

Article

Chromatographic Data in Statistical Analysis of BBB Permeability Indices

Karolina Wanat *  and Elżbieta Brzezińska 

Department of Analytical Chemistry, Faculty of Pharmacy, Medical University of Lodz, 90-419 Lodz, Poland; elzbieta.brzezinska@umed.lodz.pl

* Correspondence: karolina.wanat@umed.lodz.pl; Tel.: +48-42-677-92-11

Abstract: Blood–brain barrier (BBB) permeability is an essential phenomena when considering the treatment of neurological disorders as well as in the case of central nervous system (CNS) adverse effects caused by peripherally acting drugs. The presented work contains statistical analyses and the correlation assessment of the analyzed group of active pharmaceutical ingredients (APIs) with their BBB-permeability data collected from the literature (such as computational log BB; $K_{p,uu,brain}$, and CNS+/- groups). A number of regression models were constructed in order to observe the connections between the APIs' physicochemical properties in combination with their retention data from the chromatographic experiments (TLC and HPLC) and the indices of bioavailability in the CNS. Conducted analyses confirm that descriptors significant in BBB permeability modeling are hydrogen bond acceptors and donors, physiological charge, or energy of the lowest unoccupied molecular orbital. These molecular descriptors were the basis, along with the chromatographic data from the TLC in log BB regression analyses. Normal-phase TLC data showed a significant contribution to the creation of the log BB regression model using the multiple linear regression method. The model using them showed a good predictive value at the level of $R^2 = 0.87$. Models for $K_{p,uu,brain}$ resulted in lower statistics: $R^2 = 0.56$ for the group of 23 APIs with the participation of k_{IAM} .

Keywords: blood–brain barrier; BBB permeation; log BB; $K_{p,uu,brain}$; statistical modeling; data mining techniques; chromatographic retention data



Citation: Wanat, K.; Brzezińska, E. Chromatographic Data in Statistical Analysis of BBB Permeability Indices. *Membranes* **2023**, *13*, 623. <https://doi.org/10.3390/membranes13070623>

Academic Editor: Maria João Moreno

Received: 31 May 2023

Revised: 23 June 2023

Accepted: 24 June 2023

Published: 26 June 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

1.1. Prediction of CNS Availability

Predicting the possibility of drug penetration into the CNS is based on the behavior of drugs during pharmacokinetic distribution in the human body. This distribution is limited due to the existence of the blood–brain barrier (BBB), whose task is to maintain brain homeostasis by limiting the penetration of endogenous and exogenous chemical compounds. For this reason, the search for new, potential drugs should include their possible bioavailability in the CNS [1]. The blood–brain barrier is the most extensive “firewall”, which includes various defense mechanisms. In addition to the physical blockade, which is constituted by biological membranes connected by the tight and adherens junctions (TJs and AJs), there are also mechanisms for the active removal of xenobiotics (efflux transporters; P-glycoprotein) [2] or an enzymatic activity [3,4].

Biological experiments determining the level of drug distribution to the brain are the best source of information. However, these experiments are extremely time-consuming, expensive, and difficult to access in extensive screening studies, especially with the use of structure libraries. The most frequently used parameters in determining penetration through the BBB is log BB [5–7]. Log BB can be defined as the logarithm of the ratio of the drug concentration in the brain to the concentration of the drug in the blood in the state of dynamic equilibrium of the drug distribution between the blood and the brain, and it is considered as a convenient parameter in the assessment of BBB permeability [8]. Over

the years, the $K_{p,uu,brain}$ parameter has gained a lot of attention and its use is becoming more and more common in research [9–11]. In Table 1, the recent approaches in predicting BBB permeability *in silico* and *in vitro* are listed; computational methods seem to be more focused on the properties of the API, whereas *in vitro* modeling is based on the imitation of the BBB itself and the surrounding environment.

Table 1. *In silico* and *in vitro* approaches in predicting BBB penetration. *In silico* methods are described using molecular descriptors used in model building; *in vitro* methods by the type of BBB model used.

Prediction of BBB Penetration	
In Silico	In Vitro
Solvation free energies in various solvents of different polarities [12]	Single endothelial cell monolayer model [13]
1D, 2D, and 3D molecular descriptors and fingerprints of a molecule [14]	Stem cell modeling of the BBB: neural progenitor cells (NPCs), the precursors to neurons, astrocytes or oligodendrocytes [15]
Molecular descriptors: carboxylic acid group, polar surface area (PSA)/hydrogen-bonding ability, lipophilicity, and molecular charge [16]	Co-culture model (brain microvascular endothelial cells: BMECs with astrocytes or pericytes) [17]
1D and 2D physicochemical properties, molecular access system fingerprints -MACCS and substructure fingerprints [18]	The microfluidic BBB [13]—microfluidic devices to mimic biological environments: -observation of specific markers of TJs [19] -measuring TEER [20] -permeability assessment [21]
HPLC data (IAM, HSA and AGP columns) and molecular descriptors to model brain disposition of drugs: $K_{p,uu,brain}$ [11]	

1.2. Computational Modelling of BBB Penetration

During preliminary research, drug design, and optimization, the simplest methods are used, often not even requiring the synthesis of the designed candidates. This is the phase of designing chemical structures, the construction of which allows one to achieve the expected physicochemical properties. Therefore, only *in silico* observation of their structure and the prediction of properties is possible.

Chromatography is an analytical technique for separating mixtures of substances; the mechanism of operation of this method is based on the differences in the affinity of the components of the mixtures toward the two phases present in the analytical system. The flow of the mobile phase through the stationary phase initiates the elution of compounds bound or adsorbed on the stationary phase. This process proceeds at a different rate for each component of the mixture. To describe this phenomenon in thin-layer chromatography (TLC), the retardation factor R_f is used, while in the case of high-performance column chromatography (HPLC), the retention factor values k are collected [22,23]. In our analyses, we tried to include the chromatographic retention data (TLC and HPLC) to increase the statistical modeling capabilities. Some authors have also attempted to implement chromatographic data to predict penetration into the CNS with satisfactory results [11], but the main focus is on HPLC, where the possibilities of the thin layer are investigated in this paper. Previous experiments [24–26] have shown a close relationship between the RP-18 TLC chromatographic data and bioavailability to the central nervous system. TLC is such a simple and economical option that it is worth updating its use from time to time, as it can be provided as a first step in chemical structure analysis before moving on to more expensive methods.

The experiment presented in the paper concerns the connection of retention data obtained from various chromatographic systems with the ability of drugs to penetrate into the central nervous system ($\log BB$; CNS+/-). The aim of the work was to find a model of BBB penetration using physicochemical properties associated with retention data

from the TLC and HPLC experiments. The dependent variables used in the modeling were two computational parameters labelled as B1 and B2, representing the log BB values. These parameters were adopted on the basis of the bibliography [24–27]. The calculation parameter B2 is quantitative and corresponds to $\log BB = 0.547 - 0.016 \text{ PSA}$ [7,28]. The second parameter that describes log BB: B1 [29] was obtained from the online calculator. B2 and B1 are later described as indices of BBB permeability. Log BB was chosen as the main indicator of permeation through the BBB due to its popularity and the wide availability of its values for a large number of chemical compounds. The third dependent variable, $K_{p,uu,brain}$, was collected for only 29 APIs [30] out of all 181 drugs included in the study. An attempt was also made to build a simple regression model for this index.

The limits of the log BB index are different for many of the proposed BBB penetration prediction models. The optimal classification threshold is usually between 0 and -1 [27,31,32]. In one study, BBB+ (crossing the barrier) compounds were found to have a log BB value greater than 0 [33]. For BBB− (not crossing the barrier) compounds, a cut-off value of $\log BB < -0.3$ was established [33]. The log BB limit = -0.52 is a logical division between BBB+ and BBB−. It corresponds to a 30% ratio of the concentration of the compound in the brain to the concentration in plasma [27]. Studies conducted earlier in the department defined the BBB+ limit as $\log BB \geq -0.9$ [24–26]. Over the past 20 years, a number of different analyses have appeared to assess the bioavailability of drugs to the CNS using log BB values. The accuracy of these models ranges from 75% to 99% [34].

All mentioned BBB-permeability indices—B2, B1, and $K_{p,uu,brain}$ —were taken into account when analyzing their relationship with the properties of the tested APIs. They are all correlated with the molecular descriptors at different levels, which is presented in the correlation matrix (Table S3, Supplementary Materials). Indicators $B2 > -0.52$ and $B2 > -0.9$ were also provided in the analysis where the API group is restricted to compounds with such log BB values only.

2. Materials and Methods

2.1. Chromatographic Experiments

A total of 181 active pharmaceutical ingredients were isolated from the pharmaceutical preparations. The purity of the isolated substances was checked by TLC chromatography and densitometric scanning. The obtained API was dissolved in 99.8% methanol to give 1 mg/mL solutions that were then used for TLC and HPLC chromatography.

APIs were subjected to chromatographic in column (HPLC) and thin layer (TLC) formats. TLC chromatography was carried out on 20 cm × 20 cm glass plates by Merck, covered with silica gel and the addition of a fluorescent indicator. Merck TLC silica gel 60 F254 was used for the normal phase (NP). In the reverse phase (RP), RP-2 silanized plates were used (Merck TLC Silica gel 60 RP-2 F254, silanized). The mobile phase was buffered with ammonium acetate to pH 7.4 (LACH-NER, ammonium acetate). The mobile phase composition was water:methanol:acetonitrile with the HPLC gradient grade (60:20:20 (v/v/v)).

Plates, after developing the chromatograms, were scanned using a densitometer (Desaga CD 60) with an analytical wavelength selected individually for drugs. The values of R_f (molecular descriptors: NP, RP) were determined.

Chromatographic data also came from an experiment with a HPLC IAM (immobilized artificial membrane) column (Regis Technologies Inc.: IAM.PC.DD.2, 10 μm ; 4.6 mm × 10 mm). The column contains a standard membrane consisting of a phospholipid bilayer. Organic solvents (acetonitrile and methanol) and water (J.T. Baker HPLC) were used in HPLC. Sigma phosphate buffered saline tablets were used to prepare the phosphate buffer. The mobile phase was a mixture of 10 mM phosphate buffer at pH 7.4 and acetonitrile in the ratio of 80:20 (v/v) [35]. The mobile phase flow was set at 0.5 mL/min; detection was carried out with a UV-Vis detector. The obtained values of the retention factor (k IAM) were transformed into a logarithmic form and entered into the statistical analysis (descriptor: $\log k$ IAM).

2.2. Statistical Analyses and Molecular Descriptors (MDs)

Correlation analyses, analysis of variance (ANOVA), and multiple linear regression (MLR) were performed on a basic set of 181 active pharmaceutical ingredients (APIs). In recent years, more and more attention has been paid to data mining techniques that consist of searching for dependencies between variables in datasets. The set of the studied APIs was then analyzed using the MARSplines method, which is a non-parametric regression algorithm. The general mechanism of this method is described as multiple piecewise linear regression, built from basis functions. The boundaries of each section define the “applicability ranges” of the individual linear equations.

All statistics analyses were performed using STATISTICA 13.3 software (TIBCO Software Inc. Palo Alto, CA, United States) using the available functions. Molecular descriptors (MDs) connected to the physicochemical, pharmacokinetic, or chromatographic properties of APIs are listed in Table 2. The dataset for the 181 compounds and their structures are available in the Supplementary Materials (Tables S1 and S2).

Table 2. List of molecular descriptors (MDs) used in the statistical analyses.

MD	Description	Source
B2	Computational parameter, determines penetration through the blood–brain barrier: $\log BB = 0.547 - 0.016 PSA$	[7,28]
B1	Computational parameter, corresponds to $\log BB$	SwissADME
CNS+/-	Describes the bioavailability in the CNS	Drugbank
eL	Energy of the lowest unoccupied molecular orbital	Hyperchem
eH	Energy of the highest occupied molecular orbital	Hyperchem
eL-eH	Ionization capacity	Hyperchem
HA	Hydrogen bond acceptors	Hyperchem
HD	Hydrogen bond donors	Hyperchem
Kp,uu,brain	Unbound brain-to-plasma drug partition coefficient	[30]
log D	Distribution coefficient	ACD Labs
log k IAM	Logarithm of retention factor from HPLC IAM	HPLC
log P	Partition coefficient	Hyperchem
log U/D	Describes the extent of ionization, calculated from pKa, according to the equations: $pKa - pH$ (acids) or $pH - pKa$ (bases)	ACD Labs
MW	Molecular weight	Hyperchem
NP	The R_f from NP TLC plate	TLC
PB	Protein binding	Drugbank
PhCharg	Charge of a compound in physiological environment	Drugbank
PSA	Polar surface area of a molecule	Hyperchem
RP	The R_f from RP-2 TLC plate	TLC
Sa	Surface area of a molecule	Hyperchem
V	Molecular volume	Hyperchem

3. Results

3.1. Distribution of $\log BB$ and the Retention Values in the CNS Penetration Groups

The computational parameters B1 and B2 were examined first—these indicators were analyzed in terms of the distribution of their values in the CNS+/- groups (entering the CNS/not entering the CNS) [36] to check whether there was a significant difference in the $\log BB$ in both groups. The non-parametric Mann–Whitney U test remained due to the lack of normal distribution of B1 and B2 values in the CNS+/- groups (Table A1, Appendix A).

Among these indices, a better result was obtained with B2. Its median values in both the CNS+/- groups were statistically significantly different from each other, which was also clearly noticeable in the box plot (Figure 1). In a similar way, the distribution of the retention values from the chromatographic experiments (NP TLC, RP TLC, and HPLC IAM) was examined, together with the derivatives of these values by using the physicochemical properties—the NP and RP values were divided by the physicochemical properties of a given API. The obtained derivatives were as follows: NP/PSA; NP/PB; NP/ $\log P$; NP/Sa;

NP/V; NP/MW; NP/log D. The same descriptors were calculated for the retention values from RP TLC (descriptor RP) and HPLC IAM (descriptor log k IAM) (Table A2).

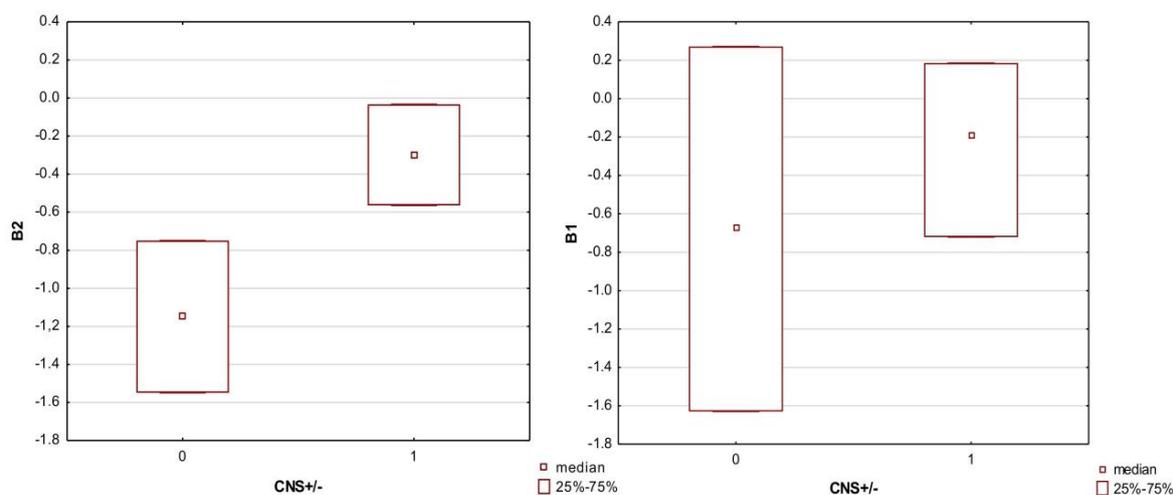


Figure 1. Boxplot of the B1 and B2 indices in the CNS+ and CNS– groups. Medians were compared due to the application of non-parametric tests. Code 0—CNS– group; code 1—CNS+ group.

The TLC descriptors showed several significant differences in the median values in the CNS+ and CNS– groups; these are all listed in Table A2. The best division visible in the boxplots was obtained for the descriptors RP, RP/PSA, and NP/PSA. The large influence of the polar surface area of the molecule PSA on the separation between the CNS+ / – groups was shown when comparing the graphs for the retention factor, RP, and its derivative, RP/PSA (Figure 2).

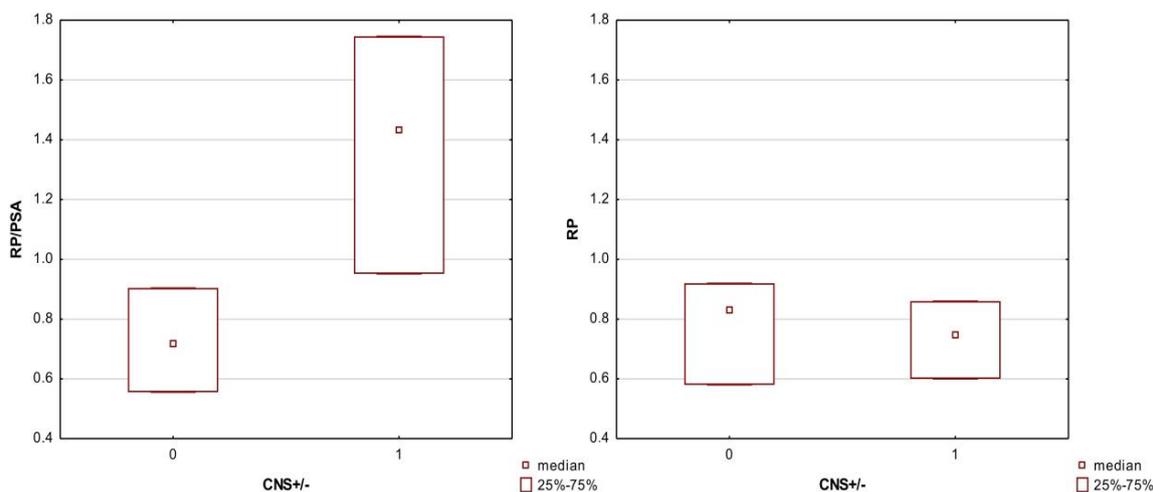


Figure 2. Boxplot of the RP and RP/PSA descriptors in the CNS+ and CNS– groups. Medians were compared because non-parametric tests were used. Code 0—CNS– group; code 1—CNS+ group.

Additionally, in the case of the HPLC descriptor log k IAM, a better division of the values between the CNS+ / – groups was obtained, when the PSA parameter was also used (Tables A3 and A4). When comparing both boxplots from Figure 3 directly, it should be noted that non-parametric tests were used for log k IAM due to the lack of normality of the distribution of this descriptor in the CNS+ / – groups; the graph shows the medians and 25–75% of cases (Figure 3B). Normal distribution, however, was characterized as log k IAM/PSA values in both groups, and here the graph shows the means along with the standard deviation (SD) (Figure 3A).

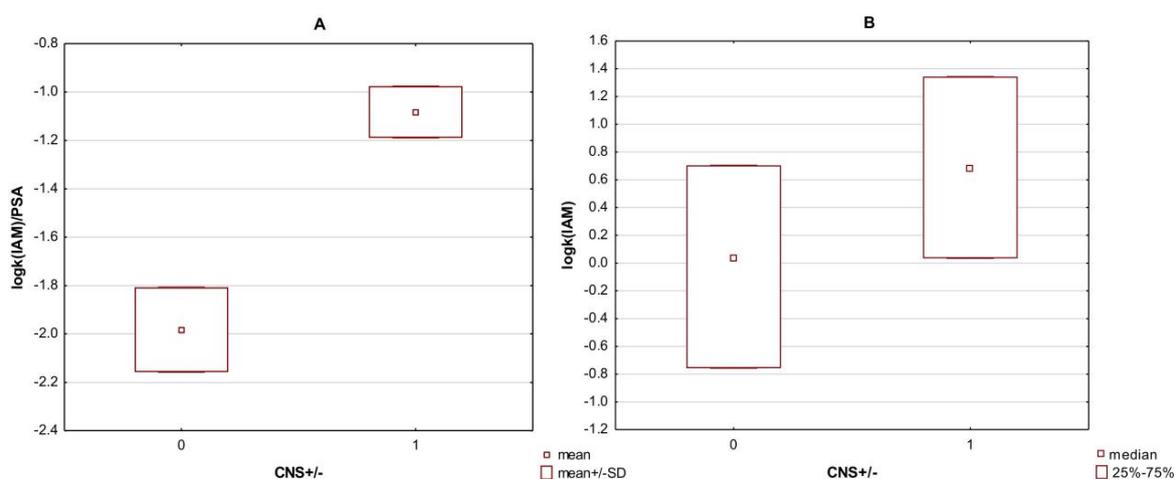


Figure 3. Boxplot of the log k IAM (B) and log k IAM/PSA (A) values in the CNS+ and CNS– groups. Medians were compared in the case of log k IAM. Mean values were compared for log k IAM/PSA. Code 0—CNS– group; code 1—CNS+ group.

The last analysis of this type was a simple ANOVA to check whether there were significant differences in the values of the TLC and HPLC descriptors in the groups of $B2 > -0.9$ and $B2 > -0.52$. A working code for the B2 values was adopted, where the limits for individual groups 1–2–3 are the mentioned -0.52 and -0.9 :

- B2 code 1—below -0.9 ;
- B2 code 2— -0.89 to -0.52 ;
- B2 code 3—above -0.52 .

The TLC descriptors that showed significant differences in the B2 code groups were NP and RP and their derivatives in combination with the percentage of plasma protein binding: NP/PB and RP/PB. The exception was that only the NP descriptor showed significant differences in the median values between group 1 and groups 2 and 3 (Table A5 and Figure 4). The rest of the parameters differed only between groups 1 and 3, that is, the extreme values of the B2 code influenced the differences in the retention values.

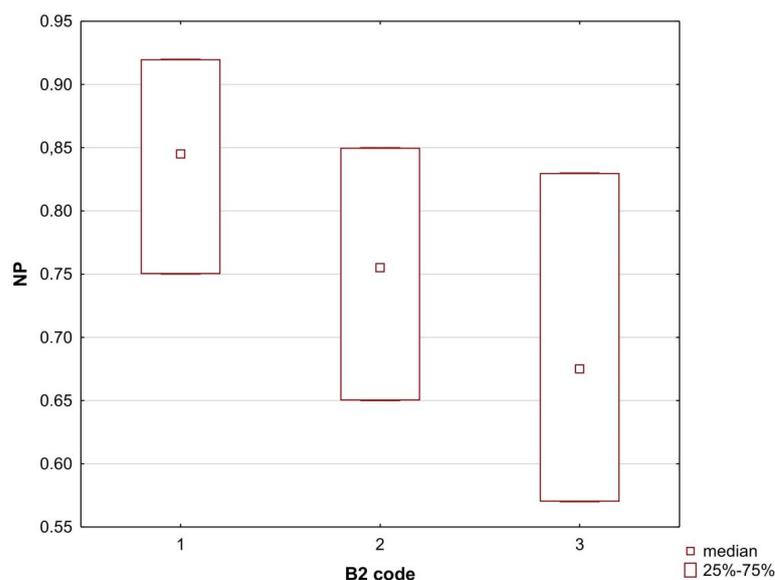


Figure 4. Boxplot of the NP median values in the B2 code groups: 1, 2, and 3. According to the Kruskal–Wallis ANOVA, the differences between medians in groups 1 and 2 and groups 1 and 3 were statistically significant; $p < 0.05$.

In the case of the IAM column, significant differences (also in the median values as non-parametric tests were used again) were revealed for the log k IAM and two derivatives: again in combination with PB: log k IAM/PB and additionally log k IAM/log P. These differences also only concerned the extreme groups B2 code: 1 and 3.

3.2. Regression Models with TLC Descriptors for B2 and B1 Indices

The next step of the study was the construction of regression models for the BBB permeation indices. For this purpose, descriptors from TLC, HPLC and their derivatives were used again in combination with their basic physicochemical properties (from Table 2). The aim of the MLR was to check whether chromatographic descriptors can be included among the factors influencing the level of log BB indices. Additionally, the usefulness of the physicochemical derivatives of the NP and RP descriptors was assessed in comparison with the standard values. Table 3 shows the aggregated MLR results for three dependent variables: B2, B2 > -0.52, and B2 > -0.9. Complete regression equations are available in Appendix A (Equations (A1)–(A10)).

Table 3. The MLR results of the B2, B2 > -0.52, and B2 > -0.9 indices for the NP and RP TLC retention data.

	B2	B2 > -0.52	B2 > -0.9
RP	R = 0.9214 R ² = 0.8643 R ² adj. = 0.8589 F(5, 158) = 200.15; <i>p</i> > 0.000 N = 164	R = 0.7374 R ² = 0.5433 R ² adj. = 0.5091 F(6, 82) = 16.18; <i>p</i> > 0.000; N = 89	R = 0.8432 R ² = 0.7112 R ² adj. = 0.6924 F(6, 104) = 42.27; <i>p</i> > 0.000 N = 111
RP derivatives	RP derivatives do not enter the model	R = 0.7287 R ² = 0.5422 R ² adj. = 0.5133 F(4, 84) = 24.24; <i>p</i> > 0.000 N = 89 derivative: RP/MW	R = 0.7842 R ² = 0.6011 R ² adj. = 0.5924 F(4, 106) = 40.30; <i>p</i> > 0.000 N = 111 derivative: RP/V
NP	R = 0.9332 R ² = 0.8713 R ² adj. = 0.8591 F(5, 158) = 240.15; <i>p</i> > 0.000 N = 151	NP does not enter the model	R = 0.8572 R ² = 0.7412 R ² adj. = 0.7156 F(7, 103) = 41.43; <i>p</i> > 0.000 N = 111
NP derivatives	R = 0.9322 R ² = 0.8734 R ² adj. = 0.8555 F(6, 144) = 160.07; <i>p</i> > 0.000 N = 151 derivative: NP/MW	R = 0.7813 R ² = 0.6122 R ² adj. = 0.5843 F(7, 81) = 18.249; <i>p</i> > 0.000 N = 89 derivative: NP/MW	R = 0.8142 R ² = 0.6481 R ² adj. = 0.6343 F(5, 105) = 38.703; <i>p</i> > 0.000 N = 111 derivative: NP/V

The regression summary of the dependent variable B2 yielded the best result (R² = 0.86 and R² = 0.87); the other indices, B2 > -0.52 and B2 > -0.90, were less successful. They explained from 54 to 74% of the variance of the variability of these indicators. In the MLR model using all B2 values, the chromatographic data entered as one of the independent variables were both RP and NP. In models of a limited number of compounds (B2 > -0.9 and B2 > -0.52), it was also possible to introduce derivatives of the retention coefficient: RP/MW, RP/V, NP/MW, and NP/V. The highest score was obtained for the NP descriptor. The resulting mathematical model explained 87% of the variance of variable B2 (Equation (1)) (Figure 5).

$$B2 = 0.652 - 0.167 HA + D - 0.215 NP + 0.007 eL + 0.037 PhCharg$$

$$R = 0.9332; R^2 = 0.8713; R^2 \text{ adj.} = 0.8624; F(5, 158) = 240.15; p > 0.000; s = 0.24; N = 151 \quad (1)$$

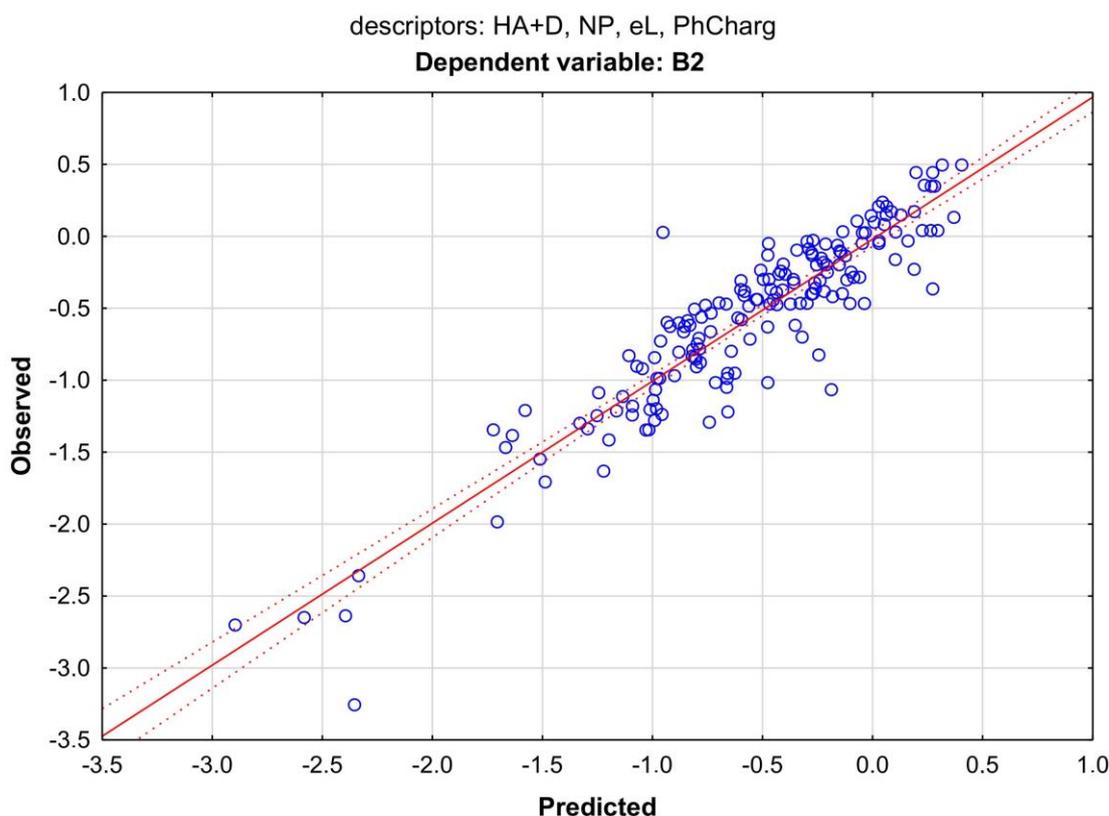


Figure 5. Scatter plot of the MLR model from Equation (1). Dependent variable: B2; predictors: HA + HD, NP, eL, and PhCharg. Coefficient of determination $R^2 = 0.87$.

In all multiple linear regression equations of the analyzed indices, B2, $B2 > -0.52$, and $B2 > -0.90$, there were MDs related to the number of possible hydrogen bonds formed by the tested compounds: HA—acceptors and HA + HD—sum of acceptors and donors. The second important parameter that appeared in this equation was the charge of the molecule in physiological conditions (i.e., at pH around 7.4—PhCharg and the energies of the lowest unoccupied molecular orbital (eL) or the difference between HOMO and LUMO (eH-eL)). For $B2 > -0.52$ and $B2 > -0.90$, other descriptors were also common: Sa (surface area of the molecule) and PB (level of plasma protein binding). These physicochemical parameters, related to the “rule of five” described above, confirm their relationship with the description of CNS bioavailability.

The MLR was performed using another computational parameter that describes log BB: B1, considered as the alternative to B2. The direct correlation between B2 and B1 was $R = 0.52$ for 143 cases of APIs; for $B2 > -0.9$ and $B2 > -0.52$, such correlation was lower, around 0.34–0.38 (Table 4).

Table 4. Correlations, $N = 143$ (missing data were removed by cases). All correlations were statistically significant, with $p < 0.05$.

N = 143	Mean	SD	B1	B2	$B2 > -0.9$	$B2 > -0.52$
B1	−0.4901	0.9472	1.0000	0.5192	0.3776	0.3350
B2	−0.6152	0.6802	0.5192	1.0000	0.7832	0.7491
$B2 > -0.9$	0.7273	0.4469	0.3776	0.7832	1.0000	0.6708
$B2 > -0.52$	0.5455	0.4997	0.3350	0.7491	0.6708	1.0000

The only model for the B1 indicator that included TLC descriptors was successfully built with the derivative of the RP retention value, namely, RP/MW. The model explained only 53% of the B1 variance, and again, the number of hydrogen bond donors and acceptors

ranked in first place among the factors determining B1 variability. Although the influence of the RP/MW descriptor stood out from other independent variables in the model, it is also together with HA + D, as the only one qualified as statistically significant, with a significance level of $p > 0.05$ (Equation (2)) (Figure 6).

$$B1 = 0.630 - 0.189 HA + D + 0.109 V - 2.091 RP/MW - 0.040 pKa + 0.021 eL + 0.125 Sa \tag{2}$$

$R = 0.7323; R^2 = 0.5314; R^2 \text{ adj.} = 0.5087; F(1, 113) = 21.61; p > 0.000; s = 0.69; N = 134$

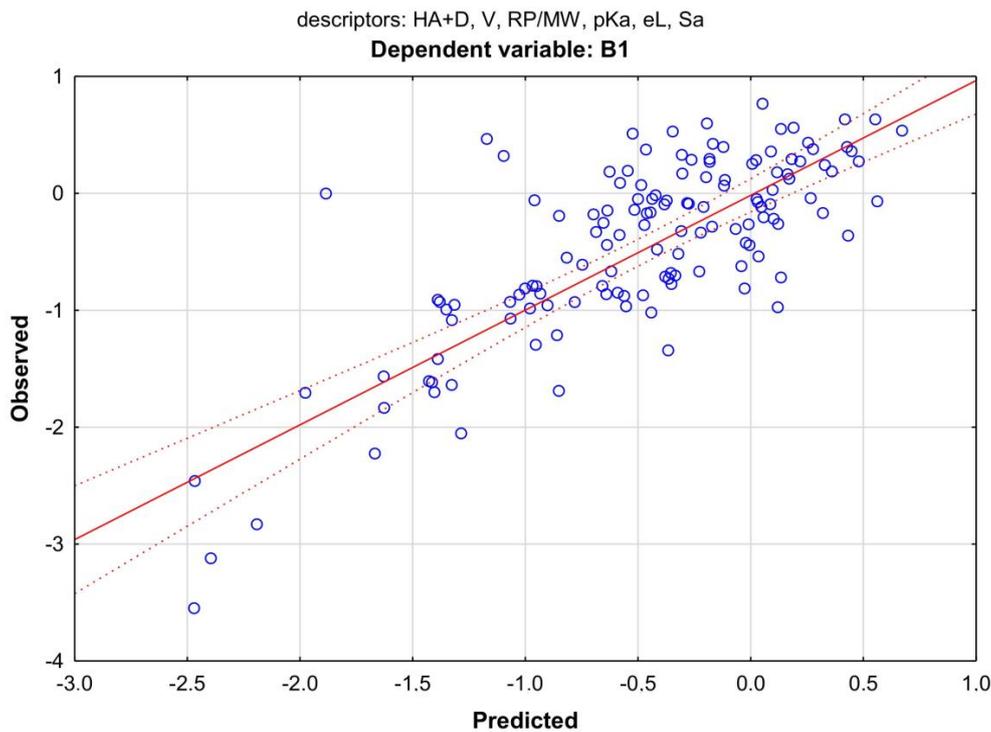


Figure 6. Scatter plot of the MLR model from Equation (1). Dependent variable: B1; predictors: HA + HD, V, RP/MW, pKa, eL, and Sa. Coefficient of determination $R^2 = 0.53$.

3.3. Regression Models with HPLC IAM Descriptors for B2, B1 $K_{p,uu,brain}$ Indices

The HPLC IAM descriptor was also included in the regression equations of the B2, B2 > -0.52, and B2 > -0.90 indices with less success, so it was possible to introduce its derivatives; only log k IAM/PB participated in building the model for B2 > -0.52. The best result was obtained for B2; the coefficient of determination was 0.89, which was the highest of all MLR models (Table 5).

Table 5. The MLR results of the B2, B2 > -0.52, and B2 > -0.9 indices for the HPLC IAM retention data.

	B2	B2 > -0.52	B2 > -0.9
log k IAM	R = 0.9521 R ² = 0.8938 R ² correct. = 0.8877 F(5, 118) = 197.87; $p < 0.000$ N = 124	R = 0.8687 R ² = 0.7644 R ² correct. = 0.7456 F(5, 60) = 37.67; $p < 0.000$ N = 66	R = 0.8933 R ² = 0.8024 R ² correct. = 0.7714 F(9, 73) = 31.85; $p < 0.000$ N = 83
log k IAM derivatives	log k IAM derivatives do not enter the model	R = 0.8742 R ² = 0.7633 R ² correct. = 0.7434 F(5, 60) = 37.15; $p < 0.000$ N = 66 derivative: log k IAM/PB	log k IAM derivatives do not enter the model

$$\begin{aligned}
 B2 &= 0.598 - 0.171 \text{ HA} + D + 0.013 \text{ eL} + 0.035 \log k \text{ IAM} + 0.016 \log U/D - 0.0007 \text{ MW} \\
 R &= 0.9513; R^2 = 0.8934; R^2 \text{ correct.} = 0.8924; F(5, 118) = 197.87; p > 0.000; s = 0.19; \\
 N &= 124
 \end{aligned}
 \tag{3}$$

Unfortunately, the log k IAM parameter entered the model (Equation (3)) with a rather high level of significance, $p = 0.142$, which may not fully prove its usefulness in predicting the B2 indicator. A similar situation occurred with the regression model for the B1 indicator (Equation (4)); here too, the log k IAM parameter entered the equation, with a p value above the assumed significance level of 0.05 ($p = 0.093$). The statistics of the model were also lower than those of the B2 index; R^2 was only 0.51.

$$\begin{aligned}
 B1 &= -0.030 - 0.172 \text{ HA} + D + 0.271 \text{ V} + 0.031 \text{ eL} + 0.004 \text{ MW} + 0.159 \log k \text{ IAM} - 0.027 \text{ pKa} \\
 R &= 0.7112; R^2 = 0.5143; R^2 \text{ correct.} = 0.4842; F(6, 117) = 20.03; p > 0.000; s = 0.70; \\
 N &= 124
 \end{aligned}
 \tag{4}$$

On the other hand, the IAM retention factor, descriptor k IAM, was the only chromatographic data that entered into the MLR model for the dependent variable $K_{p,uu,brain}$. The equation was built on a small group of compounds because only 23 of them were involved in the model. The coefficient of determination was close to the result from the B1 model, $R^2 = 0.56$, so only slightly more than 50% of the variation in $K_{p,uu,brain}$ can be explained by a model (Equation (5)) (Figure 7).

$$\begin{aligned}
 K_{p,uu,brain} &= 3.541 + 0.071 k \text{ IAM} - 0.541 \log D - 0.331 \text{ DM} - 0.836 \text{ PhCharg} \\
 R &= 0.7542; R^2 = 0.5641; R^2 \text{ correct.} = 0.4681; F(4, 18) = 5.820; p > 0.0291; s = 1.10; N = 23
 \end{aligned}
 \tag{5}$$

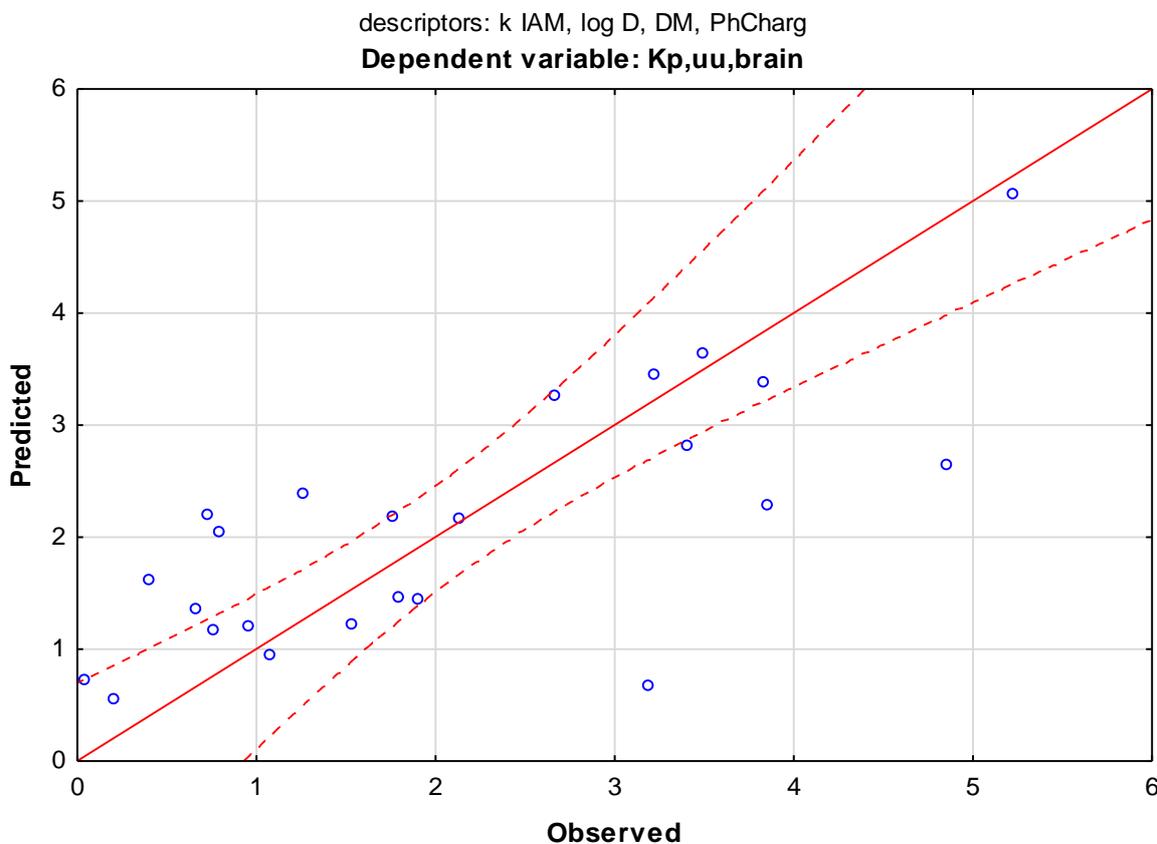


Figure 7. Scatter plot of the MLR model from Equation (5). Dependent variable: $K_{p,uu,brain}$; predictors: k IAM, log D, DM, and PhCharg. Coefficient of determination $R^2 = 0.56$.

3.4. Data Mining Models with Chromatographic Data

The MARSplines method was carried out to check whether the results of the MLR models could be improved by using a different regression algorithm. Descriptors used in model building were the same as those that occur in MLR models; dependent variables were again the B2, B1, and Kp,uu,brain indices.

The statistics of the B2 equations with the descriptors NP and RP were slightly lower than those obtained from the MLR; the coefficients of determination were 0.85 and 0.82, respectively (compared to 0.87 and 0.86 from multiple linear regression). The equation with the log k IAM descriptor had the same R² value as in the MLR model (0.89), but again, this chromatographic descriptor did not reach a significance level of *p* below 0.05 (Table 6).

Table 6. Comparison of the MLR and MARSplines models for the B2 index using the NP, RP, and log k IAM descriptors.

	NP	RP	log k IAM
MLR	R ² = 0.8743 R ² correct. = 0.8633 N = 151	R ² = 0.8574 R ² correct. = 0.8642 N = 164	R ² = 0.8912 R ² correct. = 0.8895 N = 124
MARSplines	R ² = 0.8533 R ² correct. = 0.8524 N = 169	R ² = 0.8218 R ² correct. = 0.8242 N = 169	R ² = 0.8944 R ² correct. = 0.8891 N = 116
Descriptors in the models	HA + D, NP, eL, PhCharg	HA + D, eL, RP, log P, PhCharg	HA + D, eL, log U/D, log k IAM, MW

For parameter B1, the regression result was also lower than that for the MLR (0.46 compared to 0.53) (Figure 8).

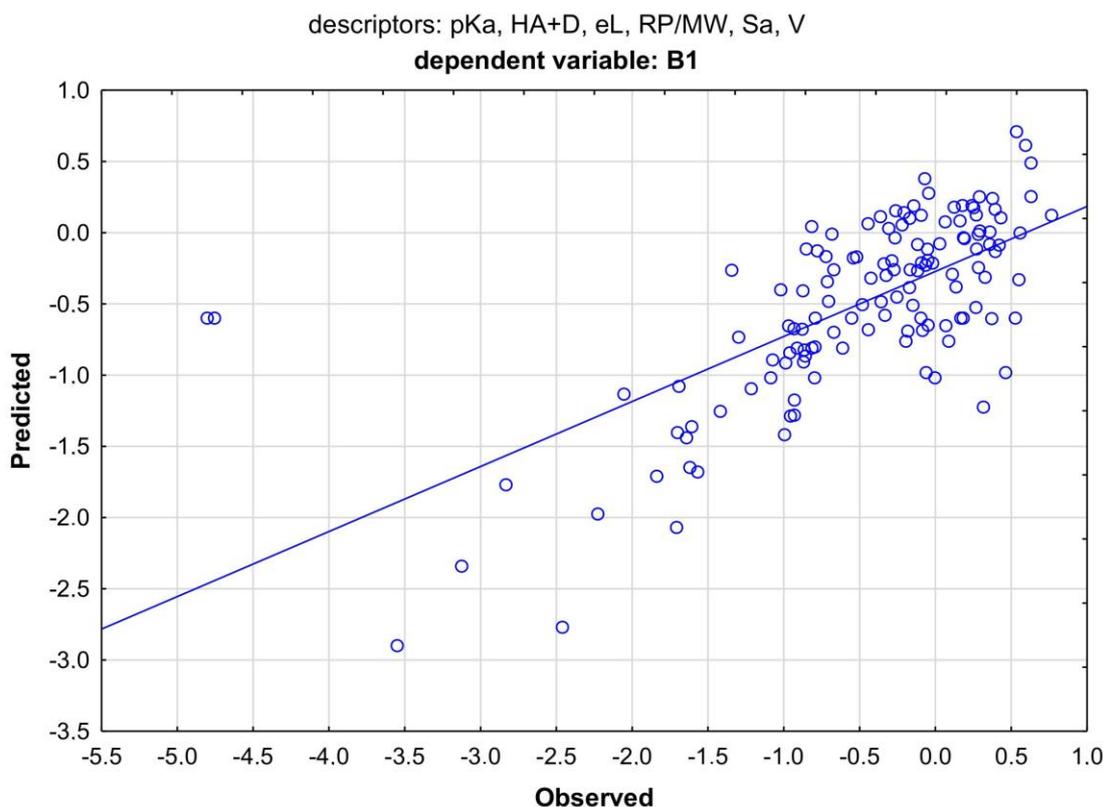


Figure 8. The MARSplines model for B1. Two basis functions were used in the model algorithm. Coefficient of determination R² = 0.46.

An attempt was made to perform a MARSplines analysis for the variable $K_{p,uu,brain}$. The results were, however, even lower than in the MLR model; the R^2 was 0.37 in the best descriptor configuration, the retention factor k_{IAM} was re-entered into the model, and was actually the only factor significantly affecting the modeling of the $K_{p,uu,brain}$ values in this small group of analyzed compounds ($N = 29$).

4. Discussion

Comparison of the median values (Mann–Whitney U test) of the computational indices B1 and B2 reflecting the log BB in the CNS+/- penetration groups showed that only the B2 values significantly differed between the '+' and '-' groups. Therefore, more attention was paid to the B2 indicator in further proceedings and regression modeling. Similar analyses concerning the distribution of values in the CNS+/- groups were performed for the chromatographic descriptors NP, RP, and log k_{IAM} and their derivatives, combined with their physicochemical properties. The descriptors alone showed significant differences in the medians in the CNS+/- groups. In the case of their derivatives, the physicochemical PSA parameter revealed a large influence because the chromatographic descriptors with its participation (RP/PSA; NP/PSA and log k_{IAM} /PSA) were characterized by the greatest difference between the medians (in the case of log k_{IAM} /PSA – means) in the CNS+ and CNS- groups, which can be seen in the boxplots (Figures 2 and 3)

Regression analyses carried out in the group of 181 APIs confirmed that physicochemical properties such as hydrogen binding capacity (HA, HD or HA + HD), the level of physiological charge (PhCharg), and the energy of the lowest unoccupied molecular orbital (eL) were the most important descriptors in regression models for the B2 index. They appeared most often in the resulting equations and usually entered them with a level of $p < 0.05$.

Statistical analysis of the chromatographic data (descriptors NP, RP, and log k_{IAM}) proved the usefulness of these retention coefficients to predict the log BB values. The TLC data entered the MLR models with good results, and the best statistics were achieved for the NP descriptor. The model with this participation explained 87% of the variance of the B2 indicator, and the descriptor itself significantly contributed to the regression modeling ($p = 0.012$). The RP descriptor turned out to be less successful: for B2, it was not possible to enter it into the model with $p < 0.05$ (Equations (A1) and (A2) in Appendix A). The situation improved for limited groups of compounds (i.e., the $B2 > -0.52$ index ($N = 89$), where RP descriptor managed to be included in the equation with a level of $p = 0.0409$).

NP derivatives combined with physicochemical parameters were also used in MLR models: the NP/MW parameter was applied in the model for the B2 and $B2 > -0.52$ indices, while NP/V appeared in the equation for $B2 > -0.9$. In all cases, these descriptors entered the equations with $p < 0.05$ (i.e., their contribution to the modeling of the regression models was significant). RP/MW and RP/V also appeared in the MLR model as the only derivatives of the TLC data. This shows the significant influence of molar mass (MW) and molecular volume (V) in combination with the retention data on log BB modeling. These physicochemical properties (MW and V) alone did not appear often in the equations.

The regression analysis for the B1 index was less successful; with the TLC data, only one model could be built that explained 53% of the B1 variance. The chromatographic descriptor that contributed to this model was RP/MW ($p = 0.026$).

Retention coefficients obtained from high-performance liquid chromatography with a column containing immobilized artificial membrane (log k_{IAM}), despite having quite promising primary correlations with the log BB indices ($R = 0.39$ for B1 and B2; Table A6) and had a smaller share in regression models (for B2 and B1), their p values did not usually exceed the assumed significance level of 0.05. However, the retention coefficient k_{IAM} was the only one to be introduced into the MLR equation for the dependent variable $K_{p,uu,brain}$. The model was built for a small group of APIs ($N = 23$), but k_{IAM} significantly contributed to the development of this model ($p = 0.0002$), along with the coefficient of distribution, log D ($p = 0.023$).

The regression method (MARSplines) in the field of data mining was applied to check whether better model statistics could be obtained by using a different regression algorithm. The results obtained using it were comparable or, unfortunately, lower than those obtained using the simplest multiple regression. The highest $R^2 = 0.85$ in the MARSplines model was again obtained for the NP descriptor; RP resulted in $R^2 = 0.82$; both parameters were statistically significant. Log k IAM did not reach the level of $p < 0.05$ as in the MLR model, although the equation was characterized by a very high coefficient of determination.

The B1 indicator also scored lower in the data mining algorithm; R^2 dropped from 0.53 (MLR) to 0.46 (MARSplines) with the RP/MW descriptor.

Attempts to model the $K_{p,uu,brain}$ indicator resulted in the creation of models with $R^2 = 0.56$ (MLR) and 0.37 (MARSplines), which were created using the retention coefficient k IAM, distribution coefficient: log D, molecule dipole moment: DM, and charge in pH 7.4: PhCharg. The chromatographic descriptor had a significant impact on the creation of both models, however, the small size of the API group with the collected $K_{p,uu,brain}$ data ($N = 29$) does not allow for further conclusions. The high correlation with the RP index also seems promising ($R^2 = 0.51$, Table A6). Further analyses of the $K_{p,uu,brain}$ indicator are planned, with more APIs participating in the study.

5. Conclusions

Retention data from the TLC experiments in combination with simple molecular descriptors (number or hydrogen acceptors and donors: HA and HD, energy of the lowest unoccupied molecular orbital: eL, physiological charge: PhCharg, molecular mass: MW etc.) allowed us to build several models of the prediction of log BB. The highest results were obtained with the NP TLC data (retardation coefficient R_f from the normal-phase TLC plate) This chromatographic descriptor is applicable in this type of statistical modeling and carries valuable information on the pharmacokinetic properties of a group of various APIs. The HPLC data seemed to show some influence on the $K_{p,uu,brain}$ index, however, these analyses should be repeated on a larger group of APIs to see if this correlation persists.

Derivatives of the NP descriptor, combined with the molar mass and volume of the molecule, NP/MW and NP/V, respectively, were also used in the log BB regression equations with good results. Polar surface area derivatives (NP/PSA and RP/PSA) showed the best distribution of values between the CNS+ and CNS− groups. Based on this information, it can be concluded that the derivatives of chromatographic data may increase the usefulness of retention values (R_f or log k) in the statistical analyses of BBB permeation.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/membranes13070623/s1>, Table S1a,b. Dataset of the 181 APIs. Table S2. The structures of 181 APIs. Table S3. Correlation matrix of molecular descriptors.

Author Contributions: Conceptualization, E.B. and K.W.; Methodology, E.B.; Software, E.B. and K.W.; Validation, E.B. and K.W.; Formal analysis, E.B. and K.W.; Investigation, K.W.; Resources, K.W.; Data curation, E.B. and K.W.; Writing—original draft preparation, K.W.; Writing—review and editing, E.B.; Visualization, K.W.; Supervision, E.B.; Project administration, E.B. and K.W.; Funding acquisition, E.B. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by an internal grant from the Medical University of Lodz number 503/3-016-03/503-31-001.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Table A1. Mann–Whitney U test; variable: CNS+ / –. The B2 results are significant with $p < 0.05$.

	Group CNS+ *	Group CNS– *	U	Z	p	Group CNS+ No. of Cases	Group CNS– No. of Cases
B2	1893	12,135	765	−7.311	0.000	120	47
B1	1968.5	6416.5	1302.5	−1.948	0.051	93	36

*—sum of ranks. $U = R1 - (n1(n1 + 1))/2$; R1—sum of rank from the first group; n1—number of cases from the first group; Z—Test statistic for the Mann–Whitney test: $Z = \frac{U - \frac{n1n2}{2}}{\sqrt{\frac{n1n2(n1+n2+1)}{12}}}$; n1, n2—number of cases from the first and second group, respectively.

Table A2. Mann–Whitney U test; variable: CNS+ / –. All presented results are significant with $p < 0.05$.

	Group CNS+ *	Group CNS– *	U	Z	p	Group CNS+ No. of Cases	Group CNS– No. of Cases
NP	10,804	5487	2158	−3.379	0.001	131	49
RP	10,774	5517	2128	−3.476	0.001	131	49
NP/PB	10,148	5078	2147	−2.953	0.003	126	48
RP/PB	10,343	4882	2342	−2.295	0.022	126	48
NP/PSA	2253	11,608	1125	−5.989	0.000	119	47
RP/MW	2952	10,251	2006	−2.094	0.036	119	43
RP/PSA	1941	11,920	813	−7.108	0.000	119	47
NP/log P	3066	10,464	1985	−2.667	0.008	118	46
RP/log P	2962	10,568	1881	−3.047	0.002	118	46

*—sum of ranks.

Table A3. Mann–Whitney U test; variable: CNS+ / –. All presented results are significant with $p < 0.05$.

	Group CNS+ *	Group CNS– *	U	Z	p	Group CNS+ No. of Cases	Group CNS– No. of Cases
k IAM	1414	6089	918	−2.893	0.004	91	31
log k IAM	1383	5998	887	−3.013	0.003	91	31
log k IAM/PB	1159	4946	753	−2.707	0.007	82	28
log k IAM/log P	627	3651	437	−2.469	0.014	73	19

*—sum of ranks.

Table A4. T-tests; grouping variable: CNS+ / –. All presented results are significant with $p < 0.05$.

	Mean ± SD CNS–	Mean ± SD CNS+	t	df	p	CNS– No. of Cases	CNS+ No. of Cases
log k IAM/PSA	−1.982 (±0.975)	−1.082 (±1.016)	−4.30	119	0.000	31	90

Table A5. p-Values for multiple comparisons; variable: NP. Independent (grouping) variable: B2 code. Kruskal–Wallis test: $H(2, N = 180) = 11.948; p = 0.0025$.

	B2 Code 1; R:112.13	B2 Code 2; R:81.300	B2 Code 3; R:82.445
B2 code 1	—	0.031	0.003
B2 code 2	0.031	—	1.000
B2 code 3	0.003	1.000	—

Table A6. Correlation matrix of the BBB permeation indices and chromatographic descriptors. Underlined correlations are statistically significant.

	Mean	SD	CNS+/-	$K_{p,uu,brain}$	B1	B2	B2 > -0.9	B2 > -0.52	NP	RP	Log k IAM
CNS+/-	0.719	0.451	1.000	0.178	<u>0.194</u>	<u>0.570</u>	<u>0.550</u>	<u>0.521</u>	-0.073	-0.064	<u>0.263</u>
$K_{p,uu,brain}$	1.880	1.468		1.000	0.301	0.274	0.180	-0.044	-0.189	<u>-0.519</u>	<u>0.474</u>
B1	-0.490	0.947			1.000	<u>0.519</u>	<u>0.378</u>	<u>0.335</u>	-0.147	<u>-0.370</u>	<u>0.397</u>
B2	-0.605	0.692				1.000	<u>0.780</u>	<u>0.750</u>	<u>-0.257</u>	<u>-0.294</u>	<u>0.390</u>
B2 > -0.9	0.729	0.446					1.000	<u>0.692</u>	<u>-0.213</u>	<u>-0.211</u>	<u>0.326</u>
B2 > -0.52	0.564	0.497						1.000	-0.130	<u>-0.187</u>	<u>0.277</u>
NP	0.662	0.265							1.000	<u>0.539</u>	-0.124
RP	0.723	0.194								1.000	<u>-0.340</u>
Log k IAM	0.450	0.923									1.000

Regression equations for RP and RP derivatives *:

$$B2 = 0.620 - \underline{0.165 HA + D} + \underline{0.042 PhCharg} + \underline{0.006 eL} - 0.202 RP + 0.012 \log P \quad (A1)$$

$$R = 0.9213; R^2 = 0.8596; R^2 \text{ adj.} = 0.8643; F(5, 158) = 200.15; p > 0.000; s = 0.15; N = 164$$

$$SDEP = 0.23; Q^2_{LOO} = 0.8234$$

$$B2 = 0.620 - \underline{0.165 HA + D} + \underline{0.042 PhCharg} + \underline{0.006 eL} - 0.202 RP + 0.012 \log P \quad (A2)$$

$$R = 0.8447; R^2 = 0.7124; R^2 \text{ adj.} = 0.6967; F(6, 104) = 42.27; p > 0.000; s = 0.16; N = 111$$

$$SDEP = 0.18; Q^2_{LOO} = 0.6544$$

$$B2 > -0.9 = 0.640 - \underline{0.133 HA} - \underline{0.621 RP/V} + 0.077 PhCharg - 0.059 Sa \quad (A3)$$

$$R = 0.7837; R^2 = 0.6029; R^2 \text{ adj.} = 0.5945; F(4, 106) = 40.30; p > 0.000; s = 0.17; N = 111$$

$$SDEP = 0.18; Q^2_{LOO} = 0.5689$$

$$B2 > -0.52 = 0.014 - \underline{0.095 HA} - \underline{0.327 RP} + \underline{0.079 PhCharg} + \underline{0.203 PB} - 0.045 V - 0.036 eH \quad (A4)$$

$$R = 0.7398; R^2 = 0.5435; R^2 \text{ adj.} = 0.5167; F(6, 82) = 16.18; p > 0.000; s = 0.18; N = 89$$

$$SDEP = 0.20; Q^2_{LOO} = 0.5042$$

$$B2 > -0.52 = 0.654 - \underline{0.115 HA} - \underline{1.331 RP/MW} - \underline{0.107 V} - 0.026 eH-eL \quad (A5)$$

$$R = 0.7278; R^2 = 0.5447; R^2 \text{ adj.} = 0.5058; F(4, 84) = 24.24; p > 0.000; s = 0.17; N = 89$$

$$SDEP = 0.18; Q^2_{LOO} = 0.5088$$

* underlined descriptors are statistically significant; $p < 0.05$.

Regression equations for NP and NP derivatives *:

$$B2 = 0.652 - \underline{0.167 HA + D} - \underline{0.215 NP} + \underline{0.007 eL} + 0.037 \text{ PhCharg} \quad (\text{A6})$$

$$R = 0.9347; R^2 = 0.8736; R^2 \text{ adj.} = 0.8648; F(5, 158) = 240.15; p > 0.000; s = 0.19; N = 151$$

$$\text{SDEP} = 0.24; Q^2_{\text{LOO}} = 0.8439$$

$$B2 = 0.711 - \underline{0.173 HA + D} + 0.037 \text{ PhCharg} + \underline{0.007 eL} - \underline{0.395 NP/MW} - 0.130 \text{ PB} + 0.014 \log P \quad (\text{A7})$$

$$R = 0.9278; R^2 = 0.8715; R^2 \text{ adj.} = 0.8617; F(6, 144) = 160.07; p > 0.000; s = 0.21; N = 151$$

$$\text{SDEP} = 0.24; Q^2_{\text{LOO}} = 0.8439$$

$$B2 > -0.9 = 0.271 - \underline{0.114 HA + D} + \underline{0.125 PhCharg} - \underline{0.025 pKa} - \underline{0.046 eH-eL} - \underline{0.040 Sa} - \underline{0.178 NP} + \text{PB} \quad (\text{A8})$$

$$R = 0.8628; R^2 = 0.7447; R^2 \text{ adj.} = 0.7188; F(7, 103) = 41.43; p > 0.000; s = 0.22; N = 111$$

$$\text{SDEP} = 0.23; Q^2_{\text{LOO}} = 0.7298$$

$$B2 > -0.9 = 0.940 - \underline{0.145 HA} - \underline{0.690 NP/V} - \underline{0.069 Sa} - \underline{0.029 pKa} + \underline{0.111 PhCharg} \quad (\text{A9})$$

$$R = 0.8123; R^2 = 0.6535; R^2 \text{ adj.} = 0.6289; F(5, 105) = 38.703; p > 0.000; s = 0.19; N = 111$$

$$\text{SDEP} = 0.22; Q^2_{\text{LOO}} = 0.6095$$

$$B2 > -0.52 = 0.279 - \underline{0.121 HA} - \underline{0.690 NP/MW} - \underline{0.029 pKa} + \underline{0.123 PhCharg} + \underline{0.196 PB} - \underline{0.048 Sa} - \underline{0.040 eH-eL} \quad (\text{A10})$$

$$R = 0.7788; R^2 = 0.6124; R^2 \text{ adj.} = 0.5836; F(7, 81) = 18.25; p > 0.000; s = 0.23; N = 89$$

$$\text{SDEP} = 0.26; Q^2_{\text{LOO}} = 0.5890$$

* underlined descriptors are statistically significant; $p < 0.05$.

References

1. Pangalos, M.N.; Schechter, L.E.; Hurko, O. Drug development for CNS disorders: Strategies for balancing risk and reducing attrition. *Nat. Rev. Drug Discov.* **2007**, *6*, 521–532. [[CrossRef](#)] [[PubMed](#)]
2. Sahi, J.; Li, Y. The role of drug transporters at the blood brain barrier. *Eur. Pharm. Rev.* **2016**, *21*, 15–19.
3. Xu, D.; Huang, S.; Wang, H.; Xie, W. Regulation of brain drug metabolizing enzymes and transporters by nuclear receptors. *Drug Metab. Rev.* **2018**, *50*, 407–414. [[CrossRef](#)] [[PubMed](#)]
4. Vázquez, M.; Fagiolino, P. The role of efflux transporters and metabolizing enzymes in brain and peripheral organs to explain drug-resistant epilepsy. *Epilepsia Open* **2022**, *7*, S47–S58. [[CrossRef](#)] [[PubMed](#)]
5. Lanevskij, K.; Japertas, P.; Didziapetris, R.; Petrauskas, A. Drug discovery interface Ionization-specific prediction of blood-brain permeability. *J. Pharm. Sci.* **2009**, *98*, 122–134. [[CrossRef](#)]
6. Abbott, N.J. Prediction of blood-brain barrier permeation in drug discovery from in vivo, in vitro and in silico models. *Drug Discov. Today Technol.* **2004**, *1*, 407–416. [[CrossRef](#)]
7. Goodwin, J.T.; Clark, D.E. In silico predictions of blood-brain barrier penetration: Considerations to “keep in mind”. *J. Pharmacol. Exp. Ther.* **2005**, *315*, 477–483. [[CrossRef](#)]

8. Faramarzi, S.; Kim, M.T.; Volpe, D.A.; Cross, K.P.; Chakravarti, S.; Stavitskaya, L. Development of QSAR models to predict blood-brain barrier permeability. *Front. Pharmacol.* **2022**, *13*, 1040838. [[CrossRef](#)]
9. Loryan, I.; Reichel, A.; Feng, B.; Bundgaard, C.; Shaffer, C.; Kalvass, C.; Bednarczyk, D.; Morrison, D.; Lesuisse, D.; Hoppe, E.; et al. Unbound Brain-to-Plasma Partition Coefficient, $K_{p,uu,brain-a}$ Game Changing Parameter for CNS Drug Discovery and Development. *Pharm. Res.* **2022**, *39*, 1321–1341. [[CrossRef](#)]
10. Loryan, I.; Sinha, V.; Mackie, C.; Van Peer, A.; Drinkenburg, W.H.; Vermeulen, A.; Heald, D.; Hammarlund-Udenaes, M.; Wassvik, C.M. Molecular properties determining unbound intracellular and extracellular brain exposure of CNS drug candidates. *Mol. Pharm.* **2015**, *12*, 520–532. [[CrossRef](#)]
11. Vallianatou, T.; Tsopelas, F.; Tsantili-Kakoulidou, A. Prediction Models for Brain Distribution of Drugs Based on Biomimetic Chromatographic Data. *Molecules* **2022**, *27*, 3668. [[CrossRef](#)]
12. Roy, D.; Hinge, V.K.; Kovalenko, A. Predicting Blood-Brain Partitioning of Small Molecules Using a Novel Minimalistic Descriptor-Based Approach via the 3D-RISM-KH Molecular Solvation Theory. *ACS Omega* **2019**, *4*, 3055–3060. [[CrossRef](#)]
13. Jiang, L.; Li, S.; Zheng, J.; Li, Y.; Huang, H. Recent progress in microfluidic models of the blood-brain barrier. *Micromachines* **2019**, *10*, 375. [[CrossRef](#)]
14. Yuan, Y.; Zheng, F.; Zhan, C.G. Improved Prediction of Blood-Brain Barrier Permeability Through Machine Learning with Combined Use of Molecular Property-Based Descriptors and Fingerprints. *AAPS J.* **2018**, *20*, 54. [[CrossRef](#)]
15. Lippmann, E.S.; Al-Ahmad, A.; Palecek, S.P.; Shusta, E.V. Modeling the blood-brain barrier using stem cell sources. *Fluids Barriers CNS* **2013**, *10*, 2. [[CrossRef](#)]
16. Zhang, D.; Xiao, J.; Zhou, N.; Zheng, M.; Luo, X.; Jiang, H.; Chen, K. A Genetic Algorithm Based Support Vector Machine Model for Blood-Brain Barrier Penetration Prediction. *Biomed Res. Int.* **2015**, *2015*, 292683. [[CrossRef](#)]
17. Bagchi, S.; Chhibber, T.; Lahooti, B.; Verma, A.; Borse, V.; Jayant, R.D. In-vitro blood-brain barrier models for drug screening and permeation studies: An overview. *Drug Des. Dev. Ther.* **2019**, *13*, 3591–3605. [[CrossRef](#)]
18. Kumar, R.; Sharma, A.; Alexiou, A.; Bilgrami, A.L.; Kamal, M.A.; Ashraf, G.M. DeePred-BBB: A Blood Brain Barrier Permeability Prediction Model With Improved Accuracy. *Front. Neurosci.* **2022**, *16*, 858126. [[CrossRef](#)]
19. De Lange, E.C.M.; vd Berg, D.J.; Bellanti, F.; Voskuyl, R.A.; Syvänen, S. P-glycoprotein protein expression versus functionality at the blood-brain barrier using immunohistochemistry, microdialysis and mathematical modeling. *Eur. J. Pharm. Sci.* **2018**, *124*, 61–70. [[CrossRef](#)]
20. Vigh, J.P.; Kincses, A.; Ozgür, B.; Walter, F.R.; Santa-Maria, A.R.; Valkai, S.; Vastag, M.; Neuhaus, W.; Brodin, B.; Dér, A.; et al. Transendothelial electrical resistance measurement across the blood-brain barrier: A critical review of methods. *Micromachines* **2021**, *12*, 685. [[CrossRef](#)]
21. Wevers, N.R.; Kasi, D.G.; Gray, T.; Wilschut, K.J.; Smith, B.; Vught, R.; Shimizu, F.; Sano, Y.; Kanda, T.; Marsh, G.; et al. A perfused human blood-brain barrier on-a-chip for high-throughput assessment of barrier function and antibody transport. *Fluids Barriers CNS* **2018**, *15*, 23. [[CrossRef](#)] [[PubMed](#)]
22. Sugano, K.; Hamada, H.; Machida, M.; Ushio, H. High Throughput Prediction of Oral Absorption: Improvement of the Composition of the Lipid Solution Used in Parallel Artificial Membrane Permeation Assay. *SLAS Discov.* **2001**, *6*, 189–196. [[CrossRef](#)] [[PubMed](#)]
23. Yoon, C.H.; Shin, B.S.; Chang, H.S.; Kwon, L.S.; Kim, H.Y.; Yoo, S.E.; Yoo, S.D. Rapid screening of drug absorption potential using the immobilized artificial membrane phosphatidylcholine column and molar volume. *Chromatographia* **2004**, *60*, 399–404. [[CrossRef](#)]
24. Sobańska, A.W.; Brzezińska, E. Application of RP-18 thin-layer chromatography and quantitative structure-activity relationship analysis for the prediction of the blood-brain barrier permeation. *J. Planar Chromatogr. Mod. TLC* **2016**, *29*, 287–298. [[CrossRef](#)]
25. Weronika Sobanska, A.W.; Robertson, J.; Brzezińska, E. RP-18 TLC chromatographic and computational study of skin permeability of steroids. *Pharmaceuticals* **2021**, *14*, 600. [[CrossRef](#)]
26. Pyzowski, J.; Brzezińska, E.; Sobańska, A.W. RP-18 chromatographic-based study of the blood-brain barrier permeability of selected sunscreens and preservatives. *J. Planar Chromatogr. Mod. TLC* **2017**, *30*, 275–284. [[CrossRef](#)]
27. Fan, Y.; Unwalla, R.; Denny, R.A.; Di, L.; Kerns, E.H.; Diller, D.J.; Humblet, C. Insights for predicting blood-brain barrier penetration of CNS targeted molecules using QSPR approaches. *J. Chem. Inf. Model.* **2010**, *50*, 1123–1133. [[CrossRef](#)]
28. Clark, D.E. Rapid calculation of polar molecular surface area and its application to the prediction of transport phenomena. 2. Prediction of blood-brain barrier penetration. *J. Pharm. Sci.* **1999**, *88*, 815–821. [[CrossRef](#)]
29. SwissADME. Available online: <http://www.swissadme.ch/> (accessed on 14 March 2023).
30. Summerfield, S.G.; Zhang, Y.; Liu, H. Examining the uptake of central nervous system drugs and candidates across the blood-brain barrier. *J. Pharmacol. Exp. Ther.* **2016**, *358*, 294–305. [[CrossRef](#)]
31. Vilar, S.; Chakrabarti, M.; Costanzi, S. Prediction of passive blood-brain partitioning: Straightforward and effective classification models based on in silico derived physicochemical descriptors. *J. Mol. Graph. Model.* **2010**, *28*, 899–903. [[CrossRef](#)]
32. Fu, X.C.; Wang, G.P.; Shan, H.L.; Liang, W.Q.; Gao, J.Q. Predicting blood-brain barrier penetration from molecular weight and number of polar atoms. *Eur. J. Pharm. Biopharm.* **2008**, *70*, 462–466. [[CrossRef](#)]
33. Cruciani, G.; Pastor, M.; Guba, W. VolSurf: A new tool for the pharmacokinetic optimization of lead compounds. *Eur. J. Pharm. Sci.* **2000**, *11*, S29–S39. [[CrossRef](#)]

34. Zhao, Y.H.; Abraham, M.H.; Ibrahim, A.; Fish, P.V.; Cole, S.; Lewis, M.L.; De Groot, M.J.; Reynolds, D.P. Predicting penetration across the blood-brain barrier from simple descriptors and fragmentation schemes. *J. Chem. Inf. Model.* **2007**, *47*, 170–175. [[CrossRef](#)]
35. Wanat, K.; Żydek, G.; Hekner, A.; Brzezińska, E. In silico plasma protein binding studies of selected group of drugs using TLC and HPLC retention data. *Pharmaceuticals* **2021**, *14*, 202. [[CrossRef](#)]
36. Drugbank. Available online: <https://www.drugbank.ca/drugs/DB01174> (accessed on 23 March 2023).

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.