Prediction of the Intrinsic Hydrogen Bond Acceptor Strength of Chemical Substances from Molecular Structure

Johannes Schwöbel,[†,‡] Ralf-Uwe Ebert,[†] Ralph Kühne,[†] and Gerrit Schüürmann*[,†,‡]

UFZ Department of Ecological Chemistry, Helmholtz Centre for Environmental Research, Permoserstrasse 15, 04318 Leipzig, Germany, and Institute for Organic Chemistry, Technical University Bergakademie Freiberg, Leipziger Strasse 29, 09596 Freiberg, Germany

Received: May 22, 2009; Revised Manuscript Received: July 4, 2009

Hydrogen bonding affects the partitioning of organic compounds between environmental and biological compartments as well as the three-dimensional shape of macromolecules. Using the semiempirical quantum chemical AM1 level of calculation, we have developed a model to predict the site-specific hydrogen bond (HB) acceptor strength from ground-state properties of the individual compounds. At present, the model parametrization is confined to compounds with one HB acceptor site of the following atom types: N, O, S, F, Cl, and Br that act as lone-pair HB acceptors, and π-electron (aromatic or conjugated) systems with the associated C atoms as particularly weak HB acceptors. The HB acceptor strength is expressed in terms of the Abraham parameter B and calculated from local molecular parameters, taking into account electrostatic, polarizability, and charge transfer contributions according to the Morokuma concept. For a data set of 383 compounds, the squared correlation coefficient r^2 is 0.97 when electrostatic potential (ESP) derived net atomic charges are employed, and the root-mean-square (rms) error is 0.04 that is in the range of experimental uncertainty. The model is validated using an extended leave-50%-out approach, and its performance is comparatively analyzed with the ones of earlier introduced ab initio (HF/6-31G**) and density functional theory (B3LYP/6-31G**) models as well as of two increment methods with respect to the total compound set as well as HB acceptor type subsets. The discussion includes an explorative model application to amides and organophosphates that demonstrates the robustness of the approach, and further opportunities for model extensions.

Introduction

Hydrogen bonding has substantial impact on the molecular structure, chemistry, and biochemistry of chemical compounds.[1−3] Predictive quantification of the hydrogen bond (HB) strength allows one to modify and optimize drug-related properties of chemicals and to assess their partition profile with respect to abiotic and biotic compartments. In this context, the well-established LFER (linear free energy relationship) approach employing Abraham parameters provides a sound opportunity to predict respective liquid−liquid partitioning properties P (such as K_{ow}) of organic compounds.[4,5] Recent publications show the broad application spectrum of the Abraham model, for example, to predict blood-brain barriers or the distribution of solutes between water and micelles.[6,7] The Abraham-type LFER is often written in the following form:

$$\log P = vV + eE + sS + aA + bB + c \qquad (1)$$

In eq 1, P is the partition property of interest, the compound descriptors V, E, S, A, and B characterize various types of intermolecular interaction, and the coefficients v, e, s, a, b, and c encode partition-relevant properties of the specific solvent system. V is the McGowan characteristic volume, E quantifies the excess molar refraction of the compound as compared to an alkane of the same volume, and S is a descriptor supposed to cover polarity and polarizability. Parameter A characterizes the HB donor strength (HB acidity), while B quantifies the HB acceptor strength (HB basicity). This communication deals with the Abraham descriptor B. Further details of the other parameters of eq 1 are described elsewhere.[8,9]

Hydrogen bonding is a directional interaction between an HB donor and acceptor site. Its strength is usually between van der Waals and covalent interactions. Recently, Gilli and co-workers reported a new method to HB strengths by a pK_a slide rule with separate scales for HB donors and HB acceptors.[10] It demonstrates the general interest in methods to predict the HB donor and acceptor strengths of chemical compounds from molecular structure. Existing approaches from 3D structure follow the supermolecular approach, evaluating the formation energies of H-bond donor−acceptor complexes,[11] and there have been also reported applications of conceptual density functional theory.[12]

For the HB acceptor strength in the Abraham scale, B, a general fragment model has been developed that provides fast predictions based on 2D topological information.[8] An updated version of this method is available through the ADME boxes software, but the implemented Absolv module including the actual method details is proprietary and not published.[13] An independent implementation of the fragment method in its original form has been developed and analyzed with respect to its prediction performance.[14]

Recently, we have introduced a quantum chemical approach to predict the HB donor and acceptor strengths of organic compounds in terms of their Abraham parameters A and B,

* Corresponding author. Tel: +49-341-235-1262. Fax: +49-341-235-1785. E-mail: gerrit.schuurmann@ufz.de.
[†] UFZ Helmholtz Centre for Environmental Research.
[‡] Technical University Bergakademie Freiberg.

10.1021/jp904812b CCC: $40.75
Published on Web 08/20/2009

employing an ab initio level of computational chemistry for the compounds as single molecules.[15,16] For both A and B, r^2 values larger than 0.90 were achieved, and almost identical statistics for separate test sets demonstrated their prediction capability. The goal of the present study is to adapt the approach to the semiempirical AM1 method. The latter reduces vastly the computation time and enables one to screen even large-scale inventories of chemical compounds with respect to B and properties based on predicted B. This combines physical reliability of quantum chemical methods with computational rapidness, the latter of which is usually the advantage of fragment models.

Following the Morokuma analysis, electrostatic effects, polarization, and charge transfer form major components of hydrogen bonding.[17] Quantification of the Coulomb interaction potential of HB acceptor sites can be based on net atomic charges, keeping in mind that the latter are not physically observable and depend on the method used for their calculation. With regard to polarization and charge transfer contributions to the HB acceptor strength, local molecular parameters encoding the site-specific electron donor and acceptor capability were used. These latter parameters were originally developed to calculate gas phase hydroxyl radical rate constants.[18−20]

In the following, we introduce a new prediction model for the HB acceptor strength B that was calibrated for the semiempirical AM1 level of quantum chemistry. Its performance, application domain and limitations are compared with existing models covering both fragment schemes and more sophisticated ab initio calculations. Moreover, opportunities for possible model extensions are discussed. Finally, explorative applications of the new model to amides and organophosphates as compounds outside its original structural domain demonstrate its robustness.

Material and Methods

Data Set. From literature sources (see Supporting Information), 383 aliphatic and aromatic molecules with experimental values for the Abraham acceptor strength parameter B have been collected that all contain only one HB acceptor site. The compound set covers various compound classes, and contains oxygen, nitrogen, sulfur, and halogen (F, Cl, Br) atoms with appropriate lone pairs, and weak carbon π bonds as HB acceptor sites. For exploratory purposes, model statistics have been checked when including an additional set of 20 monofunctional compounds with B approximated through experimental β_2^H and pK_{HB} values[21,22] (keeping in mind that only monofunctional bases are taken into account) as outlined below. Note that B had originally been termed $\Sigma\beta_2^H$.[4,22] For a final explorative model application, 28 amides and organophosphates outside the original chemical domain were used. Note that these compounds were not included in the original training set, because they contain more than one possible HB acceptor site (which could lead to secondary effects, and hence a wrong calibration of the intrinsic, site-specific HB acceptor strength).

The chemical domain and experimental B value range are shown in Table 1. The classification according to weak, moderate, strong, and very strong HB acceptors is taken from a previous publication.[16] Interestingly, organophosphates have B values substantially larger than all other compounds of the present study; similar and even larger HB acceptor strengths are likely for anionic sites. However, electrostatic interactions involving ions are not included in the Abraham approach that forms the basis for the HB acceptor strength parameter B.

TABLE 1: Chemical Compound Classes and Associated Ranges of Experimental HB Acceptor Strengths[a]

compound class	HB acceptor site	n	B_{min}	B_{mean}	B_{max}
Weak HB Acceptor Strength					
halogens (aromatically bound)	F, Cl, Br	10	0.05	0.08	0.10
olefins	C (sp^2)	24	0.07	0.09	0.19
halogens (aliphatically bound)	F, Cl, Br	35	0.05	0.11	0.16
alkynes	C (sp)	12	0.04	0.12	0.21
O-heteroaromatics	O (sp^2)	6	0.13	0.14	0.14
S-heteroaromatics	S (sp^2)	2	0.15	0.16	0.16
thiophenols	S (sp^2)	1	0.16	0.16	0.16
aromatic hydrocarbons	C (sp^2)	29	0.14	0.17	0.21
thiols	S (sp^3)	10	0.12	0.21	0.24
Moderate HB Acceptor Strength					
anisoles	O (sp^3)	7	0.29	0.31	0.34
sulfides	S (sp^3)	6	0.27	0.32	0.37
nitriles	N (sp)	14	0.32	0.36	0.40
phenols	O (sp^2)	42	0.30	0.39	0.50
aldehydes	O (sp^2)	10	0.33	0.44	0.45
Strong HB Acceptor Strength					
anilines	N (sp^2-sp^3)	20	0.41	0.47	0.59
ethers	O (sp^3)	17	0.41	0.48	0.76
ketones	O (sp^2)	27	0.49	0.52	0.56
alcohols	O (sp^3)	58	0.47	0.53	0.60
N-heteroaromatics	N (sp^2)	21	0.25	0.56	0.70
Very Strong HB Acceptor Strength					
anilides[b]	O (sp^2)	11	0.52	0.65	0.73
amines	N (sp^3)	32	0.58	0.66	0.79
amides[b]	O (sp^2)	13	0.55	0.69	0.78
phosphates[b]	O (sp^2)	4	1.00	1.11	1.21

[a] B_{min} = minimum experimental B value, B_{mean} = mean of experimental B values, B_{max} = maximum experimental B value. [b] Not included in the calibration set, because of possible subordinate HB acceptor sites: amides and anilides contain one N (sp^2-sp^3) site in addition, and phosphates 3 O(sp^3).

Computational Details. The geometries of all molecules were optimized at the semiempirical AM1 level[23] by means of MOPAC2002.[24] Different charge analysis methods were applied, namely Coulson charge densities based on zero differential overlap (ZDO), Mulliken population analysis (MUL), and electrostatic potential derived charges (ESP), which are derived from the Coulomb potential at a large number of grid points.[25] For the latter, we applied the Besler−Merz−Kollmann scheme as implemented in MOPAC.[26] The results are compared to the ones obtained with ab initio level HF/6-31G** and DFT level B3LYP/6-31G**, the latter of which were calculated with Gaussian 03[27] and with NPA net atomic charges[28] for the electrostatic contribution to B as described earlier.[16]

For the calculation of effective electron donor and acceptor energies we used an in-house Fortran 77 program, which searches for possible HB acceptor sites and quantifies their intrinsic acceptor strengths. Nonlinear parameter calibration was performed by the Levenberg−Marquardt algorithm, where the Jacobian is then calculated by a forward-difference approximation. This algorithm is included in the freely available Fortran MINPACK library.[29]

Hydrogen Bond Acceptor Model. The approach follows the recently introduced ab initio model[16] and was recalibrated for the semiempirical AM1 level. For the quantification of the HB acceptor strength B, the following components are taken into account: Electrostatic interaction (ES), polarizability (PL), and charge transfer (CT) that represents a covalent contribution. The model equation to predict the B value at HB acceptor site s reads

$$B(s) = c_{ES} \cdot Q(s) + c_{PL} \cdot \eta(s) + c_{CT} \cdot EE_{occ}(s) + c \quad (2)$$

In eq 2, Q(s), $\eta(s)$, and $EE_{occ}(s)$ quantify the ES, PL, and CT contributions to $B(s)$ as outlined below, c_{ES}, c_{PL}, and c_{CT} denote the associated regression coefficients, and c is the intercept.

The Coulomb contribution to B is characterized through the net atomic charge $Q(s)$. As discussed earlier,[16] the dominating part of this ES component is assumed to be localized at HB acceptor site s, reflecting its readiness to undergo an attractive Coulomb interaction with the (not explicitly considered) HB donor site. For the AM1 calibration of eq 2, the following three charge calculation schemes were explored: Coulson charges that are based on zero differential overlap (ZDO), Mulliken population analysis (MUL), and electrostatic potential derived charges (ESP).

The PL contribution to B is evaluated through the local hardness at the HB acceptor site, $\eta(s)$, keeping in mind that increasing hardness corresponds to decreasing polarizability. The local hardness is defined through site-specific effective donor and acceptor energies, EE_{occ} and EE_{vac}:[16]

$$\eta(s) = -\frac{1}{2}(EE_{occ}(E_{ref},s) - EE_{vac}(E_{ref},s)) \quad (3)$$

EE_{occ} and EE_{vac} characterize the site-specific molecular readiness for donating or accepting electronic charge. EE_{occ} involves occupied MO energies E_i, and an associated reference energy E_{ref},

$$EE_{occ}(E_{ref},s) = \frac{\sum_{i=\mathrm{HOMO}}^{1} E_i \cdot w_i(E_{ref},s)}{\sum_{i=\mathrm{HOMO}}^{1} w_i(E_{ref},s)} \quad (4)$$

with

$$w_i(E_{ref},s) = p_i(s) \cdot \exp\left(-\frac{E_i}{E_{ref}}\right)$$

and EE_{vac} is defined correspondingly through unoccupied MO energies E_k with reference energy E_{ref} as associated scaling factor:

$$EE_{vac}(E_{ref},s) = \frac{\sum_{k=\mathrm{LUMO}}^{\max} E_k \cdot w_k(E_{ref},s)}{\sum_{k=\mathrm{LUMO}}^{\max} w_k(E_{ref},s)} \quad (5)$$

with

$$w_k(E_{ref},s) = p_k(s) \cdot \exp\left(-\frac{E_k}{E_{ref}}\right)$$

Reference energies E_{ref} are calibrated separately for EE_{occ} and EE_{vac}, applying nonlinear regression as outlined earlier.[16] In eq 4, p_i denotes the electron population of occupied molecular orbital (MO) i, and the corresponding term in eq 5 is p_k that characterizes the electron space available in unoccupied MO k. Introducing the LCAO-MO (LCAO = linear combination of atomic orbitals) coefficients $c_{\mu i}$ (that quantify the contribution of AO μ to occupied MO i) and $c_{\rho k}$ (that quantify the contribution of AO ρ to unoccupied MO k), p_i and p_k can be calculated at atomic site r as follows:

$$p_i(r) = 2 \cdot \sum_{\mu(r)} (c_{\mu i})^2 \qquad p_k(r) = 2 \cdot \sum_{\rho(r)} (c_{\rho k})^2 \quad (6)$$

If HB acceptor site s contains more than one atomic site r, the resultant average value is

$$p_i(s) = 2 \cdot \frac{\sum_r \sum_{\mu(r)} (c_{\mu i})^2}{n_A} \qquad p_k(s) = 2 \cdot \frac{\sum_r \sum_{\rho(r)} (c_{\rho k})^2}{n_A} \quad (7)$$

In eq 7, n_A denotes the number of atomic sites that build HB acceptor site s. Taking toluene as an example, its six π electrons are delocalized over six carbon atoms, yielding $n_A = 6$ for its π-electron moiety as weak HB acceptor site.

The third term of eq 2 represents the CT contribution to the HB acceptor strength B. Here, the effective donor energy EE_{occ} at HB acceptor site s is used to characterize its readiness for donating electron charge to the donor site upon hydrogen bonding. The covalent HB component is thus modeled according to the concept of electron donor−acceptor interaction of the orbitals involved in hydrogen bonding.

Statistical Performance. The calibration quality was evaluated through calculation of the squared correlation coefficient r^2, and the prediction performance was characterized through the predictive squared correlation coefficient q^2:[30]

$$r^2 = 1 - \frac{\sum_i (y_i^{fit} - y_i^{obs})^2}{\sum_i (y_i^{obs} - y^{mean})^2} \quad (8)$$

$$q^2 = 1 - \frac{\sum_i (y_i^{pred} - y_i^{obs})^2}{\sum_i (y_i^{obs} - y^{mean})^2} \quad (9)$$

In eqs 8 and 9, y_i^{obs} denotes the observed target value (in our case: Abraham parameter B) of compound i, and y^{mean} the mean experimental value of the data set under investigation (which is the training set mean in case of r^2, and the prediction set mean in case of q^2). Moreover, y_i^{fit} in eq 8 denotes the regression-fitted (calibrated) target value of compound i, which is replaced through the predicted (not calibrated) target value y_i^{pred} in eq 9. Note that q^2 ranges from 1 (perfect agreement) to $-\infty$, with $q^2 = 0$ representing the case where the model prediction is as good as taking the experimental mean as predictor for all values. In contrast to r^2, q^2 does not correct for systematic errors and thus is preferred if existing regression models are applied to external test sets. When a regression model is derived, $q^2 = r^2$ for the training set in case of parallel calibration of all multilinear parameters (because in this case $y_i^{pred} = y_i^{fit}$).

To further characterize the model performance, the following additional parameters have been used: root-mean-square error (rms), systematic error (bias), maximum negative error (mne, largest underestimation), and maximum positive error (mpe, largest overestimation).

Model Calibration and Validation. Because the effective donor and acceptor energies EE_{occ} and EE_{vac} (through which local hardness η is defined, see eqs 3−7) depend on E_{ref} in a nonlinear manner, nonlinear regression with the Levenberg−Marquardt algorithm was applied to calibrate eq 2. To account explicitly for differences in atom type characteristics, the calibration was performed separately for each HB acceptor atom type (see Table 1), allowing a corresponding variation in the final values of the model parameters c_{ES}, c_{PL}, c_{CT}, c (intercept), E_{ref} (EE_{occ}), and E_{ref} (EE_{vac}).

The prediction capability of the resultant AM1 HB acceptor strength model was evaluated by an extended leave-50%-out procedure,[31] stratified according to atom types of the HB acceptor sites as already applied for the validation of the HF and B3LYP models.[16] Briefly, this procedure simulates an external validation by dividing the data set into two groups of (almost) the same size, group I and group II. Both groups cover (essentially) the whole target value range and chemical domain of the data set, which was achieved in the following way: For each HB acceptor atom type, the compounds were ordered according to their experimental B values, followed by their alternate allocation to group I and II, respectively. Subsequently, nonlinear calibration of eq 2 was performed separately for groups I and II and for the HB acceptor atom type subsets. Finally, the model trained with group I was used to predict all B values of group II, and vice versa. In this way, both the robustness (stability of calibration statistics) and prediction capability (q^2 statistics) can be assessed using two temporary prediction sets that are nontrivial in their size with respect to both the chemical domain and the target value range.

TABLE 2: Regression Coefficients of the Hydrogen Bond Acceptor Model to Predict Abraham Parameter *B* from Molecular Structure[a]

HB acceptor type	c_{ES}	c_{PL}	c_{CT}	C
ESP Net Atomic Charges				
C	−0.004	−0.229	+0.000	+1.717
N	−0.000	+0.004	+0.286	+4.596
O	−0.473	+0.038	+0.150	+2.279
S	−0.000	+0.142	+0.170	+1.346
F[b]	−0.491	−0.116	+0.038	+1.789
Cl	−0.275	−0.091	+0.069	+0.364
Br	−0.219	−0.012	+0.009	+0.326
ZDO Net Atomic Charges				
C	−0.000	−0.232	+0.000	+1.736
N	−0.024	+0.000	+0.286	+4.622
O	−0.723	+0.039	+0.153	+2.245
S	−0.000	+0.118	+0.182	+1.636
F	−0.491	−0.077	+0.026	+1.788
Cl	−0.322	−0.048	+0.000	+0.424
Br	−0.264	−0.038	+0.000	+0.374
MUL Net Atomic Charges				
C	−0.000	−0.232	+0.000	+1.736
N	−0.024	−0.001	+0.285	+4.622
O	−0.725	+0.031	+0.150	+2.246
S	−0.000	+0.116	+0.181	+1.637
F	−0.015	−0.078	+0.035	+1.411
Cl	−0.322	−0.048	+0.000	+0.425
Br	−0.264	−0.038	+0.000	+0.374

[a] Coefficients for eq 3. ESP = electrostatic potential, ZDO = zero differential overlap (Coulson scheme), MUL = Mulliken. [b] Regression coefficients for ESP (based on the original data set in ref 16 with 403 compounds): $c_{ES}(F) = -1.189$; $c_{PL}(F) = -0.065$; $c_{CT}(F) = +0.006$; $c(F) = +0.748$; the variation of the coefficients, compared to the data set with 383 compounds, is due to a large variation in the experimental value range (0.05 [383] vs 0.17 [403]); the regression coefficients for other elements are not shown, because their change is statistically not significant.

Results and Discussion

Model Parametrization. Nonlinear calibration of the reference energies associated with the effective donor and acceptor energies, E_{ref} (EE_{occ}) and E_{ref} (EE_{vac}), revealed that for these model parameters, a distinction is required only between π bond and lone pair HB acceptor sites. Thus, HB acceptor atoms associated with one, two or three valence-shell lone pairs (such as nitrogen vs oxygen vs halogen) can be sufficiently well described through the same E_{ref} values. The resultant AM1-calibrated E_{ref} values are

$$\begin{aligned} E_{ref}(EE_{occ},\text{lone pair}) &= -12.0\ \text{eV} \\ E_{ref}(EE_{vac},\text{lone pair}) &= +2.0\ \text{eV} \\ E_{ref}(EE_{occ},\pi\ \text{bond}) &= -4.1\ \text{eV} \\ E_{ref}(EE_{vac},\pi\ \text{bond}) &= +1.9\ \text{eV} \end{aligned} \tag{10}$$

The associated coefficients c_{ES}, c_{PL}, c_{CT}, and c are shown in Table 2, again derived by nonlinear calibration of the data set. Here, separation of the calibration according to HB acceptor atom type was necessary, but for each atom type no further distinction between different compound classes (as included in the data set) was required.

The derived parameter values are based on the following further simplifications that turned out to be adequate: First, π bond (aromatic or conjugated) systems directly attached to a non-π HB acceptor group do not provide a separate contribution to B. Second, for heteroaromatic π systems only the heteroatom lone pair contribution to B needs to be taken into account explicitly.

ES, PL, and CT Contributions to B. The regression coefficient of the electrostatic component, c_{ES}, is negative. Thus, increasingly negative charge at the HB acceptor site increases the strength of hydrogen bonding, reflecting the correspondingly increasing Coulomb attraction to the positively charged hydrogen of the (not explicitly considered) HB donor site.

The sign of the hardness coefficient c_{PL} varies with atom type. Its positive sign for S and O as mainly moderate to strong HB acceptors (see Table 1) indicates that for these lone-pair sites, B increases with increasing local hardness. This trend is in line with considering hydrogen bonding as hard−hard interaction according to the Pearson concept. In contrast, B increases with increasing polarizability (decreasing η) for halogen (F, Cl, Br) and π bonds as weak HB acceptors. The latter finding differs from the c_{PL} signs derived for the corresponding ab inito (HF/6-31G**) and DFT (B3LYP/6-31G**) models[16] that are positive for all HB acceptor sites except for π-electron C. Thus, AM1 appears to differ from ab initio calculations with respect to hardness characteristics of halogen substituents.

The coefficient of the charge transfer contribution to B, c_{CT}, is positive for all atom types and charge calculation methods. It confirms our expectation that with increasing effective donor energy at the HB acceptor site, interaction of its electron-donor orbital (lone pair or π bond) with the (not explicitly considered) unoccupied electron-acceptor orbital of the HB donor site is

TABLE 3: Hydrogen Bond Acceptor Model Statistics[a]

method[b]	n	r^2	q^2	rms	bias	mne	mpe
Semiempirical Models							
AM1 (ESP)	383	0.97	0.97	0.04	0.00	−0.17	0.15
AM1 (ZDO)	383	0.95	0.95	0.04	0.00	−0.24	0.15
AM1 (MUL)	383	0.93	0.93	0.05	0.00	−0.27	0.14
Ab Initio and DFT Models[16]							
HF/6-31G** (NPA)	383	0.96	0.96	0.04	0.00	−0.15	0.17
B3LYP/6-31G** (NPA)	383	0.95	0.95	0.04	0.00	−0.22	0.10
Fragment Methods							
Platts[8]	383	0.84	0.79	0.09	−0.04	−0.27	0.31
Absolv[13]	383	0.79	0.57	0.13	−0.09	−0.44	0.09

[a] The statistical parameters are n = number of chemicals, r^2 = squared correlation coefficient, q^2 = predictive squared correlation coefficient, rms = root-mean-square error, mne = maximum negative error (largest underestimation), mpe = maximum positive error (largest overestimation). [b] The applied quantum chemical methods are AM1 = Austin Model 1, HF = Hartree−Fock, DFT = density functional theory with the B3LYP functional. Net atomic charges for the $Q(s)$ parameter in eq 3: ESP = electrostatic potential, ZDO = zero differential overlap (Coulson scheme), MUL = Mulliken, NPA = natural population analysis.

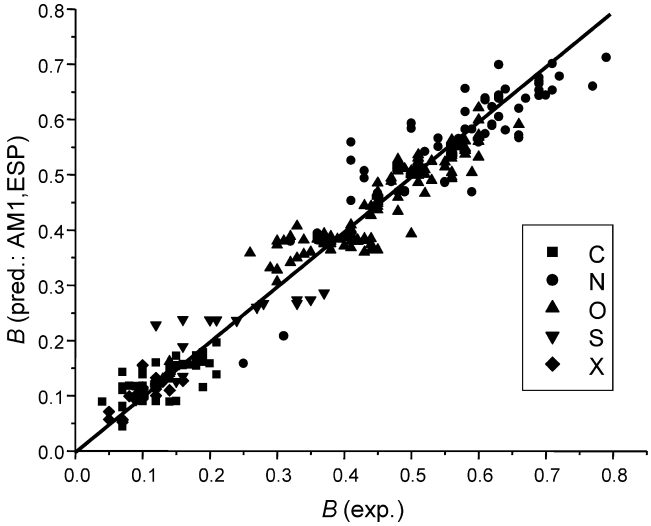

Figure 1. AM1 HB acceptor strength model, using ESP net atomic charges. Predicted vs experimental Abraham B values of carbon (squares), nitrogen (circles), oxygen (triangles up), sulfur (triangles down), and halogen (rhombs) hydrogen bond acceptors.

increasingly favored, yielding a correspondingly increasing energy stabilization through hydrogen bonding.

For the overall best model that employs ESP charges, the intercorrelation between the ES term and the PL term is 14%, and 30% for both ES-CT and PL-CT. In fact, all three terms are needed for both physical and statistical reasons to describe the HB acceptor strength sufficiently well. Interestingly, the ES-PL intercorrelation was significantly larger with HF/6-31G** and B3LYP/6-31G**,[16] suggesting a respective systematic difference between the semiempirical and ab initio level of calculation.

Overall Performance. The performance statistics of the AM1-based B prediction model are summarized in the upper part of Table 3. For the set of 383 compounds, r^2 ranges from 0.93 (Mulliken charges) to 0.97 (ESP charges). For the latter, the data distribution of predicted versus experimental B values is shown in Figure 1.

Interestingly, standard ZDO charges yield almost the same prediction quality as ESP, both of which outperform the results

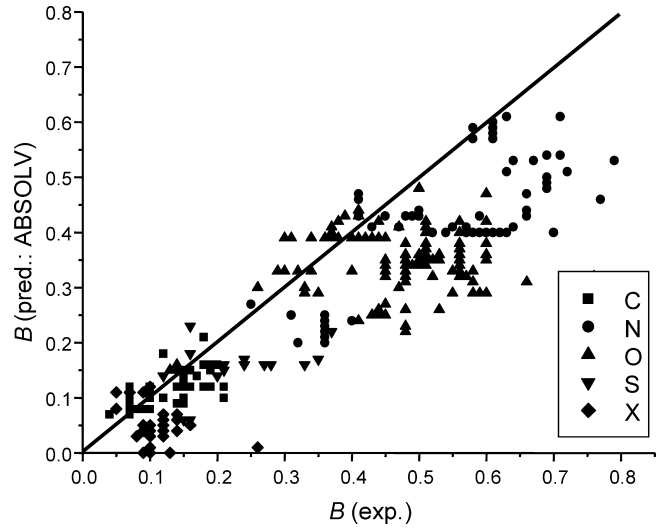

Figure 2. Absolv ADME boxes module. Predicted vs experimental Abraham B values of carbon (squares), nitrogen (circles), oxygen (triangles up), sulfur (triangles down), and halogen (rhombs) hydrogen bond acceptors.

employing Mulliken population analysis. Note that both ESP and Mulliken population analysis require additional calculation steps as compared to ZDO. Moreover, ESP net atomic charges are derived to reproduce, as well as possible, the electrostatic field around the molecule sensed grid-wise at some distance, while ZDO and Mulliken population analysis focus on different mathematical schemes to allocate the overall electron density to individual atoms.

To our surprise, AM1 yields a similar performance as HF or B3LYP (see Table 3), despite its far less computation time. However, the computational effort for semiempirical calculations is still higher than for fragment models. For example, screening of 2000 small- to medium-size organic compounds would require only 3 min with a fragment model (ca. 100 ms per molecule), about 3 h with AM1 (ca. 5 s per molecule), and about 1 month processor time when a standard ab initio level of calculation is employed (ca. 20 min per molecule). Note, however, that even the ab initio approach would require considerably less time than the experimental determination of B using chromatography.

In the lower part of Table 3, the performance statistics are shown for Platts' increment method[8] and for its commercial variant as implemented in Absolv.[13] While the former and original methods turn out to be superior for predicting B as compared to Absolv (r^2 0.84 vs 0.79, q^2 0.79 vs 0.57), both fragment schemes cannot compete with the quantum chemical methods. These statistics are based on $\Sigma\beta_2^{H}$ increments that refer to wet systems; use of $\Sigma\beta_2^{O}$ increments for dry systems (calibrated for dry octanol) yields even poorer results. In Figure 2, the data distribution of predicted vs experimental B values is shown for the 2008 version of the Absolv model, and in Figure 3 for Platts' original increment scheme.

With the Platts model, many B values are predicted to be around 0.45, but with associated experimental values ranging between 0.12 (α-pinene) and 0.69 (2,4,6-trimethylpyridine). This mis-fit suggests a lack of appropriate increments. The Absolv model yields both still inferior overall statistics and a larger bias (−0.09). A possible reason for the latter is that Absolv was calibrated also with compounds that have more than one HB acceptor site and thus an overall *effective* HB acceptor strength. For such compounds, competition between different

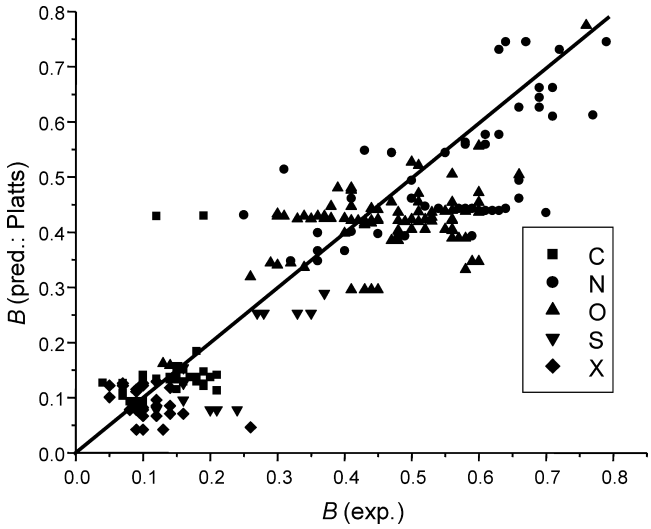

Figure 3. Platts fragment method. Predicted vs experimental Abraham *B* values of carbon (squares), nitrogen (circles), oxygen (triangles up), sulfur (triangles down), and halogen (rhombs) hydrogen bond acceptors.

TABLE 4: Model Validation by Data Set Splitting into Group I and Group II[a]

		training		external prediction	
group	group size *n*	r^2	rms	q^2	rms
I	195	0.97	0.04	0.91	0.06
II	188	0.94	0.05	0.96	0.04

[a] The semiempirical AM1 method has been applied using ESP (electrostatic potential) net atomic charges. The subsets group I and group II cover both the whole chemical domain and essentially the whole value range of *B*. The statistical parameters are: n = number of compounds, r^2 = squared correlation coefficient, rms = root-mean-square error, q^2 = predicted squared correlation coefficient.

HB acceptor sites for HB donor groups is likely to reduce the contributions of the individual sites to *effective* values that sum up to an overall *effective* HB acceptor strength. Because the Absolv increments are probably based on effective strengths, they are less suited to predict *B* for monofunctional compounds, the latter of which are driven by the site-specific *intrinsic* HB acceptor strength.

From this viewpoint, our quantum chemical models yield site-specific *intrinsic* HB acceptor strengths that are valid for compounds with only one HB acceptor site. In the case of more than one HB acceptor site, the total energy gain through parallel activation of all hydrogen bonds is less than the sum of the individual HB energies, which has to be taken into account when extending our model approach to multifunctional compounds.

Model Validation. As outlined above, the two subsets group I and group II were generated such that each of them covers both (almost) the whole target value range and the whole chemical domain of the data set. The associated calibration and prediction statistics are summarized in Table 4.

As can be seen from the table, groups I and II differ slightly in their calibration statistics, with group I results being close to the total set statistics. At the same time, both subsets yield a similar rms value of around 0.05. Moreover, application of the group I model for predicting group II *B* values yields a q^2 value similar to the group II calibration r^2 (0.96 vs 0.97), and correspondingly the group I q^2 value (achieved through application of the group II model) is similar to the group II training r^2 (0.91 vs 0.94). These findings demonstrate that our model has a good prediction power.

Method Performance for Individual HB Acceptor Types. In practice, QSAR predictions often concern only one or few compounds or compound classes to support their evaluation with respect to the target properties under investigation. In such cases, pertinent information about the class-specific method performance is required, keeping in mind that also methods with an overall good performance may vary in their prediction capability for individual compound classes.

In Table 5, the *B* prediction statistics of our semiempirical and ab initio[16] methods and of the increment methods of Platts and Absolv are summarized for the total compound set as well as for the individual HB acceptor atom types.

The observed variation of r^2 between individual HB acceptor types is partly driven by a corresponding variation in subset size. As a general trend, r^2 decreases with decreasing target value range.[30] The reason is that in eq 7, the terms $(y_i^{obs} - y^{mean})^2$ in the denominator become smaller with decreasing target value range, thus increasing the value of the quotient, and decreasing r^2. In such situations, the root-mean-square error rms is more robust in assessing the prediction quality, because it focuses on the terms $(y_i^{fit} - y_i^{obs})^2$ (training mode) or $(y_i^{pred} - y_i^{obs})^2$ (external prediction mode) that do not depend directly on the target value range.

For all atom types, rms is below 0.05 with HF/6-31G**(NPA), and below 0.06 with AM1(ESP) and B3LYP/6-31G**(NPA). This is in the same range as the experimental errors are supposed to be, making this approach a promising tool to assess the HB acceptor strength for compounds where respective experimental data are lacking.

For both AM1(ESP) and HF/6-31G**(NPA), the performance in terms of rms is slightly inferior for N sites than for most (HF) or all (AM1) other lone-pair HB acceptors. Interestingly, AM1 is superior to HF for fluorine, and here B3LYP yields the smallest subset rms. It is also surprising that for π-electron C as weak HB acceptor, the AM1 rms is again smaller than the one for HF. Note, however, that for both C and F, the experimental *B* value ranges are very small (0.17 and 0.05), which is also reflected in rather low r^2 values achieved with the different quantum chemical models. The HB acceptor strength of sulfur sites is better described with HF than with AM1, while Cl yields a lower rms (but also lower r^2) for AM1 than for HF.

Overall Table 5 suggests that among the quantum chemical methods, HF is preferred for N, O, S, and Br as well as with regard to the overall performance, and that AM1 yields the lowest rms for F, Cl, and π-electron C and is generally preferred over B3LYP.

Keeping in mind that our model is currently confined to quantifying the intrinsic HB acceptor strength, future model refinements should address steric hindrance, conformational flexibility, and (in the case of multifunctional compounds) intramolecular hydrogen bonding.

As compared to the quantum chemical *B* prediction models, the fragment schemes of Platts[8] and Absolv[13] yield significantly larger rms values for the total set as well as for all HB acceptor type subsets. Taking nitrogen sites as an example, the rms values achieved with HF, AM1, Platts, and Absolv are 0.05, 0.05, 0.09, and 0.14, respectively, and for oxygen HB acceptor sites the corresponding rms values are 0.03, 0.04, 0.10, and 0.16 (Table 5). At the same time, the Platts scheme and Absolv differ significantly in individual prediction errors, which is illustrated with cyclohexene oxide as one outlier. With the Platts scheme, the experimental *B* value of 0.48 is overestimated by 0.21 (predicted $B = 0.69$), while Absolv underestimates the HB

TABLE 5: HB Acceptor Model Statistics for Different Atom Types[a]

HB acceptor type	r^2	rms	bias	mne	mpe
AM1 (ESP)					
all[b]	0.97	0.04	0.00	−0.17	0.15
N	0.86	0.05	0.00	−0.12	0.15
O	0.88	0.04	0.00	−0.17	0.08
S	0.60	0.04	0.00	−0.08	0.11
F[c]	0.22	0.02	0.00	−0.02	0.02
Cl	0.73	0.01	0.00	−0.02	0.02
Br	0.76	0.01	0.00	−0.03	0.01
C	0.61	0.03	0.00	−0.07	0.07
HF (NPA)					
all	0.96	0.04	0.00	−0.15	0.17
N	0.86	0.05	0.00	−0.11	0.17
O	0.90	0.03	−0.01	−0.15	0.08
S	0.85	0.04	−0.01	−0.11	0.05
F[c]	0.00	0.05	0.00	−0.09	0.05
Cl	0.81	0.04	0.00	−0.07	0.06
Br	0.82	0.03	0.00	−0.05	0.04
C	0.47	0.03	0.00	−0.06	0.06
B3LYP (NPA)					
all	0.95	0.04	0.00	−0.22	0.10
N	0.87	0.04	0.00	−0.15	0.09
O	0.79	0.05	0.00	−0.22	0.08
S	0.87	0.06	−0.01	−0.12	0.10
F[c]	0.10	0.03	0.00	−0.06	0.05
Cl	0.82	0.01	0.00	−0.02	0.01
Br	0.85	0.01	0.00	−0.02	0.01
C	0.57	0.03	0.00	−0.07	0.10
Platts[8]					
all	0.83	0.09	−0.04	−0.27	0.31
N	0.58	0.09	−0.05	−0.27	0.20
O	0.32	0.10	−0.06	−0.25	0.13
S	0.53	0.11	−0.09	−0.16	0.01
F	0.19	0.06	−0.05	−0.09	0.03
Cl	0.13	0.03	0.00	−0.02	0.07
Br	0.19	0.05	−0.03	−0.09	0.03
C	0.07	0.06	0.00	−0.10	0.31
Absolv[13]					
all	0.79	0.13	−0.09	−0.44	0.09
N	0.58	0.14	−0.11	−0.30	0.06
O	0.07	0.16	−0.12	−0.44	0.09
S	0.23	0.10	−0.08	−0.18	0.07
F	0.19	0.09	−0.08	−0.13	0.02
Cl	0.33	0.05	−0.03	−0.07	0.06
Br	0.34	0.07	−0.06	−0.11	0.02
C	0.33	0.04	−0.01	−0.11	0.06

[a] The HB acceptor type subsets have the following numbers of compounds (n) and B value ranges (vr): All: $n = 383$, vr = 0.75. N: $n = 87$, vr = 0.54. O: $n = 167$, vr = 0.63. S: $n = 19$; vr = 0.25. F: $n = 9$, vr = 0.05. Cl: $n = 18$, vr = 0.07. Br: $n = 18$, vr = 0.07. C: $n = 65$; vr = 0.17. The statistical parameters are: n = number of chemicals, r^2 = squared correlation coefficient, rms = root-mean-square error, mne = maximum negative error (largest underestimation), mpe = maximum positive error (largest overestimation). AM1, HF/6-31G**, and B3LYP/6-31G** calculations have been applied for the quantum chemical models, employing either ESP (electrostatic potential) or NPA (natural population analysis) for quantifying net atomic charges. [b] Statistics for the original data set[16] that includes also 20 monofunctional compounds with B approximated through pK_{HB} and β_2^H (calculated with regression coefficients in footnote b of Table 2): $n = 403$; $r^2 = 0.96$; rms = 0.04; bias = 0.00; mne = −0.17; mpe = 0.22. [c] Statistics for H-bond acceptor F when using the original data set AM1 (HF/B3LYP): $n = 10$; vr = 0.17; $r^2 = 0.52$ (0.41/0.49); rms = 0.04 (0.05/0.04); bias = 0.00 (0.00/0.00); mne = −0.09 (−0.09/−0.06); mpe = 0.06 (0.05/0.05).

acceptor strength by 0.22 (predicted $B = 0.26$). In contrast, all three quantum chemical methods yield predicted B values close to the experimental value (AM1, 0.43; HF, 0.45; B3LYP, 0.48).

Interestingly, both fragment methods appear to perform better for halogen than for N, O, and S, with F yielding the largest rms among the halogens. Note, however, that the Absolv rms for F is already larger than its experimental value range (0.10 vs 0.05), and that for all halogens as well as for π-electron C r^2 is in the range of 0.07−0.34 for both the original Platts method and its Absolv update.

Tentative Inclusion of B Values Approximated through β_2^H or pK_{HB}. The Abraham B scale refers to solutes surrounded completely by solvent molecules and thus quantifies overall HB basicities. In contrast, both β_2^H and pK_{HB}[21] refer to a 1:1 solute−solvent complex and are related to each other through the equation[22]

$$\beta_2^H = \frac{1.1 + pK_{HB}}{4.636} \quad (11)$$

While β_2^H and B (that had originally been called $\Sigma\beta_2^H$) generally differ for multifunctional bases, their values for solutes with only one basic site appear to be usually quite similar: For a set of 81 monofunctional HB acceptors with experimental values for both β_2^H and B (details not shown), the respective r^2 was 0.90, the average unsigned and signed deviations (scatter and bias) were +0.046 and −0.0046, respectively, and the maximum deviation was 0.23 (observed for 1,8-cineole). Moreover, for 75 of the 81 compounds the deviation between β_2^H and B was smaller than 0.10. At the same time, the average B scatter per compound for the associated subset of 12 solutes with more than one B value was 0.062, with a maximum difference of 0.16. Taking dimethyl sulfide as an example, $\beta_2^H = 0.28$[32] as compared to B values of 0.27 and 0.29 reported in the literature,[33,34] although the two electron lone pairs of sulfur might suggest a difference in energy between full HB saturation and 1:1 HB complex formation.

Inclusion of experimental β_2^H as an approximation of B for 20 additional monofunctional bases (for 15 compounds derived from pK_{HB} through eq 11; see Supporting Information) yields $r^2 = 0.96$ and rms = 0.04 when applying AM1(ESP) to the accordingly augmented set of 403 compounds (see footnote b of Table 5). Interestingly, the AM1(ESP) regression coefficients remain essentially unchanged except for F as HB acceptor, where the B value range of the augmented set is 0.17 as compared to only 0.05 for the 383 compounds with original B values (see footnote b of Table 2). Note, however, that for multifunctional compounds it is not recommended to use β_2^H for modeling B.

Model Extension to Amides and Organophosphates. The quantum chemical model approach to predict B from molecular structure has been calibrated and validated (through simulated external prediction employing groups I and II as discussed above) for compounds containing only one HB acceptor site. Functional groups including two or more heteroatoms have not been considered so far. In this section, we will explore the model performance for two biochemically relevant compound classes that both contain more than one (potentially) HB-relevant site and thus are outside the initial model domain.

The first respective functional group is the amide group R′C(=O)−NR_2 that occurs frequently in amino acids, peptides and proteins. Prediction of its HB acceptor strength would allow one to forecast the HB energy gain involved in receptor binding or protein folding. The second compound class are organophosphates $(RO)_3PO$ that form components of membrane lipids and are also metabolites of insecticides acting as acetylcholine esterase inhibitors.[35]

TABLE 6: Experimental and Predicted *B* Values for Amides and Organophosphates (Model Extension, Not Included in the Original Model Domain)

compound	*B* (exp)[a]	*B* (HF)[b]	*B* (B3LYP)[b]	*B* (AM1)[c]
Amides ($n = 24$)				
formamide	0.60	0.62	0.55	0.67
acetamide	0.68	0.70	0.59	0.71
proprionamide	0.69	0.70	0.58	0.72
butanamide	0.68	0.70	0.58	0.72
isobutanamide	0.66	0.70	0.59	0.72
N-methylformamide	0.55	0.65	0.57	0.66
N-methylacetamide	0.72	0.73	0.61	0.71
N-methylpropanamide	0.71	0.72	0.60	0.71
N-butylacetamide	0.74	0.74	0.61	0.71
N,*N*-dimethylformamide	0.74	0.67	0.58	0.66
N,*N*-dimethylacetamide	0.78	0.73	0.62	0.70
N,*N*-diethylacetamide	0.78	0.75	0.63	0.72
N,*N*-dimethylpropanamide	0.78	0.73	0.61	0.71
benzamide	0.68	0.66	0.57	0.69
3-methylbenzamide	0.63	0.66	0.44	0.70
4-methylbenzamide	0.65	0.67	0.44	0.71
N,*N*-dimethylbenzamide	0.72[36]	0.67	0.60	0.68
N,*N*-diethylbenzamide	0.72[36]	0.68	0.61	0.70
formanilide	0.50	0.56	0.51	0.57
4-methylformanilide	0.52	0.57	0.52	0.58
acetanilide	0.67	0.63	0.55	0.61
2-methylacetanilide	0.70	0.64	0.55	0.62
3-methylacetanilide	0.66	0.64	0.55	0.62
4-methylacetanilide	0.67	0.65	0.55	0.62
Organophosphates ($n = 4$)				
trimethyl phosphate	1.00	1.08	0.71	0.83
triethyl phosphate	1.06	1.10	0.73	0.87
tri-*n*-propyl phosphate	1.15	1.11	0.73	0.87
tri-*n*-butyl phosphate	1.21	1.11	0.73	0.87

[a] Experimental values are taken from literature.[34,36] For the 28 compounds, the *B* value range is 0.69. [b] Quantum chemical model applying the basis set 6-31G** and NPA (natural population analysis) net atomic charges. [c] Semiempirical quantum chemical model employing ESP (electrostatic potential) to quantify net atomic charges.

For the explorative application of our quantum chemical models, we make the following simplifying assumptions: For amides, only the sp^2 carbonyl oxygen (C═O) is considered as an HB acceptor site, neglecting the lone pair of the amide N as a second site. Indeed, this assumption is in accord with the well-known fact that protonation of amides occurs only at the carbonyl oxygen, and that both the length of the central $R_2N{-}C(O)R'$ bond and its rotation barrier indicate a substantial double-bond character that can be visualized through the resonance structure $R_2N^+{=}C(-O^-)R'$.

Similarly, for the phosphate group we assume that only the (formally) double-bonded sp^2 oxygen, $(RO)_3P{=}O$, forms a potent HB acceptor site. Again, the situation can be visualized through an ionic resonance structure $(RO)_3P^+{-}O^-$, demonstrating the readiness of the partly negatively charged oxygen to act as HB acceptor. As with all other HB-active atoms with lone pairs, aromatic systems directly attached to amide or phosphate functionalities are not considered to form a separate HB acceptor site. Thus, for both functional groups the intrinsic *B* value calculated for the sp^2 oxygen is taken as the overall predicted *B* value. The influence of the other heteroatoms as well as of all other parts of the molecule is taken into account through their participation in the LCAO-MO wave functions and the resultant impact on net atomic charge, hardness, and effective donor and acceptor energies at the HB site as defined through eqs 2−7.

In Table 6, experimental and predicted *B* values are listed for 24 amides and four organophosphates. As can be seen from the table, HF/6-31G** yields surprisingly good predictions for both compound classes with an rms of 0.048. AM1 performs reasonably well for the amides but underestimates the phosphate *B* values by ca. 0.17−0.37 units. With B3LYP, still larger deviations between experimental and calculated *B* values are obtained.

In this context, two issues are important to note: First, and as mentioned already, the calibration of the models did not contain any amide or organophosphate, and thus in particular no sp^2 oxygen attached to amide carbonyl carbon or phosphorus. Second, the experimental *B* range of organophosphates (1.00−1.21) is far beyond the maximum *B* value in the calibration set (0.79). From this viewpoint, the HF prediction quality is surprisingly good, which still holds for AM1 applied to amides. Although we certainly do not recommend applying regression models outside their chemical domain, the present explorative study demonstrates the robustness of the quantum chemical approach introduced for both the ab initio[16] and AM1 levels of calculation.

With regard to the HB acceptor strength of organophosphates, a third issue deserves attention. According to 1:1 complexation values from earlier literature,[32] the four compounds should have *B* values between 0.76 and 0.79, which would be in much better agreement with B3LYP/6-31G**, and also closer to the AM1 values than to the HF/6-31G** predictions. In case the more recent organophosphate *B* values (as listed in Table 6) are correct, it could still be that our assumption to ignore the ester oxygens as additional HB acceptor sites is not adequate, and that explicit consideration of those additional sites would improve the AM1 and B3LYP performance. Thus, clarification of this issue appears to require both additional experimental and theoretical work.

Conclusions

The quantum chemical approach to predict hydrogen bond (HB) acceptor strength from local molecular parameters of the individual compounds could be transferred successfully from ab initio calculations to the fast-running semiempirical AM1 method. The new model allows one to screen large inventories of organic compounds for their site-specific intrinsic HB acceptor strength in terms of the Abraham *B* value. Despite much lower computational efforts, the AM1-based model yields an overall *B* prediction quality similar to the corresponding ab initio level of calculation. Although the model was calibrated with small molecules, the AM1 version is applicable for large compounds that go beyond the current scope of ab initio methods. While the current model parametrization is confined to HB acceptor atom types N, O, S, F, Cl, Br (lone-pair sites), and π-electron C (aromatic and conjugated systems), explorative application to amides and organophosphates demonstrates its mechanistically sound basis and robustness as well as opportunities for respective model extensions. For compounds with multiple HB acceptor sites, their competition for both intramolecular and intermolecular HB interaction is likely to reduce the individual site-specific HB acceptor strengths, which needs to be taken into account when the local parameter approach is extended to respective compounds. Additional routes to develop the model further include the consideration of steric hindrance as well as of conformational flexibility.

Acknowledgment. Financial support was provided by the European Commission through the projects NOMIRACLE (Contract No. 003956) and OSIRIS (Contract No. 037017), which is gratefully acknowledged.

Supporting Information Available: A table listing all 383 compounds together with experimental and predicted values of the Abraham parameter B and a table with 20 additional monofunctional compounds and B values approximated through β_2^H and pK_{HB}. This material is available free of charge via the Internet at http://pubs.acs.org. A copy of a program to predict the intrinsic hydrogen bond acceptor strength from information provided through MOPAC or Gaussian output files can be obtained upon request from the corresponding author.

References and Notes

(1) *An Introduction to Hydrogen Bonding*; Jeffrey, G. A., Ed.; Oxford University Press: Oxford, U.K., 1997.
(2) *Hydrogen Bonding: A Theoretical Perspective*; Scheiner, S., Ed.; Oxford University Press: Oxford, U.K., 1997.
(3) Mohajeri, A.; Nobandegani, F. F. *J. Phys. Chem. A* **2008**, *112*, 281.
(4) Abraham, M. H. *Chem Soc. Rev.* **1993**, *22*, 73.
(5) Sprunger, L. M.; Acree, W. E., Jr.; Abraham, M. H. *QSAR Comb. Sci* **2008**, *27*, 890.
(6) Abraham, M. H. *Eur. J. Med. Chem.* **2004**, *39*, 235.
(7) Sprunger, L. M.; Gibbs, J.; Acree, W. E., Jr.; Abraham, M. H. *QSAR Comb. Sci.* **2009**, *28*, 72.
(8) Platts, J. A.; Butina, D.; Abraham, M. H.; Hersey, A. J. *Chem. Inform. Comput. Sci.* **1999**, *39*, 835.
(9) Arey, J. S.; Green, W. H., Jr.; Gschwend, P. H. *J. Phys. Chem. B* **2005**, *109*, 7564.
(10) Gilli, P.; Pretto, L.; Bertolasi, V.; Gilli, G. *Acc. Chem. Res.* **2009**, *42*, 33.
(11) Thar, J.; Kirchner, B. *J. Phys. Chem. A* **2006**, *110*, 4229.
(12) Özen, A. S.; De Proft, F.; Aviyente, V.; Geerlings, P. *J. Phys. Chem. A* **2006**, *110*, 5860.
(13) Pharma Algorithms. *ADME Boxes, Version 4.1*; Pharma Algorithms Inc.: Toronto, Ontario, Canada, 2008.
(14) Schüürmann, G.; Ebert, R.-U.; Kühne, R. *Chimia* **2006**, *60*, 691.
(15) Schwöbel, J.; Ebert, R.-U.; Kühne, R.; Schüürmann, G. *J. Comput. Chem.* **2009**, *49*, 956.
(16) Schwöbel, J.; Ebert, R.-U.; Kühne, R.; Schüürmann, G. *J. Chem. Inf. Model.* **2009**, *30*, 1454.
(17) Morokuma, K. *Acc. Chem. Res.* **1977**, *10*, 294.
(18) Klamt, A. *Chemosphere* **1993**, *26*, 1273.
(19) Klamt, A. *Chemosphere* **1995**, *32*, 717.
(20) Böhnhardt, A.; Kühne, R.; Ebert, R.-U.; Schüürmann, G. *J. Phys. Chem. A* **2008**, *112*, 11391.
(21) Laurence, C.; Berthelot, M. *Perspect. Drug. Discov. Design* **2000**, *18*, 39.
(22) Abraham, M. H.; Platts, J. A. *J. Org. Chem.* **2001**, *66*, 3484.
(23) Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. *J. Am. Chem. Soc.* **1985**, *107*, 3902.
(24) Stewart, J. J. P. *MOPAC 2002*; Fujitsu Limited: Tokyo, Japan, 1999.
(25) Singh, U. C.; Kollman, P. A. *J. Comput. Chem.* **1984**, *5*, 129.
(26) Besler, B. H.; Merz, K. M.; Kollman, P. A. *J. Comput. Chem.* **1990**, *11*, 431.
(27) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A., Jr.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; and Pople, J. A. *Gaussian 03*, Revision C.02; Gaussian, Inc.: Wallingford, CT, 2004.
(28) Glendening, E. D.; Reed, A. E.; Carpenter, J. E.; Weinhold, F. *NBO Version 3.1*; Theoretical Chemistry Institute, University of Wisconsin: Madison, WI, 1993.
(29) Moré, J. J.; Garbow, T. K.; Hilstrom, K. E. *User Guide for MINPACK-1*; Argonne National Laboratory Report ANL-80-74; Argonne National Laboratory: Argonne, IL, 1980.
(30) Schüürmann, G.; Ebert, R.-U.; Chen, J.; Wang, B.; Kühne, R. *J. Chem. Inf. Model.* **2008**, *48*, 2140.
(31) Boháč, M.; Loeprecht, B.; Damborský, J.; Schüürmann, G. *Quant. Struct.-Act. Relat.* **2002**, *21*, 3.
(32) Abraham, M. H.; Grellier, P. L.; Prior, D. V.; Morris, J. J.; Taylor, P. J. *J. Chem. Soc., Perkin Trans. 2* **1990**, *1990*, 521.
(33) Laffort, P.; Héricourt, P. *J. Chem. Inf. Model.* **2006**, *46*, 1723.
(34) Abraham, M. H. *J. Phys. Org. Chem.* **1993**, *6*, 660.
(35) Schüürmann G. Ecotoxicology and structure activity studies of organophosphorus com-pounds. In *Rational Approaches to Structure, Activity and Ecotoxicology of Agrochemicals*; Fujita, T., Draber, W., Eds.; CRC Press: Boca Raton, FL, 1992; pp 485−541.
(36) Le Questel, J. Y.; Laurence, C.; Lachkar, A.; Helbert, M.; Berthelot, M. *J. Chem. Soc., Perkin Trans. 2* **1992**, *1992*, 2091.

JP904812B