2.666
13	Idazoxan	-2.9	-2.682
15	Maprotiline	-3.4	-3.249
17	Moxonidine	-3.2	-3.228
18	Naphazoline	-3.0	-3.028
19	Olanzapine	-3.2	-3.242
22	Pseudoephedrine	-3.2	-3.160
23	Rilmenidine	-2.6	-2.671
24	Tamsulosin	-3.8	-3.617
26	Tizanidine	-3.5	-3.592
28	Tramazoline	-3.2	-3.210
29	Xylometazoline	-3.2	-3.136
F-Ratio	29.935	**R^2**	0.954
P-value	1.57E-07	**Q^2**	0.923
		RMSEE	0.122
PLS equation:			
Descriptor	PLS coefficient	**SE** (PLS coefficient)	PLS - Intercept
GATS1e	-0.332009	0.171041	-5.48028
H7m	-0.497359	0.343375	
nHDon	-0.183323	0.132481	
T(N...N)	-0.183561	0.250992	
Hy	-0.133735	0.156820	
Test Set			
Primary ID	Name	logPS-brain (27)	QSPR-predicted logPS-brain
4	Clonidine	-3.0	-3.018
14	Indapamid	-2.7	-2.630
16	Mianserine	-2.5	-2.717
20	Oxymetazoline	-3.3	-3.101
21	Phenilephrine	-3.7	-3.247
25	Tetrahydrozoline	-2.9	-3.049
27	Triamterene	-3.5	-3.966
SD(predict logPS-brain)=0.535428		**$R^2_{Obs\ vs.\ Pred}$**	0.586
		R^2_{pred}	0.560
		RMSEP	0.277

Geary autocorrelations (GATS1e) lag 1 weighted by Sanderson electronegativity is calculated by applying Geary coefficient to the H-filled molecular graph. Geary coefficient is a distance-type function varying from zero to infinite. Strong spatial autocorrelation produces small values of this index. Positive autocorrelation translates in values between 0 and 1 whereas negative autocorrelation produces values larger than 1.

H-indices (such as H7m) are based on the spatial autocorrelation formulas, weighting the molecule atoms by physico-chemical properties w

together with 3D information encoded by the elements of the molecular influence matrix H (40, 41).

The number of donor atoms for H-bonds (nHDon) is a measure of the hydrogen-bonding ability of a molecule expressed in terms of number of possible hydrogen-bond donors. It is calculated by adding up the hydrogens bonded to any nitrogen and oxygen without negative charge in the molecule.

Weighted topological atom pairs T(X..Y), such as T(N..N), are sums of topological distances between all pairs of type X..Y, where X and Y referring to any heteroatom among N, O, S, P, F, Cl, Br, I.

The hydrophilic factor (Hy) is a hydrophilicity descriptor defined by Todeschini et al (42).

The hydrophilicity (Hy) and H-indices (H7m) are selected by both QSPR (logPS) and QSPR (logPS-brain) study as significant molecular determinant of the rate of brain penetration and brain/plasma equilibration rate of the examined compounds.

Predictive potential of the QSPR (logPS-brain) model was tested by use of leave-one-out cross validation of the training set (LOO-CV/Q^2: 0.923, L2O-CV/Q^2: 0.918, L3O-CV/Q^2: 0.926, RMSEE: 0.122) and evaluation of verification set ($R^2_{\text{Observed } vs. \text{ Predicted}}$: 0.586, RMSEP: 0.277, and R^2_{pred}: 0.560) (Table 3). The obtained statistical parameters indicated that the QSPR (logPS-brain) model could be used as reliable tool for evaluation of brain/plasma equilibration rate (logPS-brain) of the examined compounds.

The QSPR (logBB) model with one significant components (A=1), relating four variables (nHet, SM1_Dz(i), AVS_B(s), and TPSA(NO)), R^2: 0.714, LOO-CV/Q^2: 0.653, L2O-CV/Q^2: 0.652, L3O-CV/Q^2: 0.651, F-ratio: 17.840, P-value: 4.35E-05, and RMSEE: 0.340 (Table 4), was derived.

Significant descriptors of the QSPR (logBB) model are: nHet - number of heteroatoms/*Constitutional indices-Basic descriptors*; SM1_Dz(i)- spectral moment of order 1 from Barysz matrix weighted by ionization potential/*2D matrix-based descriptors-Barysz matrix weighted by ionization potential (Dz(i))*; AVS_B(s) - average vertex sum from Burden matrix weighted by I-State/*2D matrix-based descriptors-Burden matrix weighted by I-state (B(s))*, and TPSA(NO) - topological polar surface area using N,O polar contributions/*Molecular properties-Basic descriptors*.

Since all PLS coefficients of the QSPR (logBB) model are negative numbers (Figure 3c, Table 4) was concluded that all significant descriptors (nHet, SM1_Dz(i), AVS_B(s), and TPSA(NO)) are in negative correlation with extent of BBB permeation (logBB). Actually, increase in number of heteroatoms (nHet), spectral moment of order 1 from Barysz matrix weighted by ionization potential (SM1_Dz(i)), average vertex sum from Burden matrix weighted by I-State (AVS_B(s)), and topological polar surface area (TPSA(NO)) of a ligand will lead to decrease of logBB. Also, decrease in number of heteroatoms (nHet), spectral moment of order 1 from Barysz matrix weighted by ionization potential (SM1_Dz(i)), average vertex sum from Burden matrix weighted by I-State (AVS_B(s)), and topological polar surface area (TPSA(NO)) of a ligand will lead to increase of extent of BBB permeation (logBB).

The number of heteroatoms (nHet) counts all the atoms that are neither hydrogen nor carbon.

Barysz matrices (Dz(w)) are weighted distance matrices accounting contemporarily for the presence of heteroatoms and multiple bonds in the molecule. They were defined on the basis of a generalization of Barysz weighting scheme in terms of conventional bond orders π^* and any atomic property, such as spectral moment of order 1 from Barysz matrix weighted by ionization potential (SM1_Dz(i)) (44).

Burden matrices (B(w)) are augmented adjacency matrices derived from a H-depleted molecular graph as the following: the diagonal elements are atomic carbon-scaled properties (wi/wC); the off-diagonal elements corresponding to pairs of bonded atoms are the square roots of conventional bond orders; entries corresponding to terminal bonds are augmented by 0.1; all other matrix elements are set at 0.001. The AVS_B(s) - average vertex sum from Burden matrix is calculated on the basis of the I-State weighting scheme.

The TPSA(NO) is topological polar surface area derived only from polar fragments with nitrogen and oxygen (43). The topological polar surface area (TPSA(NO)) is parameter selected by both QSPR (logPS) and QSPR (logBB) study as significant molecular determinant of the rate of brain penetration and extent of BBB permeation of the examined compounds.

Predictive potential of the QSPR (logBB) model was tested by use of leave-one-out cross validation of the training set (LOO-CV/Q^2: 0.653, L2O-CV/Q^2: 0.652, L3O-CV/Q^2: 0.651, RMSEE:

Table 4. Statistical parameters and predicted extent of BBB permeation obtained by QSPR (logBB) study. SD - Standard Deviation, SE – Standard Error: $SE = SD/\sqrt{N}$.

Training Set		**Extent of BBB permeation**	
Primary ID	Name	logBB (27)	QSPR-predicted logBB
1	Amiloride	-2.00	-1.380
2	Brimonidine	-0.21	-0.327
3	Carvedilol	-0.40	-0.377
4	Clonidine	-0.46	-0.097
5	Clopamide	0.03	-0.731
6	Clozapine	-0.45	-0.056
7	Doxazosin	-1.20	-0.984
8	Efaroxan	0.40	0.243
9	Ephedrine	-0.02	0.366
10	Guanfacine	-0.24	-0.601
11	Harmane	-0.22	0.235
12	Harmine	-0.05	0.079
15	Maprotiline	0.97	0.651
16	Mianserine	0.47	0.494
18	Naphazoline	0.17	0.369
19	Olanzapine	0.18	0.061
20	Oxymetazoline	0.86	0.214
21	Phenilephrine	-0.06	0.016
24	Tamsulosin	-0.35	-0.647
25	Tetrahydrozoline	0.46	0.464
27	Triamterene	-0.73	-0.903
28	Tramazoline	0.21	0.273
F-Ratio	17.840	**R^2**	0.714
P-value	4.35E-05	**LOO-Q^2**	0.653
		RMSEE	0.340
PLS equation:			
Descriptor	PLS coefficient	**SE** (PLS coefficient)	PLS - Intercept
nHet	-0.219484	0.067757	-0.184448
SM1-Dz(i)	-0.233968	0.068265	
AVS_B(s)	-0.220943	0.119724	
TPSA(NO)	-0.226892	0.101973	
Test Set			
Primary ID	Name	logBB (27)	QSPR-predicted logBB
13	Idazoxan	-0.10	0.021
14	Indapamid	-0.39	-0.797
17	Moxonidine	-0.26	-0.476
22	Pseudoephedrine	-0.02	0.366
23	Rilmenidine	0.45	0.357
26	Tizanidine	-0.45	-0.422
29	Xylometazoline	0.98	0.504
SD (predicted logBB) = 0.545069		**$R^2_{Obs\ vs.\ Pred}$**	0.678
		R^2_{pred}	0.615
		RMSEP	0.296

0.340) and evaluation of verification set ($R^2_{Observed\ vs.\ Predicted}$: 0.678, RMSEP: 0.296, and R^2_{pred}: 0.615) (Table 4). The statistical parameters indicated that the created QSPR (logBB) model has acceptable accuracy for predicting extent of BBB permeation (logBB) of the α-adrenergic and imidazoline receptor ligands. In order to investigate the statistical significance of the R^2 and Q^2 and to test the model for overfitting due to the chance correlation, the response permutation test (Y

scrambling) is applied for all three created QSPR models. The R^2–intercepts were less than 0.4 while the Q^2-intercepts were less than 0.05 for the three developed QSPR models. Obtained results of the response permutation tests have proved that the formed QSPR models are not overfitted.

Correlation Study of Pharmacokinetic parameters

The HPLC retention parameters (logKw and Slope: $\log Kw_{pH\ 4.4}$, $S_{pH\ 4.4}$, $\log Kw_{pH\ 7.4}$, $S_{pH\ 7.4}$, $\log Kw_{pH\ 9.1}$, and $S_{pH\ 9.1}$) and CE migration parameters ($\mu_{effpH\ 4.4}$, $\mu_{effpH\ 7.4}$, and $\mu_{effpH\ 9.1}$) of the examined 29 ligands were used independent X-variables, while brain penetration (logPS), brain/plasma equilibration rate (logPS-brain), and extent of BBB permeation (logBB) were dependent - Y variables in the correlation study.

Because of diverse structures of the 29 α-adrenergic and imidazoline-receptors ligands, the obtained correlations between the retention/migration parameters ($\log Kw_{pH\ 4.4}$, $S_{pH\ 4.4}$, $\log Kw_{pH\ 7.4}$, $S_{pH\ 7.4}$, $\log Kw_{pH\ 9.1}$, $S_{pH\ 9.1}$, $\mu_{effpH\ 4.0}$, $\mu_{effpH\ 7.4}$, and $\mu_{effpH\ 9.1}$) and the pharmacokinetic properties (logPS, logPS-brain, and logBB) were very low. Therefore we decided to perform PCA and HCA clustering of the ligands before the correlation study.

Molecular descriptors of the examined ligands with the strongest influence on the HPLC retention parameters ($\log Kw_{pH\ 4.4}$, $\log Kw_{pH\ 7.4}$, $\log Kw_{pH\ 9.1}$) and CE migration parameters ($\mu_{effpH\ 4.4}$, $\mu_{effpH\ 7.4}$, and $\mu_{effpH\ 9.1}$) were also selected for the PCA and HCA. The HPLC retention parameters ($\log Kw_{pH\ 4.4}$ $\log Kw_{pH\ 7.4}$, $\log Kw_{pH\ 9.1}$), CE migration parameters ($\mu_{effpH\ 4.4}$, $\mu_{effpH\ 7.4}$, and $\mu_{effpH\ 9.1}$), and the selected molecular descriptors were used for the PCA and HCA study of the compounds (Table 5).

A percent of modeled variance was used in determination of the number of significant components (PCs) of the standardized data set organized in a matrix **X** (29 x 34). The PCA model with six significant principal components (PCs) described 92.31% of the total data variance. The PC1 described 57.86% of the total data variance. The studied compounds could be ordered along the PC1 as follows:

group I: carvedilol and doxazosin (objects nos 3 ad 7),
group II: clopamide, clozapine, indapamid, maprotiline, mianserin, olanzapine and tamsulosin (compounds No 5, 6,14-16, 19 and 24)
group III: harmine, oxymetazoline, triamterene and xylometazoline (compounds No 12, 20, 27 and 29),
group IV: amiloride, brimonidine, clonidine, efaroxan, guanfacine, harmane, idazoxan, moxonidine, naphazoline, tetrahydrozoline, tizanidine and tramazoline (compounds No 1, 2, 4, 8, 10, 11, 13, 17, 18, 25, 26 and 28) and
group V: ephedrine, phenilephrine, pseudoephedrine and rilmenidine (compounds No 9 and 21-23), respectively.

An efficient compression of the studied data was not possible with the PCA and obtained results required investigation of many two-dimensional plots. Therefore, hierarchical clustering analysis (HCA) was used to explore the studied data set **X** (29 x 34) and to examine the similarities between the studied compounds. HCA helped to reveal the internal data structure and its clustering tendency.

The HCA results presented in Figure 4 were based on the Euclidean distance and the Ward linkage algorithm.

The dendrogram presented in Figure 4 allows revealing four clusters: cluster A grouping clozapine, maprotiline, mianserin, olanzapine, oxymetazoline and xylometazoline (compounds No 6, 15, 16, 19, 20 and 29), cluster B, composed of carvedilol, clopamide, doxazosin, indapamid and tamsulosin (compounds No 3, 5, 7, 14 and 24), cluster C grouping amiloride, brimonidine, clonidine, guanfacine, harmane, harmine, moxonidine, tizanidine and triamterene (compounds No 1, 2, 4, 10-12, 17, 26 and 27) and cluster D including efaroxan, ephedrine, idazoxan, naphazoline, phenilephrine, pseudoephedrine, rilmenidine, tetrahydrozoline and tramazoline (compounds No 8, 9, 13, 18, 21-23, 25 and 28).

Because of very high similarity between I-II and III-IV PCA groups with A-B and C-D HCA clusters, respectively, we further examined HCA-groups A and B, as first subset, and HCA-groups C and D, as second subset, in correlation study with rate of brain penetration (logPS), extent of BBB permeation (logBB), and brain/plasma equilibration rate (logPS-brain) of the compounds.

Significant correlations were obtained between the chromatographic retention parameter (S at pH 7.4) and extent of BBB permeation (logBB) in A/B cluster ($r(S_{pH7.4}/\log BB)$: 0.677) (Figure 5a), and also between the chromatographic retention parameter (logKw at pH 9.1) and rate of brain penetration (logPS) in C/D cluster ($r(\log Kw_{pH\ 9.1}/\log PS)$: 0.684) (Figure 5b).

Table 5. Descriptors used for PCA and HCA study.

No.	Descriptor	No.	Descriptor	No.	Descriptor
1	logKw $_{pH\ 4.4}$	**12**	VE2_Dt	**24**	GGI6
2	logKw $_{pH\ 7.4}$	**13**	H_D/Dt	**25**	G1
3	logKw $_{pH\ 9.1}$	**14**	Ho_Dt	**26**	RDF045e
4	$\mu_{eff\ pH\ 4.4}$	**15**	SM3_Dt	**27**	Mor32m
5	$\mu_{eff\ pH\ 7.4}$	**16**	LogD 7,4	**28**	SpMaxA_EA(dm)
6	$\mu_{eff\ pH\ 9.1}$	**17**	F04[C-C]	**29**	SM1_Dz(Z)
7	SM04_EA(ri)	**18**	logD 9,1	**30**	AAC
8	SpMax5_Bh(i)	**19**	ALOGP	**31**	ZM1MulPer
9	SM02_EA(ri)	**20**	ATS6s	**32**	SM11_EA(dm)
10	VE2_Dz(p)	**21**	Ho_Dz(e)	**33**	ISH
11	SpMax5_Bh(e)	**22**	ATS7s	**34**	MW
		23	ATSC6m		

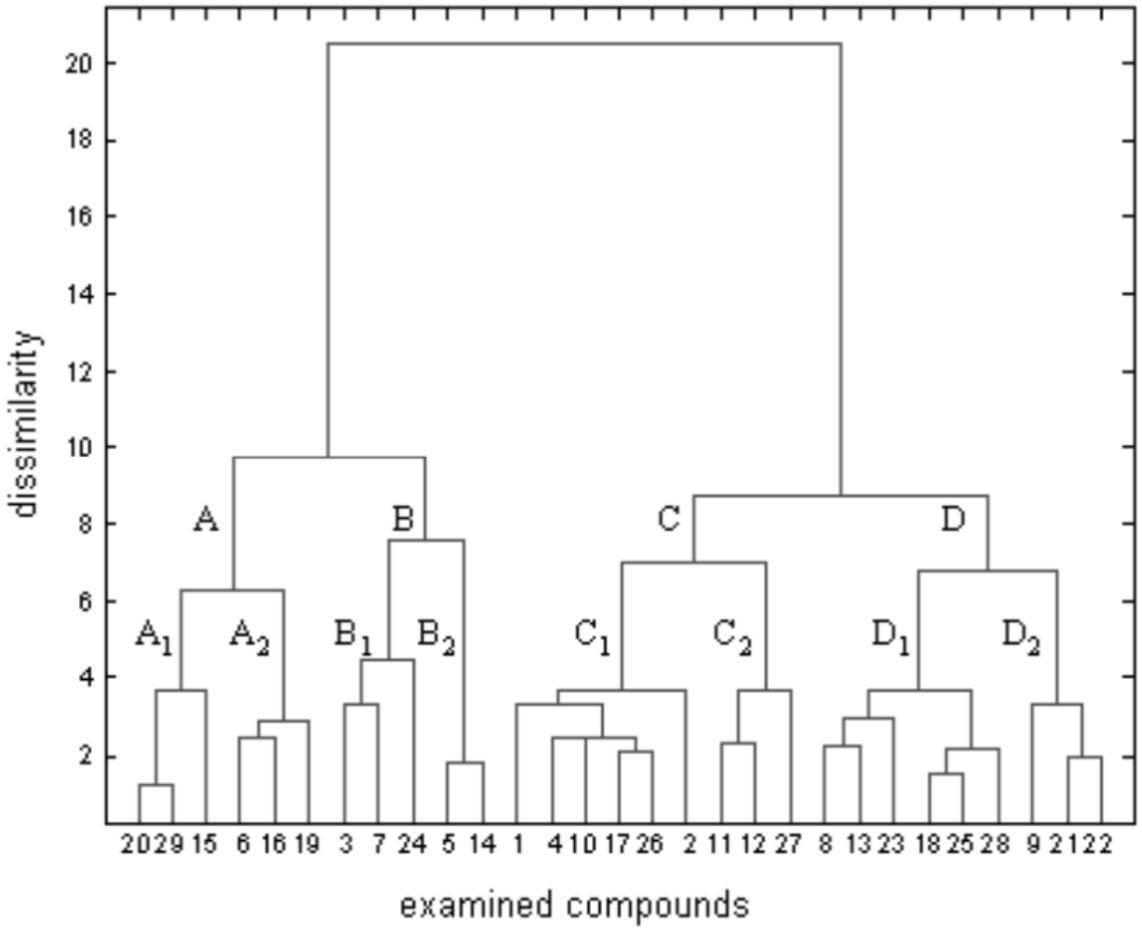

Figure 4. The dendrograms of studied compounds (objects) in the space of the 34 descriptors (listed in Table 5).

Positive sign of correlation coefficient between slope $S_{pH7.4}$ and logBB (r($S_{pH7.4}$/logBB): 0.677) in A/B cluster indicate that ligands with lower $S_{pH7.4}$ values will have lower logBB values and therefore decreased extent of BBB permeation, while ligands with higher $S_{pH7.4}$ values will have higher logBB values and therefore increased extent of BBB permeation.

Positive sign of correlation coefficient between $logKw_{pH9.1}$ and logPS ($r(logKw_{pH9.1}/logPS)$: 0.684) in C/D cluster indicate that ligands with lower $logKw_{pH9.1}$ values will have lower logPS values and therefore decreased rate of brain penetration, while ligands with higher $logKw_{pH9.1}$ values will have higher logPS values and therefore increased rate of brain penetration.

Generally, the compounds in A/B cluster have: MW > 260, $logKw_{pH4.0} > 2$, $logKw_{pH7.4} > 2$, and $logKw_{pH9.1} > 2$, while the ligands grouped in C/D cluster have: MW<260, $logKw_{pH4.0} < 2$, $logKw_{pH7.4} < 2$, and $logKw_{pH9.1} < 2$.

Therefore, in case of related α-adrenergic/imidazoline receptor ligands with MW > 260, $logKw_{pH4.0} > 2$, $logKw_{pH7.4} > 2$, and $logKw_{pH9.1} > 2$ parameters, the derived linear equation ($logBB = 0.6874*S_{7.4}+3.05$) could be applied for evaluation of their extent of BBB permeation. Also, in case of ligands with MW < 260, $logKw_{pH4.0} < 2$, $logKw_{pH7.4} < 2$, and $logKw_{pH9.1} < 2$ parameters, the derived linear equation ($logPS = 1.006*logKw_{9.1}-4.229$) could be applied for evaluation of their rate of brain penetration.

Results of the clustering and correlation studies could be readily used as time and cost efficient screening method for evaluation of brain penetration process of novel α-adrenergic/imidazoline receptor ligands and related compounds.

DISCUSSION

Rate of brain penetration (logPS), brain/plasma equilibration rate (logPS-brain), and extent of BBB permeation (logBB) were examined by QSPR study for 29 α-adrenergic and imidazoline-receptors ligands. Experimental parameters (($logKw_{pH\ 4.4}$, $S_{pH\ 4.4}$, $logKw_{pH\ 7.4}$, $S_{pH\ 7.4}$, $logKw_{pH\ 9.1}$, $S_{pH\ 9.1}$, $\mu_{effpH\ 4.4}$, $\mu_{effpH\ 7.4}$, and $\mu_{effpH\ 9.1}$) and calculated molecular descriptors of the ligands were used as independent variables in the QSPR study.

The QSPR (logPS) study indicated on negative correlation between hydrophilicity (Hy), topological polar surface area (TPSA(NO)), mean first ionization potential (Mi), H7m indices, and R3s+ indices of the α-AR/IR ligands and rate of brain penetration (logPS). Therefore increase in Hy, TPSA(NO), Mi, H7m, and R3s+ values of a ligand will lead to decrease of rate of brain penetration (logPS), and vice versa.

The QSPR (logPS-brain) study signified negative correlation between GATS1e parameter, H7m indices, nHDon, T(N..N) and hydrophilicity (Hy) of the α-AR/IR ligands and brain/plasma equilibration rate (logPS-brain). Thus increase in GATS1e, H7m, nHDon, T(N..N) and Hy, values of a ligand will lead to decrease of brain/plasma equilibration rate (logPS-brain), and opposite.

The hydrophilicity (Hy) and H7m indices are selected by both QSPR (logPS) and QSPR (logPS-brain) studies as significant molecular determinant of the compounds for rate of brain penetration and brain/plasma equilibration rate.

The QSPR (logBB) study indicated on negative correlation between nHet, SM1_Dz(i), AVS_B(s), and TPSA(NO) of the α-AR/IR ligands and extent of BBB permeation (logBB). Therefore increase in nHet, SM1_Dz(i), AVS_B(s), and TPSA(NO) values of a ligand will lead to decrease of extent of BBB permeation (logBB), and vice versa.

The topological polar surface area (TPSA(NO)) is parameter chosen by both QSPR (logPS) and QSPR (logBB) studies as significant molecular determinant of the rate of brain penetration and extent of BBB permeation of the examined compounds.

The hydrophilicity (**Hy**) and topological polar surface area (**TPSA(NO)**) are common significant descriptors of QSPR (logPS)/QSPR (logPS-brain) and QSPR (logPS)/QSPR (logBB) models respectively. The observed negative correlations of the hydrophilicity and topological polar surface area on blood-brain penetrations process were in very good agreement with previously reported QSPR studies of brain penetration of other groups of organic compounds (45).

Prognostic capacity of the created QSPR (logPS, logPS-brain, logBB) models was proved by cross- and external validation.

Therefore structures of novel test compounds could be examined by the presented QSPR procedure that includes: molecular modeling, calculation of molecular parameters of the optimized molecular models, and prediction of the logPS, logPS-brain, logBB values by use of the formed QSPR (logPS, logPS-brain, logBB) models.

After PCA and HCA grouping were obtained good correlations between the chromatographic retention parameter ($S_{pH7.4}$) and logBB in A/B cluster, and between the chromatographic retention parameter ($logKw_{pH9.1}$) and logPS in C/D cluster.

Structural diversity of the examined drugs provide wide application domain of the QSPR models, which could be used for evaluation of brain penetration of related α-adrenergic/imidazoline receptor ligands.

ACKNOWLEDGEMENT

This work was supported by the Ministry of Education and Science of the Republic of Serbia Contract No. 172033. DA, KN, and SF acknowledge COST action CM1103.

REFERENCES

1. Bousquet P, Feldman J, Schwartz J. Central cardiovascular effects of alpha-adrenergic drugs: differences between catecholamines and imidazolines. J Pharmacol Exp Ther, 1984; 230:232-236.
2. Tibirica E, Feldman J, Bousquet P. Differences in the ability of yohimbine to antagonize the hypotensive effect of clonidine in normotensive and spontaneously hypertensive anesthetized rats. J Pharmacol Exp Ther, 1988; 244:1062-1066.
3. Ernsberger P. Pharmacology of moxonidine: An I1-imidazoline receptor agonist. J Cardiovasc Pharmacol, 2000; 35:S27-S41.
4. Bousquet P, Feldman J. Drugs acting on imidazoline receptors. A review of their pharmacology, their use in blood pressure control and their potential interest in cardioprotection. Drugs, 1999; 58:799-812.
5. Schann S, Greney H, Gasparik V, Dontenwill M, Rascente C, Lacroix G, Monassier L, Bruban V, Feldman J, Ehrhardt JD, Bousquet P. Methylation of imidazoline related compounds leads to loss of α2-adrenoceptor affinity. Synthesis and biological evaluation of selective I1 imidazoline receptor ligands. Bioorg Med Chem, 2012; 20:4710-4715.
6. Eglen RM, Hudson AL, Kendall DA, Nutt DJ, Morgan NG, Wilson VG, Dillon MP. Seeing through a Glass Darkly: Casting Light on Imidazoline I Sites. Trends Pharmacol Sci, 1998; 19: 381-390.
7. Head GA, Mayorov DN. Imidazoline receptors, novel agents and therapeutic potential. Cardiovasc Hematol Agents, 2006; 4:17–32.
8. Bousquet P, Bricca G, Dontenwill M, Feldman J, Greney H, Belcourt A, Stutzmann J, Tibirica E. From the α2-adrenoceptors to the imidazoline preferring receptors. Fundam Clin Pharmacol, 1992; 6:15S–21S.
9. Piletz JE, Sletten K. Nonadrenergic imidazoline binding sites on human platelets. J Pharmacol Exp Ther, 1993; 267:1493–1502.
10. Tesson F, Limon-Boulez I, Urban P, Puype M, Vandekerckhove J, Coupry I, Pompon D, Parini A. Localization of I2-imidazoline binding sites on monoamine oxidases. J Biol Chem, 1995; 270:9856–9861.
11. Valet P, Tran MA, Damase-Michel C, De Saint-Blanquat G, Dang Tran L, Anglade F, Koenig-Berard E, Montastruc JL. Rilmenidine (S 3341) and the sympatho-adrenal system: Adrenoreceptors, plasma and adrenal catecholamines in dogs. J Auton Pharmacol, 2000; 8:319–326.
12. Mitrovic V, Patyna W, Huting J, Schlepper M. Hemodynamic and neurohumoral effects of moxonidine in patients with essential hypertension. Cardiovasc Drugs Ther, 1991; 5:967–972.
13. Kline RL, Cechetto DF. Renal effects of rilmenidine in anaesthetized rats: importance of renal nerves. J Pharmacol Exp Ther, 1993; 266:1556–1562.
14. Kimura A, Tyacke RJ, Robinson JJ, Husbands SM, Minchin MCW, Nutt DJ, Hudson AL. Identification of an imidazoline binding protein: Creatine kinase and an imidazoline-2 binding site. Brain Research 2009; 1279:21-28.
15. Halaris A, Piletz JE. Relevance of Imidazoline Receptors and Agmatine to Psychiatry, A Decade of Progress. Ann NY Acad Sci, 2003; 1009:1–20.
16. Soczewiński E, Wachtmeister CA. The relation between the composition of certain ternary two-phase solvent systems and RM values. J Chrom, 1962; 7:311-320.
17. Snyder LR, Dolan JW, Gant JR. Gradient elution in high-performance liquid chromatography : I. Theoretical basis for reversed-phase systems. J Chrom, 1979; 165:3-30.
18. Becke AD. Density-functional thermochemistry. III. The role of exact exchange. J Chem Phys, 1993; 98:5648-5652.
19. Lee C, Yang W, Parr RG, Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. Phys Rev B Condens Matter, 1988; 37:785-789.
20. Frisch MJ. et al, Gaussian 98 (Revision A.7) Gaussian, Inc., Pittsburgh PA, 1998.
21. Remko M, Swart M, Bickelhaupt FM. Theoretical study of structure, pKa, lipophilicity, solubility, absorption, and polar surface area of some centrally acting antihypertensives. Bioorg Med Chem, 2006; 14:1715-1728.
22. ChemAxon Marvin 5.5.1.0 program, Budapest, Hungary, 2011. www.chemaxon.com/products.html
23. CS Chem3D Ultra 7.0, Cambridge Soft Corporation, (Property Picker ActiveX Control), 100 Cambridge Park Dr. Cambridge, MA 02140-2317 U.S.A., 2001. http://www.cambridgesoft.com/
24. Dragon 6, TALETE srl, Via V. Pisani, 13 - 20124 Milano – Italy. http://www.talete.mi.it
25. Parr RG, Yang W, Density-Functional Theory of Atoms and Molecules, Oxford, University Press, New York USA, 1989.
26. Parr RG, Szentpály LV, Liu S. Electrophilicity index. J Am Chem Soc, 1999; 121:1922-1924.
27. ACD/i-Lab Freeware, Advanced Chemistry Development, Inc., Toronto, ON, Canada, 2012, www.acdlabs.com
28. Eriksson L, Johansson E, Kettaneh-Wold N, Trygg J, Wikstrom C, Wold S. (Eds.) Multi-and Megavariate Data Analysis. Basic Principles and Applications I, 2nd ed., Umetrics Academy, Umeå, Sweden, 2001.

29. Umetrics AB, Umea, Sweden, SIMCA P+ program, Version 12.0.0.0, May 20. 2008, www.umetrics.com
30. Tropsha A. Best Practices for QSAR Model Development, Validation, and Exploitation. Mol Inf, 2010; 29:476-488.
31. Consonni V, Ballabio D, Todeschini R. Comments on the definition of the Q2 parameter for QSAR validation. J Chem Inform Mod, 2009; 49:1669-1678.
32. Massart DL, Vandeginste BGM, Buydens LMC, De Jong S, Lewi PJ, Smeyers-Verbeke J. In: Handbook of Chemometrics and Qualimetrics: Part A, pp. 519–557, Elsevier, Amsterdam, Netherlands.
33. Vandeginste BGM, Massart DL, Buydens LMC, De Jong S, Lewi PJ, Smeyers-Verbeke J. In: Handbook of Chemometrics and Qualimetrics: Part B, pp. 57-150, Elsevier, Amsterdam, Netherlands.
34. Young RC, Mitchell RC, Brown TH, Ganellin CR, Griffiths R, Jones M, Rana KK, Saunders D, Smith IR, Development of a new physicochemical model for brain penetration and its application to the design of centrally acting H2 receptor histamine antagonists. J Med Chem, 1988; 31:656–671.
35. Kelder J, Grootenhuis PDJ, Bayada DM, Delbressine LPC, Ploemen JP. Polar Molecular Surface as a Dominating Determinant for Oral Absorption and Brain Penetration of Drugs. Pharmac Research, 1999; 16:1514-1519.
36. Wang JS, Taylor R, Ruan Y, Donovan JL, Markowitz JS, De Vane CL, Olanzapine penetration into brain is greater in transgenic Abcb1a P-glycoprotein-deficient mice than FVB1 (wild-type) animals. Neuropsychopharmacol, 2004; 29:551-557.
37. Sills GJ, Kwan P, Butler E, de Lange ECM, van der Berg DJ, Brodie MJ. P-glycoprotein-mediated efflux of antiepileptic drugs: Preliminary studies in mdr1a knockout mice. Epilepsy Behav, 2002; 3:427-432.
38. Sung JH, Yu KH, Park JS, Tsuruo T, Kim DD, Shim CK, Chung SJ. Saturable distribution of tacrine into the striatal extracellular fluid of the rat: Evidence of involvement of multiple organic cation transporters in the transport. Drug Metab Dispos, 2005; 33:440-4488.
39. Muehlbacher M, Spitzer GM, Liedl KR, Kornhuber J. Qualitative prediction of blood–brain barrier permeability on a large and refined dataset. J Comput Aided Mol Des, 2011; 25:1095–1106.
40. Consonni V, Todeschini R, Pavan M. Structure/response correlations and similarity/diversity analysis by GETAWAY descriptors. 1. Theory of the novel 3D molecular descriptors. J Chem Inf Comput Sci, 2002; 42:682-692.
41. Consonni V, Todeschini R, Pavan M, Gramatica P. Structure/response correlations and similarity/diversity analysis by GETAWAY descriptors. 2. Application of the novel 3D molecular descriptors to QSAR/QSPR studies. J Chem Inf Comput Sci, 2002; 42:693-705.
42. Todeschini R, Vighi M, Finizio A, Gramatica P. 3D-modelling and prediction by WHIM descriptors. Part 8. Toxicity and physico-chemical properties of environmental priority chemicals by 2D-TI and 3D-WHIM descriptors. SAR QSAR Environ Res, 1997; 7:173-193.
43. Ertl P, Rohde B, Selzer P. Fast calculation of molecular polar surface area as a sum of fragment-based contributions and its application to the prediction of drug transport properties. J Med Chem, 2000; 43:3714-3717.
44. Ivanciuc O. QSAR Comparative Study of Wiener Descriptors for Weighted Molecular Graphs. J Chem Inf Comput Sci, 2000; 40:1412-1422.
45. Lanevskij K, Japertas P, Didziapetris R. Improving the prediction of drug disposition in the brain. Expert Opin Drug Metab Tox, 2013; 9:473-486.

Supplement Table 1. Values of the significant descriptors used in the QSPR (logPS) model.

Primary ID	Training Set	Mi	H7m	R3s+	Hy	TPSA(NO)
1	Amiloride	1.176	0.201	0.728	6.789	156.79
2	Brimonidine	1.135	0.036	0.339	1.309	63.81
3	Carvedilol	1.123	0.079	0.141	1.564	80.32
5	Clopamide	1.136	0.009	0.268	1.159	92.50
6	Clozapine	1.127	0.097	0.169	0.278	36.36
7	Doxazosin	1.133	0.064	0.247	0.252	112.27
8	Efaroxan	1.129	0.007	0.129	-0.291	33.62
9	Ephedrine	1.136	0.003	0.284	1.394	36.84
10	Guanfacine	1.136	0.000	0.289	2.350	78.97
11	Harmane	1.111	0.000	0.193	-0.305	28.68
12	Harmine	1.117	0.001	0.138	-0.291	37.91
13	Idazoxan	1.129	0.006	0.201	-0.200	42.85
15	Maprotiline	1.117	0.008	0.053	0.143	16.61
17	Moxonidine	1.156	0.024	0.260	1.433	73.04
18	Naphazoline	1.119	0.001	0.074	0.329	26.00
19	Olanzapine	1.127	0.031	0.070	0.304	36.36
22	Pseudoephedrine	1.136	0.002	0.260	1.394	36.84
23	Rilmenidine	1.142	0.007	0.125	-0.202	33.62
24	Tamsulosin	1.133	0.142	0.222	1.698	104.46
26	Tizanidine	1.136	0.104	0.232	1.433	63.81
28	Tramazoline	1.136	0.005	0.072	1.189	38.03
29	Xylometazoline	1.134	0.013	0.046	0.265	26.00
Primary ID	Test Set	Mi	H7m	R3s+	Hy	TPSA(NO)
4	Clonidine	1.135	0.035	0.313	1.453	38.03
14	Indapamid	1.122	0.015	0.293	1.080	92.50
16	Mianserine	1.118	0.014	0.061	-0.886	6.48
20	Oxymetazoline	1.136	0.011	0.188	1.044	46.23
21	Phenilephrine	1.139	0.002	0.332	2.456	57.07
25	Tetrahydrozoline	1.128	0.004	0.071	0.366	26.00
27	Triamterene	1.144	0.012	0.130	3.956	129.62

Supplement Table 2. Values of the significant descriptors used in the QSPR (logPS-brain) model.

Primary ID	Training Set	GATS1e	H7m	nHDon	T(N..N)	Hy
1	Amiloride	1.232	0.201	8	80	6.789
2	Brimonidine	1.075	0.036	3	38	1.309
3	Carvedilol	0.998	0.079	4	8	1.564
5	Clopamide	0.531	0.009	3	13	1.159
6	Clozapine	0.922	0.097	2	22	0.278
7	Doxazosin	0.891	0.064	2	28	0.252
8	Efaroxan	0.901	0.007	1	0	-0.291
9	Ephedrine	1.196	0.003	3	0	1.394
10	Guanfacine	0.816	0.000	4	0	2.350
11	Harmane	0.971	0.000	1	3	-0.305
12	Harmine	0.909	0.001	1	3	-0.291
13	Idazoxan	0.910	0.006	1	0	-0.200
15	Maprotiline	1.389	0.008	2	0	0.143
17	Moxonidine	0.997	0.024	3	26	1.433
18	Naphazoline	1.192	0.001	2	0	0.329
19	Olanzapine	1.125	0.031	2	22	0.304
22	Pseudoephedrine	1.196	0.002	3	0	1.394
23	Rilmenidine	0.894	0.007	1	0	-0.202
24	Tamsulosin	0.776	0.142	4	7	1.698
26	Tizanidine	0.916	0.104	3	29	1.433
28	Tramazoline	1.241	0.005	3	0	1.189
29	Xylometazoline	1.241	0.013	2	0	0.265
Primary ID	Test Set	GATS1e	H7m	nHDon	T(N..N)	Hy
4	Clonidine	0.869	0.035	3	0	1.453
14	Indapamid	0.528	0.015	3	13	1.080
16	Mianserine	0.992	0.014	0	3	-0.886
20	Oxymetazoline	1.105	0.011	3	0	1.044
21	Phenilephrine	1.158	0.002	4	0	2.456
25	Tetrahydrozoline	1.195	0.004	2	0	0.366
27	Triamterene	1.268	0.012	6	62	3.956

Supplement Table 3. Values of the significant descriptors used in the QSPR (logBB) model.

Primary ID	Training Set	nHet	SM1_Dz(i)	AVS_B(s)	TPSA(NO)
1	Amiloride	9	1.058	5.490	156.79
2	Brimonidine	6	0.776	4.694	63.81
3	Carvedilol	6	0.762	4.688	80.32
4	Clonidine	5	0.662	4.676	38.03
5	Clopamide	8	0.806	5.062	92.50
6	Clozapine	5	0.709	4.537	36.36
7	Doxazosin	10	1.096	4.821	112.27
8	Efaroxan	3	0.485	4.449	33.62
9	Ephedrine	2	0.335	4.478	36.84
10	Guanfacine	6	0.748	5.206	78.97
11	Harmane	2	0.372	4.780	28.68
12	Harmine	3	0.485	4.772	37.91
15	Maprotiline	1	0.203	4.320	16.61
16	Mianserine	2	0.372	4.396	6.48
18	Naphazoline	2	0.372	4.509	26.00
19	Olanzapine	5	0.595	4.436	36.36
20	Oxymetazoline	3	0.485	4.409	46.23
21	Phenilephrine	3	0.452	4.798	57.07
24	Tamsulosin	8	0.802	4.787	104.46
25	Tetrahydrozoline	2	0.372	4.302	26.00
27	Triamterene	7	0.946	5.064	129.62
28	Tramazoline	3	0.516	4.306	38.03
Primary ID	Test Set	nHet	SM1_Dz(i)	AVS_B(s)	TPSA(NO)
13	Idazoxan	4	0.586	4.602	42.85
14	Indapamid	8	0.806	5.206	92.50
17	Moxonidine	7	0.888	4.673	73.04
22	Pseudoephedrine	2	0.335	4.478	36.84
23	Rilmenidine	3	0.485	4.202	33.62
26	Tizanidine	7	0.775	4.782	63.81
29	Xylometazoline	2	0.372	4.215	26.00

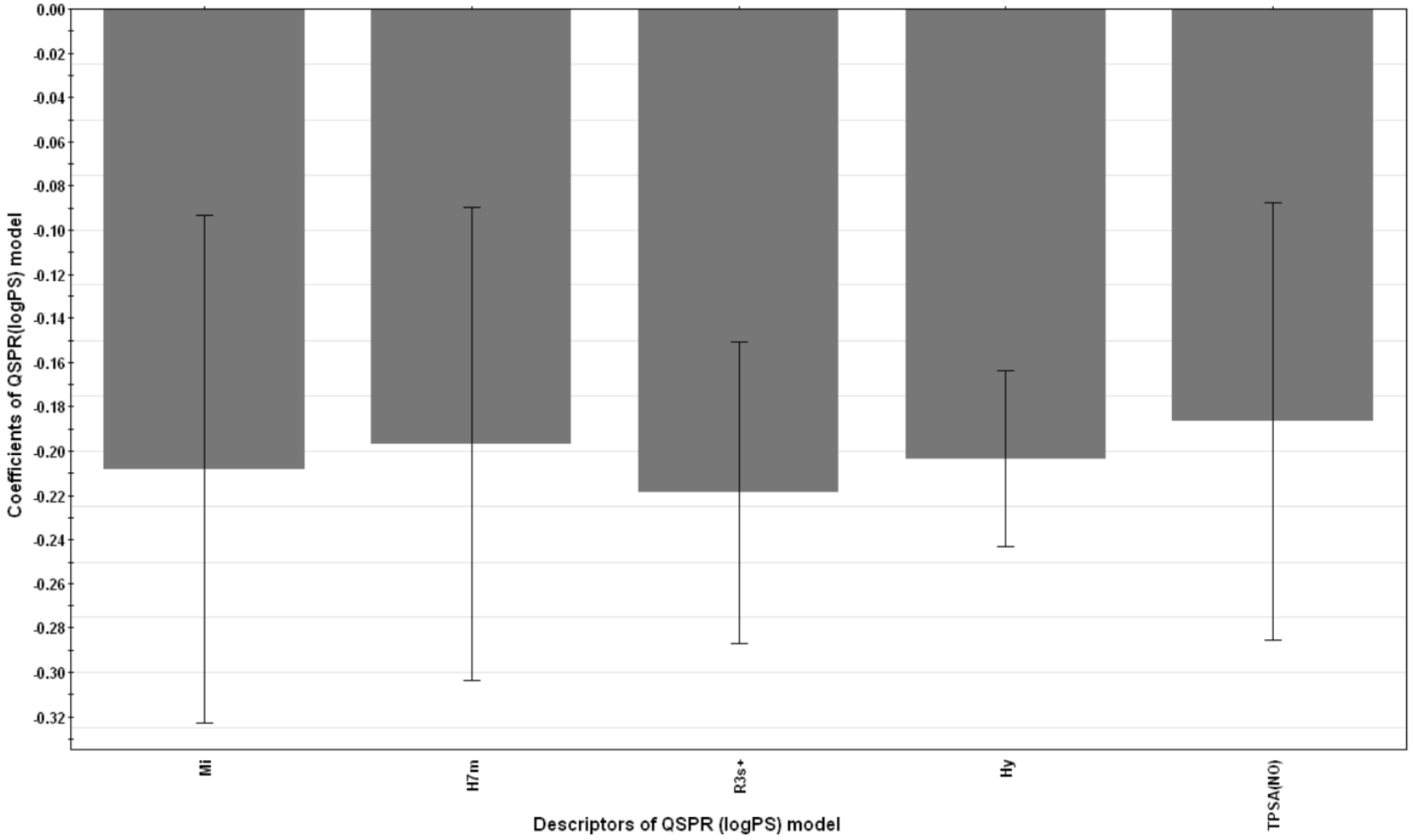

Figure 3. Coefficient Plots: **a)** QSPR(logPS) model.

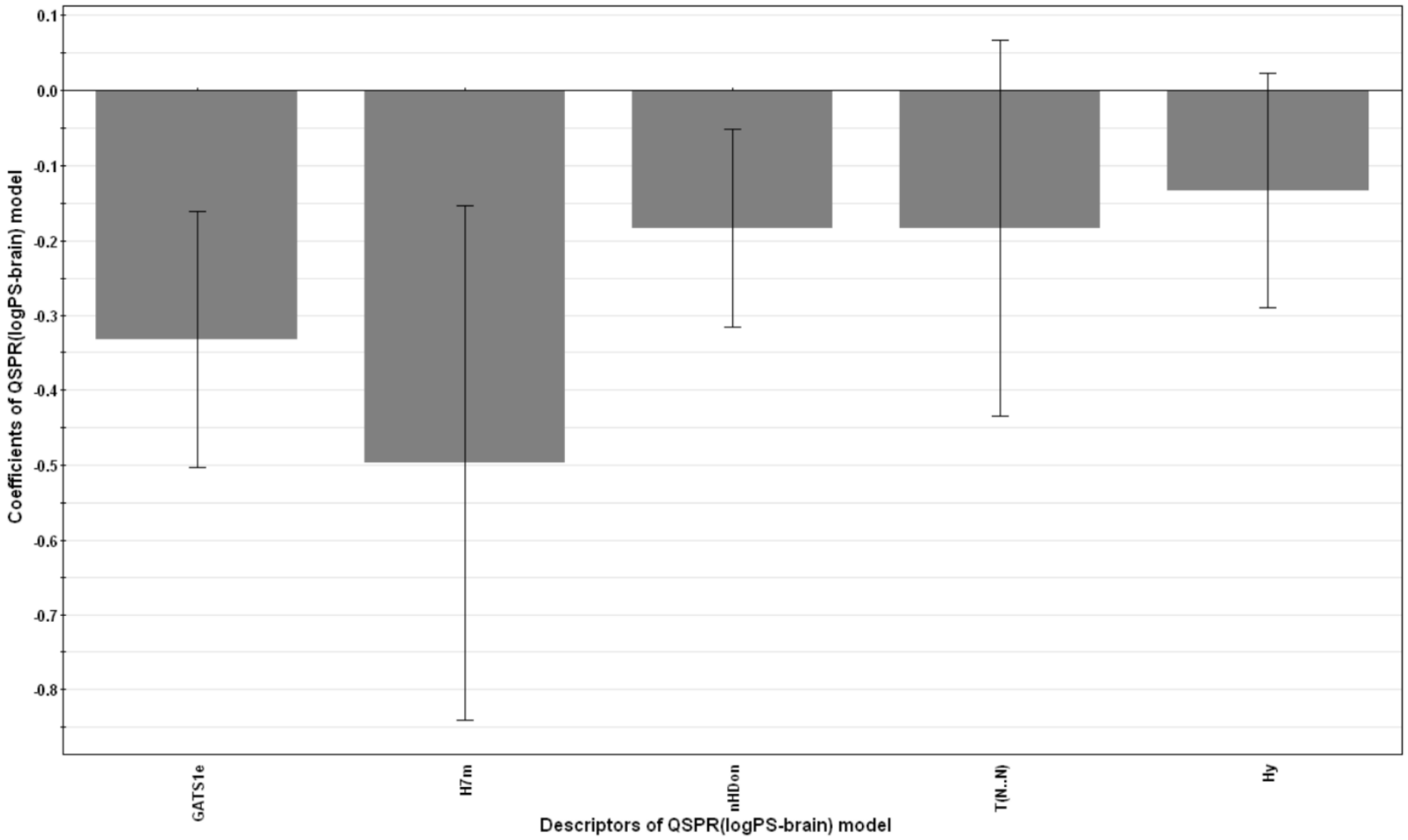

Figure 3. Coefficient Plots: **b)** QSPR(logPS-brain) model.

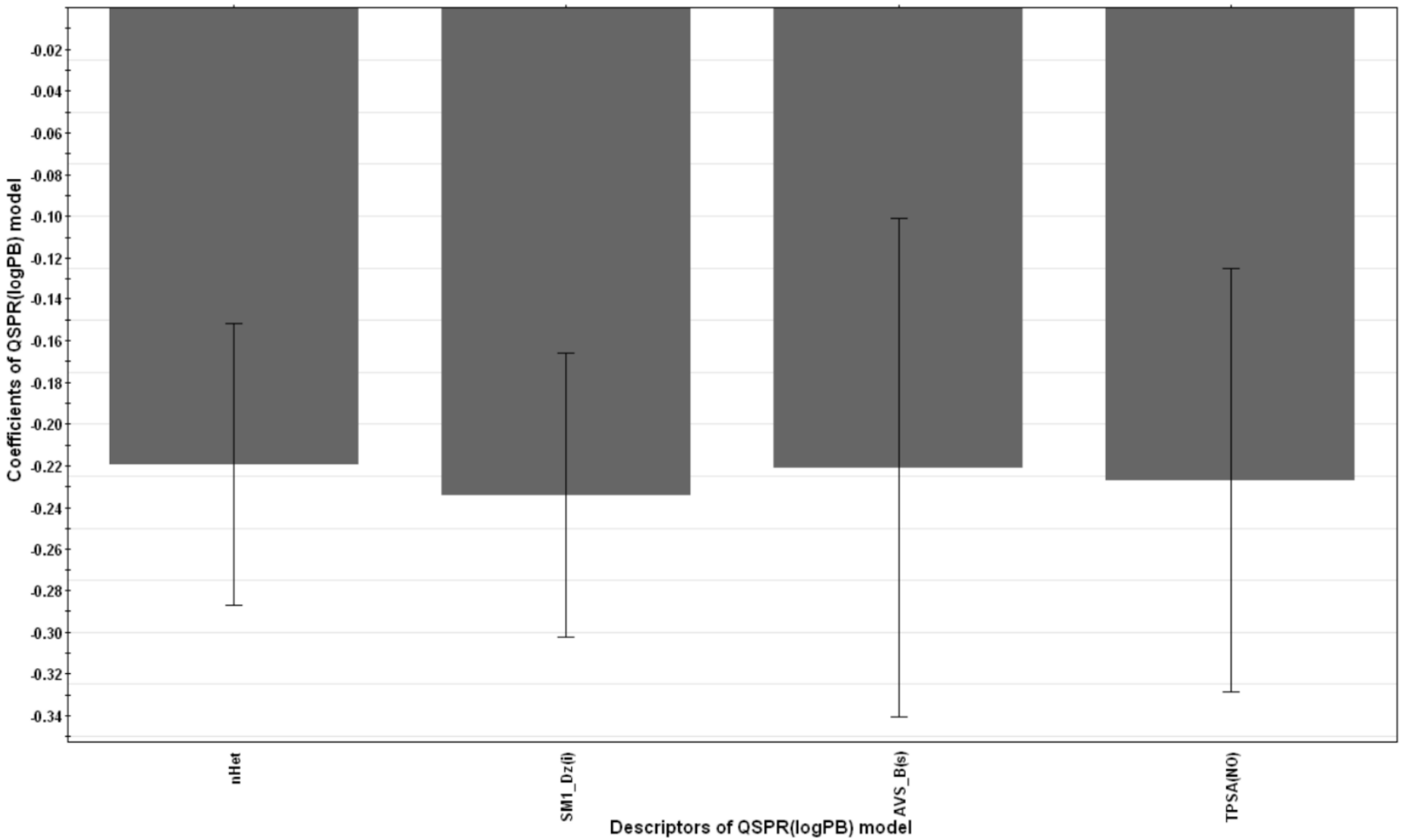

Figure 3. Coefficient Plots: **c)** QSPR(logBB) model.

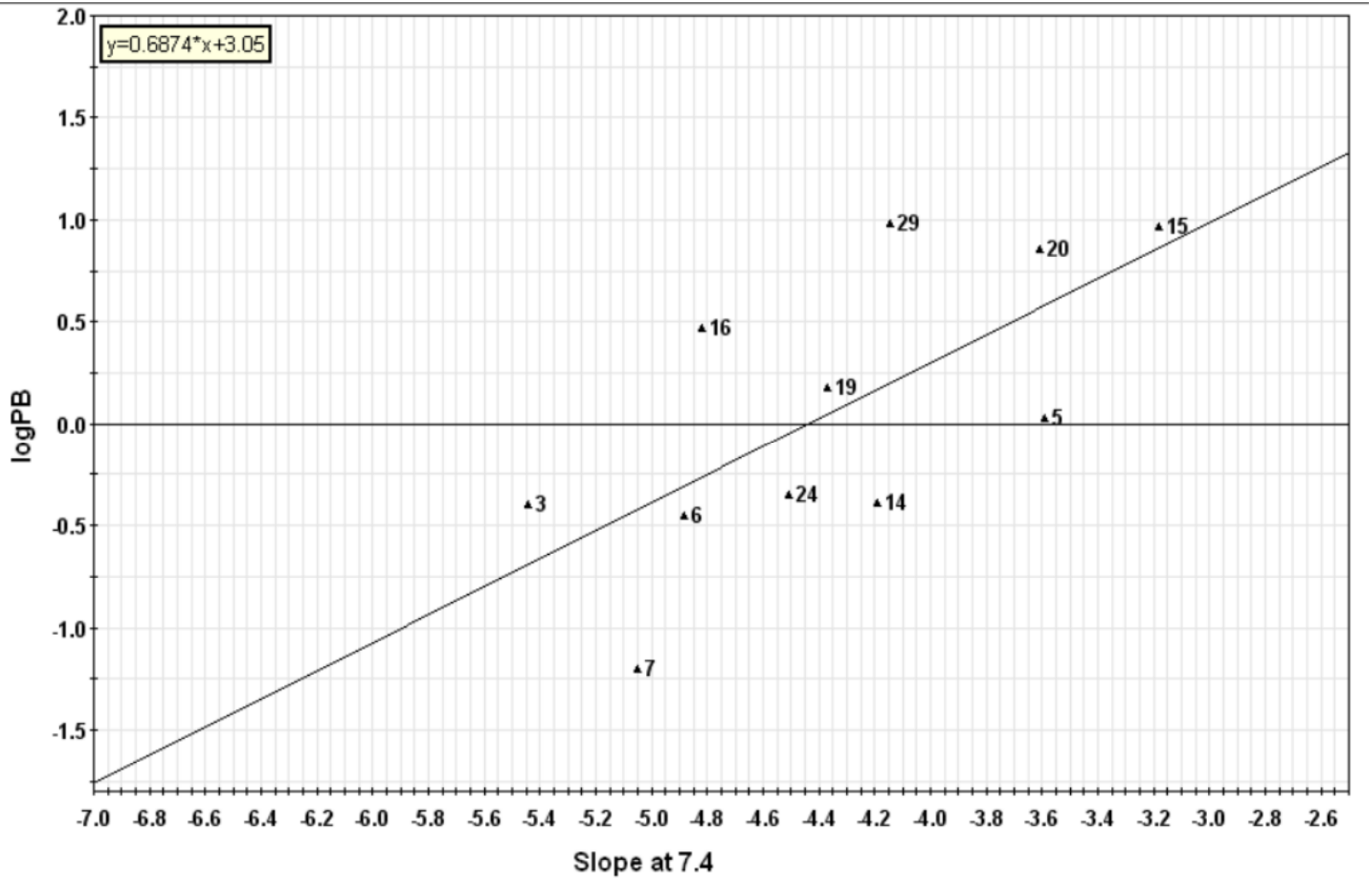

Figure 5. Linear correlation plots: **a)** HPLC retention parameter - Slope at 7.4 *vs.* extent of BBB permeation (logBB) in A/B cluster.

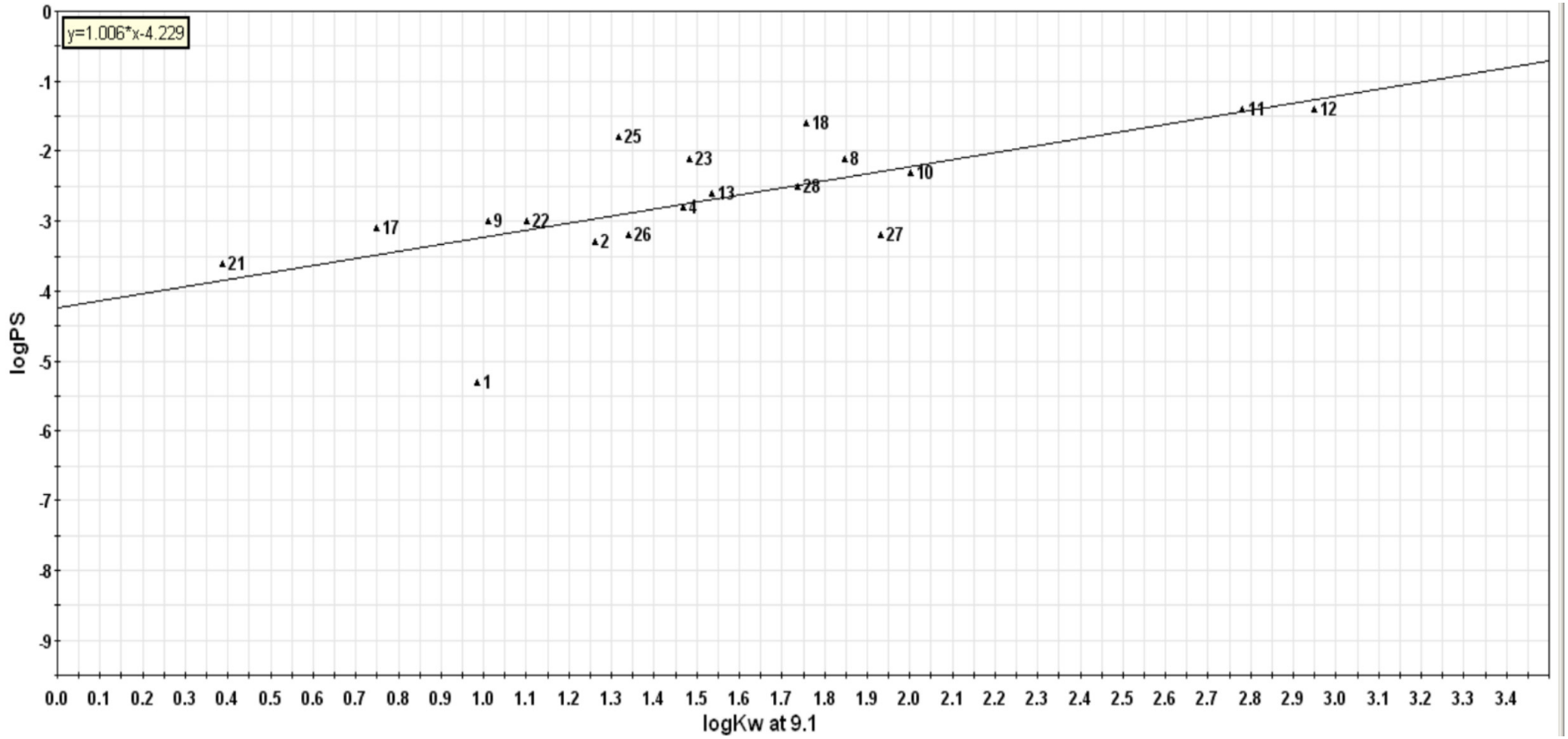

Figure 5. Linear correlation plots: **b)** HPLC retention parameter - logKw at 9.1 *vs.* rate of brain penetration (logPS) in C/D cluster.