

Stability of Complexes of Aromatic Amides with Bromide Anion: Quantitative Structure–Property Relationships

Michal Čajan,^{†,‡} Jiří Damborský,[†] Ivan Stibor,[§] and Jaroslav Koča^{†,‡,*}

Laboratory of Biomolecular Structure and Dynamics and Department of Organic Chemistry, Faculty of Science, Masaryk University, Kotlářská 2, 611 37 Brno, Czech Republic, and Department of Organic Chemistry, Institute of Chemical Technology, Technická 5, 166 28 Prague, Czech Republic

Received February 3, 2000

Most of the theoretical studies published to-date on the structural and electronic properties of supramolecules have been devoted to the neutral or cationic complexes, while little is known about anionic systems. A detailed theoretical study of the interaction between simple aromatic amides and the bromide anion has recently been published (Čajan, M.; Stibor, I.; Koča, J. *J. Phys. Chem. A* **1999**, *103*, 3778). The present work focuses on the structural and physicochemical parameters of simple aromatic amides related to their ability to form the 1:1 complex with a bromide anion. A quantitative structure–property relationships (QSPR) model for the prediction of association constants is proposed. The model based on 22 complexes and nine molecular descriptors explained 96% (84% cross-validated) of the variance in association constants. The descriptors employed in this model included parameters for the characterization of conformational behavior and the 3D structure of amide molecules, distribution of electron density on the amidic functional group, and parameters for substitution on aromatic units. The quantitative structure–property relationship approach predicted the association constants with comparable quality, but significantly lower computational demand, than molecular modeling or standard quantum chemistry calculations.

1. INTRODUCTION

Theoretical studies of supramolecular complexes provide insight into the structure, conformation, and electronic properties of these interesting chemical species. The study of electron-deficient abiotic receptor molecules designed for non-covalent binding of anion guest species is an area of ever increasing research activity.^{1,2}

Anions play numerous fundamental roles in biological and chemical processes.³ For example, a number of enzymes bind anions as either substrates or cofactors. An understanding of the structural rules underlying the engineering of these anion binding sites is of paramount importance. It can have a considerable impact on the design of molecules that may find application as therapeutic agents. For example, in phosphate binding protein, there are as many as 12 distinct hydrogen bonds fixing the position of the anion in the binding site.⁴

We have initiated a research project aimed at the understanding of factors governing the ability of the simple secondary amide bond to form 1:1 complexes with anions and designing the abiotic receptors with several amide bonds co-operating in the anion binding event.⁵ We have recently published a detailed theoretical study of these important phenomena.⁶

Association constants were measured for bromide ions and amides substituted on both sides of the amidic bond by an aromatic group bearing different substituents. Subsequently, the models of amide-bromide complexes were investigated

by molecular mechanics and quantum mechanics calculations to understand the structural determinants of complexation and to predict stabilities. The computational study of the complexes with the anions was found to be complicated due to the requirements for the methodology and the hardware. Molecular mechanics (CVFF) methods and semiempirical quantum mechanics methods (AM1, PM3) provided incorrect results, while calculations with quantum mechanics methods based on DFT and the Møller–Plesset perturbation theory were computationally very demanding.⁶ The use of such calculations for predictive purposes is therefore limited.

The structural formula of an organic compound, in principle, contains all of the information, which predetermines the chemical, biological, and physical properties of that compound⁷ or its behavior to different ambience or phenomena. Here we describe the application of quantitative structure–properties relationships (QSPR) for the prediction of the property of supramolecular species from their chemical structure. Employing a QSPR model, we will attempt to explain the stability of 22 [amide][anion] complexes based on their electronic and steric properties. We will attempt to demonstrate that the QSPR methodology provides a robust and relatively fast tool for the prediction of association constants of such complexes.

2. METHODS

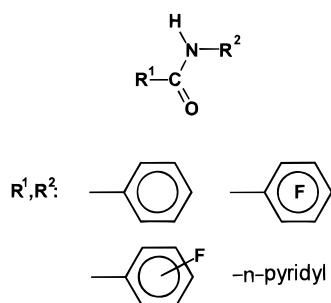
2.1. Association Constants of [Amide]Br[−] Complexes (Variable y). Association constants for 1:1 complexes [amide]Br[−] were determined by titration ¹H NMR of the solution of the corresponding amide with the addition of (tBu)₄N⁺Br[−].⁵ Stability constants were determined for 22 complexes of aromatic amides with the general structure shown in Scheme 1. The studied amides are listed in Table 1 together with the experimentally obtained stability constants.

* Corresponding author phone: +420-5-41129 310; fax: +420-5-41129 506; e-mail: jkoca@chemi.muni.cz.

[†] Laboratory of Biomolecular Structure and Dynamics, Masaryk University.

[‡] Department of Organic Chemistry, Masaryk University.

[§] Institute of Chemical Technology.

Scheme 1: General Formula for the Set of Studied Amides^a

^a The substituents R^1 and R^2 mean one of the following functional groups: phenyl, fluorophenyl, perfluorophenyl and pyridil.

2.2. Amide Property Data (Variables x). Descriptors used for QSPR analysis can be conveniently divided into three categories: (2.2.1) steric parameters, (2.2.2) electronic parameters, and (2.2.3) parameters for the description of substitution pattern. The annotation and the sources of descriptors are summarized in Table 2, while the values are summarized in Table 3. The steric and electronic descriptors are illustrated in Figure 1 and discussed in the text below.

2.2.1. Steric Descriptors. The steric descriptors were calculated for the most stable conformer of each molecule. All amides were optimized by the DFT method on B3LYP/dzvd level^{8,9} prior to calculation of the descriptors. All DFT calculations were performed with TURBOMOLE v95.0/3.0.0 (Biosym/MSI, USA). The driving force to create the complex is the interaction between the bromide anion and the amidic hydrogen. The strength of this interaction is influenced by the volume and electron affinity of the *ortho* substituents on both aromatic moieties. This effect is stronger for the C aromatic substituent than for the N substituent of the amidic group. It was included in **D(XH)**, **D(XY)**, **Dsum**, and **Tor** descriptors. On the other hand, the distance **D(YH)** ($-\text{NH}-\text{Y}-$) was not used as a descriptor because it is similar in all complexes. The amidic group with the NH-ring system is nonplanar only in one case (molecule **8**), and this effect is incorporated in the descriptor **D(XY)**. **MM**, **MS**, **MV**, **EV**, **IM1s**, **IM2s**, **IM3s**, **IM1l**, **IM2l**, **IM3l**, **Sim**, **SimQ**, **SimS**, **Rtop**, **Btop**, **Wtop**, and **FI** are the global descriptors of amide molecules, which were calculated using TSAR v3.1 (Oxford Molecular, UK). For an explanation of the abbreviations see Table 2.

2.2.2. Electronic Descriptors. The distribution of electron density in the amide and its influence on amide-anion interactions was characterized by a set of electronic descriptors. Electron density was expressed by the partial charges on the atoms around the binding site for the anion (i.e., around the amidic hydrogen) and by the selected bond orders of the central amidic group. Four population analyses were tested for the calculation of partial charges: Mulliken,¹⁰ Löwdin,^{11,12} Roby–Davidson,¹³ and ESP.¹⁴ The charges for QSPR analysis were finally derived from ESP population analysis, which gave the best results. The partial charges **qX**, **qY**, **qH**, the sum of these charges **Q**, and the total dipole **Tdip** were used for the description of the influence of electron density and the respective repulsive or attractive forces of atoms interacting with the bromide anion on the stability of the complex.

The substituents on both the aromatic moieties can activate or deactivate the amidic group through the π -electron system.

The electron density on the amidic group regulates the electronic properties of the amidic hydrogen. So, the bond orders of all bonds on the amidic group are dependent on the substitution of the aromatic moieties and on the 3D structure of the molecule. Bond orders of selected bonds, $\mu\text{-C}_{\text{Ar}}$, $\mu\text{-N}_{\text{Ar}}$, $\mu\text{-NH}$, $\mu\text{-CO}$, $\mu\text{-CN}$, and their sums, $\Sigma\mu_{\text{Ar}}$, $\Sigma\mu$, were used to include the electron density distribution on the amidic functional group.¹⁵ Mayer bond orders¹⁶ were computed on the B3LYP/dzvd/iii_iglo level by DEMON v1.0.¹⁷

2.2.3. Empirical Descriptors of the Substitution Pattern.

The properties arising from the different substitution patterns were quantified by the Hammett (σ) and Taft (E_s) constants.¹⁸ The Hammett constants for substituents in *meta* and *para* positions on both the aromatic moieties consist of $\sigma\text{-NH}$, $\sigma\text{-CO}$, and $\sigma\text{-tot}$. The steric and electronic effects of substituents in *ortho* positions are characterized by the Taft constants **E_s-NH**, **E_s-CO**, and **E_s-tot**. The sums of constants on both the aromatic moieties and the total sum of σ and E_s (**-tot**) were used to simplify the models. All **E_s** substituent constants have odd distributions. **E_s-tot** has only four different values (two of them are unique), **E_s-CO** has only three, and **E_s-NH** is essentially an indicator variable for compounds **1** and **8**. Despite this fact, the incorporation of these descriptors in the modeling could be necessary for description of nonbonding interactions in binding site.

2.3. Data Analysis. 2.3.1. Partial Least-Squares Projection to Latent Structures (PLS). Partial Least-Squares Projection to Latent Structures¹⁹ analysis was used to select the most appropriate descriptors for the prediction of association constants. It correlates the systematic variation of the matrix of dependent variables **Y** to the systematic variation in the independent variables **X**. PLS finds combinations of independent variables (here descriptors) that can predict dependent variables (here the association constant) and at the same time approximates the matrix **X** accurately. Autoscaled and centered data were used in the PLS analysis. The importance of every descriptor in the model was accessed using the VIP parameter (variable importance in the projection) and plots of the PLS variable weights (loading plots). All statistical analyses described in this article were conducted using SIMCA-S v6.1 (UmeTri, Sweden).

2.3.2. Validation. The internal validation was employed to access the quality and validity of the developed PLS models. Cross-validation routine and permutation testing secured internal validation. During the cross-validation procedure, parts of the **Y** data are kept out of the model development, predicted by the model, and compared with the actual values, providing the cross-validated Q^2 . This gives a more realistic value for the predictive power than the squared multiple regression coefficient (R^2). In the permutation testing, the model was recalculated with a randomly reordered dependent variable.

3. RESULTS

3.1. MODEL I. The PLS analysis was applied to the data matrix composed of the training set of 22 amides and 32 descriptors. The values of the dependent variable were logarithmically transformed to improve data distribution. One-component PLS model (**MODEL I**) was established explaining 41% of the variance in the stability of complexes

Table 1: Studied Molecules and Their Experimentally Determined Association Constants

| no. | compound | $K_{\text{ass}} [\text{M}^{-1}]$ | no. | compound | $K_{\text{ass}} [\text{M}^{-1}]$ |
|-----|---|----------------------------------|-----|---|----------------------------------|
| 1 | $\text{C}_6\text{F}_5\text{-CONH-}o\text{-F-C}_6\text{H}_4$ | 62 ± 16 | 12 | $p\text{-F-C}_6\text{H}_4\text{-CONH-}p\text{-F-C}_6\text{H}_4$ | 880 ± 28 |
| 2 | $o\text{-F-C}_6\text{H}_4\text{-CONH-C}_6\text{H}_5$ | 64 ± 18 | 13 | $p\text{-F-C}_6\text{H}_4\text{-CONH-}m\text{-F-C}_6\text{H}_4$ | 930 ± 38 |
| 3 | $\text{C}_6\text{H}_5\text{-CONH-C}_6\text{H}_5$ | 145 ± 12 | 14 | $\text{C}_6\text{F}_5\text{-CONH-C}_6\text{H}_5$ | 1320 ± 121 |
| 4 | $m\text{-F-C}_6\text{H}_4\text{-CONH-C}_6\text{H}_5$ | 210 ± 25 | 15 | $4\text{-pyridyl-CONH-C}_6\text{H}_5$ | 1650 ± 148 |
| 5 | $3\text{-pyridyl-CONH-C}_6\text{H}_5$ | 258 ± 22 | 16 | $\text{C}_6\text{F}_5\text{-CONH-3-pyridyl}$ | 1680 ± 174 |
| 6 | $\text{C}_6\text{H}_5\text{-CONH-}p\text{-F-C}_6\text{H}_4$ | 297 ± 10 | 17 | $\text{C}_6\text{F}_5\text{-CONH-}p\text{-F-C}_6\text{H}_4$ | 1687 ± 182 |
| 7 | $\text{C}_6\text{H}_5\text{-CONH-}m\text{-F-C}_6\text{H}_4$ | 340 ± 15 | 18 | $\text{C}_6\text{F}_5\text{-CONH-}m\text{-F-C}_6\text{H}_4$ | 2305 ± 134 |
| 8 | $\text{C}_6\text{H}_5\text{-CONH-C}_6\text{F}_5$ | 368 ± 9 | 19 | $\text{C}_6\text{F}_5\text{-CONH-4-pyridyl}$ | 2684 ± 153 |
| 9 | $p\text{-F-C}_6\text{H}_4\text{-CONH-C}_6\text{H}_5$ | 434 ± 30 | 20 | $\text{C}_6\text{F}_5\text{-CONH-2-pyridyl}$ | 3200 ± 187 |
| 10 | $\text{C}_6\text{H}_5\text{-CONH-2-pyridyl}$ | 530 ± 36 | 21 | $\text{C}_6\text{F}_5\text{-CONH-3,5-F}_2\text{-C}_6\text{H}_3$ | 4750 ± 158 |
| 11 | $m\text{-F-C}_6\text{H}_4\text{-CONH-}m\text{-F-C}_6\text{H}_4$ | 700 ± 29 | 22 | $4\text{-pyridyl-CONH-3,5-F}_2\text{-C}_6\text{H}_3$ | 5400 ± 298 |

Table 2: Molecular Descriptors—Description, Annotation and Source

| | descriptor | source | | descriptor | source |
|------|-----------------------------------|--------|-------------------------|--|-----------|
| MM | molecular mass | TSAR | Q | sum of charges on X, Y, H | DFT |
| MS | molecular surface area | TSAR | D | interatomic distance between Y and H | DFT |
| MV | molecular volume | TSAR | Tor | torsion angle for description of nonplanarity of amide | DFT |
| EV | ellipsoidal volume | TSAR | D(XY) | interatomic distance between X and Y | DFT |
| IM1s | inertia moment 1 size | TSAR | Dsum | sum of distances between X, Y, H | DFT |
| IM2s | inertia moment 2 size | TSAR | Tdip | dipole moment | DFT |
| IM3s | inertia moment 3 size | TSAR | $\sigma\text{-CO}$ | sum of σ -constants of m,p-substituents on CO ring | empirical |
| IM1l | inertia moment 1 length | TSAR | $\sigma\text{-NH}$ | sum of σ -constants of m,p-substituents on NH ring | empirical |
| IM2l | inertia moment 2 length | TSAR | $\sigma\text{-tot}$ | sum of σ -constants of all m,p-substituents | empirical |
| IM3l | inertia moment 3 length | TSAR | $E_s\text{-CO}$ | sum of E_s -constants of o-substituents on CO ring | empirical |
| Sim | combined similarity vs molecule 3 | TSAR | $E_s\text{-NH}$ | sum of E_s -constants of o-substituents on NH ring | empirical |
| SimQ | charge similarity vs molecule 3 | TSAR | $E_s\text{-tot}$ | sum of E_s -constants of all o-substituents | empirical |
| SimS | shape similarity vs molecule 3 | TSAR | $\mu\text{-CO}$ | bond order on amidic carbonyl group | DFT |
| Rtop | Randic topological index | TSAR | $\mu\text{-NH}$ | bond order on amidic -NH bond | DFT |
| Btop | Balaban topological index | TSAR | $\mu\text{-CAr}$ | bond order on bond between amidic group and NH ring | DFT |
| Wtop | Wiener topological index | TSAR | $\mu\text{-NAr}$ | bond order on bond between amidic group and CO ring | DFT |
| FI | molecular flexibility | TSAR | $\mu\text{-NC}$ | bond order on amidic OC-NH bond | DFT |
| qY | partial charge on atom Y | DFT | $\Sigma\mu$ | sum of bond orders of all bonds on amidic group | DFT |
| qX | partial charge on atom X | DFT | $\Sigma\mu_{\text{Ar}}$ | sum of bond orders on bonds between amidic functional group and both aromatic moieties | DFT |
| qH | partial charge on amidic hydrogen | DFT | | | |

($R^2 = 0.41$). The predictive power of **MODEL I**, quantified by cross-validated explained variance, was 25% ($Q^2 = 0.25$). The graph of correlation between the experimentally observed and predicted values of association constants is shown in Figure 2.

Two outliers, i.e., molecules **21** and **22**, strongly affected the model predictivity. The association constants of the compounds **21** and **22** are of extremely high value ($\sim 5000 \text{ M}^{-1}$) compared to the rest of the compounds in the analysis. In further analyses we decided to (i) construct a new model (**MODEL II**) only for compounds showing association constants lower than 3000 M^{-1} and (ii) employ the additional descriptors for the modeling of the data altogether (**MODEL III**).

3.2. MODEL II. The new matrix for PLS analysis was restricted in terms of the number of compounds (amides **21** and **22** have been eliminated) but used the same number of descriptors as in **MODEL I**. A two component PLS model with 32 descriptors was established which was characterized by $R^2 = 0.89$ and $Q^2 = 0.52$. There are a number of variables containing little or no information about amide ability for complexation of the anion. Variable selection based on VIP (variable importance in the projection) values and the position of descriptors on the plots of PLS variable weights resulted in a **MODEL II** with improved statistical quality, $R^2 = 0.91$ and $Q^2 = 0.75$. This model is based on eight descriptors. Derived PLS model can be expressed by the regression equation as follows:

$$\log K_{\text{assoc}} = 0.1396MV + 0.0514Sim - 0.3462H + 0.4348Q + 0.2987Tor + 0.5535D(XY) + 0.4963\sigma_{m,p}Tot + 0.3877E_sNH + 5.4251$$

Experimentally observed association constants are plotted against the predicted values of association constants in Figure 3. PLS variable weights are shown in Figure 4. Ten fixed permutations were used for the model validation. The validation plot for **MODEL II** is shown in Figure 5.

3.3. MODEL III. **MODEL II** provides good predictions for the complexation ability of amides that interact with the anion with association constants from 0 to about 3000 M^{-1} . However, the most stable complexes can be the most interesting ones, thus an attempt has been made to find a model for the prediction of the complexation abilities of all 22 amides. Bond orders $\mu\text{-CAr}$, $\mu\text{-NAr}$, $\mu\text{-NH}$, $\mu\text{-CO}$, and their combinations $\Sigma\mu_{\text{Ar}}$, $\Sigma\mu$ for selected bonds were used as new descriptors to characterize the distribution of electron density on the amidic group.

The matrix for PLS analysis was extended in terms of these new descriptors and included 22 amides as the training set. The variable selection resulted in **MODEL III** with $R^2 = 0.92$ and $Q^2 = 0.70$. This model is based on nine descriptors. Obtained PLS model can be expressed by the regression equation

$$\log K_{\text{assoc}} = 0.1107MV - 0.0096Sim - 0.3096H + 0.3546Q + 0.2446Tor + 0.4597D(XY) + 0.4772\sigma_{m,p}Tot + 0.3096E_sNH - 0.2344\Sigma\mu + 5.1047$$

The correlation plot of experimentally observed and predicted association constants is shown in Figure 6. The PLS variable weights (Figure 7) were calculated for important descriptors. **MODEL III** was validated with 10 fixed permutations, and the results are shown in Figure 8.

Table 3: Molecular Descriptors — Values

| D | MM | MS [Å ²] | MV [Å ³] | IM1s | IM2s | IM3s | IM1l | IM2l | IM3l | EV [Å ³] | FI | Rtop | Btop |
|----|--------|----------------------|----------------------|--------|--------|--------|-------|------|------|----------------------|------|-------|------|
| 1 | 305.19 | 239.91 | 189.94 | 129.20 | 555.11 | 655.28 | 12.03 | 2.80 | 2.37 | 335.06 | 3.99 | 11.93 | 2.05 |
| 2 | 215.24 | 212.03 | 144.49 | 55.35 | 326.92 | 380.27 | 14.60 | 2.47 | 2.13 | 321.20 | 3.08 | 11.93 | 2.05 |
| 3 | 197.25 | 205.65 | 137.62 | 41.81 | 316.80 | 357.01 | 20.44 | 2.70 | 2.39 | 552.44 | 2.90 | 11.93 | 2.05 |
| 4 | 215.24 | 218.07 | 150.31 | 61.48 | 365.54 | 420.05 | 16.17 | 2.72 | 2.37 | 436.33 | 3.08 | 11.93 | 2.05 |
| 5 | 198.24 | 211.19 | 137.14 | 41.87 | 315.53 | 354.89 | 18.22 | 2.42 | 2.15 | 396.60 | 2.85 | 11.50 | 2.11 |
| 6 | 215.24 | 213.64 | 144.91 | 41.79 | 418.38 | 458.54 | 26.89 | 2.67 | 2.45 | 741.49 | 3.08 | 11.93 | 2.05 |
| 7 | 215.24 | 214.36 | 147.46 | 56.34 | 382.16 | 434.59 | 19.84 | 2.93 | 2.57 | 625.48 | 3.08 | 11.93 | 2.05 |
| 8 | 287.20 | 242.50 | 186.52 | 110.25 | 526.68 | 600.21 | 14.71 | 3.08 | 2.70 | 512.99 | 3.80 | 11.93 | 2.05 |
| 9 | 215.24 | 212.33 | 148.95 | 44.11 | 414.48 | 453.58 | 23.77 | 2.53 | 2.31 | 582.48 | 3.08 | 11.93 | 2.05 |
| 10 | 198.24 | 208.47 | 138.45 | 42.00 | 311.25 | 349.11 | 20.07 | 2.71 | 2.41 | 549.99 | 2.85 | 11.50 | 2.18 |
| 11 | 233.23 | 220.31 | 154.15 | 79.17 | 433.78 | 507.50 | 14.63 | 2.67 | 2.28 | 373.48 | 3.25 | 11.93 | 2.05 |
| 12 | 233.23 | 223.25 | 151.90 | 43.65 | 534.62 | 576.78 | 30.15 | 2.46 | 2.28 | 709.08 | 3.25 | 11.93 | 2.05 |
| 13 | 233.23 | 223.94 | 153.95 | 60.20 | 491.22 | 547.85 | 21.25 | 2.65 | 2.34 | 541.67 | 3.25 | 11.93 | 2.05 |
| 14 | 287.20 | 240.13 | 185.01 | 111.94 | 524.60 | 609.21 | 13.50 | 2.88 | 2.48 | 404.30 | 3.80 | 11.93 | 2.05 |
| 15 | 198.24 | 210.02 | 139.81 | 41.62 | 319.35 | 355.71 | 20.08 | 2.62 | 2.35 | 517.58 | 2.85 | 11.50 | 2.22 |
| 16 | 288.19 | 243.60 | 182.89 | 112.18 | 525.14 | 607.19 | 12.81 | 2.74 | 2.37 | 347.87 | 3.75 | 11.50 | 2.20 |
| 17 | 305.19 | 243.60 | 192.24 | 112.32 | 673.77 | 757.68 | 16.39 | 2.73 | 2.43 | 455.56 | 3.99 | 11.93 | 2.05 |
| 18 | 305.19 | 238.66 | 194.74 | 129.32 | 633.84 | 712.20 | 13.78 | 2.81 | 2.50 | 406.43 | 3.99 | 11.93 | 2.05 |
| 19 | 288.19 | 236.63 | 185.30 | 112.06 | 530.82 | 605.38 | 13.56 | 2.86 | 2.51 | 407.98 | 3.75 | 11.50 | 2.12 |
| 20 | 290.21 | 237.07 | 186.62 | 114.28 | 530.81 | 602.03 | 13.90 | 2.99 | 2.64 | 460.01 | 3.96 | 12.26 | 2.09 |
| 21 | 323.18 | 253.94 | 203.13 | 147.11 | 724.70 | 806.60 | 14.14 | 2.87 | 2.58 | 438.48 | 4.18 | 11.93 | 2.05 |
| 22 | 234.22 | 221.02 | 153.76 | 75.46 | 432.07 | 501.04 | 15.80 | 2.76 | 2.38 | 434.79 | 3.20 | 11.50 | 2.02 |

| D | Wtop | Sim | SimQ | SimS | qX [au] | qH [au] | qY [au] | Q [au] | D(XH) [Å] | D(XY) [Å] | Dsum [Å] | Tor [°] | Tdip [Db] |
|----|------|--------|--------|--------|---------|---------|---------|---------|-----------|-----------|----------|---------|-----------|
| 1 | 1519 | 0.4876 | 0.1170 | 0.8583 | -0.1453 | 0.3748 | -0.2027 | 0.0268 | 2.27 | 3.92 | 8.41 | 41.00 | 2.6879 |
| 2 | 1519 | 0.9235 | 0.8795 | 0.9675 | -0.2313 | 0.3651 | 0.1708 | 0.3046 | 1.95 | 3.74 | 7.96 | 11.59 | 2.8896 |
| 3 | 1519 | 1.0000 | 1.0000 | 1.0000 | 0.1287 | 0.2967 | 0.1774 | 0.6028 | 1.99 | 3.82 | 8.07 | 16.20 | 3.4358 |
| 4 | 1519 | 0.8610 | 0.7954 | 0.9267 | 0.1714 | 0.3100 | 0.1748 | 0.6562 | 2.12 | 3.98 | 8.37 | 25.06 | 1.9716 |
| 5 | 1380 | 0.8406 | 0.7140 | 0.9673 | 0.0242 | 0.3174 | 0.1819 | 0.5235 | 2.05 | 3.87 | 8.19 | 21.59 | 1.4384 |
| 6 | 1519 | 0.9384 | 0.8973 | 0.9795 | 0.1265 | 0.2814 | 0.1914 | 0.5993 | 1.99 | 3.84 | 8.11 | 16.35 | 4.4058 |
| 7 | 1519 | 0.8959 | 0.8541 | 0.9377 | 0.1239 | 0.2879 | 0.2352 | 0.6470 | 2.07 | 3.91 | 8.28 | 23.92 | 2.9597 |
| 8 | 1519 | 0.5924 | 0.3692 | 0.8156 | 0.1250 | 0.2612 | -0.1380 | 0.2482 | 2.16 | 4.57 | 9.24 | 23.25 | 4.9407 |
| 9 | 1519 | 0.8879 | 0.8409 | 0.9350 | 0.1316 | 0.3111 | 0.1749 | 0.6176 | 2.11 | 3.97 | 8.35 | 25.02 | 2.9063 |
| 10 | 1394 | 0.8144 | 0.6715 | 0.9573 | 0.1257 | 0.3094 | -0.6641 | -0.2290 | 2.11 | 4.22 | 8.66 | 25.87 | 1.6131 |
| 11 | 1519 | 0.8613 | 0.7466 | 0.9760 | 0.1742 | 0.3018 | 0.2228 | 0.6988 | 2.07 | 3.92 | 8.29 | 21.78 | 1.6787 |
| 12 | 1519 | 0.8974 | 0.8239 | 0.9708 | 0.1362 | 0.2863 | 0.1903 | 0.6128 | 2.00 | 3.85 | 8.12 | 15.78 | 3.1549 |
| 13 | 1519 | 0.8928 | 0.8016 | 0.9840 | 0.1337 | 0.3008 | 0.2211 | 0.6556 | 2.05 | 3.91 | 8.26 | 20.79 | 1.5434 |
| 14 | 1519 | 0.6317 | 0.3966 | 0.8668 | -0.1494 | 0.3696 | 0.1592 | 0.3794 | 2.38 | 4.21 | 8.87 | 46.91 | 3.3795 |
| 15 | 1366 | 0.8370 | 0.7239 | 0.9501 | 0.1721 | 0.3012 | 0.1806 | 0.6539 | 2.19 | 4.00 | 8.46 | 31.20 | 3.2446 |
| 16 | 1380 | 0.5571 | 0.2434 | 0.8707 | -0.1511 | 0.3787 | 0.0565 | 0.2841 | 2.48 | 4.27 | 9.04 | 51.98 | 0.9025 |
| 17 | 1519 | 0.6300 | 0.3581 | 0.9018 | -0.1548 | 0.3725 | 0.1712 | 0.3889 | 2.41 | 4.22 | 8.91 | 48.79 | 2.865 |
| 18 | 1519 | 0.5469 | 0.2568 | 0.8371 | -0.1528 | 0.3722 | 0.2079 | 0.4273 | 2.54 | 4.37 | 9.22 | 54.69 | 1.5954 |
| 19 | 1366 | 0.5529 | 0.2475 | 0.8582 | -0.1484 | 0.4153 | 0.2146 | 0.4815 | 2.67 | 4.49 | 9.49 | 51.70 | 3.0044 |
| 20 | 1660 | 0.6150 | 0.3846 | 0.8454 | -0.1448 | 0.3801 | -0.6713 | -0.4360 | 2.76 | 4.68 | 9.78 | 61.81 | 3.0761 |
| 21 | 1519 | 0.5686 | 0.2768 | 0.8605 | -0.1547 | 0.3698 | 0.2250 | 0.4401 | 2.65 | 4.43 | 9.39 | 60.05 | 2.8188 |
| 22 | 1366 | 0.7733 | 0.5909 | 0.9557 | 0.1734 | 0.2867 | 0.2399 | 0.7000 | 2.18 | 4.04 | 8.53 | 30.37 | 2.6273 |

| D | σ -CO | σ -NH | σ -tot | E _s -CO | E _s -NH | E _s -tot | μ -NH | μ -C _{Ar} | μ -N _{Ar} | μ -CN | μ -CO | $\Sigma\mu$ | $\Sigma\mu_{Ar}$ |
|----|--------------|--------------|---------------|--------------------|--------------------|---------------------|-----------|------------------------|------------------------|-----------|-----------|-------------|------------------|
| 1 | 0.46 | 0 | 0.46 | -0.92 | -0.46 | -0.14 | 0.9245 | 1.0052 | 0.9327 | 1.2080 | 1.9361 | 6.0065 | 1.9379 |
| 2 | 0 | 0 | 0 | -0.46 | 0 | -0.46 | 0.9258 | 1.0516 | 0.9729 | 1.1848 | 1.8334 | 5.9683 | 2.0244 |
| 3 | 0 | 0 | 0 | 0 | 0 | 0 | 0.9585 | 1.0723 | 0.9726 | 1.1675 | 1.8585 | 6.0293 | 2.0449 |
| 4 | 0.06 | 0 | 0.06 | 0 | 0 | 0 | 0.9551 | 1.0468 | 0.9584 | 1.1787 | 1.8611 | 6.0000 | 2.0052 |
| 5 | 0.73 | 0 | 0.73 | 0 | 0 | 0 | 0.9537 | 1.0711 | 0.9574 | 1.1804 | 1.8773 | 6.0398 | 2.0285 |
| 6 | 0 | 0.34 | 0.34 | 0 | 0 | 0 | 0.9498 | 1.0802 | 0.9616 | 1.1559 | 1.8598 | 6.0074 | 2.0418 |
| 7 | 0 | 0.06 | 0.06 | 0 | 0 | 0 | 0.9457 | 1.0793 | 0.9412 | 1.1606 | 1.8707 | 5.9974 | 2.0205 |
| 8 | 0 | 0.46 | 0.46 | 0 | -0.92 | -0.92 | 0.9609 | 1.0751 | 0.8961 | 1.1119 | 1.9220 | 5.9661 | 1.9712 |
| 9 | 0.34 | 0 | 0.34 | 0 | 0 | 0 | 0.9573 | 1.0773 | 0.9610 | 1.1746 | 1.8600 | 6.0301 | 2.0383 |
| 10 | 0 | 0.90 | 0.9 | 0 | 0 | 0 | 0.9272 | 1.0838 | 0.9542 | 1.1737 | 1.8521 | 5.9910 | 2.0380 |
| 11 | 0.06 | 0.06 | 0.12 | 0 | 0 | 0 | 0.9564 | 1.0425 | 0.9588 | 1.1537 | 1.8675 | 5.9788 | 2.0013 |
| 12 | 0.34 | 0.34 | 0.68 | 0 | 0 | 0 | 0.9514 | 1.0903 | 0.9568 | 1.1540 | 1.8379 | 5.9903 | 2.0471 |
| 13 | 0.34 | 0.06 | 0.40 | 0 | 0 | 0 | 0.9560 | 1.0747 | 0.9626 | 1.1462 | 1.8692 | 6.0087 | 2.0373 |
| 14 | 0.46 | 0 | 0.46 | -0.92 | 0 | -0.92 | 0.9543 | 0.9830 | 0.9800 | 1.2228 | 1.9623 | 6.1024 | 1.9630 |
| 15 | 0.83 | 0 | 0.83 | 0 | 0 | 0 | 0.9442 | 1.0468 | 0.9350 | 1.1863 | 1.8909 | 6.0031 | 1.9818 |
| 16 | 0.46 | 0.73 | 1.19 | -0.92 | 0 | -0.92 | 0.9504 | 0.9889 | 0.9884 | 1.2153 | 1.9775 | 6.1204 | 1.9772 |
| 17 | 0.46 | 0.34 | 0.80 | -0.92 | 0 | -0.92 | 0.9434 | 0.9918 | 0.9498 | 1.2108 | 1.9497 | 6.0454 | 1.9416 |
| 18 | 0.46 | 0.06 | 0.52 | -0.92 | 0 | -0.92 | 0.9381 | 0.9867 | 0.9570 | 1.2040 | 1.9526 | 6.0383 | 1.9437 |
| 19 | 0.46 | 0.83 | 1.29 | -0.92 | 0 | -0.92 | 0.9514 | 1.0033 | 0.9902 | 1.1957 | 1.9902 | 6.1308 | 1.9935 |
| 20 | 0.46 | 0.90 | 1.36 | -0.92 | 0 | -0.92 | 0.9253 | 0.9854 | 0.9558 | 1.2231 | 1.9385 | 6.0281 | 6.0281 |
| 21 | 0.46 | 0.12 | 0.58 | -0.92 | 0 | -0.92 | 0.9395 | 1.0022 | 0.9389 | 1.2024 | 1.9407 | 5.0842 | 1.9410 |
| 22 | 0.83 | 0.12 | 0.95 | 0 | 0 | 0 | 0.9484 | 1.0292 | 0.9253 | 1.1727 | 1.8991 | 5.0263 | 1.9545 |

4. DISCUSSION

The concept of quantitative structure—property relationships (QSPR) is to transform searches for compounds with desired properties using chemical intuition and experience into a mathematically quantified and computerized form.¹⁵ QSPR requires a set of descriptors, which gives information about the investigated properties. Such information may not

only lead to a better understanding of the system but may also allow for transferring information and predictions from one series of compounds to another.²⁰ This work presents a description and prediction of the stability of 1:1 complexes of the aromatic amides with bromide anion focused on the prediction of their stability. Theoretical chemistry tools and descriptors are employed in the constructed QSPR models.

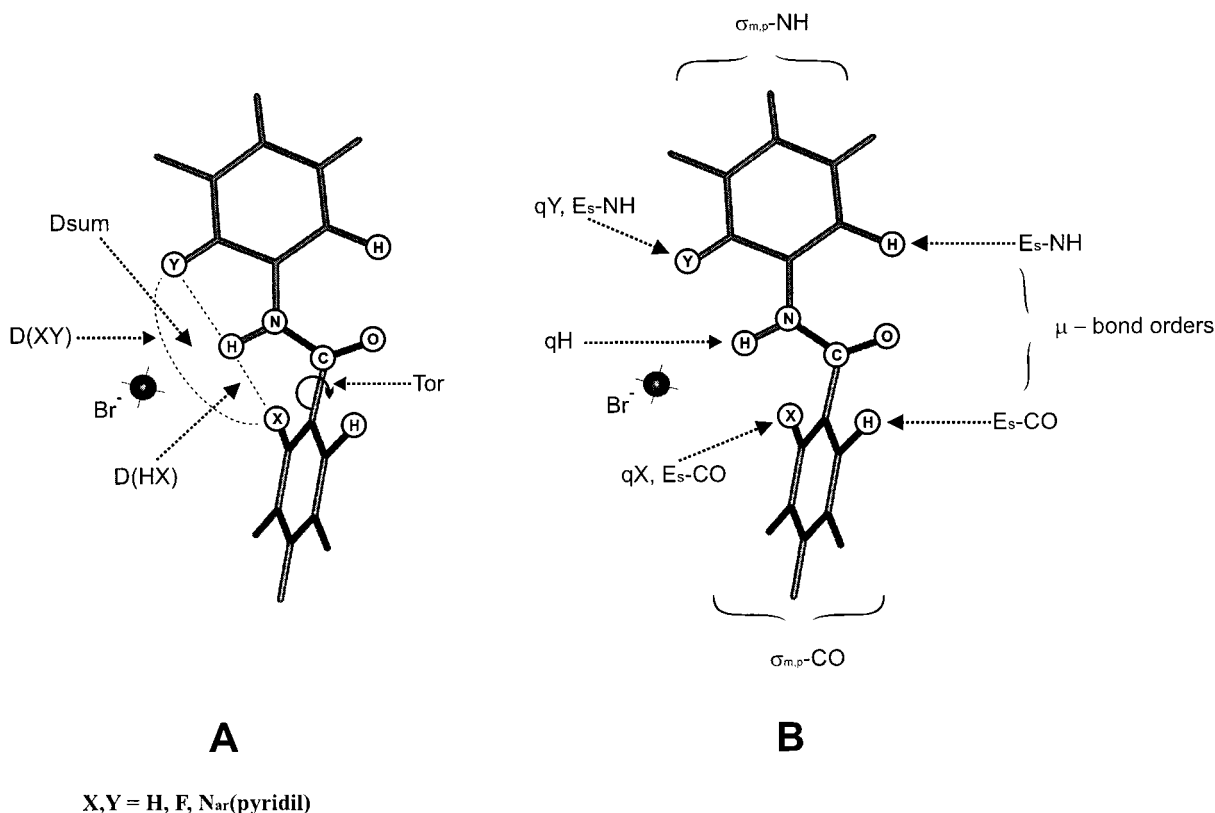


Figure 1. Graphical interpretation of selected descriptors: (A) steric descriptors of geometrical properties of the amide binding site and (B) electronic descriptors of the amide binding site and empirical descriptors of the substituents on both aromatic rings.

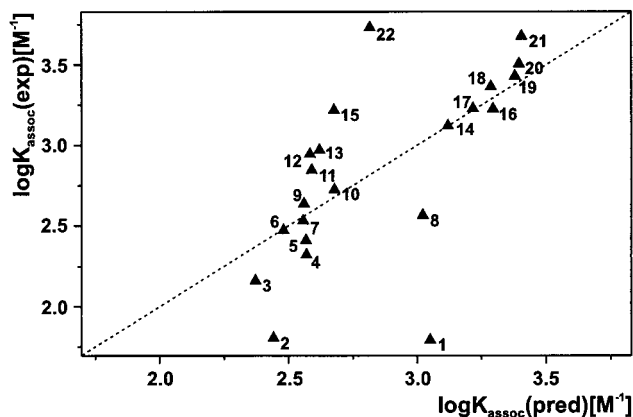


Figure 2. Plot of experimentally observed versus predicted association constants. **MODEL I** based on 32 descriptors was used for predictions (one component, $n = 22$, $R^2 = 0.41$, $Q^2 = 0.25$).

The computing of interaction energies or even association constants for such complexes, using molecular modeling or standard quantum chemistry methods, is a considerable problem. The QSPR approach can provide results of comparable quality but significantly lower computational demand. The objective of this study was to find the structural and physicochemical parameters of aromatic amides related to their ability to form a complex with bromide anion and to construct a QSPR model for the prediction of association constants.

Overall 39 descriptors were applied in QSPR. These descriptors were designed according to the knowledge of the structure and properties of the complexes.^{5,6,21} Three models were developed for the description of association constants for [amide] Br^- complexes. The initial model, **MODEL I**, served for the characterization and classification of the input

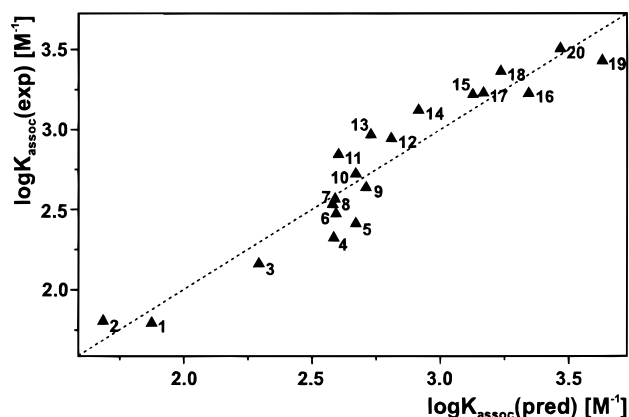


Figure 3. Plot of experimentally observed versus predicted association constants. **MODEL II** based on eight descriptors was used for predictions (two components, $n = 20$, $R^2 = 0.91$, $Q^2 = 0.75$).

data. This model, characterized by $R^2 = 0.41$ and $Q^2 = 0.25$ for 22 compounds and 32 descriptors, was further refined and resulted in **MODELS II** and **III**.

The later two models attempted to eliminate redundant descriptors and to deal with the outlying compounds **21** and **22** (Figure 2). **MODEL II** eliminated the outliers from the training set, while new descriptors have been included to **MODEL III** in order to keep all compounds in the model. The descriptors used in **MODEL I** and **MODEL II** primarily described global properties of the anion binding site and the properties of the amidic hydrogen. Amides **21** and **22** have similar structural properties to other molecules but show very high association constants. Extreme complexation of these compounds with bromide anions is probably related to the electronic properties of the aromatic rings. The predictive

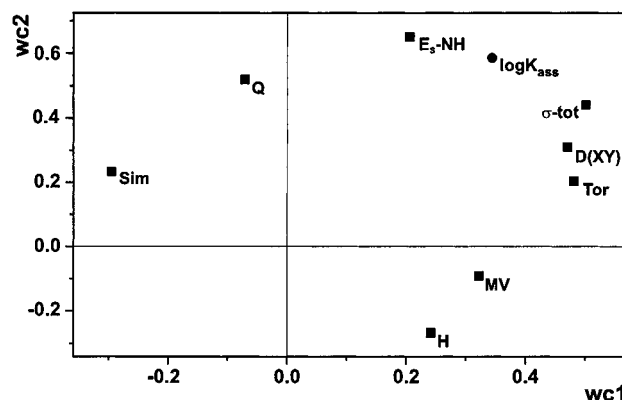


Figure 4. Loading plot of the first two components from PLS analysis applied to descriptors fixed in **MODEL II**.

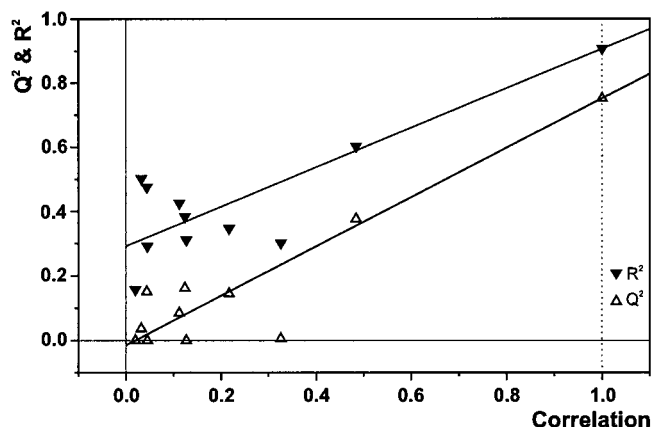


Figure 5. Validation plot of **MODEL II**. Ten fixed permutations were used for validation. The Y-axis represents R^2 (filled triangles) and Q^2 (empty triangles) for the original model (right-hand part of the plot) and every model with permuted y . X-axis designates the correlation coefficient between original and permuted y . Good models have the R^2 -intercept below 0.4 and Q^2 -intercept below 0.05.

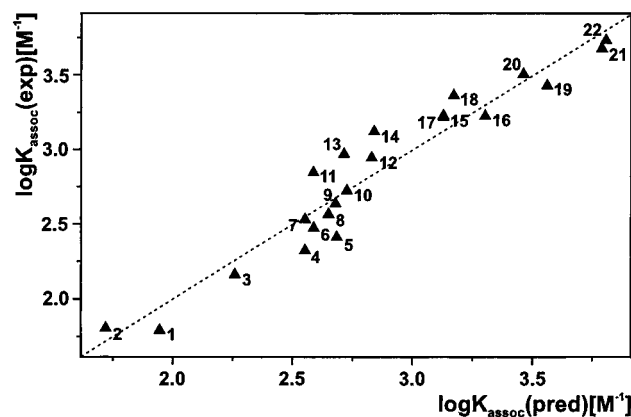


Figure 6. Plot of experimentally observed versus predicted association constants. **MODEL III** based on 10 descriptors was used for prediction (two components, $n = 22$, $R^2 = 0.92$, $Q^2 = 0.70$).

power of **MODEL II** is characterized by $R^2 = 0.91$ and $Q^2 = 0.75$, for 20 compounds and eight descriptors. This model provided good predictions for compounds with association constants below 3000 M^{-1} . **MODEL III** used the same descriptors as **MODEL II** and employed one additional descriptor to characterize the electronic properties of the central amidic group. **MODEL III** showed $R^2 = 0.92$ and $Q^2 = 0.70$ (22 compounds, nine descriptors).

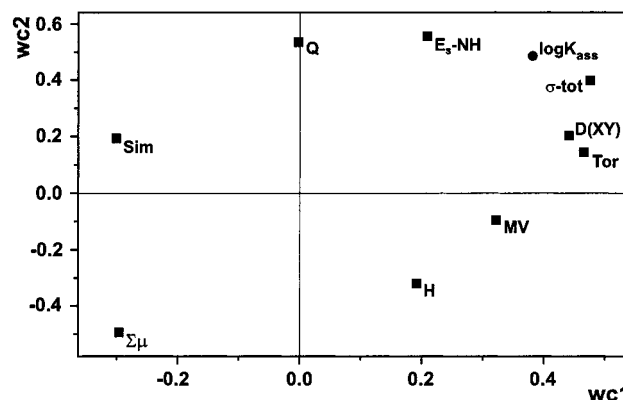


Figure 7. Loading plot of first two components from PLS analysis applied to descriptors fixed in **MODEL III**.

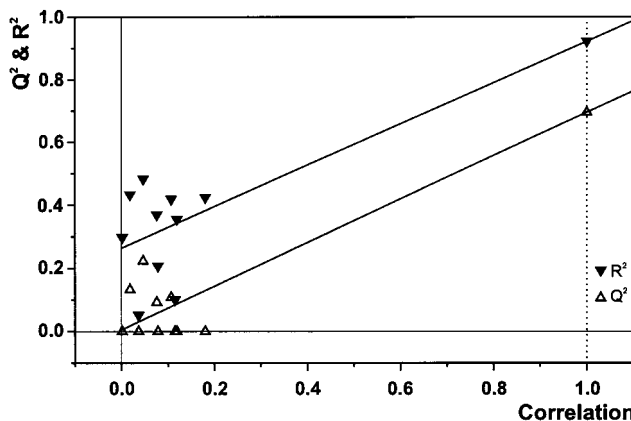


Figure 8. Validation plot of **MODEL III**. Ten fixed permutations were used for validation. See Figure 5 for explanation of plot interpretation.

On the basis of previous theoretical studies,^{6,22} and also supported by X-ray analysis of hydrogen bonding of anions in amides,²¹ we propose that the conformational behavior of the amide molecule is one of the most important parameters influencing the stability of the complexes. A planar conformation is not a global minimum on the conformational energy hypersurface, and the total π -electron system conjugation is slightly disrupted by repulsion between substituents in *ortho* positions of the aromatic rings on one side and atoms of the amidic group on the other side. On the other hand, this repulsion opens the space around the amidic hydrogen and allows the anion to interact with the amide molecule. The steric "opening" effect predominates over the activation of the amidic hydrogen by electron-acceptor properties of substituents on aromatic moieties, which are the most significant for the compounds with fluorine atoms in *ortho* positions on the CO ring. The above observation is expressed by the **Tor** and **D(XY)** descriptors. The former includes information about conformation around the bond between the amidic carbonyl and the aromatic ring. The latter one gives information about the conformational behavior of both the aromatic moieties simultaneously.

The substituents in *ortho* positions of both aromatic units may also influence the interaction with the anion by electrostatic forces. This is confirmed by the fact that the sum of partial charges Q on the atoms around the binding site is one of the significant electronic descriptors. The amidic hydrogen creates a hydrogen bond with the anion and the atoms in X and Y positions (Figure 1) have a complementary

effect on the interaction with the anion. When the fluorine atom is bonded in position X or Y, the intramolecular hydrogen bond between this atom and the amidic hydrogen is preferred to the intermolecular hydrogen bond with the bromide anion. On the other hand, the fluorine atom in either *ortho* position of both the aromatic moieties may interact with the carbonyl oxygen and this repulsion leads to a conformational movement between the planes of the amidic group and the aromatic rings. The acidity of the amidic hydrogen and its ability to interact with the anion is expressed by the partial charge **qH**.

The aromatic amides contain a large conjugated π -electron system. Both the aromatic moieties of the studied amides were substituted in various positions of the rings. The empirical descriptors based on the Hammett (σ) and Taft (E_s) constants were used for the characterization of the electronic and steric effects of these substituents. The ability of the amidic hydrogen to interact with the anion may be influenced by the electron-acceptor properties of substituents on aromatic rings through out the π -electron conjugation. While substituents in *meta* and *para* positions only interact with the amidic hydrogen through the π -electron system, substituents in *ortho* positions also interact by steric and Coulombic effects. The influence of *meta* and *para* substituents is expressed by the σ -**tot** descriptor. The properties of substituents in *ortho* positions on both aromatic rings are described by the Taft E_s constants. The descriptor **E_s-NH** represents the sum of Taft E_s constants on NH connected ring. The **E_s** descriptors of *ortho* substituents in combination with the conformational behavior of the molecule on the amidic group partially include the steric and Coulombic properties of the anion binding site. The **E_s-NH** descriptor is an indicator variable only for compounds **1** and **8**, but its elimination from the data set results in worse statistics for both the final models (MODEL II: $R^2 = 0.82$, $Q^2 = 0.50$, 20 compounds, seven descriptors; MODEL III: $R^2 = 0.80$, $Q^2 = 0.52$, 22 compounds, eight descriptors). There is no other descriptor in the data set that can describe the substituents in *ortho* positions on NH connected aromatic ring. An additional descriptor $\Sigma\mu$ was incorporated in **MODEL III** for a better description of the electronic properties of the amide molecule. This descriptor quantifies the sum of bond orders on the central amidic group. Such parameters complement information about the distribution of electronic density in the molecule provided by the empirical descriptor σ -**tot** and partially by the descriptors **E_s-NH**. Global steric descriptors, the molecular volume **MV** and the similarity index **Sim**, complement the set of steric and electronic descriptors.

The quality of QSPR models could be compared with our previous results from the DFT quantum chemical calculations of the interaction energies applied to the identical system.⁶ The predictions based on the interaction energies are substantially worse than the predictions generated by the QSPR models. Moreover, the QSPR modeling required significantly less computational resources compared to the DFT and ab initio quantum chemical calculations. Developed

QSPR models will be used for predictions of association constants of other complexes, which are not included in the tested set.

ACKNOWLEDGMENT

The authors thank the Academic Supercomputer Center in Brno for providing them with access to computer facilities. The authors are grateful to Mr. Laurence Benjamin, B. A., for language corrections. The research has partially been supported by the Grant Agency of the Czech Republic under Grant 203/00/1011.

REFERENCES AND NOTES

- (1) *Supramolecular Chemistry of Anions*; Bianchi, A., Bowman-James, K., García-España, E., Eds.; Wiley-VCH: New York, 1997.
- (2) Schmidtchen, F. P.; Berger, M. Artificial Organic Host Molecules for Anions. *Chem. Rev.* **1997**, *97*, 1609–1646.
- (3) Frausto da Silva, J. J. R.; Williams, R. J. P. The Uptake of Elements by Biological Systems. *Struct. Bonding (Berlin)* **1976**, *29*, 67–121.
- (4) Luecke, H.; Quiñocho, F. A. High Specificity of a Phosphate Transport Protein Determined by Hydrogen Bonds. *Nature* **1990**, *347*, 402–406.
- (5) Štibor, I.; Haffed, D. S. M.; Lhoták, P.; Hodačová, J.; Koča, J.; Čajan, M. From the Amide Bond Activation to Simultaneous Recognition of Anion–Cation Couple. *Gazz. Chim. Ital.* **1997**, *127*, 673–685.
- (6) Čajan, M.; Štibor, I.; Koča, J. Computational Studies on the [Amide]Br[−]. *J. Phys. Chem. A* **1999**, *103*, 3778–3783.
- (7) Katritzky, A. R.; Lobanov, V. S.; Karelson, M. QSPR: The Correlation and Quantitative Prediction of Chemical and Physical Properties from Structure. *Chem. Soc. Rev.* **1995**, *24*, 279–287.
- (8) Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti Correlation-Energy Formula into a Functional of the Electron Density. *Phys. Rev. B* **1988**, *37*, 785–789.
- (9) Godbout, N.; Salahub, D. R.; Andzelm, J.; Wimmer, E. Optimization of Gaussian-type Basis Set for Local Spin Density Functional Calculations. I. Boron through Neon, Optimization Technique and Validation. *Can. J. Chem.* **1992**, *70*, 560–571.
- (10) Mulliken, R. S. J. Electronic Population Analysis on LCAO-MO Molecular Wave Functions I. – IV. *Chem. Phys.* **1955**, *23*, 1833–1840, 1841–1846, 2338–2342, 2343–2346.
- (11) Löwdin, P. Q. On the Orthogonality Problem. *Adv. Quantum Chem.* **1970**, *5*, 185–199.
- (12) Cusachs, L. C.; Politzer, P. On the Problem of Defining the Charge on an Atom in a Molecule. *Chem. Phys. Lett.* **1968**, *1*, 529–531.
- (13) Roby, K. R. Quantum Theory of Chemical Valence Concepts. I. Definition of the Charge on an Atom in a Molecule and of Occupation Numbers for Electron Density Shared between Atoms. *Mol. Phys.* **1974**, *27*, 81–104.
- (14) Davidson, E. R. Electronic Population Analysis of Molecular Wave Functions. *J. Chem. Phys.* **1967**, *46*, 3320–3324.
- (15) Karelson, M.; Lobanov, V. S.; Katritzky, A. R. Quantum-Chemical Descriptors in QSAR/QSPR Studies. *Chem. Rev.* **1996**, *96*, 1027–1043.
- (16) Mayer, I. Charge, Bond Order and Valence in the Ab Initio SCF Theory. *Chem. Phys. Lett.* **1983**, *97*, 270–274.
- (17) Malkin, V. G.; Malkina, I. L.; Salahub, D. R. deMon/1.0, A Gaussian Density Functional Program; Université de Montréal: Montréal, 1994.
- (18) Hansch, C.; Leo, A.; Hoepman D. *Exploring QSAR: Hydrophobic, Electronic, and Steric Constants*; American Chemical Society: Washington, DC, 1995.
- (19) Geladi, P.; Kowalski, B. R. Partial Least-Squares Regressions: A Tutorial. *Anal. Chim. Acta* **1986**, *185*, 1–17.
- (20) Livingstone, D. *Data Analysis for Chemists*; Oxford University Press: Oxford, 1995.
- (21) Kavallieratos, K.; Bertao, C. M.; Crabtree, R. H. Hydrogen Bonding in Anion Recognition: A Family of Versatile, Nonpreorganized Neutral and Acyclic Receptors. *J. Org. Chem.* **1999**, *64*, 1675–1683.
- (22) McDonald, N. A.; Duffy, E. M.; Jorgensen, W. L. Monte Carlo Investigations of Selective Anion Complexation by bis(phenylurea)-p-tert-butylcalix[4]arene. *J. Am. Chem. Soc.* **1998**, *120*, 5104–5111.

CI000382S