

Theoretical Scales of Hydrogen Bond Acidity and Basicity for Application in QSAR/QSPR Studies and Drug Design. Partitioning of Aliphatic Compounds

Alexander A. Oliferenko,^{†,‡} Polina V. Oliferenko,^{†,‡} Jonathan G. Huddleston,[§] Robin D. Rogers,[§] Vladimir A. Palyulin,[†] Nikolai S. Zefirov,[†] and Alan R. Katritzky^{*,‡}

Department of Chemistry, Moscow State University, Moscow, 119899 Russia, Florida Center for Heterocyclic Compounds, Department of Chemistry, University of Florida, Gainesville, Florida 32611-7200, and Department of Chemistry and Center for Green Manufacturing, The University of Alabama, Tuscaloosa, Alabama 35487

Received December 17, 2003

Phenomenological analysis of existing hydrogen bond (HB) donor and acceptor scales and apparent physical considerations have enabled the establishment of new quantitative scales of hydrogen bond basicity and acidity. Chemical structures represented by molecular graphs and the orbital electronegativities of Hinze and Jaffe are utilized as an input data. The scales obtained correlate well with several experimental solvent polarity scales such as $\Sigma\beta_2^H$ and $\Sigma\alpha_2^H$, $pK(HB)$, and $E_T(30)$. To demonstrate the applicability of the new quantities, we have applied them to seven equilibrium partitioning data sets: octanol–water, hexadecane–water, chloroform–water, gas–water, gas–octanol, gas–hexadecane, and gas–chloroform partition coefficients. The hydrogen bond descriptors when supplemented by a cavity-forming term and a dipolarity term show high performance in correlations of the partition coefficients of aliphatic compounds. These new HB descriptors can be used in studying hydrogen bonding and fluid phase equilibria as well as scoring functions in ligand docking and descriptors in ADME evaluations.

INTRODUCTION

Quantitative characterization of the ability of chemicals to act as hydrogen bond donors (HBD) and/or acceptors (HBA) is crucial in the description of diverse processes occurring in condensed media such as dissolving, partitioning, solubilization, etc. Drug action and bioavailability critically depends on aqueous solubility, blood–tissue distribution, and specifically on hydrogen binding to receptor active sites and transport proteins.¹

Historically, quantitative scales of solvent hydrogen bond acidity and basicity originated in the solvatochromic comparison method developed by Taft, Kamlet, and co-workers in 1970s.^{2–4} These scales were derived from model spectroscopic experiments with different solvents and judiciously selected probe molecules. Such methodology is, in effect, an application of extrathermodynamical correlations (LFERs) as originally elaborated by Hammett⁵ and Taft⁶ for chemical rates and equilibria. The aim of these LFERs developed in 1970s by the Taft group was to account quantitatively for solvent effects on chemical reactivity. Therefore, the first LFER scales of hydrogen bond acidity and basicity characterize solvents rather than solutes. Systematic experimental and theoretical studies on solute descriptors were undertaken by Kamlet, Taft, and co-workers,^{7,8} by Abraham,⁹ by Laurence and Berthelot,^{10,11} and by Raevsky.¹² As distinct from the classical solvent solvatochromic parameters, which

characterize the properties of a single solute in an array of different solvents, the new solute descriptors characterize the properties of many different solutes in one solvent. This gave a working methodology applicable to various processes of chemical, environmental, and pharmaceutical interest. Within the approaches mentioned, the solute HB descriptors, which are of special interest to us, were derived from model electron spectroscopy (Kamlet and Taft, Abraham) or FT-IR spectroscopy (Laurence and Berthelot) experiments involving variable acids (hydrogen bond donor) and constant bases (hydrogen bond acceptors) and vice versa. All these studies refer to 1:1 hydrogen bond complexes, a situation which rarely occurs in real fluid phases. For the QSAR, ADME, and molecular modeling purposes, some “effective”, i.e. those accounting for the formation of multiple hydrogen bonds, HBA and HBD terms are needed, and such overall or summation values were “back-calculated” by Abraham using known terms of his general solvation equation⁹ and known experimental solubility data. The general solvation equation of Abraham is expressed in the form of a linear regression, which includes five terms as maximum

$$\log SP = c + rR_2 + s\pi_2^H + a\alpha_2^H + b\beta_2^H + vV_x^H \quad (1)$$

where R_2 is the excess molar refraction; π_2^H is the solute dipolarity–polarizability; α_2^H and β_2^H are the solute HB acidity and basicity, respectively; and V_x^H is the McGowan characteristic volume.¹³ SP means some physical or biological property, which is closely related to the solvation free energy. The regression coefficients (r , s , a , b , and v) are not merely adjustment constants, because each of them has

* Corresponding address phone: (352)392-0554; fax: (352)392-9199; e-mail: katritzky@chem.ufl.edu.

[†] Moscow State University.

[‡] University of Florida.

[§] The University of Alabama.

physical meaning, reflecting a certain feature of the condensed phase in which the solute is distributed.⁹ Some modification and refinement (based on gas chromatography data) of the latter approach was made by Carr and co-workers¹⁴ who aimed to isolate pure patterns of the intermolecular interactions from the Abraham solute descriptors.

Rather accurate correlations of solubility and partitioning properties have been produced by Abraham and co-workers, using the LSER (linear solvation free energy relationships), owing to the experimental origin of the latter. Up to date, the Abraham method is the best means for predicting fluid phase equilibria and related biological properties. However, the obvious advantages of using experimentally derived quantities bring with it the disadvantage of their limited availability. To circumvent this difficulty and to give a working tool to organic chemists, Abraham and Platts developed a system of structural hydrogen bond constants,¹⁵ resembling the substituent constant theories of Hammett and Taft. In the framework of their approach, they assume a strong additivity and transferability of frequently occurring atomic and polyatomic groups, for which the characteristic values are taken as the original alpha and beta values of simple model compounds.

A more advanced automated procedure based on the statistical analysis of fragment contributions to molecular structure was developed by the Abraham group in collaboration with Glaxo Wellcome¹⁶ and then implemented in the ABSOLV computer program.¹⁷ The applicability of this method was demonstrated by the prediction of partition coefficients in various solvent systems.^{18,19} The first such study, utilizing the principle of fragmental contributions, goes back to Hickey and Passino-Reader.²⁰ In 1991 they compiled an inventory of molecular fragments, which covered a rather wide range of organic compounds and proposed an empirical, nonautomated procedure for counting the fragment contributions. Theoretical attempts^{21–23} have also been undertaken to analyze the Abraham descriptors using quantum chemical calculations. These analyses, based on DFT and Bader's AIM theory, showed that simple theoretical predictors of $\sum \alpha_2^H$, $\sum \beta_2^H$, or of the polarity/polarizability parameter π_2^H do not exist. Alternatives to the LSER descriptors have been developing for the past decade. Of the recent attempts we mention a comparative study carried out by the groups of Abraham and Klamt²⁴ of the Abraham LSER descriptors and the solvation descriptors derived from the COSMO-RS theory.²⁵ The general conclusions were as follows: (i) although the two sets of descriptors vary in the manner they reflect the chemical information hidden therein, both are adequate for use; (ii) that solvation effects, especially, interphase partitioning, can probably be described by five independent terms. The latter conclusion requires verification, because one can find only four and three "good" principal components in the Abraham and Klamt data sets, respectively. Actually, PCA (principal component analysis) performed over the two data sets shows one principal component, covering 3% of the total variance in the Abraham data set, and two components covering 2 and 3% in the Klamt data set. It is believed that factors, covering such small proportions of the total variance are insignificant from the physical standpoint. Jorgensen and Duffy²⁶ derived another set of theoretical solvation descriptors from a virtual experi-

ment, Monte Carlo statistical mechanics simulations in which the HBD and HBA parts are given as counts of HB donor and acceptor sites averaged over strong and medium interactions between the solute molecule and explicit TIP4P water molecules. Explicit crystallographic data retrieved from the Cambridge Crystallographic Database were used by Klebe et al.²⁷ for derivation of knowledge-based HB donor and acceptor descriptors, which showed good performance in 3D QSAR studies.

The first recorded attempt to elaborate purely theoretical LSER (TLSER) scales, including those of HB donor and acceptor ability, was undertaken by Famini and Wilson²⁸ in the early 1990s. These workers selected some semiempirically calculated features as descriptors for regression analysis. Their descriptors, taking account of HB acidity and basicity, are specific atomic charges and reduced HOMO and LUMO energies. The TLSER method was extensively applied to various physical and chemical equilibrium problems,^{28,29} however, it did not and probably could not achieve as high level of statistical fit as Abraham's method. Nevertheless, the work of Famini and Wilson gave an impetus to further research in the field.

To model macroscopic distribution and reactivity patterns, Politzer and co-workers^{30,31} studied *ab initio* derived molecular electrostatic potentials (MEP) and found that some global and site-specific MEP quantities were good as descriptors in statistical fits. Using a generalization of their approach, known as GIPF (general interaction properties function), the authors^{30,31} successfully modeled such physical and biological properties as boiling points, critical constants, partition coefficients, dibenzo-p-dioxin toxicity, anticonvulsant, and anti-HIV activities. Other authors also used MEP-derived quantities.^{32,33} However, from the point of view of computational expense the MEP-based solvation descriptors are hardly available for complex molecules and especially biomolecules: a reliable molecular electrostatic potential can be built on the *ab initio* level of theory with the use of rather advanced basis set, 6-31G* at least.

Simplified HB parameters such as hydrogen bond donor and acceptor counts, atomic charges on the acidic hydrogen atoms, and some others are currently used in many QSAR programs, but they are of limited appropriateness despite their wide application in QSAR/QSPR studies. The obvious reason for the continuing usage of such characteristics is the absence of stronger alternatives. Comprehensive analysis of such descriptors has been done by Dearden and Ghafourian.³⁴

The aim of the current study is to analyze electronic, basically electronegativity-driven effects, contributing to the HB donor–acceptor interactions, and based on this analysis to derive from theory some scales or descriptors, characterizing basicity and acidity features of both solutes and solvents on hydrogen bonding. These descriptors appear in the Methodology section as functions of orbital energies and topological distance. As the previous workers³⁵ found significant differences in the solubility properties of aliphatic and aromatic solutes, we decided to first study only aliphatic species. Moreover, since we treat changes in HB basicity as a result of purely polar influences, we deliberately limit the scope of the article to nonconjugated aliphatic compounds. To obtain an approximation to a theoretical solvation equation we have supplemented the designed HB descriptors with other necessary terms. As a cavity-forming (and

dispersion) term, we take (i) the AM1 calculated alpha polarizability; (ii) the dipolarity term approximated by a specially designed parameter; and (iii) the HOMO and LUMO energies, which are considered to contribute to general basicity/acidity and the dispersion interaction. Our intention is to derive theoretical quantities, simulating HB donor–acceptor effects, which can be used as solvation descriptors for rapid, large-scale calculation of novel or yet unmeasured compounds, in particular for high-throughput virtual screening of drug candidates. Although we did not set out to reproduce any experimental solvent polarity scales on a theoretical basis, it turns out that our new HB scales correlate well with (i) Abraham's descriptors,⁹ (ii) the ΣpK_x –(HB) scale designed by Berthelot et al.,⁴⁵ and (iii) Reichardt's polarity parameter $E_T(30)$:³⁶ the correlation of our basicity term B^* with $\Sigma \beta_2^H$ is 97.9% and with $\Sigma pK_x(\text{HB})$ is 97.4%; the HB acidity term A^* shows apparent correlations with $\Sigma \alpha_2^H$ (93.3%) and $E_T(30)$ (90.3%). The relevancy and practical utility of these descriptors is thus far confirmed only for a representative set of aliphatic compounds as it is demonstrated by their correlations with free energy-related properties such as octanol–water ($\log K_{ow}$), hexadecane–water ($\log K_{hw}$), chloroform–water ($\log K_{cw}$), gas–water ($\log L_w$), gas–octanol ($\log L_{oct}$), gas–hexadecane ($\log L_h$), and gas–chloroform ($\log L_c$) partition coefficients.

METHODOLOGY

Basicity. Analysis of any scale of intrinsic, gas-phase (proton affinities), or hydrogen bond basicity (Taft and Kamlet, Abraham, Gutmann) reveals two main contributions to this property: (i) the specific basicity of electronegative heteroatoms, bearing lone electron pairs and (ii) the non-specific basicity of unsaturated species due to different H– π interactions.³⁷ Accordingly, either the lone pairs, or the π -systems of molecules, or both of them can act as hydrogen atom-attracting sites in processes of proton transfer or hydrogen bonding. Here, we assume that the hydrogen atom-attracting sites are formed through the interaction of intramolecular sites of different electronegativity. Such reasoning is in line with the general concept of electronegativity equilibration (see ref 38 and references therein). Some qualitative consideration of the role of electronegativity and hardness in hydrogen bonding were recently discussed by Ouvrard et al.³⁹

In the current treatment of hydrogen bond basicity, we consider only aliphatic species for which the first type of basicity is operative; i.e., that due to the specific basicity of electronegative heteroatoms.

Heteroatoms with Lone Pairs. The relative hydrogen bond basicity of heteroatoms can be derived from their electronic parameters, in particular from the lone pair electronegativity. As basic parameters we use orbital electronegativity (EN) and hardness values proposed by Hinze and Jaffe⁴⁰ and further developed by Bergmann and Hinze.⁴¹ For the halogen lone pairs we used the values from ref 38 which were derived from topological analysis of the electron density Laplacian.⁴² We assume that the propensity of heteroatoms to donate the electron density depends inversely on the electronegativity of their lone pair orbitals. Importantly, only heteroatoms of high ability to retain their own charge density (characterized by the lone pair hardness)

Table 1. Characteristics of Hybrid Orbitals, Atomic Inherent Basicities, and Acidities, and Reference Electronegativities of Heteroatoms

atom	χ_{lp}	η_{lp}	χ^σ	η^σ	χ^{ref}	B^{in}	$D_{1H}, \text{\AA}$	A^{in}
F ^a	18.160	8.180	13.240	8.630	7.3	0.225		
Cl ^a	12.270	5.460	10.520	5.610	8.0	0.148		
Br ^a	11.360	4.510	9.030	4.730	6.0	0.107		
I ^a	10.350	3.960	8.660	4.460	8.0	0.096		
N	4.570	9.580	11.490	7.456	10.0	1.048	1.012	0.195
O	7.880	10.825	15.250	9.150	10.5	0.687	0.958	0.566
S	6.590	7.045	10.140	5.361	10.0	0.356	1.335	0.015
C			8.060	6.565			1.096	−0.145

^a Lone pair electronegativity and hardness values are taken from ref 38.

exhibit a tendency to hydrogen bonding. Hence, we define the inherent basicity B^{in} of an atom possessing lone pairs as follows

$$B^{in} = \frac{\eta^{lp}}{\chi^{lp} n^*} \quad (2)$$

where χ^{lp} is the lone pair electronegativity, η^{lp} is the corresponding hardness value, and n^* is the effective principal quantum number (used to account for the inherent basicities of the elements belonging to the Periodic Table rows III, IV, etc.).

As seen from Table 1, B^{in} ranges from 0.096 for iodine to 1.048 for nitrogen in the sp^3 hybridization state, thus B^{in} serves as an initial basicity approximation. The hydrogen bond basicity of sp^3 carbon and other atoms not possessing lone pairs or π -electrons is zero.

Inductive Effect. A simple function of the electronegativities of the constituent atoms and distances from the basic site to them can account for influences of the environment. A similar approach was used in our previous modeling of the Taft σ^* inductive constants.⁴³ Based on the concept of σ – π separation, we apply such a function to the inductive effects over the molecular sigma-bonded network. We consider the extent of electronic charge withdrawn by an atom as proportional to its orbital sigma-bond electronegativity reduced by reference electronegativity. Initially, as a permanent reference, we took 8.87 V for the electronegativity of the sp^2 hybridization carbon atom following Bergmann and Hinze.⁴¹ This choice seems reasonable because saturated alkyl groups have some $-I$ -effect (electron-donating), whereas the acetylenic group demonstrates weak electron-withdrawing behavior ($+I$ -effect). However, we encountered the problem of rather strong nontransferability among families of compounds: the partial correlations for amines, and alcohols had markedly different slopes. Despite the well-known faith of chemists in the transferability principle, it is violated even for alkyl substituents. The interactions of the same alkyl group, say, ethyl with hydroxyl, amine, or a halogen are different because of the different valence state EN of the carbon atom and different ENs and polarizabilities of the substituents. This family dependent behavior of hydrogen bond bases was noted by Kamlet et al.⁴ and recently discussed by Popelier and Smith⁴⁴ and Berthelot et al.⁴⁵ As an atom can be characterized by its unique ability to attract the electron density from its neighbors, we introduce special reference values for each atom type to account for a distinct intensity of the neighborhood influence. Kamlet et al.⁴

introduced their coordinate covalency parameter ξ for much the same reason. The reference EN values are given in Table 1 along with the inherent atomic basicities. Therefore, we write the sigma contribution, B^σ , to hydrogen bond basicity by the following formula

$$B^\sigma = \sum_i^D \frac{\chi_i^\sigma - \chi_1^{ref}}{\eta_i^\sigma d_i^2 n^*} \quad (3)$$

where χ_1^{ref} is the reference electronegativity of the basic center, χ_i^σ , η_i^σ , and d_i are the electronegativity and hardness of the sigma orbital, and topological distance for i th substituent atom, respectively. The summation runs over all atoms positioned at particular topological distances from the basic center to D , where D is the diameter (the shortest path of maximum length) of the corresponding molecular graph. By topological distance, we mean not a distance in the real 3D space but just a count of edges separating one vertex of the corresponding molecular graph.⁴⁶ Only the shortest walks are considered. The total basicity consists of the heteroatom inherent characteristics minus the influence exerted by the sigma-bonded network:

$$B^* = B^{in} - B^\sigma = \frac{\eta^{lp}}{\chi^{lp} n^*} - \sum_i^D \frac{\chi_i^\sigma - \chi_1^{ref}}{\eta_i^\sigma d_i^2 n^*} \quad (4)$$

To account for the steric hindrance of alkyl groups to donor attack, we introduce empirical corrections to be multiplied by B^* calculated by eq 4. For acyclic O-, S-, and N-containing compounds with two alkyl substituents (ethers, dialkylamines, dialkyl sulfides, etc.) the steric correction (designated below as s) is equal to 0.9. In the case of trialkylamines the final B^* value is multiplied by 0.8. Cyclic bases such as dioxanes, morpholine, quinuclidine, piperidine, *N*-methylsubstituted piperidines, piperazine, and *N*-methyl-substituted piperazines are considered to be free of steric hindrance to hydrogen bonding and are not assigned steric corrections.

Multifunctional Compounds. Because the contributions of different basic centers to the overall basicity are non-equivalent (in general case), we multiply the current B^* value of each electron-rich atom by an attenuation factor, which reflects (i) the competition for H-bonding and (ii) steric hindrance to the intervening HB donor if the centers are closely located (two to four bond separation). In the general case each HB acceptor atom should be characterized by its own attenuation factor $a = 0.9^n \cdot 0.7^m$, where the factors 0.9 and 0.7 are empirical constants assigned to bonding situations where the neighboring basic sites are located at topological distances 4 or 3 and 2, respectively. The exponents n and m are the counts of the corresponding neighboring centers. In cases such as 2,2,2-trifluoroethanol, when the calculated basicities differ by a factor of at least two, the strongest site is freed of the attenuation factor. No attenuation is applied to directly linked acceptor atoms such as in hydrazine, disulfide, and their derivatives. In triethanolamine we take into account three effective oxygen HB acceptor (and donor) sites but do not take into account the nitrogen atom because of the presumably intramolecular hydrogen bond O—H \cdots N.

Acidity. Inherent Atomic Property. Hydrogen bond acidity is a specific quantity though it resembles in some features the gas-phase acidity (deprotonation energy) and aqueous acidity (pK_a). For example, while hydrogen iodide is very acidic both in the gas-phase and water, it is a poor hydrogen bond former. The boiling point, a direct indicator of intermolecular association, is low (-35°C) for hydrogen iodide as compared with the lower molecular weight but strongly associated hydrogen fluoride (bp = 19.5°C).

Hydrogen bonding has a large electrostatic component, and high electronegativity of the acidic center (the atom bonded to the acidic hydrogen) is critical to provide a significant positive charge on the hydrogen atom. Again, as in the case of basicity, electronegativity is very important for modeling HB-related acidity phenomena. Therefore, for acidity modeling we assume a linear dependence on the reduced orbital electronegativity of the bonded orbitals, not of the lone pair as in the case of basicity. Every hydrogen atom in the molecule is considered as partially acidic to a degree depending on the orbital electronegativity of the acidic center. The formula for the intrinsic acidity is written as follows

$$A^{in} = \frac{\chi_i^\sigma - \chi_1^{ref}}{\eta_i^\sigma D_{1H}^2} \quad (5)$$

where χ_i^σ and η_i^σ are electronegativity and hardness of the sigma orbital of the first atom, bonded to the hydrogen respectively, χ_1^{ref} is the reference electronegativity of this atom, and D_{1H} is the geometric distance between the hydrogen and the nearest atom. The intrinsic acidities of various atoms bonded to the hydrogen atom are given in Table 1 as well as distances between the hydrogen and heteroatoms.

Inductive Effect. Polar influences exerted by the environment are accounted for in a manner similar to the basicity treatment (vide supra). Thus, the hydrogen bond acidity term, encompassing just the sigma-bonded molecular network, is measured as follows

$$A^\sigma = \sum_{i=2}^N \frac{\chi_i^\sigma - \chi_i^{ref}}{\eta_i^\sigma d_{iH}^2} \quad (6)$$

where χ_i^σ and η_i^σ are the electronegativity and hardness of the sigma orbital of the acidic center and of the i th atom, respectively; χ_i^{ref} is the reference electronegativity of the i th atom, and d_{iH} is the topological distance between i th and the hydrogen atom.

The intrinsic acidity and the contributions from all the neighboring atoms provide a final A^* :

$$A^* = A^{in} + A^\sigma = \frac{\chi_1^\sigma - \chi_1^{ref}}{\eta_1^\sigma D_{1H}^2} + \sum_{i=2}^N \frac{\chi_i^\sigma - \chi_i^{ref}}{\eta_i^\sigma d_{iH}^2}$$

A^* values of multifunctional molecules are obtained by summation of the corresponding values from all the acidic sites. Water and ammonia are assumed to contain two donating hydrogens, while primary amines are assumed to contain only one donor. For hydrocarbon and halohydrocar-

bons, the most “acidic” hydrogens, i.e. those with the maximum numerical value, are taken.

COMPUTATIONAL DETAILS

As we have mentioned above, the main source of parameters used are the works of Hinze and co-workers,^{40,41} introducing the notion and numerical values of orbital electronegativity and hardness. Because such parameters for all halogen atoms in specific valence states are not listed in refs 41 and 42, we used available parameters calculated by Oliferenko et al.³⁸ The full list of the parameters used in this study is given in Table 1S of the Supporting Information. All necessary geometrical features of the molecules discussed have been obtained through AM1⁴⁷ semiempirical Hamiltonian implemented into HyperChem software.⁴⁸ Characteristic internuclear separations in simple hydrides are taken from the handbook.⁴⁹

The computational procedure starts from the atom of the highest inherent basicity (acidity). If two or more candidates occur, the calculation continues until a hit is found. Because of symmetry, multiple hits are possible; in such a case, one of them is taken as the first atom. When all polar influences are summed up and all necessary corrections are applied for the first atom, the procedure is repeated until all HB basic (acidic) sites are treated.

To demonstrate the performance of the model, we show the calculation protocols for the HB acidity of ethylene glycol and basicity of diethanolamine and halothane.

Ethylene Glycol. Intrinsic acidity of oxygen atom is $A_O^{in} = (\chi_O^\sigma - \chi_O^{ref})/\eta_O^\sigma D_{OH}^2 = (15.25 - 10.5)/(9.15 \cdot 0.958^2) = 0.566$; the neighbor contribution is

$$A_O^\sigma = \frac{\chi_{C_1}^\sigma - \chi_C^{ref}}{\eta_{C_1}^\sigma d_{C_1H}^2} + \frac{\chi_{C_2}^\sigma - \chi_C^{ref}}{\eta_{C_2}^\sigma d_{C_2H}^2} + \frac{\chi_{O_2}^\sigma - \chi_O^{ref}}{\eta_{O_2}^\sigma d_{O_2H}^2} = \frac{8.06 - 9.2}{6.565 \cdot 2^2} + \frac{8.06 - 9.2}{6.565 \cdot 3^2} + \frac{15.25 - 10.5}{9.15 \cdot 4^2} = -0.030$$

Then the HB acidity of one oxygen site is $A_O^* = A_O^{in} + A_O^\sigma = 0.566 - 0.03 = 0.536$; the resulting ethylene glycol acidity is calculated by the summation over all donor sites: $A^* = 0.44 + 2 \cdot A_O^* = 1.472$. The additive constant 0.44 is described below.

Diethanolamine. First, we treat the nitrogen atom which intrinsic basicity is $B_N^{in} = \eta_N^{lp}/\chi_N^{lp} n^* = 9.58/4.57 \cdot 2 = 1.048$; the neighboring influence is calculated as the successive sum of all atomic contributions:

$$B_N^\sigma = 2 \cdot \left(\frac{\chi_{C_1}^\sigma - \chi_N^{ref}}{\eta_{C_1}^\sigma \cdot d_{C_1}^2 \cdot 2} + \frac{\chi_{C_2}^\sigma - \chi_N^{ref}}{\eta_{C_2}^\sigma \cdot d_{C_2}^2 \cdot 2} + \frac{\chi_{O}^\sigma - \chi_N^{ref}}{\eta_{O}^\sigma \cdot d_{O}^2 \cdot 2} \right) = 2 \cdot \left(\frac{8.06 - 10}{6.565 \cdot 1^2 \cdot 2} + \frac{8.06 - 10}{6.565 \cdot 2^2 \cdot 2} + \frac{15.25 - 10}{9.15 \cdot 3^2 \cdot 2} \right) = -0.306$$

Then the nitrogen basicity is $B_N^{in} - B_N^\sigma = 1.048 - (-0.306) = 1.354$. The final nitrogen basicity is obtained by multiplying 1.354 by an appropriate steric hindrance coefficient: $B_N^* = B_N \cdot s_N = 1.354 \cdot 0.9 = 1.218$.

A similar calculation procedure is applied to the equivalent oxygen atoms:

$$B_O^{in} = \frac{\eta_O^{lp}}{\chi_O^{lp} \cdot n^*} = \frac{10.825}{7.88 \cdot 2} = 0.687$$

$$B_O^\sigma = \frac{\chi_{C_1}^\sigma - \chi_O^{ref}}{\eta_{C_1}^\sigma \cdot d_{C_1}^2 \cdot 2} + \frac{\chi_{C_2}^\sigma - \chi_O^{ref}}{\eta_{C_2}^\sigma \cdot d_{C_2}^2 \cdot 2} + \frac{\chi_N^\sigma - \chi_O^{ref}}{\eta_N^\sigma \cdot d_N^2 \cdot 2} + \frac{\chi_{C_3}^\sigma - \chi_O^{ref}}{\eta_{C_3}^\sigma \cdot d_{C_3}^2 \cdot 2} + \frac{\chi_{C_4}^\sigma - \chi_O^{ref}}{\eta_{C_4}^\sigma \cdot d_{C_4}^2 \cdot 2} + \frac{\chi_O^\sigma - \chi_O^{ref}}{\eta_O^\sigma \cdot d_O^2 \cdot 2} = \frac{8.06 - 10.5}{6.565 \cdot 1^2 \cdot 2} + \frac{8.06 - 10.5}{6.565 \cdot 2^2 \cdot 2} + \frac{11.49 - 10.5}{7.456 \cdot 3^2 \cdot 2} + \frac{8.06 - 10.5}{6.565 \cdot 4^2 \cdot 2} + \frac{8.06 - 10.5}{6.565 \cdot 5^2 \cdot 2} + \frac{15.25 - 10.5}{9.15 \cdot 6^2 \cdot 2} = -0.237$$

The basicity of each oxygen atom is $B_O^* = 1.048 - (-0.237) = 0.924$. The overall basicity of the molecule is obtained as the sum of all the components multiplied by the attenuation factor 0.9 and 0.9² for oxygen and nitrogen centers, respectively: $B^* = B_N^* \cdot a_N + 2 B_O^* \cdot a_O + 0.55 = 1.218 \cdot 0.81 + 2 \cdot 0.924 \cdot 0.9 + 0.55 = 3.199$.

Halothane (1,1,1-Trifluoro-2-chloro-2-bromoethane). The intrinsic basicities of the fluorine, chlorine, and bromine atoms are calculated as follows:

$$B_F^{in} = \frac{\eta_F^{lp}}{\chi_F^{lp} \cdot n_F^*} = \frac{8.18}{18.16 \cdot 2} = 0.225$$

$$B_{Cl}^{in} = \frac{\eta_{Cl}^{lp}}{\chi_{Cl}^{lp} \cdot n_{Cl}^*} = \frac{5.46}{12.27 \cdot 3} = 0.148$$

$$B_{Br}^{in} = \frac{\eta_{Br}^{lp}}{\chi_{Br}^{lp} \cdot n_{Br}^*} = \frac{4.51}{11.36 \cdot 3.7} = 0.107$$

The contributions from neighboring atoms are

$$B_F^\sigma = \frac{\chi_{C_1}^\sigma - \chi_F^{ref}}{\eta_{C_1}^\sigma \cdot d_{C_1}^2 \cdot 2} + \frac{\chi_{C_2}^\sigma - \chi_F^{ref}}{\eta_{C_2}^\sigma \cdot d_{C_2}^2 \cdot 2} + 2 \frac{\chi_F^\sigma - \chi_F^{ref}}{\eta_F^\sigma \cdot d_F^2 \cdot 2} + \frac{\chi_{Cl}^\sigma - \chi_F^{ref}}{\eta_{Cl}^\sigma \cdot d_{Cl}^2 \cdot 2} + \frac{\chi_{Br}^\sigma - \chi_F^{ref}}{\eta_{Br}^\sigma \cdot d_{Br}^2 \cdot 2} = \frac{8.06 - 7.3}{6.565 \cdot 1^2 \cdot 2} + \frac{8.06 - 7.3}{6.565 \cdot 2^2 \cdot 2} + 2 \cdot \frac{13.24 - 7.3}{8.63 \cdot 2^2 \cdot 2} + \frac{10.52 - 7.3}{5.61 \cdot 3^2 \cdot 2} + \frac{9.03 - 7.3}{4.73 \cdot 3^2 \cdot 2} = 0.283$$

$$B_{Cl}^\sigma = \frac{\chi_{C_1}^\sigma - \chi_{Cl}^{ref}}{\eta_{C_1}^\sigma \cdot d_{C_1}^2 \cdot 3} + \frac{\chi_{C_2}^\sigma - \chi_{Cl}^{ref}}{\eta_{C_2}^\sigma \cdot d_{C_2}^2 \cdot 3} + \frac{\chi_{Br}^\sigma - \chi_{Cl}^{ref}}{\eta_{Br}^\sigma \cdot d_{Br}^2 \cdot 3} + 3 \cdot \frac{\chi_F^\sigma - \chi_{Cl}^{ref}}{\eta_F^\sigma \cdot d_F^2 \cdot 3} = \frac{8.06 - 8}{6.565 \cdot 1^2 \cdot 3} + \frac{8.06 - 8}{6.565 \cdot 2^2 \cdot 3} + \frac{9.03 - 8}{4.73 \cdot 3^2 \cdot 3} + 3 \cdot \frac{13.24 - 8}{8.63 \cdot 3^2 \cdot 3} = 0.089$$

$$B_{Br}^\sigma = \frac{\chi_{C_1}^\sigma - \chi_{Br}^{ref}}{\eta_{C_1}^\sigma \cdot d_{C_1}^2 \cdot 3.7} + \frac{\chi_{C_2}^\sigma - \chi_{Br}^{ref}}{\eta_{C_2}^\sigma \cdot d_{C_2}^2 \cdot 3.7} + \frac{\chi_{Br}^\sigma - \chi_{Br}^{ref}}{\eta_{Br}^\sigma \cdot d_{Br}^2 \cdot 3.7} + 3 \cdot \frac{\chi_F^\sigma - \chi_{Br}^{ref}}{\eta_F^\sigma \cdot d_F^2 \cdot 3.7} = \frac{8.06 - 6}{6.565 \cdot 1^2 \cdot 3.7} + \frac{8.06 - 6}{6.565 \cdot 2^2 \cdot 3.7} + \frac{10.52 - 6}{5.61 \cdot 3^2 \cdot 3.7} + 3 \cdot \frac{13.24 - 6}{8.63 \cdot 3^2 \cdot 3.7} = 0.236$$

Table 2. Molecular Descriptors and Solvent Scales Discussed in the Paper^a

no.	compound	B^*	A^*	α , au	HOMO, eV	LUMO, eV	P^*	$\Sigma\lambda_p K_x(\text{HB})^b$	$E_T(30)$	$\Sigma\beta_2^H$	$\Sigma\alpha_2^H$
1	methane	0	0.26	8.87	-13.308	4.66	0.579	0		0	0
2	ethane	0	0.21	17.30	-11.766	4.117	0.579	0		0	0
3	propane	0	0.19	25.26	-11.324	3.92	0.579	0		0	0
4	butane	0	0.18	33.26	-11.172	3.829	0.579	0		0	0
5	2-methylpropane	0	0.17	32.87	-11.291	3.833	0.579	0		0	0
6	pentane	0	0.17	41.27	-11.111	3.775	0.579	0	31	0	0
7	2-methylbutane	0	0.17	40.83	-11.198	3.747	0.579	0	30.9	0	0
8	2,2-dimethylpropane	0	0.15	40.18	-11.534	3.905	0.579	0		0	0
9	cyclohexane	0	0.31	48.07	-10.937	3.656	0.579	0	30.9	0	0
10	fluoromethane	0.717	0.43	9.266	-13.111	3.448	0.906			0.1	0
11	tetrafluoromethane	0.425	0.00	11.639	-15.319	1.163	1.886			0	0
12	fluoroethane	0.703	0.38	18.075	-11.575	3.555	0.683			0.1	0
13	tetrachloromethane	0.595	0.00	34.628	-12.378	-1.117	4.963		32.4	0	0
14	chloromethane	0.695	0.37	13.368	-11.338	1.598	1.675			0.08	0
15	dichloromethane	0.701	0.48	19.377	-11.390	0.595	2.271	0.45	40.7	0.05	0.1
16	chloroform	0.654	0.59	26.591	-11.771	-0.304	3.867		39.1	0.02	0.15
17	chloroethane	0.695	0.32	22.387	-11.153	1.497	1.123			0.1	0
18	1,2-dichloroethane	0.780	0.37	27.801	-11.416	0.684	1.749	0.86	41.3	0.11	0
19	1,1,1-trichloroethane	0.652	0.36	35.813	-11.992	-0.266	2.232		36.2	0.09	0
20	pentachloroethane	0.652	0.63	47.373	-11.869	-0.681	3.628			0.06	0.17
21	hexachloroethane	0.592	0.00	54.591	-12.182	-0.967	4.254			0.06	0
22	chloropropane	0.693	0.31	30.546	-11.135	1.518	0.902			0.1	0
23	<i>t</i> -butyl chloride	0.693	0.20	38.690	-11.248	1.327	0.783			0.12	0
24	bromomethane	0.572	0.42	15.451	-11.553	0.570	1.881		37.6	0.1	0
25	dibromomethane	0.521	0.58	23.791	-10.958	-0.052	3.182		39.4	0.1	0.1
26	bromoform	0.456	0.74	33.352	-11.072	-0.747	4.484		37.7	0.06	0.15
27	<i>i</i> -propylbromide	0.530	0.33	33.742	-10.675	0.730	0.984			0.14	0
28	<i>t</i> -butylbromide	0.509	0.24	41.929	-10.680	0.659	0.846				
29	bromodichloromethane	0.633	0.69	28.503	-12.285	-0.923	4.073			0.07	0.13
30	iodomethane	0.643	0.29	18.056	-10.513	0.516	2.458			0.13	0
31	diiodomethane	0.668	0.33	28.814	-10.564	-0.270	4.337		36.5	0.23	0.05
32	halothane	0.475	0.71	33.914	-11.752	-0.815	2.426			0.05	0.15
33	water	1.237	1.53	3.378	-12.464	4.418	1.546		63.1	0.35	0.82
34	methanol	1.423	0.92	12.317	-11.135	3.778	0.894	1.95	55.4	0.47	0.43
35	ethanol	1.469	0.86	20.551	-10.876	3.565	0.704	2.11	51.9	0.48	0.37
36	1-propanol	1.490	0.89	28.625	-10.846	3.489	0.617	2.14	50.7	0.48	0.37
37	2-propanol	1.516	0.88	28.221	-10.895	3.492	0.617		48.4	0.56	0.33
38	1-butanol	1.501	0.89	36.680	-10.846	3.425	0.568		50.2	0.48	0.37
39	2-butanol	1.536	0.87	36.297	-10.899	3.477	0.568		47.1	0.56	0.33
40	2-methyl-1-propanol	1.510	0.88	36.278	-10.874	3.456	0.568		48.6	0.48	0.37
41	2-methyl-2-propanol	1.562	0.86	35.606	-10.992	3.438	0.568		43.3	0.6	0.31
42	2-methyl-2-butanol	1.583	0.85	43.581	-10.814	3.454	0.536		41.1	0.6	0.31
43	3-methyl-1-butanol	1.513	0.88	44.253	-10.804	3.384	0.536		49	0.48	0.37
44	4-methyl-2-pentanol	1.559	0.86	51.905	-10.851	3.440	0.513			0.56	0.33
45	cyclopentanol	1.509	0.86	41.891	-10.791	3.404	0.491		47	0.56	0.32
46	2-chloroethanol	1.573	0.89	24.104	-11.642	0.779	1.252		55.5	0.57	0.47
47	2-bromoethanol	1.460	0.90	26.961	-11.533	0.122	1.355				
48	2,2,2-trifluoroethanol	1.327	0.99	20.181	-12.593	1.486	1.194		59.8	0.25	0.57
49	ethylene glycol	2.153	1.47	21.945	-11.217	3.157	0.865		54.1	0.78	0.58
50	propylene glycol	2.212	1.43	29.404	-11.272	3.012	0.745		54.9	0.85	0.77
51	<i>trans</i> -1,2-cyclohexanediol	2.331	1.37	54.141	-10.551	3.141	0.563			0.809	
52	dimethyl ether	1.503	0.37	21.372	-10.612	3.250	0.732			0.41	0
53	dichloromethyl ether	1.597	0.51	31.344	-11.891	0.483	1.828				0
54	diethyl ether	1.586	0.31	38.321	-10.393	2.981	0.583	1.76	34.5	0.45	0
55	2-methoxyethanol	2.233	0.93	30.444	-10.953	3.201	0.833		52.3	0.84	0.3
56	2-ethoxyethanol	2.278	0.92	38.052	-10.905	3.049	0.737		51	0.83	0.3
57	dimethoxymethane	1.828	0.48	31.146	-11.322	3.365	0.782		35.8	0.532	0
58	dimethoxyethane	2.312	0.37	39.247	-10.929	3.251	0.703	3.54	38.2	0.534	0
59	1,4-dioxane	2.571	0.37	39.792	-10.205	2.826	0.671	3.04		0.64	0
60	1,3-dioxane	2.035	0.47	36.991	-11.206	3.050	0.671				0
61	1,3,5-trioxane	1.946	0.51	32.039	-11.334	3.014	0.884	2.70	36		0
62	tetrahydropyran	1.722	0.29	44.010	-10.355	3.176	0.511	1.95	36.6	0.55	0
63	tetrahydrofuran	1.701	0.30	34.526	-10.180	3.110	0.539	1.99	37.4	0.48	0
64	ammonia	1.598	0.79	5.283	-10.418	4.223	1.77			0.62	0.14
65	methylamine	1.746	0.55	14.677	-9.755	3.813	1.031	2.25		0.58	0.16
66	ethylamine	1.783	0.53	22.782	-9.682	3.651	0.816	2.23		0.61	0.16
67	dimethylamine	1.759	0.51	23.787	-9.385	3.480	0.785	2.3		0.66	0.08
68	trimethylamine	1.743	0.27	32.574	-9.117	3.192	0.662	2.32		0.67	0
69	diethylamine	1.826	0.47	40.317	-9.289	3.227	0.625	2.28	35.4	0.69	0.08
70	triethylamine	1.832	0.20	54.244	-9.741	2.848	0.544		32.1	0.79	0
71	N,N'-dimethylhydrazine	1.974	0.76	29.045	-9.368	2.779	0.803				
72	1,2-diaminoethane	2.749	0.69	26.926	-10.179	3.326	1.02	4.53	42		
73	1,3-diaminopropane	2.787	0.66	34.442	-10.163	3.348	0.884	4.63			
74	diethanolamine	3.199	1.56	43.036	-10.135	2.896	0.806				
75	triethanolamine	3.080	1.43	61.223	-9.862	2.439	0.745				

Table 2 (Continued)

no.	compound	B^*	A^*	α , au	HOMO, eV	LUMO, eV	P^*	$\Sigma\lambda pK_s(\text{HB})^b$	$E_T(30)$	$\Sigma\beta_2^H$	$\Sigma\alpha_2^H$
76	hydrogen sulfide	0.906	0.41	21.945	-11.217	3.157	2.371			0.16	
77	methylmercaptane	1.005	0.23	19.752	-8.934	0.874	1.553			0.24	0
78	dimethyl sulfide	1.048	0.24	29.075	-8.481	0.934	1.348			0.29	0
79	ethylmercaptane	1.029	0.35	28.559	-8.907	0.833	1.115			0.24	0
80	diethyl sulfide	1.092	0.19	47.017	-8.443	0.863	0.936	1.34		0.32	0
81	dimethyl disulfide	1.397	0.25	48.425	-8.188	-1.517	1.198			0.28	0
82	1,2-dithioethane	1.459	0.14	37.291	-9.792	1.459	1.592				
83	piperazine	3.082	0.56	42.284	-10.029	2.893	0.739	4.37			
84	N,N'-dimethylethylenediamine	2.784	0.59	44.301	-9.876	3.087	0.763	4.62			
85	tetramethylenediamine	2.761	0.22	60.591	-9.491	2.887	0.64	4.23			0
86	1-methylpiperidine	2.115	0.25	46.051	-9.052	3.206	0.535	2.24		0.75	0
87	piperidine	1.984	0.46	46.720	-9.277	3.317	0.552	2.37	35.5	0.69	0.1
88	N-ethylpiperidine	1.845	0.19	51.861	-9.744	2.959	0.496			0.63	0
89	quinclidine	2.169	0.18	57.290	-10.003	3.067	0.466				0
90	morpholine	2.827	0.50	40.095	-10.241	2.941	0.706		41	0.91	0.06
91	1,4-dimethylpiperazine	3.364	0.25	61.904	-8.592	2.931	0.617	4.03		1.256	0
92	18-crown-6	5.482	0.38	122.20	-10.416	1.931	0.671			1.713	0
93	[2.2.2]cryptand	7.999	0.34	184.48	-8.948	2.385	0.644			2.895	0

^a The last two columns contain Abraham's experimental values for comparison. ^b Berthelot, M.; Graton, J.; Ouvrard, C.; Laurence, C. Hydrogen-bond basicity of solutes in hydroxylic solvents from octanol–water partition coefficients. *J. Phys. Org. Chem.* **2002**, *15*, 218–228.

Therefore, the basicity values of the halogen basic centers are

$$B_F^* = 0.225 - 0.283 = -0.057$$

$$B_{Cl}^* = 0.148 - 0.089 = 0.059$$

$$B_{Br}^* = 0.107 - 0.236 = -0.129$$

As the basicity of the chlorine center is more than twice as large compared to that of the fluorine and bromine centers, it is not corrected by an attenuation factor. The attenuation factors are $0.7^2 \cdot 0.9^2 = 0.397$ and $0.7^1 \cdot 0.9^3 = 0.51$ for the fluorine and bromine atoms, respectively. Therefore, the overall basicity of haloethane comes to be $B^* = B_{Cl}^* + 3 \cdot B_F^* \cdot 0.397 + B_{Br}^* \cdot 0.51 + 0.55 = 0.475$.

RESULTS AND DISCUSSION

Basicity. Because of computational reasons, the original B^* values calculated for the current data set started from -0.125, the value corresponding to CF_4 . However, it would be desirable to set up a real zero of the scale because CF_4 still possesses some HB basicity. Hydrocarbons are the compounds to which zero HB basicity should be assigned because they have no specific basicity sources such as lone pairs or π -electrons. To approach zero we tried a set of converging B^* values calculated for CF_4 , C_2F_6 , C_3F_8 , $(\text{CF}_3)_3\text{CF}$, and $\text{C}(\text{CF}_3)_4$. The perfluorinated hydrocarbons were used because they showed the weakest basicity over the whole data set. Given these B^* values plotted against linear function $y = x$, this series demonstrates a clear exponential decay to the value of -0.550, with perfluorinated neopentane being the least basic species ($B^* = -0.529$). Therefore, we took this value as the origin of the B^* scale. Finally, 0.55 was added to all the calculated B^* values, giving rise to a normalized scale. Thus HB basicity of a molecule is written as $B^* = 0.55 + \Sigma_i B_i a_i$.

Haloalkanes have rather low B^* values, not exceeding 0.780 for 1,2-dichloroethane which is the most basic within our scheme. O-bases range from 1.237 for water to 2.571 for 1,4-dioxane. The values for N-bases vary from 1.598

(ammonia) to 3.199 (diethanolamine). The highest B^* values are inherent in macrocyclic bases 18-crown-6 (5.482) and [2.2.2]-cryptand (7.999). As it is clearly seen from Table 2, the closer neighboring basic centers are located in multifunctional compounds where they strongly attenuate each another. Thus, 1,2-dioxy- and 1,2-diaminocompounds possess lower basicity values than the corresponding 1,3-diheterocompounds, cf. ethylene glycol vs 1,3-propylene glycol and ethylenediamine vs propylenediamine.

Acidity. The problem of the origin is solved in a manner similar to that for the B^* scale. As branched alkanes seem to be less acidic, we calculated A^* values of the less "acidic" hydrogen (bonded to the tertiary carbon atom) for the following series: propane, 2-methylpropane, 3-ethylpentane, 4-propylpentane, and 5-butylnonane. As in the case of the B^* scale, an exponential fit was used, which gave -0.44 as an approximate limit. Then all the A^* scale values was raised by the constant 0.44: $A^* = 0.44 + \Sigma_i A_i^*$.

The acidity scale A^* varies over a smaller range compared to both the basicity term described above and the Abraham experimental acidity term: from zero for CF_4 and CCl_4 to 1.56 for diethanolamine. In the framework of our design for HB donor ability, OH-acids are expectedly the strongest HB donors that is in line with the Abraham theory built on experimental data⁹ and quantum-chemical calculations.^{23–29}

Correlation with Existing Solvent Polarity Scales. As mentioned in the Introduction, our scales B^* and A^* correlate well with experimental scales such as the Abraham basicity and acidity terms $\Sigma\beta_2^H$ and $\Sigma\alpha_2^H$,⁹ the statistically modified $pK_s(\text{HB})$ by Berthelot et al.,⁴⁵ and the $E_T(30)$ polarity scale developed by Reichardt.³⁶ This means that the theoretical B^* and A^* values are able to reveal an essential part of the chemical information hidden in these experimental scales. The correlations cannot be perfect because the experimental solvent polarity scales inevitably exhibit a mixture of effects including dipole–dipole and dispersion interactions. By contrast, theory-derived scales should demonstrate a pure effect due to acidity, basicity, etc.

One-parameter correlation equations, where $\Sigma\beta_2^H$, $\Sigma\alpha_2^H$, $pK_s(\text{HB})$, and $E_T(30)$ are the dependent variables and B^* and A^* are predictors, are listed in Table 3. Here r and N

Table 3. Correlations with Experimental Solvent Polarity Scales of the form $y = a_0 + a_1x$

parameter	statistics	solvent polarity scales			
		$\Sigma\beta_2^H$	$\Sigma\alpha_2^H$	$pK_s(\text{HB})$	$E_T(30)$
B^*	a_0	-0.080		-0.120	
	a_1	0.365		1.427	
	r	0.979		0.974	
	F	1759		582	
	N	79		34	
A^*	a_0		-0.148		28.645
	a_1		0.540		21.817
	r		0.933		0.903
	F		540		172
	N		82		41

correspond to the correlation coefficient and the number of points, respectively.

As can be seen in Table 3, the correlations for $\Sigma\beta_2^H$ and $\Sigma\alpha_2^H$ are characterized with rather small intercepts and slopes less than unity because the variation ranges of B^* and A^* are wider than those of $\Sigma\beta_2^H$ and $\Sigma\alpha_2^H$. The statistical quality of correlations in terms of correlation coefficient r and Fisher's statistics F vary from excellent ($\Sigma\beta_2^H$ against B^* and $pK_s(\text{HB})$ against B^*) to satisfactory ($\Sigma\alpha_2^H$ against A^* and $E_T(30)$ against A^*).

Modeling Equilibrium Solvation Properties. We express equilibrium solubility values in the logarithmic form of the Ostwald solubility coefficients

$$\log L = \log \left(\frac{C_l}{C_g} \right)$$

where C_l and C_g are molar concentrations of a chemical in the liquid phase and the gas phase, respectively. These $\log L$ values are a more appropriate characteristic than bulk solubility, because $\log L$ is straightforwardly connected with the corresponding free energies of solvation (transfer of one mole of a substance from the gas phase to the liquid phase if the standard states of unit concentrations in the gas phase and in solution are chosen):

$$\Delta G_s = -2.3RT \log L = -2.3RT \log \left(\frac{C_l}{C_g} \right)$$

Therefore, now we are in a position to apply our theoretical descriptors to a real, free energy-driven process. As attested by many authors,^{4,9,28} hydrogen bond acidity and basicity are very important but not the sole descriptors necessary for modeling phase and chemical equilibria. Polarizability/dipolarity, molecular volume, and cavity-formation terms are also needed. In the framework of the general solvatochromic comparison method (Taft-Kamlet-Abboud, Abraham), special free energy-related experimental or theoretical (like the McGowan characteristic volume V_x^{13}) features such as π^* , π_2^H (dipolarity/polarizability indices for solvents and solutes, respectively), R_2 (excess molar refraction), and $\log L_{hw}$ (gas-hexadecane partition coefficients) have been derived. However, we tend to use only theoretical descriptors because such an approach has been proved⁵⁰ as successful one. We use these new hydrogen bond descriptors along with the descriptors calculated by our new in-house CODESSA PRO QSAR software⁵¹ such as polarizability, dipole moment, and LUMO energy. All the descriptors used in the below listed

correlations have been autoscaled by subtracting the mean values and dividing by the standard deviations.

Octanol-Water Partition Coefficients ($\log K_{ow}$). $\log K_{ow}$ is a highly important experimental measure of hydrophobicity/lipophilicity, it finds numerous applications in QSAR and biologically-oriented molecular modeling.⁵² This entity is a measure of the free energy of transfer of one mole of a compound from one liquid phase (water) to another (octanol). Our four-parameter correlation for 90 organic molecules is as follows:

$$\log K_{ow} = 0.839 - 1.711B^* - 0.17A^* + 1.283\alpha + 0.169P^* \quad (7)$$

$$N = 90, R = 0.985, R^2 = 0.970, F = 682, s = 0.233$$

As distinct from r in Table 3, R in eq 7 identifies the multiple regression coefficient. Eq 7 shows as most significant the negative contributions of descriptors B^* and a rather large positive one from molecular polarizability α . A certain positive contribution goes from the so-called integral polarity descriptor P^* , which we define as the sum of all bond dipole moments in a molecule scaled by the differences in atomic polarizabilities of immediately bonded atoms and reduced by the number of bonds in a molecule, N :

$$P^* = \frac{\sum_i^N d_i(a_{i1} - a_{i2})}{N}$$

Here d_i is the bond dipole moment tabulated by Minkin, Osipov, and Zhdanov,⁵³ and a_{i1} and a_{i2} are the hybrid polarizabilities of the two atoms constructing the i th bond, as calculated by Miller⁵⁴ (numerical values for both these quantities are given in Table 2S of the Supporting Information). Descriptor P^* is turned out to be useful in the correlations listed below as an additional dipolarity/polarizability term. Unlike gas-water partition coefficients $\log L_w$ (vide infra), the HB basicity term B^* has the negative sign in eq 7 because of higher HB acidity of water comparing to octanol. The correlation plot of predicted against observed $\log K_{ow}$ values is given in Figure 1, and the numerical values are given in Table 4. We should note that partial correlations for separate groups of compounds are also statistically relevant: for 31 haloalkanes, 28 O-compounds, and 17 N-compounds the R^2 are respectively 0.942, 0.949, and 0.965. The standard deviations s are 0.218, 0.177, and 0.219.

Hexadecane-Water Partition Coefficients ($\log K_{hw}$). The distribution of a solute between two liquid phases, water and hexadecane (typical alkane), reflects convincingly the chemical nature of the compound. In fact, $\log K_{hw}$ can be considered as an additional characteristic of hydrophobicity. For 55 aliphatic compounds we derived eq 8:

$$\log K_{hw} = 0.085 - 2.225B^* - 0.775A^* + 1.585\alpha - 0.369LUMO \quad (8)$$

$$N = 55, R = 0.990, R^2 = 0.980, F = 614, s = 0.291$$

Besides the HB and cavity-formation terms eq 8 contains the LUMO energy, as an indicator both of polarity (inversely) and polarizability (directly). The negative signs of the HB descriptors (B^* , A^* , and LUMO because the latter character-

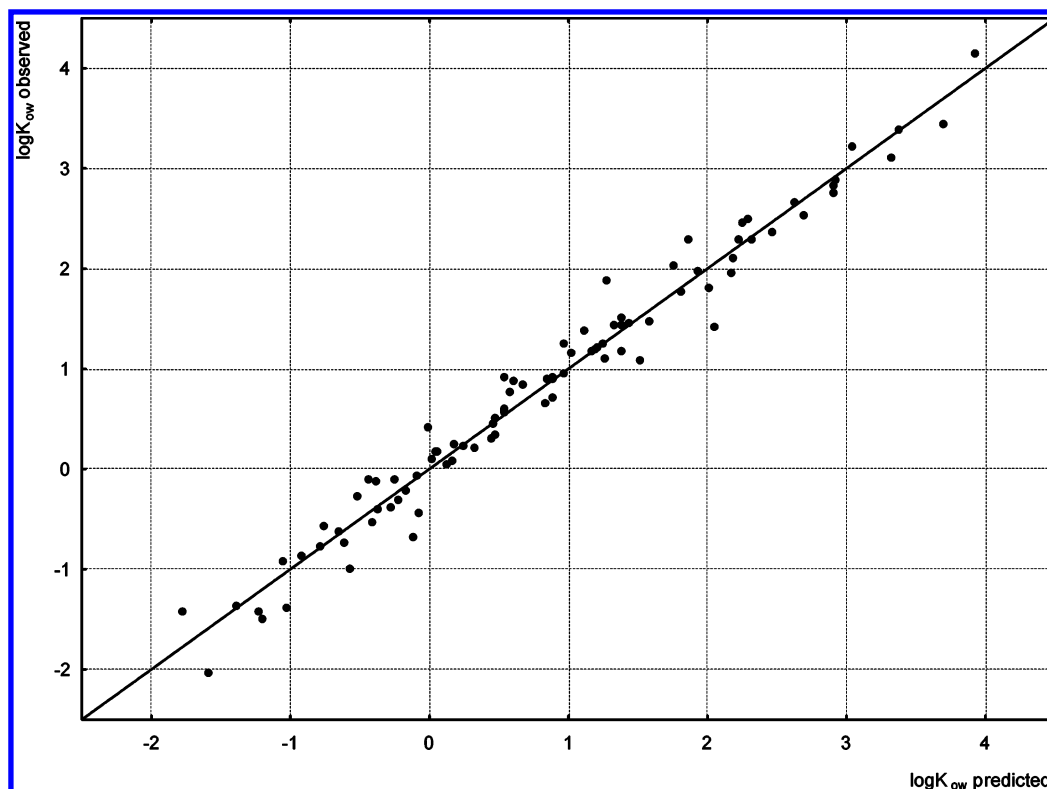


Figure 1. Plot of observed vs predicted octanol–water partition coefficients.

izes polarity of molecules) as well as the positive signs of the cavity-forming term (α) clearly demonstrate that the two phases are very different: preferential transfer to the alkane phase is inherent only for compounds of low polarity and polarizability.

Chloroform–Water. This partitioning system is very interesting because chloroform is almost unique with marked HB acidity but virtually zero basicity. This system has been also included into the “critical quartet” of water–solvent distributions by Leahy et al.⁵⁵ Eq (9) differs from eq (8) only by the values of the coefficients and the greater intercept.

$$\log K_{cw} = 1.086 - 1.645B^* - 0.819A^* + 1.295\alpha - 0.228LUMO \quad (9)$$

$$N = 34, R = 0.986, R^2 = 0.971, F = 246, s = 0.312$$

Gas–Water Partition Coefficients ($\log L_w$). The four-parameter regression model was derived from 60 experimental $\log L_w$ points. Three of the four terms of eq 10 contribute positively to the target property: (i) the hydrogen bond acidity and basicity terms A^* , B^* and (ii) and the integral polarity descriptor P^* . The HOMO energy, bearing a small negative contribution, serves as an additional measure of covalent HB basicity. No cavity-formation term enters eq 10 because of the low importance of such a term in partitioning in water.

$$\log L_w = 2.442 + 2.265B^* + 1.040A^* - 0.198HOMO + 0.144P^* \quad (10)$$

$$N = 69, R = 0.984, R^2 = 0.968, F = 484, s = 0.382$$

The linear fit of the predicted and observed Ostwald solubility coefficients for water is demonstrated in Figure

2. The only strong outlier in this fit is 2-chloroethanol, a multifunctional compound difficult to model.

Gas–Octanol Partition Coefficients ($\log L_{oct}$). In this case, we treated 47 experimental $\log L_{oct}$ points referring to water-saturated octanol. All the terms of eq 11 bear the positive sign, this bears witness to the approximately equal importance of HB, molecular volume, and dipolarity in partitioning to amphiphilic octanol, with the cavity-formation term being the most important.

$$\log L_o = 3.363 + 0.497B^* + 0.880A^* + 1.626\alpha + 0.375P^* \quad (11)$$

$$N = 47, R = 0.958, R^2 = 0.918, F = 118, s = 0.379$$

The smaller contribution of basicity term B^* to eq (11) is explained by markedly smaller HB acidity of octanol as compared to water.

Gas–Hexadecane Partition Coefficients ($\log L_h$). The Ostwald solubility in hexadecane is an important scale in LSER: $\log K_{hw}$ is used by Abraham as a permanent solute descriptor, and Leahy et al.⁵⁵ included this partition ratio into their “critical quartet”.

$$\log L_h = 2.467 + 0.237A^* + 1.438\alpha + 0.283HOMO + 0.466P^* \quad (12)$$

$$N = 65, R = 0.950, R^2 = 0.902, F = 137, s = 0.306$$

Eq 12 contains significant contributions from α polarizability and marked contributions from A^* and HOMO (partially responsible for general basicity) terms despite the poor, negligible acidity and basicity of hexadecane. This may be due either to entropic effects, facilitating the transfer to the liquid phase, and the intense dispersion interaction due

Table 4. Experimental and Predicted Values of Partition Coefficients for Octanol–Water, Hexadecane–Water, Chloroform–Water, Gas–Water, Gas–Octanol, Gas–Hexadecane, and Gas–Chloroform Systems

substance		log K_{ow}		log K_{hw}		log K_{cw}		log L_w		log L_o		log L_h		log L_c	
		exp	calc	exp ^a	calc	exp ^f	calc	exp	calc	exp	calc	exp	calc	exp ^f	calc
1	methane	1.09 ^a	1.52	1.14	1.00			-1.46 ^h	-1.74	-0.38 ^k	-0.15	-0.32 ⁱ	-0.38		
2	ethane	1.81 ^a	2.01	1.83	1.82			-1.34 ^h	-1.60	0.42 ^k	0.36	0.49 ⁱ	0.52		
3	propane	2.36 ^a	2.47	2.49	2.48			-1.44 ^h	-1.58	0.97 ^k	0.90	1.05 ⁱ	1.13		
4	butane	2.89 ^a	2.93	3.13	3.10			-1.52 ^h	-1.58	1.53 ^k	1.46	1.61 ⁱ	1.68		
5	2-methylpropane	2.76 ^a	2.91	3.11	3.09			-1.7 ^h	-1.63			1.41 ⁱ	1.62		
6	pentane	3.39 ^a	3.38	3.87	3.70			-1.7 ^h	-1.59	1.95 ^k	2.03	2.16 ⁱ	2.21		
7	2-methylbutane			3.76	3.68			-1.75 ^h	-1.62			2.01 ⁱ	2.16		
8	2,2-dimethylpropane	3.11 ^a	3.33	3.66	3.64			-1.84 ^h	-1.73			1.82 ⁱ	2.02		
9	cyclohexane	3.44 ^a	3.70	3.91	3.93	4.16	4.19	-0.9 ^h	-1.18	2.71 ^k	2.85	2.96 ⁱ	2.78	3.26	3.15
10	fluoromethane	0.51 ^a	0.47												
11	tetrafluoromethane	1.18 ^a	1.38	1.49	1.79			-2.29 ^g	-1.84	-0.95 ^k	0.07	-0.8 ^g	-0.28		
12	fluoroethane	1.26 ^b	0.97									0.56 ⁱ	0.78		
13	tetrachloromethane	2.83 ^a	2.91	2.88	3.66	3.31	3.99	-0.05 ^h	-0.58	2.79 ^k	2.96	2.82 ⁱ	3.34	3.25	3.53
14	chloromethane	0.91 ^a	0.88	0.76	0.51	1.42	1.37	0.44 ^h	0.40	1.39 ^k	1.14	1.16 ^g	0.97	1.82	2.10
15	dichloromethane	1.25 ^a	1.25	1.06	0.92	2	1.60	0.88 ^h	0.82	2.27 ^k	2.08	2.02 ⁱ	1.69	2.69	2.56
16	chloroform	1.97 ^a	1.93	1.69	1.50	2.28	1.96	0.82 ^h	1.21	2.8 ^k	3.45	2.48 ⁱ	2.86	3.07	3.03
17	chloroethane	1.43 ^a	1.32	1.22	1.27			0.47 ^h	0.23	1.9 ⁱ	1.50	1.68 ⁱ	1.32		
18	1,2-dichloroethane	1.48 ^a	1.58	1.26	1.59	2.13	2.22	1.28 ^h	0.58	2.78 ^k	2.28	2.57 ^g	1.92	3.44	3.03
19	1,1,1-trichloroethane	2.49 ^a	2.30	2.59	2.65	3.1	3.03	0.14 ^h	0.27	2.7 ^k	2.96	2.73 ⁱ	2.51	3.24	3.50
20	pentachloroethane	3.22 ^a	3.05	3.25	3.00			1.1 ^h	1.27			4.27 ^g	4.11		
21	hexachloroethane	4.14 ^a	3.93							4.47 ^k	4.16				
22	chloropropane	2.04 ^a	1.76	1.96	1.90	2.46	2.53	0.2 ^h	0.14	2.24 ^k	1.97	2.1 ⁱ	1.74	2.66	2.99
23	<i>t</i> -butyl chloride	2.45 ^b	2.25	3.02	2.76			-0.8 ^g	-0.21			2.22 ^g	2.12		
24	bromomethane	1.19 ^a	1.19	1.03	1.03			0.6 ^g	0.30			1.63 ^g	1.18		
25	dibromomethane	2.3 ^c	1.86	1.45	1.52			1.47 ^h	0.96	3.07 ^k	2.90	2.89 ⁱ	2.56		
26	bromoform	2.67 ^a	2.63	2.22	2.14			1.66 ^h	1.46			3.78 ⁱ	3.85		
27	<i>i</i> -propylbromide	2.1 ^a	2.18	2.04	2.57			0.35 ^g	0.00			2.39 ⁱ	2.12		
28	<i>t</i> -butylbromide	2.54 ^b	2.69	3.24	3.39			-0.22 ^h	-0.31			2.62 ^g	2.53		
29	bromodichloromethane	1.41 ^a	2.05	2.1	1.61										
30	iodomethane	1.51 ^a	1.38	1.46	1.36	2.13	2.05	0.65 ^g	0.34	2.16 ^k	1.57	2.11 ⁱ	1.78	2.78	2.47
31	diiodomethane	2.3 ^a	2.23									3.86 ⁱ	3.35		
32	halothane	2.3 ^a	2.32					0.08 ^g	1.03			2.18 ⁱ	2.77		
33	water	-1.38 ^d	-1.02	-4.38	-4.42	-3.1	-3.05	4.62 ^h	4.61	3.24 ⁱ	3.46	0.26 ⁱ	0.76	1.54	1.50
34	methanol	-0.74 ^a	-0.61	-2.77	-2.65	-1.33	-1.29	3.75 ^h	3.33	2.84 ^k	2.45	0.97 ⁱ	0.97	2.41	2.17
35	ethanol	-0.3 ^a	-0.22	-2.19	-1.96	-0.87	-0.70	3.67 ^h	3.25	3.2 ^k	2.85	1.49 ⁱ	1.44	2.8	2.65
36	1-propanol	0.25 ^a	0.18	-1.53	-1.47	-0.3	-0.32	3.56 ^h	3.38	3.68 ^k	3.50	2.03 ⁱ	1.95	3.26	3.09
37	2-propanol	0.05 ^a	0.12	-1.72	-1.53	-0.35	-0.36	3.48 ^h	3.39	3.38 ^k	3.46	1.81 ⁱ	1.91	3.13	3.09
38	1-butanol	0.88 ^a	0.61	-0.86	-0.89	0.42	0.16	3.44 ^h	3.37	4.19 ^k	4.06	2.6 ⁱ	2.45	3.88	3.53
39	2-butanol	0.61 ^a	0.54	-1.05	-0.96	0.3	0.11	3.43 ⁱ	3.39	3.8 ^k	4.02	2.34 ^g	2.40	3.69	3.52
40	2-methyl-1-propanol	0.76 ^a	0.58	-0.89	-0.93	0.34	0.13	3.3 ^g	3.37	3.93 ^k	4.03	2.41 ^g	2.41	3.64	3.51
41	2-methyl-2-propanol	0.35 ^a	0.47	-1.32	-1.03	-0.02	0.06	3.31 ⁱ	3.40	3.5 ^k	3.96	2.01 ⁱ	2.33	3.26	3.51
42	2-methyl-2-butanol	0.89 ^a	0.89	-0.62	-0.48			3.25 ^h	3.44			2.63 ^g	2.87		
43	3-methyl-1-butanol	1.16 ^a	1.02	-0.23	-0.34			3.24 ^g	3.38	4.52 ^k	4.60	3.01 ^g	2.93		
44	4-methyl-2-pentanol	1.43 ^b	1.39	0.44	0.15			2.74 ^h	3.40			3.18 ^g	3.39		
45	cyclopentanol	0.71 ^a	0.89	-0.79	-0.47			4.03 ^g	3.32			3.24 ⁱ	2.75		
46	2-chloroethanol	-0.06 ^a	-0.09			-0.4	-0.29	4.51 ⁱ	3.47	4.3 ^k	3.43				
47	2-bromoethanol	0.23 ^b	0.24												
48	2,2,2-trifluoroethanol	0.41 ^a	-0.01	-1.93	-1.50			3.15 ⁱ	3.13	3.56 ⁱ	3.26	1.23 ⁱ	1.29		
49	ethylene glycol	-1.36 ^a	-1.39					5.68 ⁱ	6.29	4.32 ⁱ	4.81	2.06 ⁱ	1.92		
50	propylene glycol	-0.92 ^e	-1.05			-2.9	-2.42								
51	<i>trans</i> -1,2-cyclohexanediol	0.08 ^d	0.16												
52	dimethyl ether	0.1 ^a	0.01					1.39 ⁱ	1.92	1.37 ^k	1.72	1.09 ⁱ	1.24		
53	dichloromethyl ether	0.57 ^b	0.54									2.02 ^g	2.16		
54	diethyl ether	0.89 ^a	0.85	0.85	0.43	1.88	1.52	1.17 ⁱ	1.94	2.19 ^k	2.81	2.06 ⁱ	2.29	3.05	3.73
55	2-methoxyethanol	-0.77 ^a	-0.78					4.96 ^g	4.91			2.49 ^g	2.16		
56	2-ethoxyethanol	-0.1 ^a	-0.43	-2.03	-2.23			4.91 ^g	4.97			2.81 ^g	2.62		
57	dimethoxymethane	0.18 ^d	0.05					2.18 ⁱ	2.75	2.36 ⁱ	2.86				
58	dimethoxyethane	-0.21 ^d	-0.16					3.55 ⁱ	3.41	3.34 ⁱ	3.37	2.66 ⁱ	2.31		
59	1,4-dioxane	-0.27 ^a	-0.51	-0.81	-1.39	0.73	0.14	3.71 ^g	4.02	3.17 ^k	3.50	2.8 ⁱ	2.51	4.44	4.44
60	1,3-dioxane	0.18 ^b	0.06												
61	1,3,5-trioxane	-0.43 ^e	-0.07												
62	tetrahydropyran	0.95 ^a	0.97	0.77	0.58	1.99	1.69	2.29 ⁱ	2.14	3.22 ^k	3.21	2.99 ⁱ	2.62	4.28	4.08
63	tetrahydrofuran	0.46 ^a	0.46	0.09	-0.07	1.31	1.14	2.57 ⁱ	2.17	2.86 ^k	2.54	2.64 ⁱ	2.07	3.86	3.59
64	ammonia					-1.38	-1.71	3.15 ^g	3.52			0.68 ^g	1.00	1.77	1.85
65	methylamine	-0.57 ^a	-0.76	-2.04	-2.28	-1.02	-0.75	3.35 ^h	3.12	1.9 ^k	1.90	1.3 ^g	1.28	2.32	2.49
66	ethylamine	-0.13 ^a	-0.38	-1.62	-1.69	-0.35	-0.26	3.3 ^h	3.12			1.68 ⁱ	1.71	2.95	2.95
67	dimethylamine	-0.38 ^a	-0.28	-1.55	-1.48	-0.44	-0.09	3.15 ^h	3.05	2 ^k	2.37	1.6 ⁱ	1.82	2.71	3.01
68	trimethylamine	0.22 ^a	0.33	-0.73	-0.23	0.51	1.04	2.37 ^h	2.34	2.53 ⁱ	2.37	1.62 ⁱ	2.24	2.86	3.50
69	diethylamine	0.58 ^a	0.54	-0.6	-0.28	0.79	0.91	2.98 ⁱ	3.06			2.4 ⁱ	2.82	3.78	3.94
70	triethylamine	1.45 ^a	1.43	0.72	1.38	1.86	2.39	2.22 ^h	2.20			3.08 ⁱ	3.40	4.22	4.72
71	N,N'-dimethylhydrazine	-0.54 ^b	-0.41												
72	1,2-diaminoethane	-2.04 ^e	-1.59												
73	1,3-diaminopropane	-1.43 ^b	-1.23												

Table 4 (Continued)

substance	log K_{ow}		log K_{hw}		log K_{cw}		log L_w		log L_o		log L_h		log L_c	
	exp	calc	exp ^a	calc	exp ^f	calc	exp	calc	exp	calc	exp	calc	exp ^f	calc
74 diethanolamine	-1.43 ^e	-1.77												
75 triethanolamine	-1 ^e	-0.57												
76 hydrogen sulfide							0.51 ^j	1.06			2.06 ⁱ	1.90		
77 methylmercaptane	0.65 ^a	0.84					0.91 ⁱ	0.98	1.7 ⁱ	1.35				
78 dimethyl sulfide	1.1 ^d	1.26	2.03	1.41			0.63 ^h	1.16	1.73 ⁱ	2.02	2.24 ⁱ	2.47		
79 ethylmercaptane	1.18 ^a	1.17	1.33	1.19			0.84 ^h	1.34			2.17 ⁱ	2.29		
80 diethyl sulfide	1.95 ^a	2.17	3.01	2.74	3.64	3.31	1.07 ^g	1.04			3.1 ⁱ	3.42	4.71	4.19
81 dimethyl disulfide	1.77 ^a	1.81					1.35 ⁱ	1.89			3.55 ⁱ	3.74		
82 1,2-dithioethane	1.21 ^b	1.21												
83 piperazine	-1.5 ^b	-1.20												
84 N,N'-dimethylethylenediamine	-0.62 ^b	-0.66												
85 tetramethylenediamine	0.3 ^b	0.44												
86 1-methylpiperidine	0.92 ^d	0.54					2.95 ⁱ	2.98	3.64 ^k	3.42				
87 piperidine	0.84 ^d	0.67			0.92	1.08	3.74 ⁱ	3.32	4.04 ^k	3.94			4.67	4.36
88 N-ethylpiperidine	1.88 ^a	1.27												
89 quinuclidine	1.38 ^b	1.11												
90 morpholine	-0.86 ^d	-0.92			-0.33	-0.51	5.26 ^h	4.90	4.4 ⁱ	3.97			4.93	4.59
91 1,4-dimethyl piperazine	-0.4 ^d	-0.38					5.56 ⁱ	5.46	5.16 ⁱ	5.15				
92 18-crown-6	-0.68 ^d	-0.11												
93 [2.2.2]cryptand	-0.1 ^d	-0.24												

^a Abraham, M. H.; Chadha, H. S.; Whiting, G. S.; Mitchell, R. C. Hydrogen bonding. 32. An Analysis of Water-Octanol and Water-Alkane Partitioning and the $\Delta\log P$ Parameter of Seiler. *J. Pharm. Sci.* **1994**, 83, 8, 1085–1100. ^b Howard, Ph. H.; Meylan, W. M. *Handbook of Phys. properties of organic chemicals*; 1997. ^c *Handbook of chem. and physics*; 1999–2000. Mackay, D.; Shin, W. Y.; Ma, K. C. *Illustrated handbook of phys-chem. properties and environmental fate for organic chemicals*; Lewis Pub./CRC Press: Boca Raton, FL, 1992. ^d Duffy, E. M.; Jorgensen, W. L. Prediction of Properties from Simulations: Free Energies of Solvation in Hexadecane, Octanol, and Water. *J. Am. Chem. Soc.* **2000**, 122, 2878–2888. ^e Sangster. *Octanol–water partition coeff*; 1997. ^f Abraham, M. H.; Platts, J. A.; Hersey, A.; Leo, A. J.; Taft, R. W. Correlation and Estimation of Gas-Chloroform and Water-Chloroform Partition Coefficients by a Linear Free Energy Relationship Method. *J. Pharm. Sci.* **1999**, 670–679. ^g Bordner, A. J.; Cavasotto, C. N.; Abagyan, R. A. Accurate Transferable Model for Water, *n*-Octanol, and Hexadecane Solvation Free Energies. *J. Phys. Chem. B* **2002**, 11009–11015. ^h Abraham, M. H.; Andonian-Haftvan, J.; Whiting, G. S.; Leo, A.; Taft, R. S. Hydrogen Bonding. Part 34. The Factors that Influence the Solubility of Gases and Vapors in Water at 298 K, and a New Method for its Determination. *J. Chem. Soc., Perkin Trans. 2* **1994**, 1771–1791. ⁱ Katritzky, A. R.; Tatham, D. B.; Maran, U. The Correlation of the Solubilities of Gases and Vapors in Methanol and Ethanol with their Molecular Structures. *J. Chem. Inf. Comput. Sci.* **2001**, 41, 358–363. ^j Giesen, D. J.; Gu, M. Z.; Cramer, C. J.; Truhlar, D. G. A Universal Organic Solvation Model. *J. Org. Chem.* **1996**, 61, 8720–8721. ^k Abraham, M. H.; Le, J.; Acree, Jr., W. E.; Car, P. W.; Dallas, A. J. The Solubility of Gases and Vapors in dry Octan-1-ol at 298 K. *Chemosphere* **2001**, 855–863.

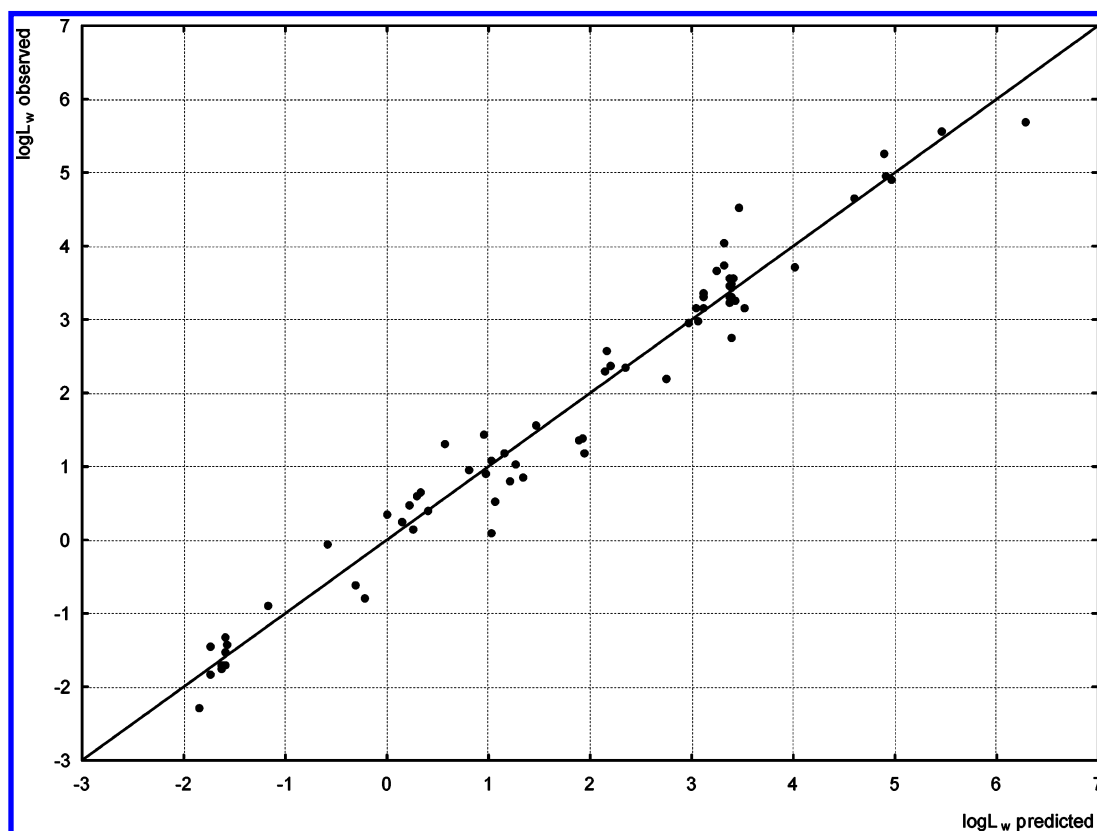


Figure 2. Plot of observed vs predicted gas–water partition coefficients.

Table 5. Correlation Matrix of the Descriptors Used in Eqs 7–13

	B^*	A^*	α	HOMO	LUMO
A^*	0.265				
α	0.724	0.114			
HOMO	0.460	0.080	0.383		
LUMO	0.210	0.202	0.029	0.108	
P^*	0.322	0.108	0.167	0.331	0.767

to the high hexadecane polarizability, which can likely disguise the specific solvation.

Gas–Chloroform Partition Coefficients ($\log L_c$). The transfer from the gas phase to chloroform occurs favorably for basic, highly polarizable molecules as can be seen from the data listed in Table 4.

$$\log L_c = 3.637 + 0.736B^* + 1.143\alpha - 0.228LUMO \quad (13)$$

$$N = 32, R = 0.934, R^2 = 0.872, F = 64, s = 0.319$$

Notably, eq 13 is virtually the inverse of eq 12: it contains B^* and LUMO while eq 12 contains A^* and HOMO. The presence of B^* and LUMO descriptors in eq 13 again attests the marked HB acidity of chloroform.

Looking at eqs 7–13, a general conclusion can be drawn that liquid phase–liquid phase equilibria are better modeled with QSPR/QSAR methods than gas–liquid phase. This is probably because of the experimental difficulties: the determination of a gas–solvent partition coefficient requires an exact value of solute's vapor pressure, which is frequently not available, especially for solids. Nevertheless, the above listed QSPR models of important solvation processes are statistically significant and physically relevant. To support the hypothesis of the models statistical significance the problem of mutual descriptor correlations should be addressed. The respective correlation matrix is presented in Table 5 which demonstrates a significant correlation for only two of the descriptor pairs correlation coefficients, r , of 0.724 for B^* and α polarizability and 0.767 for P^* and LUMO. The real multicollinearity of the QSAR equations being discussed is even smaller when expressed in terms of the

VIF, variance inflation factor:⁵⁶ only in one case is the VIF value equal to 3.2, in all other cases the VIF lies in the range from 1 to 2.

Performance of Our Descriptor Set in Comparison with Existing Methods. To make a qualified comparison one has to apply different methods (descriptor sets) to one and the same data set. We carefully inspected available modern data sets of Abraham's solvation descriptors and COSMO-derived descriptors of Klamt²⁴ (both presented in the same paper) and TLSER descriptors²⁹ and found that the maximum overlaps of our present data set of aliphatic compounds with the Abraham–Klamt data set and TLSER data set are 52 and 47 compounds, respectively. Based on these overlapped data sets we carried out the following comparative studies.

The numerical data are given in Table 6 for two partitioning systems, $\log K_{ow}$ and $\log L_w$. In Table 6 the Abraham descriptors are given first, then subsequently the COSMO-derived values, TLSER, and lastly our values. Numerical data of the former three descriptor sets are given in the Supporting Information Table 3S. As it is clearly seen, the Abraham general solvation equation expectedly demonstrates the best correlation results in both cases. This is true both in the sense of the whole regression equation features such as R^2 and the standard deviation (SD) and the individual performance of each descriptor characterized by the VIFs and the Student t -criteria. However, all four descriptor sets provide stable correlations since the individual performance of each of the descriptors expressed in terms of the VIF and t -crit. is quite good. At the same time, the explanatory power (estimated with respect to the values of R^2 and SD) of the COSMO descriptors and the TLSER descriptors is inferior to the Abraham solvation descriptors and the descriptors described in this paper. Although the performance of any descriptor set (LSER method) depends on a particular data set being used, it is clear that in the case of aliphatic compounds the descriptor set under development is ranked second after the Abraham solvation descriptors, with the COSMO descriptors and the TLSER descriptors being ranked third and fourth, respectively.

Table 6. Comparative Performance of Different Descriptor Sets on an Overlapping Data Set of Aliphatic Compounds

	N		R^2		SD		Fisher stat.		Student t -crit.		VIF	
	$\log P_{a/w}$	$\log L_w$	$\log P_{a/w}$	$\log L_w$	$\log P_{a/w}$	$\log L_w$	$\log P_{a/w}$	$\log L_w$	$\log P_{a/w}$	$\log L_w$	$\log P_{a/w}$	$\log L_w$
E	52	46	0.988	0.993	0.162	0.189	735	1068	4.979	4.846	2.315	2.754
S									−8.336	9.392	2.586	3.350
A									−2.870	19.478	1.530	1.650
B									−43.153	39.686	1.161	1.451
V									28.694	4.774	1.377	1.399
sig2	52	46	0.917	0.926	0.413	0.582	176	176	−10.846	9.643	1.494	2.768
sig3									−7.382		1.467	
Hb don3												
HB acc3										5.023		2.798
CSA									9.562	3.286	1.023	1.034
Vm	47	44	0.870	0.906	0.431	0.636	70	73	8.198	−2.614	1.149	1.100
pi										3.928		1.301
eB									−2.342	2.031	2.299	2.527
q−									−6.858	6.202	2.513	3.397
eA												
q+	52	46	0.980	0.971	0.202	0.369	591	344	−3.341	8.189	1.303	1.501
B^*									−30.640	17.199	1.589	1.889
A^*									−4.400	15.619	1.505	1.488
α									22.545		1.269	
HOMO										−3.568		1.526
P^*	47	44	0.976	0.970	0.186	0.355	420	314	6.527	2.204	1.157	1.157

CONCLUSIONS

Two theoretical scales, quantifying the hydrogen bond ability of compounds, are proposed: hydrogen bond basicity, B^* , and hydrogen bond acidity, A^* . Both these scales are based on clear physical considerations and use molecular topology and orbital valence-state energetic characteristics as parameters. Convincing correlations are shown with seven free energy-related partition properties. The relative performance of our descriptor set compared to other solvation descriptor sets has been estimated using a common data set. The next important step of our work will be the generalization of the theory to include aromatic and conjugated compounds. These new HB descriptors should significantly improve the performance of QSAR/QSPR software and be useful in molecular modeling and drug design. The latter becomes possible if one considers A_i^* and B_i^* values of each atom in the molecule rather than the overall molecular features. Two sets of such atomic values, $\{A_i^*\}$ and $\{B_i^*\}$, define two scalar potential fields of hydrogen bonding which can be used as molecular fields in CoMFA and/or scoring functions in ligand docking.

ACKNOWLEDGMENT

Financial support of this study from the U.S. Environmental Protection Agency is gratefully acknowledged. We thank Professor Mati Karelson (Tartu) for comments on this manuscript. We are also grateful to the reviewers of this paper for their valuable suggestions.

Supporting Information Available: Orbital electronegativity and hardness parameters (Table 1S), bond dipole moments and atomic polarizabilities (Table 2S), and numerical data of the descriptors belonging to the three descriptor systems (Table 3S). This material is available free of charge via the Internet at <http://pubs.acs.org>.

REFERENCES AND NOTES

- Jeffrey, G. A.; Saenger, W. *Hydrogen bonding in biological structures*; Springer: Berlin, 1991.
- Kamlet, M. J.; Taft, R. W. The solvatochromic comparison method. I. The beta-scale of solvent hydrogen-bond acceptor (HBA) basicities. *J. Am. Chem. Soc.* **1976**, *98*, 377–383.
- Kamlet, M. J.; Taft, R. W. The solvatochromic comparison method. II. The alpha-scale of solvent hydrogen-bond donor (HBA) acidities. *J. Am. Chem. Soc.* **1976**, *98*, 2886–2894.
- Kamlet, M. J.; Abboud, J. L. M.; Abraham, M. H.; Taft, R. W. Linear solvation energy relationships. 23. A comprehensive collection of the solvatochromic parameters, π^* , α , and β , and some methods for simplifying the generalized solvatochromic equation. *J. Org. Chem.* **1983**, *48*, 2877–2887.
- Hammett, L. P. *Physical organic chemistry; reaction rates, equilibria, and mechanisms*; McGraw-Hill: New York, 1970.
- Taft, R. W. In *Steric effects in organic chemistry*; Newman, M. S., Ed.; Wiley: New York, 1956.
- Taft, R. W.; Gurka, D.; Joris, L.; Schleyer, P. von R.; Rakshys, J. W. Studies of Hydrogen-Bonded Complex Formation with p-Fluorophenol. V. Linear Free Energy Relationships with OH Reference Acids. *J. Am. Chem. Soc.* **1969**, *91*, 4801–4808.
- Kamlet, M. J.; Abraham, M. H.; Doherty, R. M.; Taft, R. W. Solubility properties in polymers and biological media. 4. Correlation of octanol/water partition coefficients with solvatochromic parameters. *J. Am. Chem. Soc.* **1984**, *106*, 464–466.
- Abraham, M. H. Physicochemical and biological processes. *Chem. Soc. Rev.* **1993**, 73–83.
- Laurence, C.; Berthelot, M.; Helbert, M.; Sraidi, K. First measurement of the hydrogen bond basicity of monomeric water, phenols, and weakly basic alcohols. *J. Phys. Chem.* **1989**, *93*, 3799–3802.
- Laurence, C.; Berthelot, M. Observations on the strength of hydrogen bonding. *Perspect. Drug Discovery Des.* **2000**, *18*, 39–60.
- Raevsky, O. A. Quantification of noncovalent interactions on the basis of the thermodynamic hydrogen bond parameters. *J. Phys. Org. Chem.* **1997**, *10*, 405–413.
- Abraham, M. H.; McGowan, J. C. *Chromatographia* **1987**, *23*, 243–247.
- Weckwerth, J. D.; Vitha, M. F.; Carr, P. W. The development and determination of chemically distinct solute parameters for use in linear solvation energy relationships. *Fluid Phase Equil.* **2001**, *183–184*, 143–157.
- Abraham, M. H.; Platts, J. A. Hydrogen bond structural group constants. *J. Org. Chem.* **2001**, *66*, 3484–3491.
- Platts, J. A.; Butina, D.; Abraham, M. H.; Hersey, A. Estimation of molecular linear free energy relation descriptors using a group contribution approach. *J. Chem. Inf. Comput. Sci.*, **1999**, *39*, 835–845.
- ABSOLV program, Sirius Analytical Instruments Limited, UK. <http://www.sirius-analytical.com>.
- Platts, J. A.; Abraham, M. H.; Butina, D.; Hersey, A. Estimation of molecular linear free energy relation descriptors using a group contribution approach. 2. Prediction of partition coefficients. *J. Chem. Inf. Comput. Sci.* **2000**, *40*, 71–80.
- Dearden, J. C.; Bentley, D. The components of the “critical quartet” LOGKOW values assessed by four commercial software packages. *SAR QSAR Environ. Res.* **2002**, *13*, 185–197.
- Hickey, J. P.; Passino-Reader, D. R. Linear solvation energy relationships: “Rules of thumb” for estimation of variable values. *Environ. Sci. Technol.* **1991**, *25*, 1753–1760.
- Lamarche, O.; Platts, J. A. Theoretical prediction of the hydrogen-bond basicity $pK(\text{HB})$. *Chem. Eur. J.* **2002**, *8*, 457–466.
- Lamarche, O.; Platts, J. A.; Hersey, A. Theoretical prediction of the polarity/polarizability parameter π_{H}^2 . *Phys. Chem. Chem. Phys.* **2001**, *3*, 2747–2753.
- Platts, J. A.; Theoretical prediction of hydrogen bond donor capacity. *Phys. Chem. Chem. Phys.* **2000**, *2*, 973–980.
- Zissimos, A. M.; Abraham, M. H.; Klamt, A.; Eckert, F.; Wood, J. A comparison between the two general sets of linear free energy descriptors of Abraham and Klamt. *J. Chem. Inf. Comput. Sci.* **2002**, *42*, 1320–1331.
- Klamt, A. Conductor-like screening model for real solvents: A new approach to the quantitative calculation of solvation phenomena. *J. Phys. Chem.* **1995**, *99*, 2224–2235.
- Duffy, E. M.; Jorgensen, W. L. Prediction of Properties from Simulations: Free Energies of Solvation in Hexadecane, Octanol, and Water. *J. Am. Chem. Soc.* **2000**, *122*, 2878–2888.
- Bohm, M.; Klebe, G. Development of new hydrogen bond descriptors and their application to comparative molecular field analysis. *J. Med. Chem.* **2002**, *45*, 1585–1597.
- Famini, G. R.; Wilson, L. Y. Using Theoretical Descriptors in Linear Solvation Energy Relationships. In *Quantitative Treatments of Solute/Solvent Interactions*; Politzer, P., Murray, J. S., Eds.; Elsevier: Amsterdam, 1994; pp 213–242.
- Cronce, D. T.; Famini, G. R.; De Soto, J. A.; Wilson, L. Using theoretical descriptors in quantitative structure-property relationships: some distribution equilibria. *J. Chem. Soc., Perkin Trans. 2* **1998**, 1293–1301.
- Murray, J. S.; Politzer, P. In *Quantitative Treatments of Solute/Solvent Interactions*; Politzer, P., Murray, J. S., Eds.; Elsevier, Amsterdam, 1994; pp 243–289.
- Politzer, P.; Murray, J. S.; Peralta-Inga, Z. Molecular surface electrostatic potential in relation to noncovalent interactions in biological systems. *J. Comput. Chem.* **2001**, *22*, 676–684.
- Eisfeld, W.; Maurer, G. Study on the correlation and prediction of octanol/water partition coefficients by quantum chemical calculations. *J. Phys. Chem. B* **1999**, *103*, 5716–5729.
- Haeblerlein, M.; Brinck, T. Prediction of water–octanol partition coefficients using theoretical descriptors derived from the molecular surface area and the electrostatic potential. *J. Chem. Soc., Perkin Trans. 2* **1997**, 289–294.
- Dearden, J. C.; Ghafourian, T. Hydrogen bonding parameters for QSAR: Comparison of indicator variables, hydrogen bond counts, molecular orbital and other parameters. *J. Chem. Inf. Comput. Sci.* **1999**, *39*, 231–235.
- Kamlet, M. J.; Doherty, R. M.; Abraham, M. H.; Carr, P. W.; Doherty, R. F.; Taft, R. W. Linear solvation energy relationships. 41. Important differences between aqueous solubility relationships for aliphatic and aromatic solutes. *J. Phys. Chem.* **1987**, *91*, 1996–2004.
- Reichardt, C. *Solvents and Solvent Effects in Organic Chemistry*; VCH: Weinheim, 1988.
- Nishio, M.; Hirota, M. *The CH–[pi] interaction: evidence, nature, and consequences*; Wiley: New York, c1998.

- (38) Oliferenko, A. A.; Palyulin, V. A.; Pisarev, S. A.; Neiman, A. V.; Zefirov, N. S. Novel point charge models: Reliable instrument for molecular electrostatics. *J. Phys. Org. Chem.* **2001**, *14*, 355–369.
- (39) Ouvrard, C.; Berthelot, M.; Laurence, C. The first basicity scale of fluoro-, chloro-, bromo-, and iodo-alkanes: some cross-comparisons with simple alkyl derivatives of other elements. *J. Chem. Soc., Perkin Trans. 2* **1999**, 1357–1362.
- (40) Hinze, J.; Jaffé, H. H. Electronegativity. I. Orbital electronegativity of neutral atoms. *J. Am. Chem. Soc.* **1961**, *84*, 540–546.
- (41) Bergmann, D.; Hinze, J. Electronegativity and molecular properties. *Angew. Chem.* **1996**, *35*, 150–163.
- (42) Bader, R. F. W.; Johnson, S.; Tang, T.-H.; Popelier, P. L. A. The electron pair. *J. Phys. Chem.* **1996**, *100*, 15398–15415.
- (43) Oliferenko, A. A.; Krylenko, P. V.; Palyulin, V. A.; Zefirov, N. S. A New Scheme for Electronegativity Equalization as a Source of Electronic Descriptors: Application to Chemical Reactivity. *SAR QSAR Environ. Res.* **2002**, *13*, 297–305.
- (44) Popelier, P. L. A.; Smith, P. J. Protonation energies and the Laplacian of the electron density: A critical examination. *Phys. Chem. Chem. Phys.* **2001**, 4208–4212.
- (45) Berthelot, M.; Graton, J.; Ouvrard, C.; Laurence, C. Hydrogen-bond basicity of solutes in hydroxylic solvents from octanol–water partition coefficients. *J. Phys. Org. Chem.* **2002**, *15*, 218–228.
- (46) Bonchev, D.; Rouvray, D. H. *Chemical topology: applications and techniques*; Gordon and Breach Science Publishers: Amsterdam, Netherlands, 2000.
- (47) Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; and Stewart, J. J. P. AM1: A New General Purpose Quantum Mechanical Molecular Model. *J. Am. Chem. Soc.* **1985**, *107*, 3902–3909.
- (48) HyperChem 7.0, HyperCube, 2002. <http://www.hypercube.com>
- (49) Gordon, A. J. *The chemist's companion: a handbook of practical data, techniques, and references*; Wiley: New York, 1972.
- (50) Katritzky, A. R.; Maran, U.; Lobanov, V. S.; Karelson, M. Structurally Diverse Quantitative Structure–Property Relationship Correlations of Technologically Relevant Physical Properties. *J. Chem. Inf. Comput. Sci.* **2000**, *40*, 1–18.
- (51) <http://www.codessa-pro.com>.
- (52) Leo, A.; Hansch, C.; Elkins, D. Partition coefficients and their uses. *Chem. Rev.* **1971**, *71*, 525–616.
- (53) Minkin, V. I.; Osipov, O. A.; Zhdanov, Yu. A. *Dipole Moments in Organic Chemistry*; Khimia: Moscow, 1968.
- (54) Miller, K. J. Additivity methods in molecular polarizability. *J. Am. Chem. Soc.* **1990**, *112*, 3–8542.
- (55) Leahy, D. E.; Morris, J. J.; Taylor, P. J.; Wait, A. R. Model solvent system for QSAR. Part 2. Fragment values (f-values) for the “critical quartet”. *J. Chem. Soc., Perkin Trans. 2* **1992**, 723–731.
- (56) Neter, J.; Wasserman, W.; Kutner, M. H. *Applied linear statistical models: Regression, analysis of variance, and experimental designs*; Irwin: Homewood, IL, 1985.

CI0342932