


Article

# Binding Constants of Substituted Benzoic Acids with Bovine Serum Albumin

Diliara Khaibrakhmanova, Alena Nikiforova and Igor Sedov \* 

Chemical Institute, Kremlevskaya 18, Kazan Federal University, 420008 Kazan, Russia; dilusechka@gmail.com (D.K.); alnikiforova22@gmail.com (A.N.)

\* Correspondence: igor\_sedov@inbox.ru; Tel.: +7-9600503916

Received: 4 February 2020; Accepted: 19 February 2020; Published: 20 February 2020



**Abstract:** Experimental data on the affinity of various substances to albumin are essential for the development of empirical models to predict plasma binding of drug candidates. Binding of 24 substituted benzoic acid anions to bovine serum albumin was studied using spectrofluorimetric titration. The equilibrium constants of binding at 298 K were determined according to 1:1 complex formation model. The relationships between the ligand structure and albumin affinity are analyzed. The binding constant values for m- and p-monosubstituted acids show a good correlation with the Hammett constants of substituents. Two- and three-parameter quantitative structure–activity relationship (QSAR) models with theoretical molecular descriptors are able to satisfactorily describe the obtained values for the whole set of acids. It is shown that the electron-density distribution in the aromatic ring exerts crucial influence on the albumin affinity.

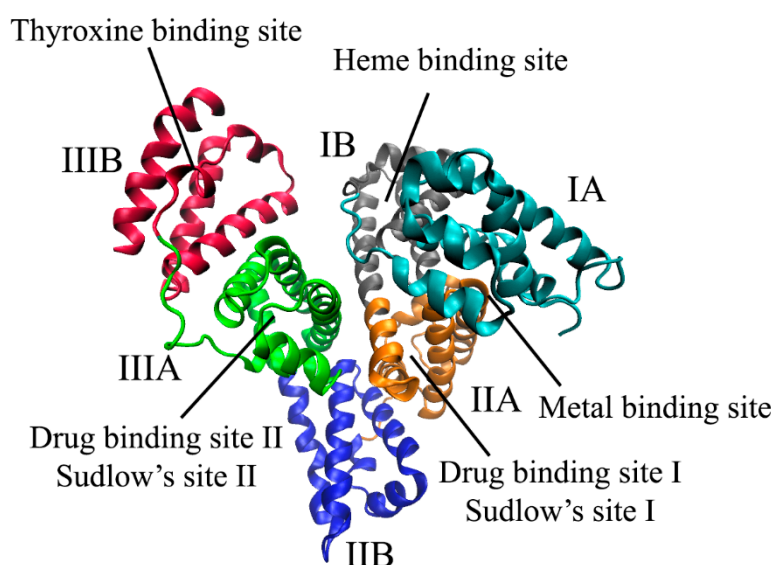
**Keywords:** albumin; binding constant; benzoic acids; spectrofluorimetry; Hammett constants; QSAR

## 1. Introduction

Experimental *in vivo* study of plasma protein binding is an important step in development of novel drug substances. The strength of binding affects the distribution of a drug in the body along with its pharmacodynamic parameters, biological activity and toxicity [1]. The most abundant protein in blood plasma is serum albumin. It can be extracted from plasma and used to study binding processes *in vitro*. In such experiments, the strength of binding is characterized quantitatively using thermodynamic binding constants. The measured values of these constants can be used to estimate the fractions of plasma-bound and free drug in the blood, which helps to reduce the number of costly *in vivo* experiments [2].

However, accurate determination of the binding constant values is complicated by many factors. In particular, the albumin molecule contains several binding sites with different affinity to different species [3] (Figure 1). Furthermore, albumin partially dimerizes in solution, and each binding site of the dimeric form may change its affinity to the ligands [4]. For bovine serum albumin, the dimeric form is predominant at concentrations above 5  $\mu\text{M}$  ( $K = 10 \pm 2 \mu\text{M}$  [5]).

Moreover, none of the commonly used experimental techniques allows accurate determination of separate binding constants for each site. The experimental data are always analyzed using some simplified binding models, and the obtained binding constant values are in fact some complicated functions of the site-specific constants. Thus, the results are heavily dependent on the method of measurement and data processing. The existing literature data for the same ligand from different authors often contradict each other and are unreliable (see below for the examples).



**Figure 1.** Positions of binding sites in a bovine serum albumin molecule.

Even though the method-dependence of results is inevitable, the affinities of different compounds to albumin can be still compared if the experiments are conducted at similar conditions. In other words, a set of compounds should be studied using the same technique with albumin from the same source taken at the same concentration, and the data should be analyzed according to the same binding model. Only such uniformly obtained data can be used to develop quantitative structure–property relationships or other models for a priori estimation of affinity in virtual screening [1]. However, most of the extremely numerous works providing binding constant values are limited to a single ligand or a small number of ligands.

Serum albumin is rather nonselective and can bind ligands with different molecular size, shape and functional groups. In the present work, we limit our study with a series of substituted benzoic acids. Variation of positions and types of substituents in the aromatic ring is an easy way to create a set of compounds with different electronic structure and hydrogen-bonding properties, which was widely used in the studies of organic reactivity [6]. In addition, many derivatives of benzoic acid are biologically and pharmacologically active. Analysis of the influence of the nature and position of aromatic ring substituents on the toxicities and activities of these compounds led to the development of the first quantitative structure–activity relationships (QSARs) [7]. Thus, the substituted benzoic acids are interesting and convenient objects for accumulation of experimental albumin binding data and studying the links between the structure and affinity of ligands.

Several previous works were devoted to binding of substituted benzoic acids to albumin. Matsushita et al. [8] determined the binding constants of 14 *p*-substituted benzoic acids to human and bovine serum albumin using ultracentrifugation method. The results were analyzed using multiple linear regressions with the ligand properties such as molecular volume and dissociation constant  $pK_a$ . It was concluded that large ligand size and strong acidity favor binding. Moriguchi studied binding of a series of 14 *o*-, *m*-, and *p*-substituted benzoic acids to bovine serum albumin using spectrophotometry [9]. The logarithms of the obtained binding constants values for *m*- and *p*-substituted acids correlated with  $pK_a$  values of the acids, showing stronger binding for stronger acids. However, *o*-substituted benzoic acids did not follow this correlation, which was attributed to the ortho-effect. Authors suggested that these results prove that binding is due to the interactions between carboxylic group and cationic groups of amino acid residues in albumin. Two recent works made by Hao Zhang et al. are devoted to spectrofluorimetric determination of binding constants of 111 substituted benzoic acids to human serum albumin [10] and 114 substituted benzoic acids to bovine serum albumin [11] and development of the structure–affinity relationships. Despite such a large

number of acids, only combinations of three different substituents ( $-\text{OH}$ ,  $-\text{OCH}_3$ , and  $-\text{CH}_3$ ) in aromatic ring were considered. Correlations of the obtained values with any single ligand parameter including  $\text{pK}_a$  were quite poor. Some observations on the influence of substituents were made, in particular, it was concluded that *o*-hydroxylation increases the affinity, while *p*-hydroxylation decreases it.

In the present work, binding of 24 substituted (mostly monosubstituted) benzoic acids to bovine serum albumin at 298 K is studied. Correlations of the obtained binding constants with various ligand properties are considered, and the possibility of their a priori estimation using QSAR approach is analyzed.

## 2. Methodology

### 2.1. Binding Centers of Albumin

Albumins from the plasma of different mammals are globular proteins consisting of three homologous domains (I, II, III). Each of them contains two subdomains, A and B [12]. Human serum albumin (HSA) and bovine serum albumin (BSA) (Figure 2) attract the most attention from researchers, however, the processes of binding to albumin from the blood plasma of other animals, such as rabbit, dog and horse, have also been studied [13–15]. All these albumins have minor differences in structure.



**Figure 2.** Comparison of the spatial structures of bovine (red) and human (yellow) serum albumin.

Most of the drug molecules are thought to bind with two high affinity binding sites in albumin subdomains IIA and IIIA [16], called Sudlow I and Sudlow II, respectively. According to the numerous publications, Sudlow I is preferred by bulky heterocyclic molecules and Sudlow II is preferred by aromatic carboxylates [17]. Since the primary amino acid sequence of bovine serum albumin is 75.6% identical to that of human serum albumin [18], their binding centers are also similar. The amino acid residues in Sudlow II are the same, however, for Sudlow I only 57% identity was shown [18]. Bovine serum albumin has arginine residues in 195 and 199 positions instead of lysine in Sudlow I site. This change may cause some difference in binding properties.

## 2.2. Spectrofluorimetric Method and Its Pitfalls

Spectrofluorimetry is a common method to study protein–ligand interactions. Tryptophan, tyrosine and phenylalanine can contribute to the intrinsic fluorescence of proteins. When protein–ligand interactions occur, fluorescence quenching is usually observed due to the changes in the microenvironment of amino acid residues. This phenomenon can be used to determine the binding constant values. In such studies, the excitation wavelength 295 nm is often used. At this wavelength, fluorescence is predominantly caused by tryptophan residues. Bovine serum albumin has two tryptophans: Trp-212 located in Sudlow I binding site and Trp-134 in subdomain IA while human serum albumin has only one Trp-214 in Sudlow I [19]. However, binding to each of the sites can quench the fluorescence. The quenching is monitored by following the fluorescence intensity at the wavelengths near the emission maximum (340–355 nm).

In fluorescence experiments, the concentration of ligand is most commonly varied, and the signal change is analyzed. There is no way to obtain the data on binding with separate sites of albumin from such experiments. A very simple model of 1:1 binding is usually used. Thus, we should avoid a large excess of ligand, especially in the case of strong binding. The most common way to calculate the binding constant is using the Stern–Volmer equation relating the fluorescence of protein in the absence and in the presence of ligand ( $F_0$  and  $F$ , respectively) to the total concentration of added ligand  $[L]_{\Sigma}$  and the binding constant  $K$ :

$$\frac{F_0}{F} = K[L]_{\Sigma} + 1, \quad (1)$$

Derivation of this equation relies on several assumptions. As mentioned above, the first one is 1:1 complex stoichiometry. Second, there should be only one type of fluorophore in protein. This is not true for bovine serum albumin, but always neglected. Third, bound ligand should quench fluorescence intensity down to zero. This is not true for all ligands. Fourth, the equilibrium concentration of the bound ligand should be negligible. This is a problem only in the case of strong binding and can be overcome by correcting the obtained values, e.g., using numerical simulations of Stern–Volmer plots for various large constant values.

If all these assumptions are correct, then for the binding process



with the binding constant

$$K = \frac{[\text{AlbL}]}{[\text{Alb}][\text{L}]}, \quad (2)$$

the total albumin concentration in solution is given by

$$[\text{Alb}]_{\Sigma} = [\text{Alb}] + K[\text{Alb}][\text{L}]_0, \quad (3)$$

thus

$$\frac{F_0}{F} = \frac{[\text{Alb}]_{\Sigma}}{[\text{Alb}]} = K[L]_{\Sigma} + 1. \quad (4)$$

The use of the methods based on fluorescence quenching and particularly the Stern–Volmer equation was an object of severe criticism [20,21]. In addition to the above-mentioned pitfalls, a common mistake leading to unreliable constant values is the neglect of the inner filter effect caused by significant absorbance of solution at the excitation or emission wavelength. Moreover, fluorescence can be quenched by a dynamic collisional quenching mechanism not related to the formation of the protein–ligand complex. The shift of the fluorescence peak upon binding can also lead to incorrect results. In the present work, we avoid most of the known pitfalls. We also believe that the rest of the possible systematic errors and the limitations of the simplified binding model will affect the whole

series of benzoic acids in the same manner so we at least can compare the relative strength of their binding with albumin.

### 3. Experimental

#### 3.1. Chemicals and Instruments

All the chemicals were commercial products and were used without further purification. Fluorescence measurements were performed using Cary Eclipse fluorescence spectrophotometer (Agilent Technologies, Santa Clara, CA, USA).

#### 3.2. Spectrofluorimetric Titration Procedure

In a typical experiment, a fresh solution of bovine serum albumin (PanEko, Moscow, Russia, defatted, purity > 98%) with a concentration of about 1  $\mu\text{M}$  in phosphate buffer (pH 7.4, 0.05 M) was prepared. Ligands (acids) were dissolved in the same buffer and then diluted to concentrations in the range of 20–50  $\mu\text{M}$ . According to the  $pK_a$  values of the studied acids, all of them exist almost entirely in ionized form at such pH value. All solutions were thoroughly degassed by stirring under vacuum. Next, 2.3 mL of protein solution was poured into a quartz spectrofluorometric cuvette and allowed to settle for 40 min to saturate the walls of the cell with adsorbed protein molecules. Then, the protein solution was replaced with a fresh one and titrated with 10  $\mu\text{L}$  portions of ligand solution using an automated syringe pump equipped with a 250  $\mu\text{L}$  Hamilton microsyringe with a long steel cannula. The endpoint of titration usually corresponded to  $[\text{L}]_{\Sigma} : [\text{Alb}]_{\Sigma} \approx 2:1$ . The cuvette holder had a built-in Peltier thermostat keeping temperature equal to 298 K and a magnetic stirrer. After addition of each portion of ligand solution, the fluorescence spectrum with the excitation wavelength  $\lambda_{\text{ex}} = 295$  nm and emission wavelengths  $\lambda_{\text{em}} = 310\text{--}360$  nm was recorded. Additional spectrofluorimetric measurements of absorbance of solutions before and after titration at  $\lambda_{\text{ex}}$  and  $\lambda_{\text{em}}$  have shown that the influence of the inner filter effect is negligible. The fluorescence intensities at the emission maximum, 347 nm, were taken for further calculations. The studied acids do not have a significant fluorescence signal at this wavelength (p-aminobenzoic acid was not included in the study because it does).

#### 3.3. Evaluation of the Binding Constants

In order to use Stern–Volmer equation (Equation (4)), the fluorescence intensity should be measured for several solutions with the same concentration of protein but at different concentrations of ligand [21]. However, during titration the protein solution is being diluted at each step. Thus, the ratio of the fluorescence intensity to the current total concentration of albumin should be used instead. The equation used to calculate the binding constant has the following form:

$$\frac{\varepsilon_0}{\varepsilon} = K[\text{L}]_{\Sigma} + 1, \quad (5)$$

where  $\varepsilon_0 = \frac{F_0}{[\text{Alb}]_{\Sigma 0}}$ ,  $\varepsilon = \frac{F}{[\text{Alb}]_{\Sigma}}$ ,  $[\text{Alb}]_{\Sigma 0}$  is the initial total albumin concentration and  $[\text{Alb}]_{\Sigma}$  is the total albumin concentration after dilution with ligand solution. In a control titration with buffer solution containing no ligand, the values of  $\varepsilon$  remained equal to  $\varepsilon_0$ .

The plots of  $\frac{\varepsilon_0}{\varepsilon}$  against  $[\text{L}]_{\Sigma}$  allow to determine the value of the binding constant, which is equal to the slope of the linear correlation. These values were averaged over 4–5 titrations with freshly prepared albumin and ligand solutions.

The constants of binding of the studied compounds to BSA appear to be not very large, so that no more than 10% of ligand is bound in solution. Nevertheless, we have written a program for numerical simulations of Stern–Volmer plots for titration with a given binding constant and concentrations of BSA and ligand solutions. (see the Supplementary Materials for the source code and example of simulation of 4-nitrobenzoic acid titration in Wolfram Mathematica nb format). It helps to find the value of  $K$  providing the best fit of the simulated curve to the experimental data. The difference

between the apparent and corrected values of  $K$  was not great: for the largest apparent constant  $7.6 \times 10^4$  (4-nitrobenzoic acid) the corrected value equals  $8.1 \times 10^4$ . Table 1 lists the final corrected values for all the studied systems. The standard state of the concentrations in the expression (Equation (2)) for the binding constants is 1 M.

**Table 1.** Measured values of the binding constants of substituted benzoic acids with bovine serum albumin (BSA) at 298 K.

Ligand	$K \cdot 10^{-4}$
benzoic acid	$1.9 \pm 0.1$
2-hydroxybenzoic acid	$4.0 \pm 0.2$
3-hydroxybenzoic acid	$1.9 \pm 0.1$
4-hydroxybenzoic acid	$1.0 \pm 0.1$
2,4-dihydroxybenzoic acid	$3.1 \pm 0.2$
2-nitrobenzoic acid	$2.1 \pm 0.2$
3-nitrobenzoic acid	$7.8 \pm 0.3$
4-nitrobenzoic acid	$8.1 \pm 0.2$
2,4-dinitrobenzoic acid	$6.5 \pm 0.2$
2-aminobenzoic acid	$2.3 \pm 0.2$
3-aminobenzoic acid	$1.4 \pm 0.2$
2-fluorobenzoic acid	$3.3 \pm 0.2$
3-chlorobenzoic acid	$5.8 \pm 0.3$
4-chlorobenzoic acid	$5.1 \pm 0.1$
2-bromobenzoic acid	$2.3 \pm 0.2$
3-bromobenzoic acid	$3.5 \pm 0.3$
4-bromobenzoic acid	$5.4 \pm 0.2$
2-iodobenzoic acid	$3.3 \pm 0.3$
3-iodobenzoic acid	$5.1 \pm 0.2$
4-iodobenzoic acid	$4.1 \pm 0.3$
2-cyanobenzoic acid	$2.6 \pm 0.1$
3-cyanobenzoic acid	$3.9 \pm 0.2$
4-cyanobenzoic acid	$6.4 \pm 0.2$
4-methylbenzoic acid	$1.4 \pm 0.1$

## 4. Discussion

### 4.1. Literature Albumin Binding Data for the Studied Compounds

For some of the studied benzoic acids, the binding constants have been previously measured by different authors using different methods and reported in the literature [8–11,22–24]. Their results (for binding with both BSA and HSA) are listed in Table 2. The pH of buffer solutions used in these studies varied from 7.0 to 8.50. It can be seen that the values from different works strongly disagree with each other. As discussed above, the binding constant values are heavily method dependent. There is no point in direct comparison with our results, even if there are some occasional agreements. The same trend of decrease of  $K$  values in a row for 2-hydroxy > 3-hydroxy > 4-hydroxybenzoic acid observed in the present work and in two previous studies [9,11] provides some evidence of the consistency of results, but consideration of the structure–affinity relationships should be limited to the data obtained using exactly the same experimental procedure.



**Table 2.** Binding constants of the studied benzoic acids with BSA and human serum albumin (HSA) previously reported in literature (at 298 K if not otherwise mentioned). The letters in brackets indicate the method of measurement (CE—affinity capillary electrophoresis, SP—spectrophotometry, SF—spectrofluorimetry, UC—ultracentrifugation, ED—equilibrium dialysis, CS—capacitive sensing technique).

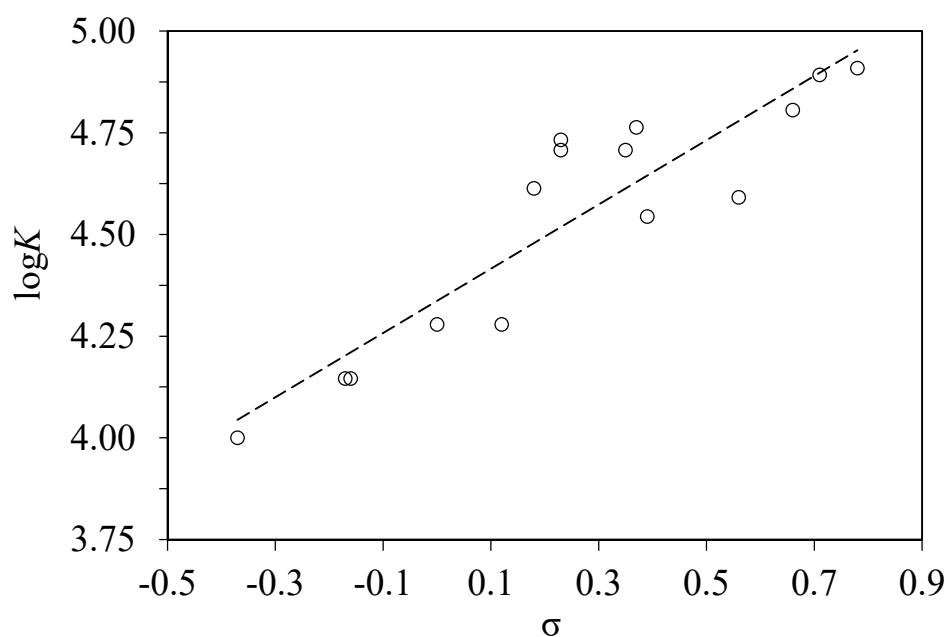
Acid	$K \cdot 10^{-4}$	
	BSA	HSA
2-hydroxybenzoic acid	3.92 (CE) [22], 7.38 (SF) [22], 10.0 (SP, 310 K) [9], 1.95 (SF) [11]	5.99 (CE) [22], 1.89 (CS) [22], 190 (SF) [23], 1.95 (SF) [10]
3-hydroxybenzoic acid	5.01 (SP, 310 K) [9], 0.85 (SF) [11]	0.34 (SF) [10]
4-hydroxybenzoic acid	0.79 (SP, 310 K) [9], <0.1 (SF) [11], 0.68 (UC, 288 K) [8]	<0.1 (SF) [10], 0.45 (UC, 288 K) [8]
2,4-dihydroxybenzoic acid	<0.1 (SF) [11]	0.16 (SF) [10]
4-nitrobenzoic acid	6.03 (UC, 288 K) [8], 3.09 (ED, 310 K) [24]	8.13 (UC, 288 K) [8]
2-aminobenzoic acid	0.63 (SP, 310 K) [9]	
3-aminobenzoic acid	0.32 (SP, 310 K) [9]	
3-chlorobenzoic acid	11.22 (ED, 310 K) [24]	
4-chlorobenzoic acid	14.13 (UC, 288 K) [8], 14.96 (ED, 310 K) [24]	14.8 (UC, 288 K) [8]
2-bromobenzoic acid	3.98 (UC, 310 K) [9]	
3-bromobenzoic acid	63.1 (UC, 310 K) [9]	
4-bromobenzoic acid	20.0 (UC, 310 K) [9], 20.42 (ED, 310 K) [24]	
4-cyanobenzoic acid	1.66 (UC, 288 K) [8]	2.75 (UC, 288 K) [8]
4-methylbenzoic acid	7.94 (SP, 310 K) [9], <0.1 (SF) [11], 6.76 (UC, 288 K) [8], 3.16 (ED, 310 K) [24]	0.35 (SF) [10], 3.63 (UC, 288 K) [8]
benzoic acid	5.01 (SP, 310 K) [9], 2.24 (UC, 288 K) [8], 1.15 (ED, 310 K) [24]	2.29 (UC, 288 K) [8]

#### 4.2. Correlation of the Binding Constants with Hammett Substituent Constants

We have not found a good correlation of the binding constants for the whole set of 24 substituted benzoic acid with any single physicochemical parameter of ligand. However, the values of  $\log K$  for a subset of *m*- and *p*-monosubstituted acids correlate with the Hammett constants  $\sigma$  of the substituents:

$$\begin{aligned} \log K &= 0.79\sigma + 4.34 \\ \text{RMS} &= 0.13, n = 15, r^2 = 0.8257. \end{aligned} \quad (6)$$

This correlation is shown in Figure 3. The interactions involving *o*-substituted benzoic acids is heavily influenced by the steric effects of substituents. Their behavior in any physical or chemical process cannot be interpreted in terms of Hammett constants, which are not defined for *o*-substituents as well as for disubstituted derivatives [25].



**Figure 3.** Correlation of the logarithms of binding constants  $\log K$  of *m*- and *p*-monosubstituted benzoic acids to BSA with the Hammett constants  $\sigma$  [25] of the substituents.

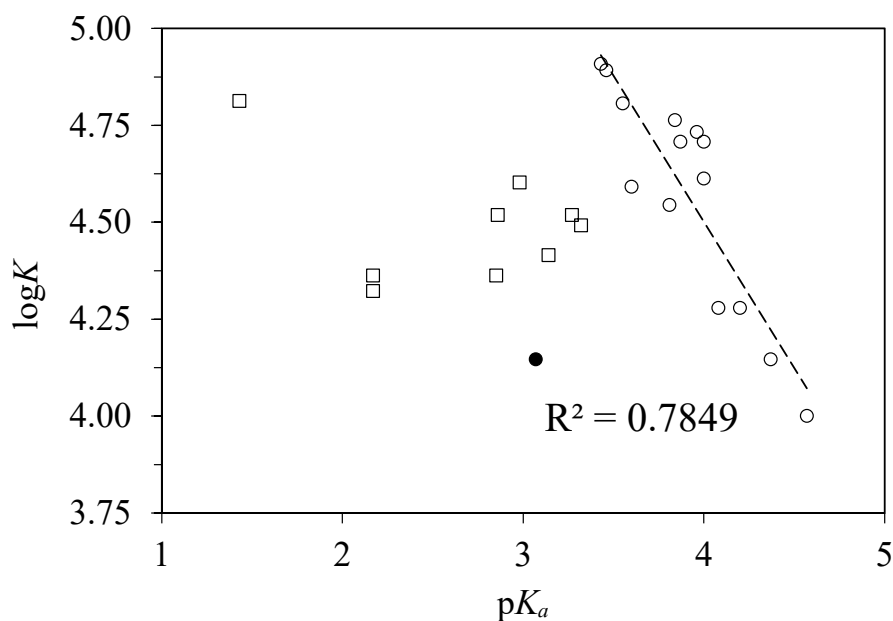
The Hammett constants describe the electron donating or withdrawing ability of a substituent in comparison with a hydrogen atom. Equilibrium and rate constants of various chemical processes involving substituted benzoic acids derivatives and other aromatic compounds have been shown to correlate with these constants [26,27]. The applications of this approach to protein–ligand binding are quite limited. Previously, the binding constants of *p*-substituted benzamides to trypsin were shown to have some correlation with the Hammett constants of substituents [28]. Enthalpies and entropies of binding of *p*-substituted acetanilides to BSA have shown a similar correlation [29].

The noncovalent interactions of *m*- and *p*-substituted benzoic acids with BSA molecule seem to follow the generic pattern observed for interactions with small organic molecules.

#### 4.3. Correlations with Other Experimentally Derived Ligand Properties

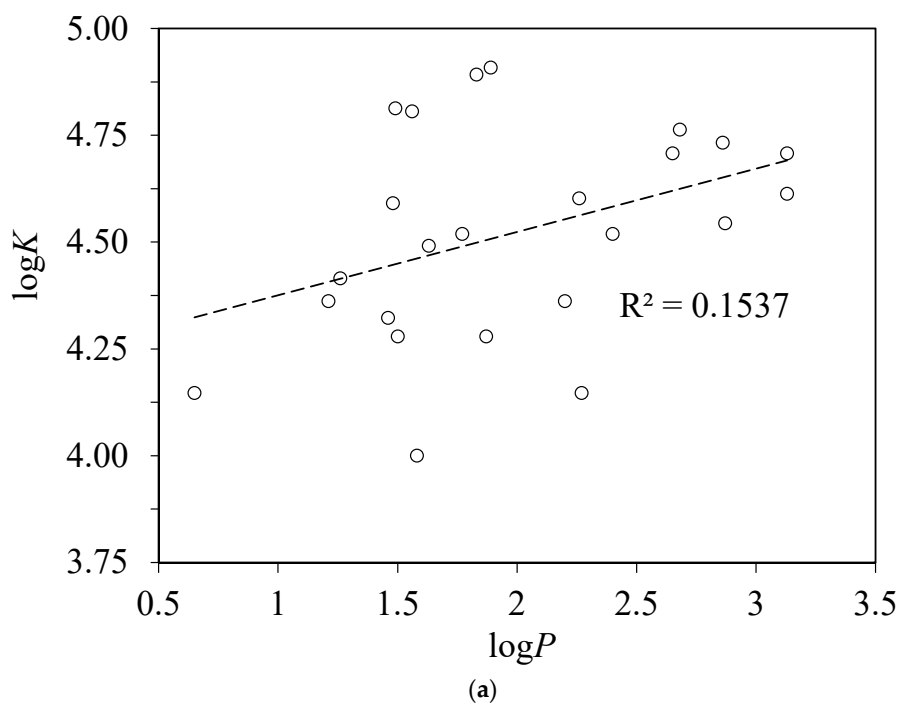
Since the Hammett constants of *m*- and *p*-substituted benzoic acids are derived from the logarithms of their dissociation constants  $pK_a$  in water [25], it is not surprising that the binding constants to BSA also correlate with the  $pK_a$  values (see Figure 4). This correlation does not hold for *o*-substituted acids. Another exception is 3-aminobenzoic acid for which the Hammett constant was derived using the dissociation constant in alcohol–water mixture different from that in pure water [30].



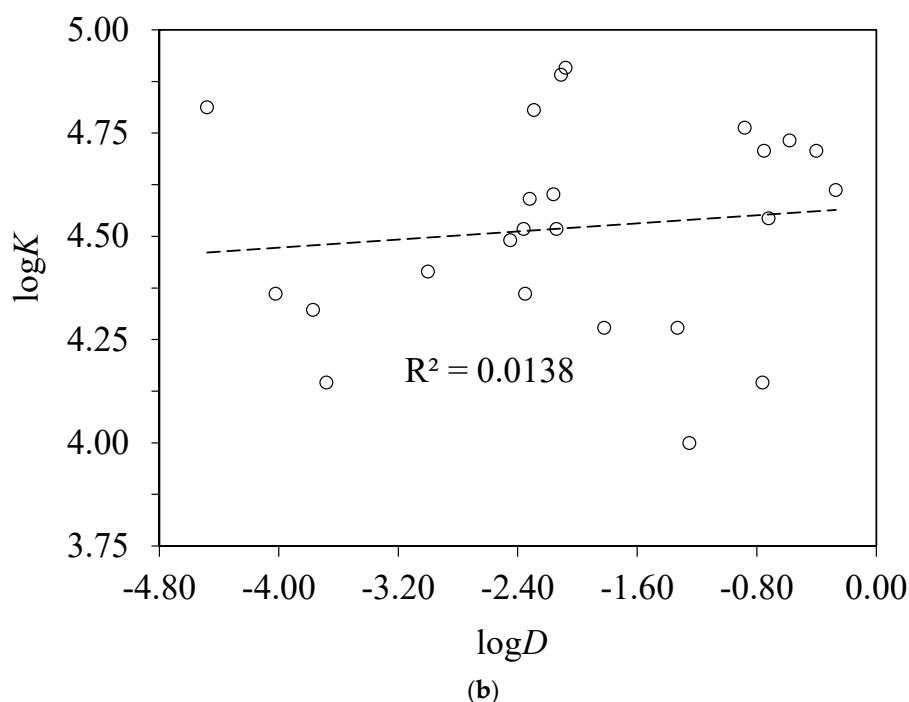


**Figure 4.** Plot of the logarithms of binding constants  $\log K$  of substituted benzoic acids to BSA against their  $pK_a$  values [31]. Empty circles represent m- and p-substituted acids, filled circle represents 3-aminobenzoic acid, empty squares represent o- and o,p-substituted acids.

There were attempts to compare or correlate the albumin binding constants of different compounds with their octanol–water partition coefficients  $\log P$  [1,2,24,32,33]. The quality of such correlations was not good, and it was not good in our case as well. For acidic ligands, it may be more correct to consider the distribution coefficient  $\log D$  of both ionized and nonionized forms between octanol and water, which is related to the distribution coefficient of nonionized form  $\log P$  through equation  $\log D = \log P - \log(1 + 10^{pH-pK_a})$ . Plots of  $\log K$  against  $\log P$  and  $\log D$  are shown in Figure 5.



**Figure 5.** Cont.

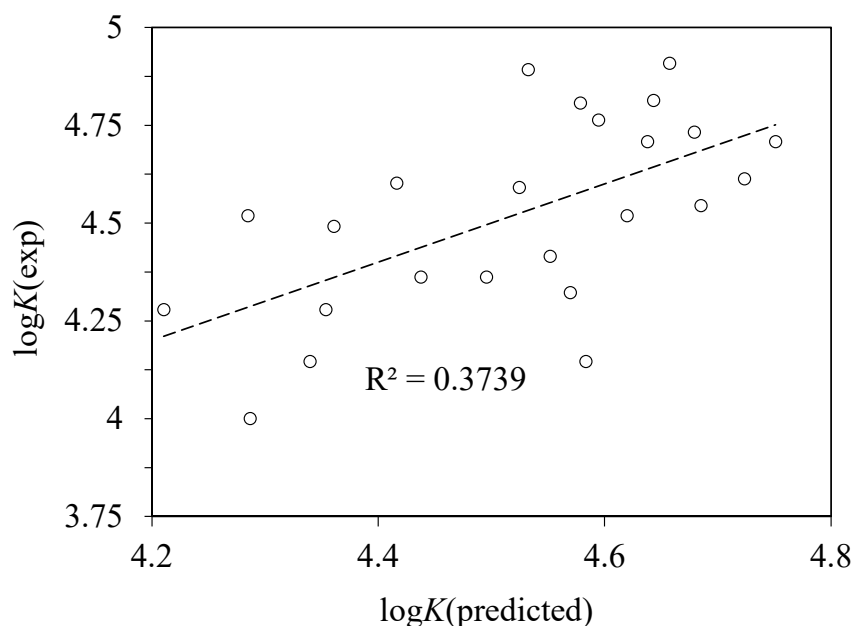


**Figure 5.** Plots of the logarithms of binding constants  $\log K$  of substituted benzoic acids to BSA against their octanol–water partition coefficients  $\log P$  [34] (a) and  $\log D$  (b).

In the works [2,32], the affinities of a large series of compounds, predominantly drugs, to immobilized HSA determined using HPLC method were correlated with a set of Abraham molecular descriptors of these compounds ( $E, S, A, B, V$ ) [35], which show a good predictive ability in describing partition of many different compounds between different solvents and biological fluids. We performed a linear regression with these five descriptors, which led to a low-quality correlation with  $r^2 = 0.374$  (Figure 6). One may argue that the acidity descriptor  $A$  is irrelevant when we consider binding of anionic forms of acids which have lost their acidic proton. The analysis performed in the work [32] showed that only the parameters  $S, B,$  and  $V$  are significant for description of binding of a large set of structurally diverse molecules to an immobilized HSA stationary phase in HPLC experiments. Regression of our results with these three descriptors led to a lower coefficient of correlation  $r^2 = 0.348$ . Moreover, we have considered all the possible linear regressions involving from one to five Abraham descriptors, and all of them had even lower values of  $r^2$ . Adding  $\log P$  and  $pK_a$  parameters to these regressions does not significantly improve the correlation quality. For the whole set of 24 ligands, the best possible correlation with one to seven of these parameters has  $r^2 = 0.407$ .

#### 4.4. QSAR Modeling of Albumin Binding

An alternative approach is based on QSAR modeling in which regressions with theoretically calculated descriptors are used. Previously, Colmenarejo et al. [36] developed a nonlinear QSAR model with six descriptors to describe the binding affinities of a set of 95 structurally diverse compounds to an immobilized HSA stationary phase. The  $r^2$  value for the best model was 0.83. Xue et al. [37] examined linear QSAR models for the same set of compounds and found a seven-parameter model with  $r^2 = 0.86$ .



**Figure 6.** Plot of the experimental logarithms of binding constants  $\log K$  against those predicted using a multiple linear regression with Abraham molecular descriptors E, S, A, B, and V [35].

We considered linear QSAR models with two sets of descriptors. One of them was generated using PaDEL-Descriptor software [38] and included 1410 2D, 1D and 0D descriptors. Another set was generated by eDragon software [39] and included more than 1600 3D, 2D, 1D and 0D descriptors. The descriptors having the same value for all the ligands and strictly intercorrelated descriptors were eliminated. This left 874 descriptors from the first set and 1159 descriptors from the second set.

With the first set, we initially considered all the possible regressions with two different descriptors (381,501 combinations). The highest correlation coefficient achieved was  $r^2 = 0.739$ :

$$\log K = 3.501\text{BCUTw-11} + 1.324\text{MATS3c} - 37.295. \quad (7)$$

RMS = 0.124,  $n = 24$ ,  $r^2 = 0.739$

The first 40 top scoring correlations included BCUTw-11 descriptor. Next, all the possible regressions with three descriptors, one of which was always BCUTw-11, were considered. The best one had  $r^2 = 0.851$  and is given by the following equation (Figure 7):

$$\log K = 3.601\text{BCUTw-11} + 0.0458\text{MPC7} - 0.5258\text{SP-4} - 37.706. \quad (8)$$

RMS = 0.094,  $n = 24$ ,  $r^2 = 0.851$

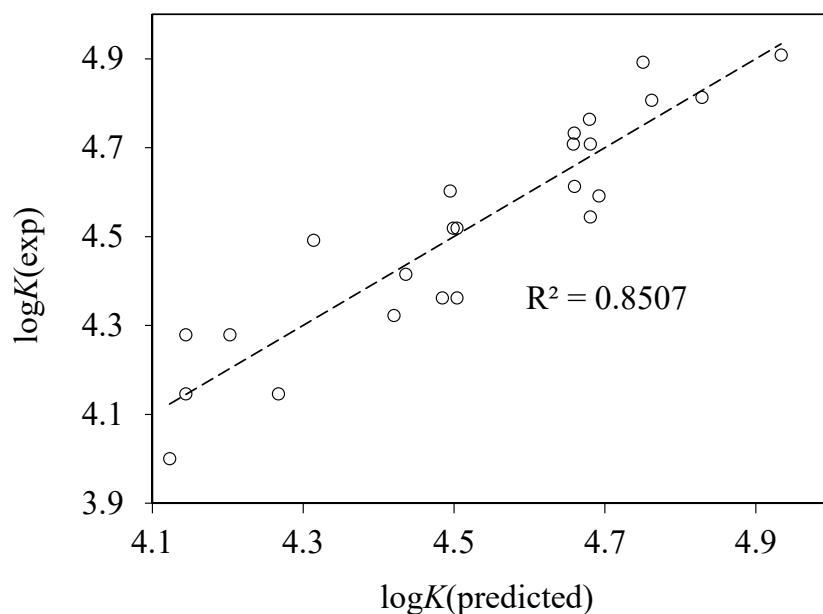
Additionally, a search for correlations with three different descriptors was performed using a genetic algorithm. No better correlations were found. The use of four descriptors leads to a further increase in the quality of correlations (up to  $r^2 = 0.914$  for the following correlation):

$$\log K = 3.476\text{BCUTw-11} - 4.748\text{AATSC8c} - 1.435\text{GATS3c} - 5.809 \text{SpMin1\_Bhs} - 25.099. \quad (9)$$

RMS = 0.071,  $n = 24$ ,  $r^2 = 0.914$

The importance of the BCUTw-11 descriptor arises from its calculation algorithm. This descriptor is a weighted version of the Burden matrix [40] which takes into account the connectivity, charge and polarizability of atoms in a molecule [41]. It incorporates the information about the electron density in benzene ring. Despite BCUTw-11 for *m*- and *p*-monosubstituted benzoic acids being not strictly correlated with the Hammett constant of substituents, its values are in general lower in the presence of electron donating and higher in the presence of electron accepting groups. Some other applications

of descriptors from BCUT family in description of protein–ligand interactions have been previously demonstrated [42–44].

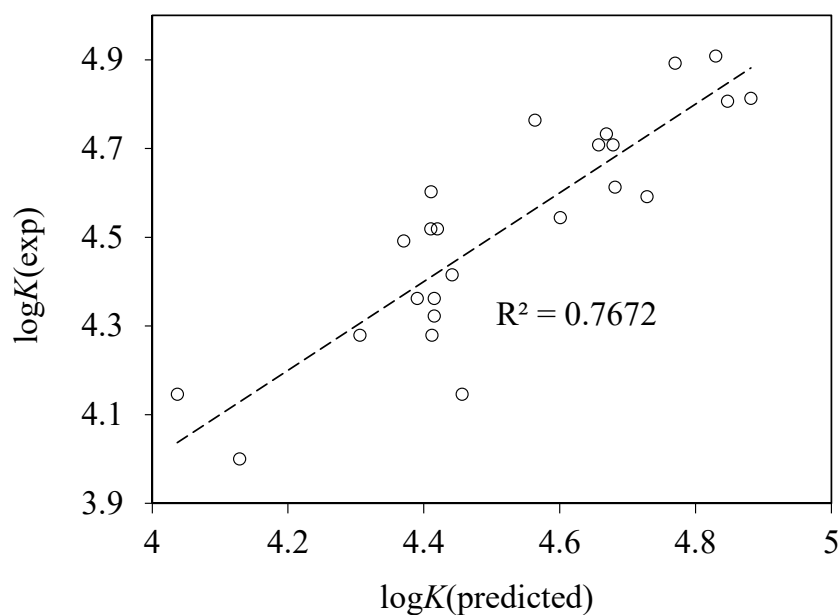


**Figure 7.** Plot of the experimental vs. predicted logarithms of binding constants  $\log K$  obtained using the best three-parameter model (Equation (8)) with parameters from PaDEL descriptor set.

With the second set of descriptors, we have also performed an exhaustive search of two-parameter regressions (671,061 combinations). The first 36 correlations with the largest  $r^2$  values included either HATS7v or HATS7u descriptors. The best correlation (Figure 8) is given by:

$$\log K = -10,274 \text{ HATS7v} + 0.26\text{DP13} + 4.415. \quad (10)$$

$$\text{RMS} = 0.117, n = 24, r^2 = 0.767$$



**Figure 8.** Plot of the experimental vs. predicted logarithms of binding constants  $\log K$  obtained using the best two-parameter model (Equation (10)) with parameters from eDragon descriptor set.

The best of three-parameter models with HATS7v as one of the descriptors is given by equation:

$$\log K = -9.095\text{HATS7v} + 0.202\text{Mor06e} + 0.307\text{DP14} + 4.781. \quad (11)$$
$$\text{RMS} = 0.1, n = 24, r^2 = 0.829$$

## 5. Conclusions

Binding constants of substituted benzoic acids to BSA are sensitive to the nature and position of the substituents. The electron-density distribution in aromatic ring exerts crucial influence on the affinity of the acid anions, which is confirmed by correlation of the binding constants with the Hammett constants of substituents in m- and p-positions. While the Hammett constants or any other single ligand parameter cannot be used to describe binding of o-isomers, QSAR models with even as few as two descriptors are quite successful, and a three-descriptor model allowed to achieve the  $r^2$  value of 0.85 for 24 different acids. Further accumulation of the data on albumin binding constants of more chemically diverse sets of substances are necessary to develop general QSAR models, which can be useful for virtual screening of the affinity of drug candidates to blood plasma. However, only the constant values obtained using a uniform experimental methodology and data analysis procedure can be used for the development of such models. Numerous results from the previously published studies of albumin binding are unreliable and often inconsistent with each other even for the same system due to the strong dependence on the measurement technique.

**Supplementary Materials:** The following are available online at <http://www.mdpi.com/1424-8247/13/2/30/s1>, Script: Numerical simulation of Stern-Volmer plot for titration of 4-nitrobenzoic acid.

**Author Contributions:** Conceptualization, I.S.; Investigation, D.K., A.N. and I.S.; Methodology, I.S.; Writing—original draft, D.K. and I.S. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work is supported by the Russian Science Foundation under grant 19-73-00209.

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

## References

1. Lexa, K.W.; Dolgih, E.; Jacobson, M.P. A Structure-Based Model for Predicting Serum Albumin Binding. *PLoS ONE* **2014**, *9*, e93323. [[CrossRef](#)] [[PubMed](#)]
2. Valko, K.; Nunhuck, S.; Bevan, C.; Abraham, M.H.; Reynolds, D.P. Fast Gradient HPLC Method to Determine Compounds Binding to Human Serum Albumin. Relationships with Octanol/Water and Immobilized Artificial Membrane Lipophilicity. *J. Pharm. Sci.* **2003**, *92*, 2236–2248. [[CrossRef](#)] [[PubMed](#)]
3. Kragh-Hansen, U. Molecular aspects of ligand binding to serum albumin. *Pharmacol. Rev.* **1981**, *33*, 17–53.
4. Nakano, N.I.; Shimamori, Y.; Yamaguchi, S. Binding capacities of human serum albumin monomer and dimer by continuous frontal affinity chromatography. *J. Chromatogr. A* **1982**, *237*, 225–232. [[CrossRef](#)]
5. Levi, V.; González Flecha, F.L. Reversible fast-dimerization of bovine serum albumin detected by fluorescence resonance energy transfer. *Biochim. Biophys. Acta BBA Proteins Proteom.* **2002**, *1599*, 141–148. [[CrossRef](#)]
6. Lowry, T.H.; Richardson, K.S. *Mechanism and Theory in Organic Chemistry*, 3rd ed.; Harper & Row: New York, NY, USA, 1987; ISBN 978-0-06-044084-8.
7. Kubinyi, H. QSAR: Hansch Analysis and Related Approaches. In *Methods and Principles in Medicinal Chemistry*; VCH: Weinheim, Germany; New York, NY, USA, 1993; ISBN 978-3-527-30035-8.
8. Matsushita, Y.; Kijima (née Shimizu), Y.; Moriguchi, I. Comparison of binding characteristics of human and bovine serum albumins with benzoates over a wide range of concentration. *Chem. Pharm. Bull.* **1987**, *35*, 2589–2593. [[CrossRef](#)]
9. Moriguchi, I. Protein Bindings. II. Binding of Aromatic Carboxylic Acids to Bovine Serum Albumin. *Chem. Pharm. Bull.* **1968**, *16*, 597–600. [[CrossRef](#)]
10. Zhang, Y.; Wu, S.; Qin, Y.; Liu, J.; Liu, J.; Wang, Q.; Ren, F.; Zhang, H. Interaction of phenolic acids and their derivatives with human serum albumin: Structure–affinity relationships and effects on antioxidant activity. *Food Chem.* **2018**, *240*, 1072–1080. [[CrossRef](#)]

11. Yuan, S.; Zhang, Y.; Liu, J.; Zhao, Y.; Tan, L.; Liu, J.; Wang, Q.; Zhang, H. Structure-affinity relationship of the binding of phenolic acids and their derivatives to bovine serum albumin. *Food Chem.* **2019**, *278*, 77–83. [[CrossRef](#)]
12. Rosenoer, V.M.; Oratz, M.; Rothschild, M.A. (Eds.) *Albumin Structure, Function and Uses*, 1st ed.; Pergamon Press: Oxford, UK; New York, NY, USA, 1977; ISBN 978-0-08-019603-9.
13. Kaneko, K.; Fukuda, H.; Chuang, V.T.G.; Yamasaki, K.; Kawahara, K.; Nakayama, H.; Suenaga, A.; Maruyama, T.; Otagiri, M. Subdomain IIIA of Dog Albumin Contains a Binding Site Similar to Site II of Human Albumin. *Drug Metab. Dispos.* **2008**, *36*, 81–86. [[CrossRef](#)]
14. Majorek, K.A.; Porebski, P.J.; Dayal, A.; Zimmerman, M.D.; Jablonska, K.; Stewart, A.J.; Chruszcz, M.; Minor, W. Structural and immunologic characterization of bovine, horse, and rabbit serum albumins. *Mol. Immunol.* **2012**, *52*, 174–182. [[CrossRef](#)]
15. Zhang, Y.-X.; Chen, C.-W.; Wang, M.; Wei, S.-S.; Guan, H.; Chi, T.-T.; Qi, X.-Z.; Hu, W.-T. Purification and Characterization of Albumin from Frog Skin of *Duttaphrynus melanostictus*. *Protein J.* **2011**, *30*, 464–470. [[CrossRef](#)] [[PubMed](#)]
16. Sudlow, G.; Birkett, D.J.; Wade, D.N. The characterization of two specific drug binding sites on human serum albumin. *Mol. Pharmacol.* **1975**, *11*, 824–832. [[PubMed](#)]
17. Wenskowsky, L.; Schreuder, H.; Derdau, V.; Matter, H.; Volkmar, J.; Nazaré, M.; Opatz, T.; Petry, S. Identification and Characterization of a Single High-Affinity Fatty Acid Binding Site in Human Serum Albumin. *Angew. Chem. Int. Ed.* **2018**, *57*, 1044–1048. [[CrossRef](#)] [[PubMed](#)]
18. Goncharov, N.V.; Terpilovskii, M.A.; Shmurak, V.I.; Belinskaya, D.A.; Avdonin, P.V. Comparative analysis of esterase and paraoxonase activities of different serum albumin species. *J. Evol. Biochem. Physiol.* **2017**, *53*, 271–281. [[CrossRef](#)]
19. Moriyama, Y.; Ohta, D.; Hachiya, K.; Mitsui, Y.; Takeda, K. Fluorescence behavior of tryptophan residues of bovine and human serum albumins in ionic surfactant solutions: A comparative study of the two and one tryptophan(s) of bovine and human albumins. *J. Protein Chem.* **1996**, *15*, 265–272. [[CrossRef](#)]
20. Stella, L.; Van de Weert, M.; Burrows, H.D.; Fausto, R. Fluorescence spectroscopy and binding: Getting it right. *J. Mol. Struct.* **2014**, *1077*, 1–3. [[CrossRef](#)]
21. Van de Weert, M.; Stella, L. Fluorescence quenching and ligand binding: A critical discussion of a popular methodology. *J. Mol. Struct.* **2011**, *998*, 144–150. [[CrossRef](#)]
22. Li, Z.-M.; Wei, C.-W.; Zhang, Y.; Wang, D.-S.; Liu, Y.-N. Investigation of competitive binding of ibuprofen and salicylic acid with serum albumin by affinity capillary electrophoresis. *J. Chromatogr. B* **2011**, *879*, 1934–1938. [[CrossRef](#)]
23. Matias, I.; Ceballos, A.; Gonzalez-Velasco, F.; Cachaza, J.M. Thermodynamics of the binding of salicylate to human serum albumin: Evidence of non-competition with imidazole. *J. Pharm. Pharmacol.* **1989**, *41*, 123–125. [[CrossRef](#)]
24. Henneberger, L.; Goss, K.-U.; Endo, S. Equilibrium Sorption of Structurally Diverse Organic Ions to Bovine Serum Albumin. *Environ. Sci. Technol.* **2016**, *50*, 5119–5126. [[CrossRef](#)]
25. Hansch, C.; Leo, A.; Taft, R.W. A survey of Hammett substituent constants and resonance and field parameters. *Chem. Rev.* **1991**, *91*, 165–195. [[CrossRef](#)]
26. Hammett, L.P. *Physical Organic Chemistry; Reaction Rates, Equilibria, and Mechanisms*; McGraw-Hill: New York, NY, USA, 1970; ISBN 978-0-07-025905-8.
27. Lewis, M.; Bagwill, C.; Hardebeck, L.K.E.; Wireduaah, S. The use of hammett constants to understand the non-covalent binding of aromatics. *Comput. Struct. Biotechnol. J.* **2012**, *1*, e201204004. [[CrossRef](#)] [[PubMed](#)]
28. Talhout, R.; Engberts, J.B.F.N. Thermodynamic analysis of binding of *p*-substituted benzamidines to trypsin: Thermodynamics of binding of benzamidines to trypsin. *Eur. J. Biochem.* **2001**, *268*, 1554–1560. [[CrossRef](#)] [[PubMed](#)]
29. Dearden, J.C.; Tomlinson, E. Physico-chemical studies of analgesics. The protein-binding of some *p*-substituted acetanilides. *J. Pharm. Pharmacol.* **1970**, *22*, 53S–59S. [[CrossRef](#)] [[PubMed](#)]
30. McDaniel, D.H.; Brown, H.C. An Extended Table of Hammett Substituent Constants Based on the Ionization of Substituted Benzoic Acids. *J. Org. Chem.* **1958**, *23*, 420–427. [[CrossRef](#)]
31. Albert, A.; Serjeant, E.P. *The Determination of Ionization Constants: A Laboratory Manual*; Springer: Dordrecht, The Netherlands, 1984; ISBN 978-94-010-8948-7.



32. Kamble, S.; Loadman, P.; Abraham, M.H.; Liu, X. Structural properties governing drug-plasma protein binding determined by high-performance liquid chromatography method. *J. Pharm. Biomed. Anal.* **2018**, *149*, 16–21. [[CrossRef](#)]
33. Barbato, F.; di Martino, G.; Grumetto, L.; La Rotonda, M.I. Retention of quinolones on human serum albumin and  $\alpha$ 1-acid glycoprotein HPLC columns: Relationships with different scales of lipophilicity. *Eur. J. Pharm. Sci.* **2007**, *30*, 211–219. [[CrossRef](#)]
34. Hansch, C.; Leo, A.; Hoekman, D.H. (Eds.) *Exploring QSAR*; ACS Professional Reference Book; American Chemical Society: Washington, DC, USA, 1995; ISBN 978-0-8412-2987-7.
35. ACD/Absolv. The Absolv Prediction Module Calculates Abraham Solvation Parameters and Is the Result of Collaboration between ACD/Labs and Prof. MH Abraham. 2014. Available online: <http://www.acdlabs.com/products/percepta/predictors/absolv/> (accessed on 15 February 2020).
36. Colmenarejo, G.; Alvarez-Pedraglio, A.; Lavandera, J.-L. Cheminformatic Models to Predict Binding Affinities to Human Serum Albumin. *J. Med. Chem.* **2001**, *44*, 4370–4378. [[CrossRef](#)]
37. Xue, C.X.; Zhang, R.S.; Liu, H.X.; Yao, X.J.; Liu, M.C.; Hu, Z.D.; Fan, B.T. QSAR Models for the Prediction of Binding Affinities to Human Serum Albumin Using the Heuristic Method and a Support Vector Machine. *J. Chem. Inf. Comput. Sci.* **2004**, *44*, 1693–1700. [[CrossRef](#)]
38. Yap, C.W. PaDEL-descriptor: An open source software to calculate molecular descriptors and fingerprints. *J. Comput. Chem.* **2011**, *32*, 1466–1474. [[CrossRef](#)] [[PubMed](#)]
39. Tetko, I.V.; Gasteiger, J.; Todeschini, R.; Mauri, A.; Livingstone, D.; Ertl, P.; Palyulin, V.A.; Radchenko, E.V.; Zefirov, N.S.; Makarenko, A.S.; et al. Virtual Computational Chemistry Laboratory-Design and Description. *J. Comput. Aided Mol. Des.* **2005**, *19*, 453–463. [[CrossRef](#)] [[PubMed](#)]
40. Burden, F.R. Molecular identification number for substructure searches. *J. Chem. Inf. Model.* **1989**, *29*, 225–227. [[CrossRef](#)]
41. Pearlman, R.S.; Smith, K.M. Metric Validation and the Receptor-Relevant Subspace Concept. *J. Chem. Inf. Comput. Sci.* **1999**, *39*, 28–35. [[CrossRef](#)]
42. González, M.P.; Terán, C.; Teijeira, M.; Besada, P.; González-Moa, M.J. BCUT descriptors to predicting affinity toward A3 adenosine receptors. *Bioorg. Med. Chem. Lett.* **2005**, *15*, 3491–3495. [[CrossRef](#)]
43. Pirard, B.; Pickett, S.D. Classification of Kinase Inhibitors Using BCUT Descriptors. *J. Chem. Inf. Comput. Sci.* **2000**, *40*, 1431–1440. [[CrossRef](#)]
44. Hao, M.; Li, Y.; Wang, Y.; Zhang, S. A classification study of human  $\beta$  3-adrenergic receptor agonists using BCUT descriptors. *Mol. Divers.* **2011**, *15*, 877–887. [[CrossRef](#)]



© 2020 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 (<http://creativecommons.org/licenses/by/4.0/>).