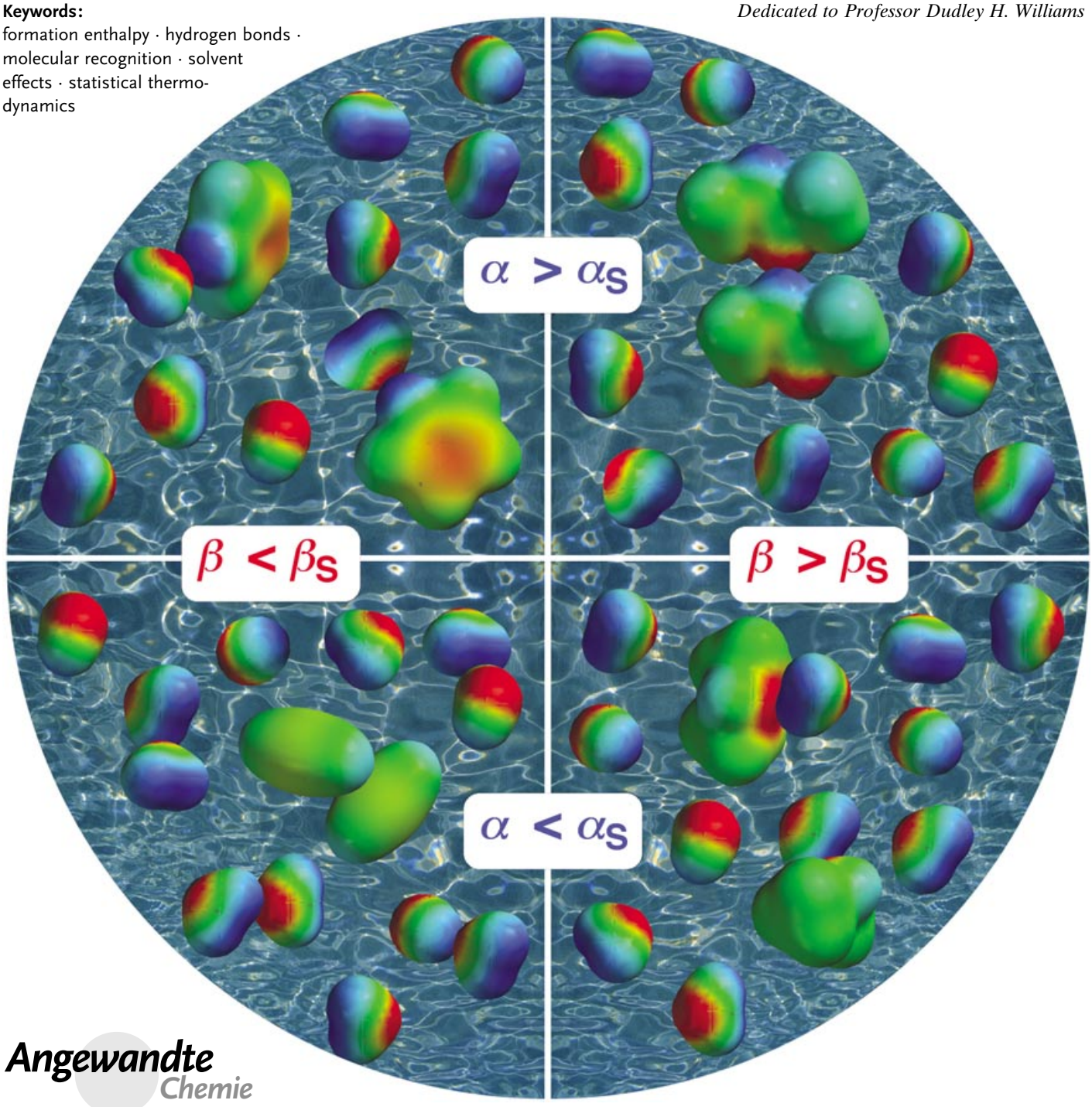
Molecular Recognition

Quantifying Intermolecular Interactions: Guidelines for the Molecular Recognition Toolbox

*Christopher A. Hunter**

Keywords:
formation enthalpy · hydrogen bonds · molecular recognition · solvent effects · statistical thermodynamics

Dedicated to Professor Dudley H. Williams



DOI: 10.1002/anie.200301739

Molecular recognition events in solution are affected by many different factors that have hampered the development of an understanding of intermolecular interactions at a quantitative level. Our tendency is to partition these effects into discrete phenomenological fields that are classified, named, and divorced: aromatic interactions, cation–π interactions, CH–O hydrogen bonds, short strong hydrogen bonds, and hydrophobic interactions to name a few.[1] To progress in the field, we need to develop an integrated quantitative appreciation of the relative magnitudes of all of the different effects that might influence the molecular recognition behavior of a given system. In an effort to navigate undergraduates through the vast and sometimes contradictory literature on the subject, I have developed an approach that treats theoretical ideas and experimental observations about intermolecular interactions in the gas phase, the solid state, and solution from a single simplistic viewpoint. The key features are outlined here, and although many of the ideas will be familiar, the aim is to provide a semiquantitative thermodynamic ranking of these effects in solution at room temperature.

From the Contents

1. Introduction

Let us first deal with the issue of free energy (Gibbs free energy, G), enthalpy, and entropy. Clearly, these three parameters are intimately related, but the behavior of a system is determined by the free energy, and the experimental values are simple to interpret: is the complex more stable or not? Although the concepts of entropy and enthalpy are useful for understanding the various factors that contribute to the free energy of an intermolecular interaction, interpretation of the experimental values of ΔH and ΔS is problematic, even for simple systems. The enthalpy and entropy changes observed for intermolecular interactions can fluctuate wildly when relatively small changes are made to the system, but in such a way that the two parameters compensate one another, and the associated variation in free energy is small.[2] A phase change is the classic scenario where there are large enthalpy and entropy changes that compensate each other precisely. The values of ΔH and ΔS tell us about changes in the internal structure or organization of the system and are not necessarily related to the overall stability. The thermodynamics of the hydrophobic effect illustrate the point.[3] The free energy change on dissolving a nonpolar molecule in water is large and positive across the entire temperature range from 0 to 100 °C and does not fluctuate significantly. However, the associated entropy and enthalpy changes vary dramatically, so that the process is dominated by entropy at low temperatures and enthalpy at high temperatures. This is not to say that enthalpy and entropy are not useful parameters. The information they contain about changes in structure is important, for example, in understanding the binding of agonists and antagonists to biological receptors, but it is something quite different from overall stability or binding affinity.[4]

We will therefore focus on free energy as the key observable that is important for understanding intermolecular interactions in solution. The contributions from factors such as the unfavorable entropy change associated with bringing two molecules together to form a complex or the favorable enthalpy change associated with the formation of a hydrogen bond will all be considered as free energy terms. Methods for parsing free energy in this manner have been relatively successful, suggesting that free energy contributions are indeed additive to a first approximation.[5]

2. Theoretical Considerations

The basic theory of intermolecular interactions in the gas phase separates the enthalpy of interaction between two molecules into four components: repulsion between the electron densities at close distances of approach; induction interactions between the permanent charge distribution of one molecule and an induced change in the charge distribution of the other molecule; dispersion interactions between

[*] Prof. C. A. Hunter
Centre for Chemical Biology
Krebs Institute for Biomolecular Science
Department of Chemistry
University of Sheffield
Sheffield S3 7HF (UK)
Fax: (+44) 114-273-8673
E-mail: c.hunter@shef.ac.uk

 DOI: 10.1002/anie.200301739

mutually induced dipoles, and electrostatic interactions between the permanent charge distributions of the two molecules.[6,7] We will consider each of these factors in turn.

2.1. Repulsion

The repulsive interactions between electron densities simply define the molecular volume, and for molecules in van der Waals contact, we can ignore the contribution of the repulsion term to differences in intermolecular interaction energies.

2.2 Induction

There is some evidence for induction in complexes of organic molecules, but the magnitude of the effect is relatively small.[8] For example, urea forms linear hydrogen-bonded polymers in organic solvents with two hydrogen bonds between each molecule (Figure 1).[9] The first step in assembly of the polymer is formation of the urea dimer, which polarizes the molecules, so that K_2 for adding a third urea unit to the chain is larger than K_1. However, the magnitude of the effect is a small fraction of the hydrogen-bond energy (1 kJ mol^{-1} change in 20–25 kJ mol^{-1} for the urea–urea hydrogen bond, see Section 5). Therefore to a first approximation, we can neglect the effects of induction.

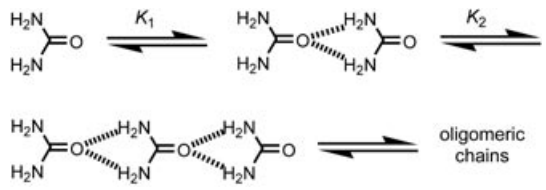


Figure 1. Urea forms a linear hydrogen-bonded polymer in benzene. The value of K_1 is 400 M^{-1}, and the value of K_2 is 900 M^{-1}. The difference is attributed to polarization of the molecules by the hydrogen bonds in the dimer that make the second set of hydrogen bonds in the trimer stronger.

2.3 Dispersion

The magnitude of the dispersion interaction between two atoms i and j (E_{ij}) depends on the atomic polarizabilities (α_i and α_j) and the number of valence electrons (N_i and N_j). If the atoms are separated by a distance r_{ij}, E_{ij} can be estimated by using the Slater–Kirkwood method [Eq. (1) and Eq. (2); c is a constant].[10]

$$E_{ij} = \frac{B_{ij}}{r_{ij}^6} \quad (1)$$

$$B_{ij} = c\,\frac{\alpha_i\,\alpha_j}{\sqrt{\alpha_i/N_i} + \sqrt{\alpha_j/N_j}} \quad (2)$$

Table 1 shows the values of the constant, B_{ij}, for some atoms of the second and third periods. The larger, softer atoms have larger B_{ij} values as expected, but this does not necessarily equate to larger dispersion interactions. Larger atoms are further apart, when they are in van der Waals contact, and the dispersion energy depends on r_{ij}^{-6} [Eq. (1)]. Moreover in solution, there are competitive dispersion interactions with the solvent, and because large atoms make more solvent contacts than small atoms, they displace more solvent when they form a complex (Figure 2). In solution, the surfaces of all molecules are fully coated by other molecules, and the equilibrium in Figure 2 simply represents a rearrangement of the molecular surfaces that are in contact. The relevant parameter for comparing dispersion interactions in solution is therefore the interaction energy per unit surface area of contact. Table 2 shows that there is remarkably little

Table 1: Relative values of B_{ij} for some atoms of the second and third period.[a]

	C	N	O	F	P	S	Cl
C	1.0	0.7	0.6	0.4	1.8	1.6	1.3
N	0.7	0.6	0.5	0.4	1.3	1.2	1.0
O	0.6	0.5	0.4	0.3	1.0	0.9	0.8
F	0.4	0.4	0.3	0.2	0.8	0.7	0.6
P	1.8	1.3	1.0	0.8	3.3	2.9	2.4
S	1.6	1.2	0.9	0.7	2.9	2.6	2.2
Cl	1.3	1.0	0.8	0.6	2.4	2.2	1.8

[a] Relative to the C···C interaction. The values were calculated by using Equation (2), and atomic polarizabilities, α_i and α_j, were taken from reference [10b].

Table 2: Relative values of the dispersion interaction energy per unit surface area of contact for some atoms of the second and third period.[a]

	C	N	O	F	P	S	Cl
C	1.0	1.0	1.1	1.1	1.5	1.4	1.1
N	1.0	0.9	1.0	1.1	1.5	1.3	1.1
O	1.1	1.0	1.0	1.1	1.6	1.4	1.2
F	1.1	1.1	1.1	1.1	1.6	1.5	1.3
P	1.5	1.5	1.6	1.6	2.1	1.8	1.5
S	1.4	1.3	1.4	1.5	1.8	1.6	1.4
Cl	1.1	1.1	1.2	1.3	1.5	1.4	1.2

[a] The dispersion energy was calculated by using Equation (1) and atomic radii were taken from reference [10b]. The surface area of contact was taken to be proportional to the surface area of the smaller of the two atoms.

Chris Hunter was educated at the University of Cambridge and graduated with a PhD in 1989. He was a lecturer at the University of Otago from 1989 till 1991. In 1991 he moved to the University of Sheffield, where he is currently Professor of Chemistry. He has research interests in various aspects of molecular recognition, design, synthesis, measurement, and theory.

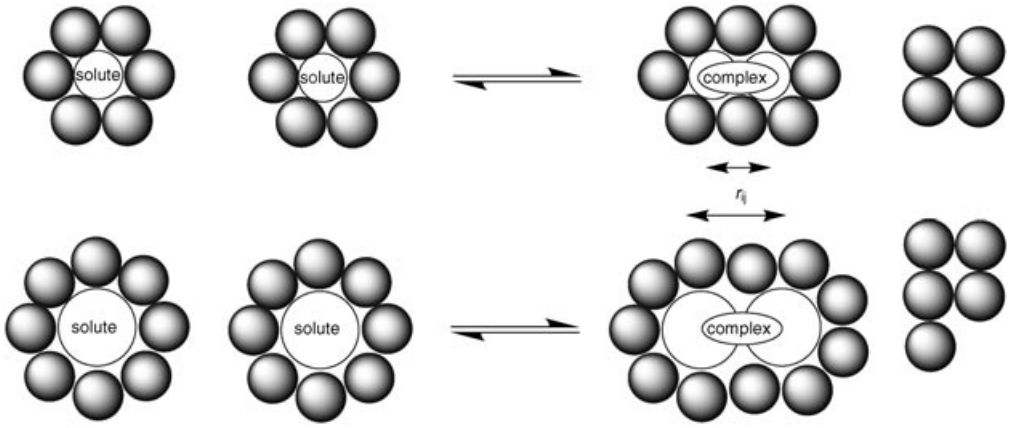


Figure 2. Although B_{ij} is larger for larger more polarizable atoms, the nuclear separation r_{ij} is also larger, and the amount of solvent that is displaced is larger. These effects cancel each other out to a first approximation. The dispersion interaction energy per unit surface area of contact is almost constant and independent of atom type, and the total contact surface area does not change for the equilibrium shown, so the change in dispersion energy for the interaction of two molecules in solution is small.

variation of this parameter with atom type.[11] The interactions with third-period atoms are slightly larger than for the second period, but to a first approximation, we can assume that the change in the dispersion energy is negligible, when two shape complementary molecules interact in solution.[12] If Cl···Cl interactions were significantly more favorable than C···C interactions, carbon tetrachloride would be a poor solvent for nonpolar organic molecules, which is not the case. The dispersion interactions per unit surface area for C···C, C···Cl and Cl···Cl have very similar values, as expected from the experimental behavior. Since the dispersion energy for the interaction of two methyl groups is about 2 kJ mol^{-1}, the small variations in Table 2 do not translate to large differences in free energy.[13]

2.4. Electrostatics

Since repulsion, induction, and dispersion contribute only negligibly to intermolecular interactions, this leaves us with the prospect of using only electrostatics to explain everything. This not only makes life very simple, but there is good experimental evidence for the dominant role of electrostatics in intermolecular interactions. Physical organic chemists have measured the association constants (K) for a huge number of intermolecular interactions in the gas phase and in solution. For simple molecules, these data can be analyzed as simple pairwise hydrogen-bonding interactions between two functional groups, and remarkably, the results can be accounted for by a single simple relationship [Eq. (3)].[14]

$$\log K = c_1\,\alpha_2^{\mathrm{H}}\,\beta_2^{\mathrm{H}} + c_2 \qquad (3)$$

c_1 and c_2 are constants that depend on the solvent, and α_2^{H} and β_2^{H} are functional group constants that relate to the hydrogen-bond donor and hydrogen-bond acceptor properties of the molecules.

Equation (3) is equivalent to an expression of the electrostatics of the hydrogen-bonding interaction, where the free energy of interaction varies as the product of the positive charge on the hydrogen-bond donor (α_2^{H}) and the negative charge on the hydrogen-bond acceptor (β_2^{H}).[15] Indeed, the values of α_2^{H} and β_2^{H} have been correlated with a variety of computed molecular properties, such as atomic charge and electrostatic potential.[16] The constant c_1 is solvent dependent: it increases as the polarity of the medium decreases, as expected for electrostatic interactions. The constant c_2 is -1.0 ± 0.1 and is relatively insensitive to solvent, which implies that it is a fundamental property of the interaction between any two molecules.

3. The Cost of Restricting Relative Molecular Motions

If we consider the interaction between two nonpolar entities, such as neon atoms, where $\alpha_2^{\mathrm{H}} = \beta_2^{\mathrm{H}} = 0$, Equation (3) gives an association constant of 0.1 M^{-1}. In other words, complexation between two neon atoms becomes significant at concentrations above 10 M, even though there are no polar groups to form hydrogen bonds. The reason is that 10 M is approximately the concentration of the bulk liquid state, where by definition there are extensive intermolecular contacts. The c_2 term therefore represents the adverse free energy associated with bringing two molecules together to form a noncovalent complex in solution. The magnitude of the effect is only +6 kJ mol^{-1}, which is significantly lower than the +60 kJ mol^{-1} associated with the adverse free energy for formation of a covalent complex.[17] The difference reflects the looseness of noncovalent associations. The formation of noncovalent complexes is intrinsically much more favorable than the formation of covalent complexes, because the relative molecular motions are not restricted to the same extent. This value of c_2 is consistent with an empirical analysis of small molecules binding to biological receptors in water that found a value of +5.4 kJ mol^{-1} for the free energy cost of restricting relative molecular motions.[18]

One consequence is that the upper limit for the effective molarities one might expect to see for systems that make multiple intermolecular interactions is rather low, of the order of 10 M. The effective molarities of 10^{10} M that have been reported for intramolecular reactions clearly do not translate to noncovalent interactions.[19] Although the adverse free energy associated with bringing molecules together to make a

noncovalent complex is something we would generally consider to be "an entropic effect", the magnitude of the effect bears no relation to measured values of the entropy of complexation.[4] Experimental entropy measurements include additional contributions from desolvation and changes in internal structure.

4. Molecular Electrostatic Potential Surfaces

Equation (3) holds for a remarkably wide range of functional groups, and this implies that, apart from some notable exceptions such as aromatic stacking, we can treat all intermolecular interactions as a form of hydrogen bonding.[14] The reason is that the maximum in the electrostatic potential on the van der Waals surface of a molecule is usually located near a hydrogen atom and the minimum is usually over a lone pair or an area of π-electron density (Figure 3a). The dominant electrostatic interactions between two molecules

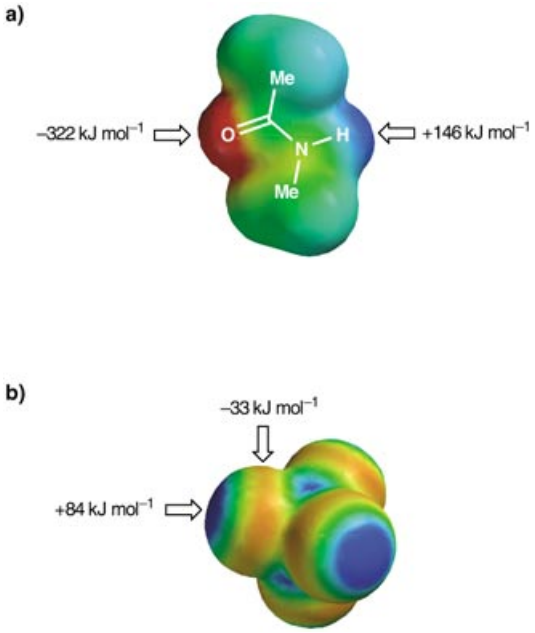


Figure 3. Molecular electrostatic potential surfaces plotted on the van der Waals' surface of the molecule calculated by using AM1 and a positive point charge in a vacuum as the probe. a) *N*-methyl acetamide. Positive regions are shown in blue ($>+150$ kJ mol^{-1}), negative regions are shown in red (<-150 kJ mol^{-1}), and green is neutral. The maximum in the electrostatic potential, E_{max}, lies over the NH group, and the minimum, E_{min}, lies over the carbonyl oxygen atom, the two primary hydrogen-bonding sites in the molecule. b) Carbon tetrachloride. Positive regions are shown in blue ($>+35$ kJ mol^{-1}), negative regions are shown in red (<-35 kJ mol^{-1}), and green is neutral.

are pairwise interactions between these maxima and minima that may generally be considered to be of a hydrogen-bonding nature. Even at a relatively low level of theory, the calculated maxima and minima in the molecular electrostatic potential surfaces correlate well with the experimentally determined values of α_2^H and β_2^H (Figure 4).[20] The intercept for the β_2^H plot is close to the origin, but the α_2^H plot crosses the electrostatic potential axis at $+70$ kJ mol^{-1}, and all functional groups with lower E_{max} values have been assigned an α_2^H value of zero. The reasons for this can be found in the properties of the solvent used to develop the α_2^H and β_2^H scales (carbon tetrachloride).

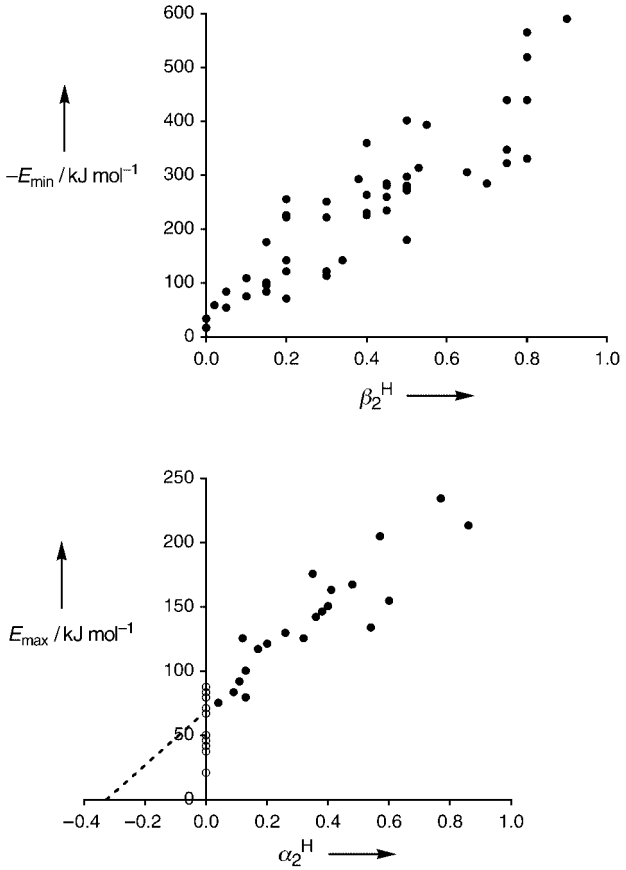


Figure 4. The maxima (E_{max}) and minima (E_{min}) in the AM1 molecular electrostatic potential surfaces of a range of simple molecules containing only one functional group plotted against the corresponding experimentally determined values of α_2^H and β_2^H from reference [14]. Electrostatic potentials in sterically inaccessible crevices on the surface of a tetrahedrally coordinated central atom (e.g. phosphorus) were ignored. In some cases, the locations of the maxima and minima in the molecular electrostatic potential surface do not necessarily correspond to conventional hydrogen-bonding sites (e.g. for carbon tetrachloride shown in Figure 3). The open circles correspond to functional groups that do not form hydrogen-bonded complexes in carbon tetrachloride and have been assigned a value of $\alpha_2^H = 0$. The dotted line is an extrapolation of the α_2^H data to encompass nonpolar hydrogen-bond donors that have values of E_{max} lower than that of carbon tetrachloride. The best fit straight lines give $E_{min} = -535$ ($\beta_2^H + 0.06$ kJ mol^{-1}) and $E_{max} = 211$ ($\alpha_2^H + 0.33$ kJ mol^{-1}).

The molecular electrostatic potential surface of carbon tetrachloride is shown in Figure 3b, and although there are no hydrogen-bond donors, the surface is quite strongly positive. In third period elements, the lone pair electrons are diffuse and relatively far from the nucleus, so that the positive nuclear charge is not well screened. Combined with the

strongly electron withdrawing nature of the CCl_3 group, this gives the chlorine atoms of carbon tetrachloride electrostatic properties equivalent to a weak hydrogen-bond donor. The value of E_{max} for carbon tetrachloride is $+84\ kJ\,mol^{-1}$, which is very close to the point at which the α_2^H plot crosses the E_{max} axis in Figure 4. All less polar hydrogen-bond donors are unable to compete with the solvent for solute hydrogen-bonding sites, and so it is impossible to detect interactions with weak hydrogen-bond donors in this solvent. The origin of the α_2^H scale has been set by the properties of carbon tetrachloride, and if we extrapolate the experimental α_2^H data, it is clear that the origin should be at $\alpha_2^H = -0.33$. This would allow us to incorporate all functional groups, including relatively nonpolar hydrogen-bond donors such as C–H groups. The value of E_{min} for carbon tetrachloride is $-33\ kJ\,mol^{-1}$, and so the β_2^H scale is only displaced by 0.06, which is a relatively small discrepancy.

5. Desolvation

The correlations in Figure 4 provide an important clue how to treat interactions in solution. There is a competition between solute–solute, solvent–solvent, and solute–solvent interactions (Figure 5), and if we change our frame of reference, we can produce a universal hydrogen-bond scale that can be used to predict the free energy of hydrogen-bonding interactions ($\Delta\Delta G_{H\text{-bond}}$ in $kJ\,mol^{-1}$) in any solvent [Eq. (4)].

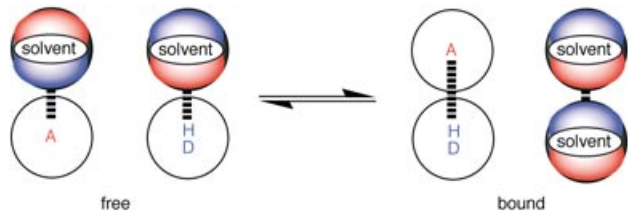


Figure 5. Intermolecular interactions in solution are a competition between solute–solvent interactions in the free state, and solute–solute and solvent–solvent interactions in the bound state. For simple functional groups, the primary mode of interaction is hydrogen-bond contacts between the maxima (blue) and minima (red) in the electrostatic potential surfaces of the molecules. A represents a hydrogen-bond acceptor solute and DH a hydrogen-bond donor solute.

$$\begin{aligned}\Delta\Delta G_{H\text{-bond}} &= -(\alpha\beta + \alpha_S\beta_S) + (\alpha\beta_S + \alpha_S\beta) \\ &= -(\alpha-\alpha_S)(\beta-\beta_S)\end{aligned} \quad (4)$$

α and β are hydrogen-bond donor and hydrogen-bond acceptor constants for the solute molecules, and α_S and β_S are the corresponding hydrogen-bond donor and hydrogen-bond acceptor constants for the solvent.

The new hydrogen-bond parameters, α and β, correspond to normalized dimensionless versions of E_{max} and E_{min} [Eq. (5) and Eq. (6)]. The normalization constant of $52\ kJ\,mol^{-1}$ was obtained from the slopes of the correlations in Figure 4 and the gas phase value of c_1 in Equation (3) such that Equation (4) gives $\Delta\Delta G_{H\text{-bond}}$ in units of $kJ\,mol^{-1}$. Where experimental values of α_2^H and β_2^H are available, they can be converted to the new α and β scale using Equation (5) and (6) which are based on the correlations shown in Figure 4. If experimental data are not available, α and β can be estimated from calculated AM1 molecular electrostatic potential surfaces by using Equation (5) and (6). The scatter in the correlations in Figure 4 provides an indication of the accuracy that can be achieved by using E_{max} and E_{min} values. Some classes of functional groups, for example, amides, behave consistently poorly, and the accuracy of the α and β values obtained from electrostatic potential calculations is of the order $\pm 20\,\%$. A list of values of α and β for common functional groups is provided in Table 3.

$$\alpha = E_{max}/52\ kJ\,mol^{-1} = 4.1(\alpha_2^H + 0.33) \quad (5)$$

$$\beta = -E_{min}/52\ kJ\,mol^{-1} = 10.3(\beta_2^H + 0.06) \quad (6)$$

Equation (4) is the equivalent of the first term in Equation (3), and to estimate the overall free energy of complexation, the adverse free energy associated with bimolecular complexation ($\approx +6\ kJ\,mol^{-1}$) should be added to $\Delta\Delta G_{H\text{-bond}}$. In this scheme, α and β represent the positive and negative parts of a continuous scale, so that the magnitudes of repulsive interactions can also be estimated. For example, the adverse free energy associated with the interaction between two hydrogen-bond donors forced into van der Waals contact is given by $\alpha_1\alpha_2$.

The general features of the hydrogen-bond interaction space for all uncharged functional groups are illustrated in Figure 6. For a given solvent, functional group interactions are partitioned into four quadrants defined by the hydrogen-

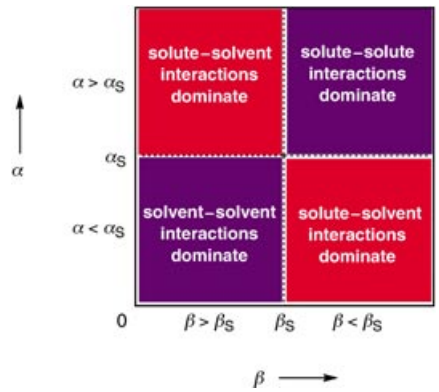


Figure 6. The generalized profile for hydrogen-bond interactions between neutral functional groups in solution [Eq. (4)]. The hydrogen-bond parameters introduced in this work are denoted α and β for the solute molecules and α_S and β_S for the solvent. For a given solvent, the functional group interaction space is partitioned into four quadrants. In the two red quadrants, $\Delta\Delta G_{H\text{-bond}}$ is positive, and the functional group interactions are unfavorable. In the two blue quadrants, $\Delta\Delta G_{H\text{-bond}}$ is negative, and the functional group interactions are favorable. The values of α_S and β_S set the boundaries between these quadrants and define how the space is partitioned.

Table 3: Hydrogen-bond parameters for common functional groups and solvents in order of increasing hydrogen-bond strength.

H-bond donors	$\alpha^{[a]}$	$\alpha^{[b]}$	H-bond donors	$\alpha^{[a]}$	$\alpha^{[b]}$	H-bond donors	$\alpha^{[a]}$	$\alpha^{[b]}$
alkane		0.4	alkyl ketone	1.5		carbamate	2.8	
alkene		0.7	amine		1.5	sulfonamide		2.8
alkyl ether		0.9	aldehyde		1.6	amide	2.9	
benzene		1.0	furan		1.7	urea	3.0	
alkyl thioether		1.0	thiol		1.7	pyrrole	3.0	
alkyl iodide		1.2	acetonitrile	1.7		sulfinamide		3.2
alkyl fluoride		1.2	thiophenol	1.8		thioamide		3.3
alkyl chloride		1.3	nitromethane	1.8		carboxylic acid	3.6	
alkyl bromide		1.3	dichloromethane	1.9		imidazole		3.7
aryl chloride		1.3	alkyne	1.9		2,2,2-trifluoroethanol	3.7	
carbon tetrachloride	1.4		1,1,2,2-tetrachloroethane		2.0	phenol	3.8	
aryl ether		1.4	ammonia		2.0	trifluoroacetic acid		3.9
aryl fluoride		1.4	aniline	2.1		phosphoric acid		4.0
pyridine		1.4	chloroform	2.2		hexafluoropropan-2-ol	4.5	
1,1,1-trichloroethane		1.5	alcohol	2.7		perfluoro-*tert*-butanol	4.9	
alkyl ester		1.5	water	2.8				

H-bond acceptors	$\beta^{[a]}$	$\beta^{[b]}$	H-bond acceptors	$\beta^{[a]}$	$\beta^{[b]}$	H-bond acceptors	$\beta^{[a]}$	$\beta^{[b]}$
alkane		0.3	alkyne	2.7		alcohol	5.8	
carbon tetrachloride	0.6		isothiocyanate	2.7		imine	5.8	
perfluoroalkane		0.7	thiol	2.7		ketone	5.8	
chloroform	0.8		hexafluoropropan-2-ol		3.1	sulfonamide	5.8	
dichloromethane	1.1		alkyl selenide	3.4		thioamide	5.8	
alkene	1.1		thioether	3.6		ammonia	6.1	
1,1,2,2-tetrachloroethane		1.3	nitroalkane	3.7		sulfone	6.3	
1,1,1-trichloroethane		1.4	aryl ether	3.7		pyridine	7.0	
aryl chloride	1.6		disulfide	3.7		carbamate	7.3	
aryl bromide	1.6		trifluoracetic acid		3.8	amine	7.8	
aryl iodide	1.6		pyrrole	4.1		sulfinamide	8.3	
aryl fluoride	1.6		2,2,2-trifluoroethanol		4.2	amide	8.3	
alkyl chloride	2.2		water	4.5		urea	8.3	
alkyl bromide	2.2		aldehyde	4.7		phosphinate diester	8.9	
alkyl iodide	2.2		nitrile	4.7		phosphonate diester	8.9	
benzene	2.2		sulfate diester	4.7		sulfoxide	8.9	
furan	2.2		thiocyanate	4.7		amidine	8.9	
thiophenol	2.2		carboxylic acid	5.3		phosphoric acid		9.3
perfluoro-*tert*-butanol		2.3	alkyl ether	5.3		phosphine oxide	9.9	
phenol	2.7		aniline	5.3				
alkyl fluoride	2.7		ester	5.3				

[a] Values based on the literature values of α_2^H and β_2^H. [b] Values based on the molecular electrostatic potential surface.

bond donor and hydrogen-bond acceptor properties of the solvent. The values of α_S and β_S set the boundaries where the exchange process in Figure 5 is associated with zero free energy change. There are two red quadrants in Figure 6, where hydrogen-bonding interactions are unfavorable, because the solute–solvent interactions dominate. In the two blue quadrants, hydrogen-bonding interactions between solute molecules are favorable. The top right quadrant corresponds to interactions between the most polar functional groups where solute–solute interactions dominate. The bottom left quadrant is the solvophobic zone. Here, functional group interactions are favorable, because the solvent–solvent interactions are stronger than the solute–solvent interactions.

Figure 7 shows the functional group interaction profiles for different solvent environments.[21] The positions of a selection of functional groups on the new universal hydrogen-bond scale are shown for calibration. Figure 7a shows the functional interaction energies for $\alpha_S = \beta_S = 0$. This corresponds to a condensed phase with no solvent, that is, the solid state, or to a completely nonpolar solvent, that is, a noble gas, and provides a measure of the intrinsic functional group interaction energies. Under these conditions all interactions are attractive, because there is no solvent competition. Figure 7b illustrates why dimethyl sulfoxide (DMSO) is one of the best solvents. The α_S and β_S boundary lines intersect in the bottom right corner of the interaction space, and so almost all functional group interactions are unfavorable. Chloroform has been widely used for molecular recognition studies on model systems, because there are no solvophobic effects, and there is a large zone of favorable solute–solute interactions (Figure 7c). The functional group interaction profile for carbon tetrachloride, the solvent used to determine the original α_2^H and β_2^H parameters, is not shown but is very similar to the chloroform plot. Dimethyl ether also partitions the interaction space into a repulsive and an attractive zone with no solvophobic zone, but the profile is quite different from chloroform (Figure 7d).[21] Chloroform is a good solvent

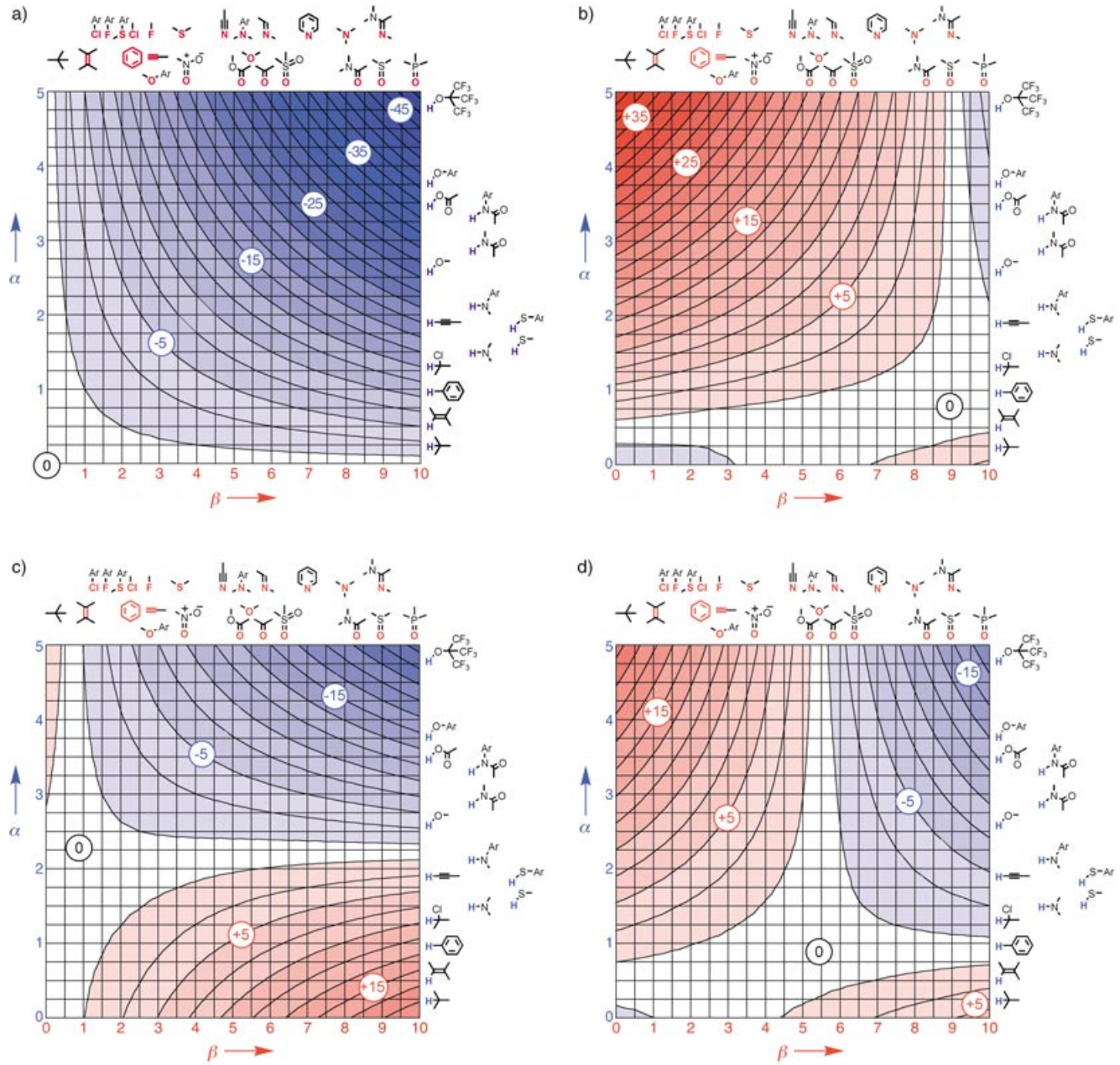


Figure 7. Functional group interaction profiles a) in the solid state or a noble gas solvent ($\alpha_S = \beta_S = 0$), b) in DMSO ($\alpha_S = 0.8$, $\beta_S = 8.9$), c) in chloroform ($\alpha_S = 2.2$, $\beta_S = 0.8$) and d) in dimethyl ether ($\alpha_S = 0.9$, $\beta_S = 5.3$). $\Delta\Delta G_{\text{H-bond}}$ (kJ mol^{-1}) calculated by using Equation (12) is plotted against the range of values of α and β found for neutral functional groups. The structures of a representative set of functional groups are illustrated at appropriate points on the α and β scale as orientation for the reader. The contour lines are drawn at 2 kJ mol^{-1} intervals. Blue represents a favorable interaction ($\Delta\Delta G_{\text{H-bond}} < 0$), and red represents an unfavorable interaction ($\Delta\Delta G_{\text{H-bond}} > 0$). The zero point on each plot corresponds to ($\alpha = \alpha_S$, $\beta = \beta_S$) where desolvation perfectly balances the solute–solute interactions.

for studying weak hydrogen-bond acceptors, and ether is a good solvent for studying weak hydrogen-bond donors.

The solvent dependence of hydrogen-bonding interactions has been studied experimentally in relation to Equation (3).[22] The solvent dependence of the constant c_1 in Equation (3) can be accounted for using Equation (4) to estimate $\Delta\Delta G_{\text{H-bond}}$ and hence the association constant, K, in the relevant solvents [Eq. (7)].

$$\lg K = -\frac{\Delta\Delta G_{\text{H-bond}} + 6}{RT} \qquad (7)$$

Figure 8 shows a plot of $\lg K$ for the favorable solute–solute interactions in the top right quadrant of the functional group interaction space for the gas phase, carbon tetrachloride and 1,1,1-trichloroethane versus the product of the original H-bond parameters, α_2^H and β_2^H. The slopes of these plots predict values for c_1 that are in excellent agreement with experiment.

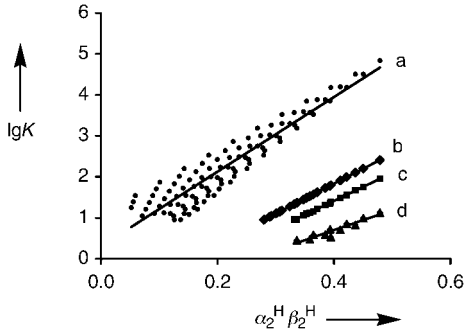


Figure 8. Lg *K* predicted by using Equation (7) plotted against the corresponding values of $\alpha_2^H\beta_2^H$ for functional group interactions in a) the gas phase (this is approximated by using $\alpha_S=0$, $\beta_S=0$, and the additional contribution from dispersion interactions is ignored), b) carbon tetrachloride ($\alpha_S=1.6$, $\beta_S=0.6$), c) 1,1,1-trichloroethane ($\alpha_S=1.5$, $\beta_S=1.4$), and d) chloroform ($\alpha_S=2.2$, $\beta_S=0.8$). The gas-phase plot (a) diverges significantly at low $\alpha_2^H\beta_2^H$, and the choice of log *K* cut off affects the slope of the line of best fit. It is clear however that the slope is significantly steeper than for the other solvents. The slopes of these plots give values of c_1 in Equation (3) that are identical to the experimental values:[22] a) $c_1(\text{expt})=c_1(\text{calcd})=9.1$, b) $c_1(\text{expt})=c_1(\text{calcd})=7.3$, c) $c_1(\text{expt})=c_1(\text{calcd})=6.8$, d) the experimental value of c_1 has not been determined for chloroform, but the prediction of Equation (7) is that $c_1=4.9$.

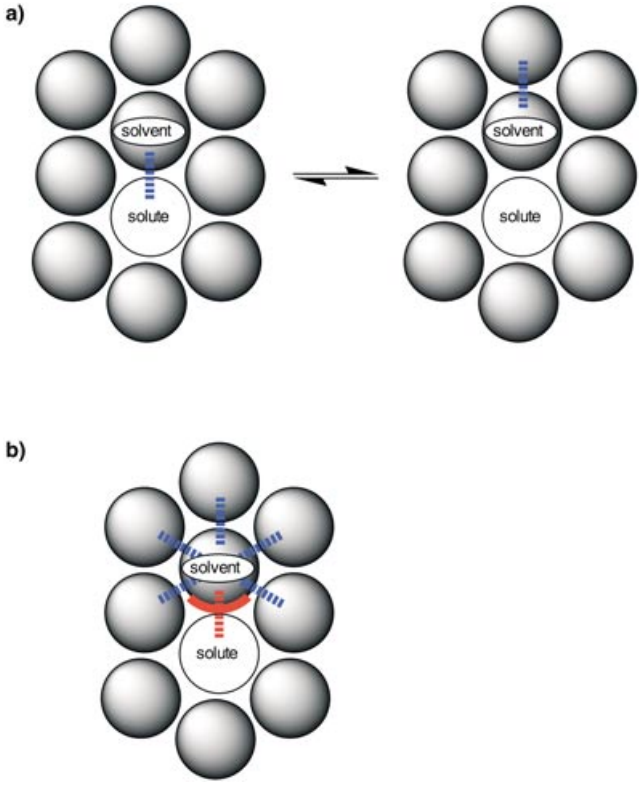


Figure 9. a) A solvent molecule that solvates a solute can choose between hydrogen-bonding to the solute or hydrogen-bonding to the surrounding bulk solvent. At equilibrium, there is a Boltzmann distribution of the two states. b) A solvent molecule that solvates a solute is sterically blocked from interacting freely with the surrounding bulk solvent. A fraction of the surface of the solvent is always forced to interact with the solute (red zone), even if the interaction with bulk solvent is more favorable.

6. Solvophobic Interactions

Water is a special solvent, because the values of both α_S and β_S lie in the middle of the hydrogen-bond scales. The interaction space is equally partitioned into the four different quadrants described above, and both solvophobic and solute driven interactions are important. The analysis above assumes that the solute–solvent complexes in Figure 5 are 100% bound, but for a nonpolar solute in a polar solvent, this is not true. The solvent molecules can choose between a weak interaction with the solute and a strong interaction with another solvent molecule (Figure 9a). Interactions involving the solvent molecules are therefore better treated as the Boltzmann weighted average of the interaction with the solute and interaction with the bulk solvent. In the limit, very polar solvents, like water, form a cage around very nonpolar solutes, like hydrocarbons, to minimize solvent–solute interactions.[3] This dramatically attenuates the magnitudes of the interaction energies in the solvophobic quadrant. The driving force is the formation of solvent–solvent interactions, but if these are largely formed already in the free state, the gain in solvent–solvent interactions in the bound state is relatively small. There is another important factor to consider in the solvophobic quadrant. The molecules in the liquid state are close packed, and so the solvent molecules are forced to interact with the solute to some extent. It is never possible to completely desolvate the solute, even if the Boltzmann factor predicts complete population of the solvent-solvent interaction. Although more attractive solvent–solvent interactions may be available, they are sterically blocked by the solute (Figure 9b). Equation (4) must therefore be modified to account for these effects in the interactions between the first solvation shell and bulk solvent. Implementation of the concepts outlined above is rather involved and is explained in detail in the Appendix.

In practice, the only significant difference caused by using Equation (12) in the Appendix is that the solvophobic quadrant is compressed and flattened. There is little impact on the other quadrants, where the interactions between the solute and solvent are highly populated states. The functional group interaction profile for water is shown in Figure 10. There is a broad flat attractive region on the free energy surface for hydrophobic interactions which is consistent with the empirical observation that hydrophobic interaction energies are largely dictated by the surface area of contact and are relatively insensitive to the precise identities of the hydrophobic groups involved.[3]

7. Solvents with Multiple Hydrogen-Bond Donors and Hydrogen-Bond Acceptors

The analysis above is restricted to solvents that have a single type of hydrogen-bond donor and hydrogen-bond acceptor. To treat systems with more functional groups or mixed solvents, we need to think about the possible solvent–

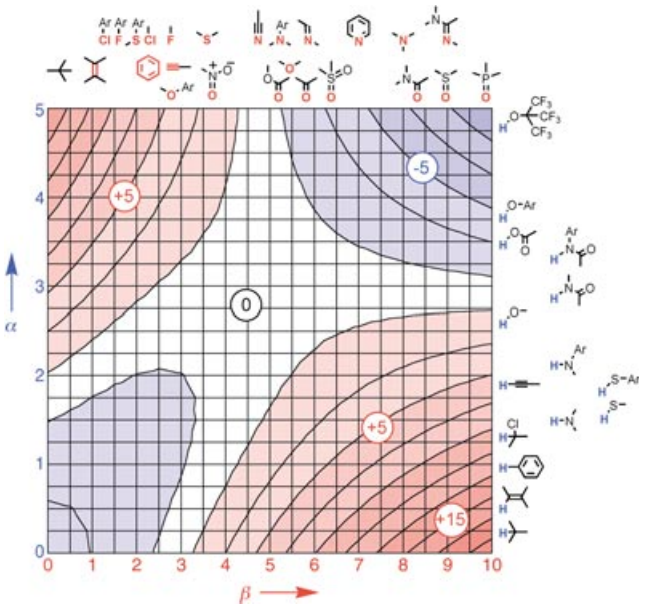


Figure 10. The functional group interaction profile in water ($\alpha_S = 2.8$, $\beta_S = 4.5$). $\Delta\Delta G_{\text{H-bond}}$ (kJ mol^{-1}) calculated by using Equation (12) (see Appendix) is plotted against the range of values of α and β found for neutral functional groups. The structures of a representative set of functional groups are illustrated to calibrate the α and β scale. The contour lines are drawn at 2 kJ mol^{-1} intervals. Blue represents a favorable interaction ($\Delta\Delta G_{\text{H-bond}} < 0$), and red represents an unfavorable interaction ($\Delta\Delta G_{\text{H-bond}} > 0$). The zero point corresponds to $\alpha = \alpha_S$ and $\beta = \beta_S$.

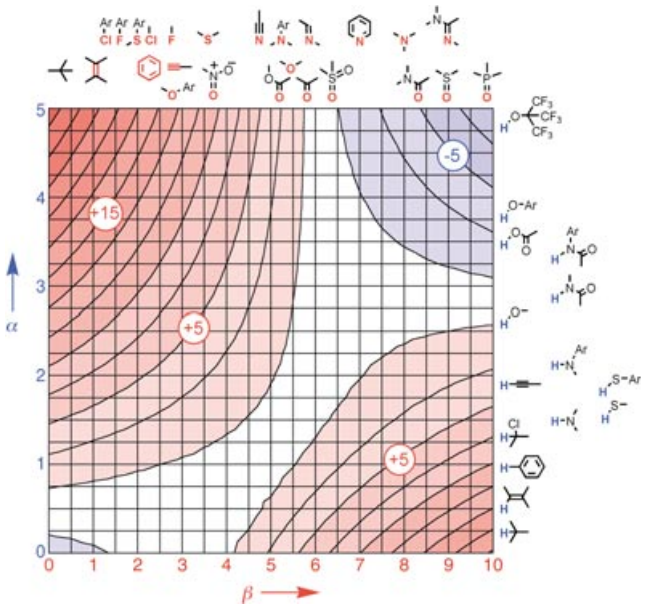


Figure 11. The functional group interaction profile in methanol ($\alpha_S = 0.9$ and 2.7, $\beta_S = 5.8$). $\Delta\Delta G_{\text{H-bond}}$ (kJ mol^{-1}) was calculated by using Equation (12) (see Appendix) for every possible combination of solvent–solute interactions and the maximum value is plotted against the range of values of α and β found for neutral functional groups. The structures of a representative set of functional groups are illustrated to calibrate the α and β scale. The contour lines are drawn at 2 kJ mol^{-1} intervals. Blue represents a favorable interaction ($\Delta\Delta G_{\text{H-bond}} < 0$), and red represents an unfavorable interaction ($\Delta\Delta G_{\text{H-bond}} > 0$).

solute interactions that might be present. The solvent molecules will choose to interact with the solute in the most thermodynamically favorable manner, as in Figure 9. In other words, the solvent will maximize its interactions with the solute and consequently minimize the solute–solute interactions. Functional group interaction energies can therefore be evaluated by considering all possible solvent–solute hydrogen-bond interactions and selecting the arrangement that maximizes $\Delta\Delta G_{\text{H-bond}}$ in Equation (12).

Let us consider methanol as a simple example. This has two types of hydrogen-bond donor, the O–H and the C–H group, and one type of hydrogen-bond acceptor, the oxygen atom. The functional group interaction profile that maximizes $\Delta\Delta G_{\text{H-bond}}$ is shown in Figure 11. At each point on the profile, we compare the free energy of the solute–solute interaction using the solvent O–H as the donor with the free energy of the solute–solute interaction using the solvent C–H as the donor and take the least favorable of the two free energies. The hydrogen-bond properties of methanol can be considered to be similar to those of dimethyl ether and water (Figure 7c and Figure 10), and Figure 11 is effectively the combination of the ether and water profiles that maximizes the free energy of interaction. The solvophobic zone in the water profile has been reduced by the ability of the methyl group to compete with weak hydrogen-bond donors, so that the left hand side of the profile looks like that of dimethyl ether. The attractive solute–solute interactions in the dimethyl ether profile have been reduced by the ability of the hydroxy group to compete with good hydrogen-bond donors, so that the right-hand side of the profile looks like that of water. The conclusion is that the potential for attractive functional group interactions will always be reduced in solvents that have more effective and versatile competitors.

A similar approach could be applied to the analysis of intermolecular interactions in mixtures of solvents. For example, a hypothetical 1:1 mixture of water and dimethyl ether would behave in the same way as methanol, and the functional group interaction profile would look very like that in Figure 11.

8. Solubility

Solubility is a parameter that is governed by many different factors and so is difficult to predict in practice. Nevertheless, the functional group interaction profiles can provide some insight into the solubility properties of organic molecules. Functional groups that make favorable interactions in a given solvent are likely to reduce solubility, whereas unfavorable functional group interactions should lead to increased solubility. In other words, the unfavorable red zones in Figure 6 can be used to predict good solubility, and the favorable blue zones can be used to predict poor solubility. Let us consider a simple example, the solubility of dichloromethane. In Figure 7b–d and Figure 11, the ClC–H···Cl interaction lies in the red zone, and so dichloromethane has high solubility in DMSO, chloroform, ether, and methanol. In

Figure 10 on the other hand, the ClC–H···Cl interaction lies in the blue solvophobic zone, and dichloromethane consequently has low solubility in water. Similarly, Figure 7c and 7d suggest that chloroform should be a better solvent for primary and secondary amines than dimethyl ether. Chloroform is a better hydrogen-bond donor than an amine, and so the most favorable pairwise interaction in a mixture of chloroform and amine is the Cl_3C–H···N interaction. In contrast, dimethyl ether is a weaker hydrogen-bond acceptor and a weaker hydrogen-bond donor than an amine, so in ether, the most favorable interaction is the amine–amine interaction.

9. Relationship of α and β with Chemical Structure

Hydrogen-bonding properties have often been correlated with pK_a data. Within a particular functional group class, such as pyridines or phenols, trends in α and β values correlate well with pK_a values, but over the entire range of functional groups shown in Figure 7, pK_a is a relatively poor predictor of α and β.[23] The correlations in Figure 4 are not reproduced by using the corresponding pK_a values in place of the maxima and minima in the molecular electrostatic potential surface. For example, thiols are much more acidic than alcohols, because they can stabilize the charged anionic state formed on deprotonation, but they are much worse hydrogen-bond donors, because they are less polar. Similarly, pyridine is much more basic than DMSO, but DMSO is a significantly better hydrogen-bond acceptor.

There are no special effects observed for highly polarizable functional groups, providing some justification for the assumptions in Section 2 about the limited role of dispersion and induction. Functional group interactions involving elements of the third period are all relatively weak, as expected from the electrostatic potentials. The trends in the values of α and β shown in Table 3 can be rationalized based on conventional electrostatic arguments. The value of α is determined to a large extent by the net positive charge on the hydrogen-bond donor hydrogen atom. There are more factors that influence the value of β. At a simplistic level, the electrostatic potential at the van der Waals surface of a hydrogen-bond acceptor (E_{min}) is determined by the effective nuclear charge ($n+$), the average separation of the lone pair from the nucleus (r_e), and the van der Waals radius (r_{VDW}) as illustrated in Figure 12a.

9.1 Nuclear Charge

β increases along the series F < O < N. Figure 12b shows a plot of β versus nuclear charge for trialkyl amine, dialkyl ether and alkyl fluoride (the nuclear charge is taken as the net charge without the valence electrons). There is a clear correlation, but what is most striking is the value of the intercept, $n+ = 8$. This is neon, which as a non-polar noble gas should have a β value of zero. We can also extrapolate the data to a hypothetical neutral carbon lone pair ($n+ = 4$), which would have a β value of 10.5. Although r_{VDW} and r_e also change along this series, and inductive effects perturb the effective value of $n+$, Figure 12b suggests that the dominant parameter that distinguishes the hydrogen-bond acceptor properties of the second period elements is the nuclear charge.

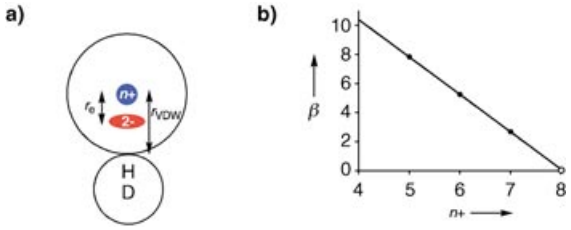


Figure 12. a) The electrostatic potential experienced by a hydrogen-bond donor (DH) at the surface of a hydrogen-bond acceptor is determined by the effective charge on the nucleus ($n+$, blue), the location of the lone pair (2−, red, r_e from the nucleus), and the van der Waals radius (r_{VDW}). b) A plot of β versus $n+$ for alkyl-N, alkyl-O, and alkyl-F atoms. The nuclear charge was taken as the net charge without the valence electrons. The open circle represents neon ($n+ = 8$), which is completely nonpolar, and so is expected to have a β value of zero.

9.2 Van der Waals Radius

β increases along the series Se < S < O. Going down the Periodic Table, the van der Waals radius (r_{VDW}) increases and the lone pairs become more diffuse, so the electrostatic potential is effectively smeared out over a larger area on the surface of the atom (Figure 12a).

9.3 Lone Pair Location

β increases along the series $N(sp) < N(sp^2) < N(sp^3)$. Here the nuclear charge and van der Waals radii do not change significantly, and the dominant factor is the average separation of the lone pair electrons from the nucleus (r_e), which decreases in orbitals with more s character (Figure 12a). In many functional groups, more than one of the effects discussed here is important, so for example, sp^2 and sp^3 oxygen hydrogen-bond acceptors have similar β values, because the effect of changing r_e is balanced by the increased polarization of the π-electrons (see Section 9.5).

9.4. Electronegativity of Substituents

If we compare oxygen hydrogen-bond acceptors, β for phosphine oxide is larger than β for sulfoxide. This is due to polarization of the bonding electrons towards the oxygen as the electronegativity of the other nucleus decreases. Similarly, hydrogen-bond donor ability increases with the electronegativity of the atom bonded to the hydrogen, so α increases along the series C–H < S–H < N–H < O–H. Hybridization state affects electronegativity, so α increases along the series $C(sp^3)$–H < $C(sp^2)$H < C(sp)–H. Inductive effects transmit the polarization due to electronegativity differences further through the bonding framework, so hydrogen-bonding properties are modulated by polarizing substituents in a predict-

able manner. Representative data for the effects of methyl and trifluoromethyl groups on the properties of alcohols and ketones are shown in Table 4.[22b,24] The mutual electron-withdrawing effects of the fluorine substituents in perfluoroalkanes leads to hydrogen-bond acceptor properties that are comparable to those of simple alkanes, and the special properties of these compounds can be in part attributed to this effect (Table 3).

Table 4: Substituent effects on the hydrogen-bonding properties of alcohols and ketones.[a]

		R_xH_yC-OH		R_xH_yC-Ac	
		$R=CF_3$	$R=CH_3$	$R=Cl$	$R=CH_3$
x	y	α	α	β	β
0	3	2.9	2.9	5.7	5.7
1	2	3.7	2.7	4.5	5.8
2	1	4.5	2.7	3.6	5.8
3	0	4.9	2.7	3.1	5.7

[a] The values of α and β were calculated by using Equation (5) and (6) and the literature values[22b,24] of α_2^H and β_2^H for these compounds.

9.5 Delocalization

The associated changes in electron density modulate the values of α and β accordingly. Thus the amide oxygen atom is a better hydrogen-bond acceptor than a ketone, and the amide NH group is a poorer hydrogen-bond acceptor and a better hydrogen-bond donor than an amine.

9.6 Through-Space Effects

Nearby functional groups can perturb the electrostatic potential surface. The α values for phenol and carboxylic acid are a good example. A carbonyl group is more strongly electron-withdrawing than an aromatic ring, but phenol is a better hydrogen-bond donor than carboxylic acid. The reason is that when a hydrogen-bond acceptor interacts with the OH group, there are long-range through-space interactions with the adjacent CH group of the phenol and the carbonyl group of the carboxylic acid (Figure 13). These secondary electrostatic interactions are most important when the hydrogen-bond donors and hydrogen-bond acceptors are close (that is, separated by one atom) and oriented in the same direction as in a carboxylic acid. Secondary electrostatic interactions have been extensively studied in relation to heterocycle base-pairing in chloroform, where each secondary interaction contributes 2–3 kJ mol^{-1} to the free energy of complexation.[25]

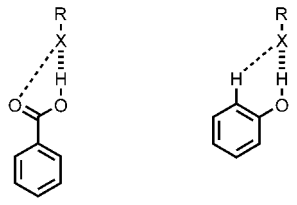


Figure 13. Through-space interactions can significantly perturb the electrostatic potential surface if polar groups are close in space. The secondary electrostatic interactions (dashed line) make phenol a better hydrogen-bond donor than carboxylic acid, which has repulsive secondary electrostatic interactions.

9.7 Configuration

The value of E_{min} for amines is strongly dependent on the degree of pyramidalization at the nitrogen center, as illustrated in Figure 14a. The pyramidal geometry exposes the lone pair electron density and improves the hydrogen-bond acceptor properties.

9.8 Conformation

Neighboring functional groups can perturb the electrostatic potential surface of a functional group (see Section 9.6),

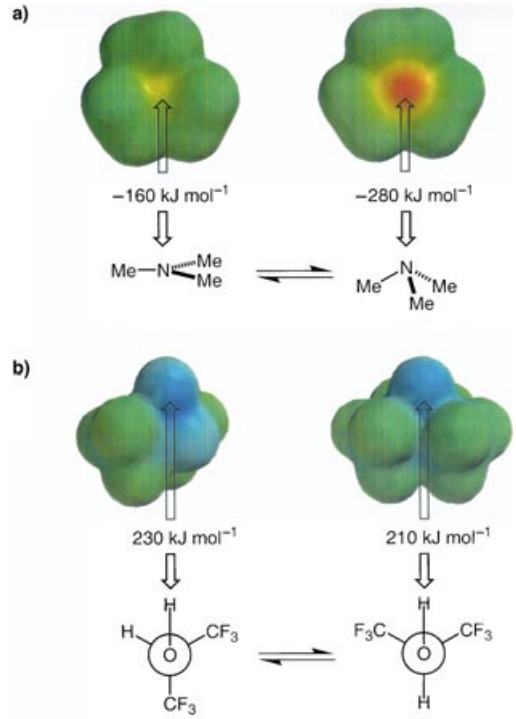


Figure 14. Changes in three-dimensional structure can have a significant influence on the molecular electrostatic potential surface. a) A pyramidal amine is a significantly better hydrogen-bond acceptor than a trigonal amine. The molecular electrostatic potential surfaces of trimethylamine calculated by using AM1 are shown. Negative regions are shown in red (< -280 kJ mol^{-1}), and neutral regions in green. The difference between the two conformations is 120 kJ mol^{-1} which corresponds to a difference of more than 2 on the β scale. b) The molecular electrostatic potential surfaces of hexafluoropropan-2-ol in two different conformations calculated by using AM1. Positive regions are shown in blue ($> +230$ kJ mol^{-1}), and neutral regions in green. Although the color differences are smaller than in a), the conformation in which the hydroxy group is encumbered by two trifluoromethyl groups is a significantly weaker hydrogen-bond donor (by 0.4 on the α scale).

and through-space effects depend on the precise arrangement of the functional groups. Figure 14b shows how changing the conformation of an alcohol affects the value of E_{max}. In one conformation, the OH group is exposed, but in the other, it is encumbered by a trifluoromethyl group, and this lowers the electrostatic potential on the molecular surface.

9.9 Proton Transfer

So far, we have dealt exclusively with neutral molecules, but if the pK_a values are appropriate, then proton transfer will take place. Hydrogen-bonding interactions between charged functional groups have not been experimentally quantified in the same way, presumably due to the low solubility of salts in carbon tetrachloride and complications in the analysis of the extent of counterion association.[26] However, the interactions between a deprotonated acid and a protonated base are likely to be significantly larger than those between the neutral species, and this possibility must always be considered when using this approach.[27]

10. Conclusions

The analysis presented here allows us to extrapolate from the experimental data on the thermodynamics of hydrogen-bonding interactions in carbon tetrachloride to interactions between a wide range of functional groups in any solvent or the solid state. The functional group interaction profiles provide a benchmark for estimating the magnitudes of intermolecular interactions in the condensed state.

There are limitations. For example, the scatter in the correlations in Figure 4, which we used to set our origin and scale, necessarily introduces approximations. In this analysis, all interactions are treated as a type of hydrogen bond at a single well-defined point on the molecular surface. The conclusions may not extrapolate to functional group combinations that interact over a larger surface area, such as stacked aromatic systems, where the integration of lower electrostatic potentials over a larger area may lead to relatively strong interactions. In such systems, the desolvation term will also be affected by the geometric complementarity and surface area of the solvent–solute interaction.[12] Analyzing the behavior of more complicated complexes that contain multiple functional group interactions introduces additional factors not considered here. The stability of the complex will be reduced by non-perfect molecular shape complementarity, failure to optimize functional group interaction geometries and restriction of conformational flexibility. Cooperative effects between different functional group interaction sites can lead to more favorable free energies of binding due to secondary functional group interactions, long-range electrostatic interactions, induction caused by bond polarization or conformational changes.

Thus applications of the interaction profiles to predict binding free energies by summing the individual interactions in complicated molecules is problematic. Nevertheless, there have been some significant successes in the use of simple QSAR parameters for predicting binding free energies of molecular complexes,[5,18,28] and so there may be potential in such an approach. At the very least, speculation about whether this interaction or that interaction is likely to be the most important in any given system can now be calibrated with a sensible estimate of the associated free energy contributions that are possible.

There are some interesting aspects of analyzing the full functional group interaction space. Consider hydrogen-bonding interactions in water for example. The best hydrogen-bond donors that will not readily ionize in water are NH donors with electron withdrawing substituents, and the best hydrogen-bond acceptors that will not readily protonate in water are oxygen acceptors with electron-donating groups. This is clearly one reason that peptides and the nucleic acid bases have been selected as the biological building blocks of choice. They combine excellent hydrogen-bonding properties with the reliability of not engaging in acid–base reactions, and Figure 10 shows that they make weakly attractive interactions in water. Hopefully, the interaction profiles in this paper will encourage synthetic chemists to think about molecular designs based on the potential interaction free energies that are available from some of the unexplored functional group and solvent combinations.

Appendix: Treatment of the Solvophobic Zone

The equilibrium in Figure 9a can be used to determine the Boltzmann factors that reflect the extent to which the solute–solvent interactions are present in the solvated free state of the hydrogen-bond donor and hydrogen-bond acceptor, - χ_α [Eq. (8)] and χ_β [Eq. (9)].

$$\chi_\alpha = \frac{1}{1+e^{-(\alpha\beta_S-\frac{1}{2}\alpha_S\beta_S)/RT}} \tag{8}$$

$$\chi_\beta = \frac{1}{1+e^{-(\alpha_S\beta-\frac{1}{2}\alpha_S\beta_S)/RT}} \tag{9}$$

The factor of $^1/_2$ accounts for the fact that when the solvent does not interact with the solute, it effectively interacts with itself, because it becomes part of the bulk solvent. A similar analysis must be applied to both sides of the equilibrium in Figure 5, and so in the bound state, we also consider to what extent the released solvent interacts with the bulk solvent. Otherwise, the free energy for a water molecule interacting with another water molecule in a solvent of water would be nonzero. Thus χ_S reflects the extent to which solvent–solvent interactions are formed in the bound state [Eq. (10)].

$$\chi_S = \frac{1}{1+e^{-(\alpha_S\beta_S-\frac{1}{2}\alpha_S\beta_S)/RT}} = \frac{1}{1+e^{-\frac{1}{2}\alpha_S\beta_S/RT}} \tag{10}$$

If we now consider desolvation of the hydrogen-bond donor in preparation for complexation, the number of solute–solvent interactions lost is χ_α, and therefore in principle, χ_α solvent–solvent interactions should be formed. However, the fraction of solvent–solvent interactions that are populated is χ_S, and so $(1-\chi_S)$ solvent–solvent interactions are never made.

If we allow for the fact that the released solvent becomes part of the bulk solvent and therefore interacts with itself, the number of solvent–solvent interactions that are formed on desolvation of the hydrogen-bond donor is $^1/_2\{\chi_\alpha-(1-\chi_S)\}$. The free energy for the equilibrium in Figure 5 can thus be evaluated allowing for the Boltzmann distributions of the solute–solvent and solvent–solvent interactions illustrated in Figure 9a, but assuming that the solute–solute state that we are interested in is fully bound [Eq. (11)].

$$\begin{aligned}\Delta\Delta G_{\text{H-bond}} &= -\alpha\beta-\frac{1}{2}\{\chi_\alpha-(1-\chi_S)\}\alpha_S\beta_S-\frac{1}{2}\{\chi_\beta-(1-\chi_S)\}\alpha_S\beta_S \\ &\quad +\chi_\alpha\alpha\beta_S+\chi_\beta\alpha_S\beta \\ &= -\alpha\beta+\chi_\alpha\alpha\beta_S+\chi_\beta\alpha_S\beta-(\frac{1}{2}\chi_\alpha+\frac{1}{2}\chi_\beta+\chi_S-1)\alpha_S\beta_S\end{aligned} \tag{11}$$

The number of solvent–solvent interactions that can be made in the free state is reduced by the steric effects of close contact with the solute, as illustrated in Figure 9b. The solvent and solute are forced to interact to some extent, even if the Boltzmann distribution suggests that they should not. If ϕ is the number of solvent–solvent interactions that are sterically blocked by solute contacts, then only a fraction of the solvent-solute interactions $(1-\phi)$ can participate in the equilibrium leading to the Boltzmann distribution discussed above. Thus the free energy of formation of a hydrogen bond in solution is given by a composite of the simple expression in Equation (4) (see Section 5) and the more complex Equation (11) [Eq. (12)].

$$\begin{aligned}\Delta\Delta G_{\text{H-bond}} &= -\phi(\alpha-\alpha_S)(\beta-\beta_S)-(1-\phi)\{\alpha\beta-\chi_\alpha\alpha\beta_S-\chi_\beta\alpha_S\beta \\ &\quad +(\frac{1}{2}\chi_\alpha+\frac{1}{2}\chi_\beta+\chi_S-1)\alpha_S\beta_S\}\end{aligned} \tag{12}$$

For which the Equations (13)–(15) are applicable.

$$\chi_\alpha=\frac{1}{1+e^{-(1-\phi)(\alpha\beta_S-\frac{1}{2}\alpha_S\beta_S)/RT}} \tag{13}$$

$$\chi_\beta=\frac{1}{1+e^{-(1-\phi)(\alpha_S\beta-\frac{1}{2}\alpha_S\beta_S)/RT}} \tag{14}$$

$$\chi_S=\frac{1}{1+e^{-\frac{1}{2}(1-\phi)\alpha_S\beta_S/RT}} \tag{15}$$

Equation (12) was used to generate all of the interaction profiles shown here. The only difference between the results obtained by using Equation (4) is that the solvophobic zone is compressed and flattened. The Boltzmann factors come directly from the hydrogen-bond parameters α and β, but the value of ϕ is an independent variable between zero and one that must be obtained by different methods. For the purposes of discussion, ϕ is estimated using water. If we assume close packing of spherical molecules, then about $^1/_{12}$ of the surface area of a water molecule that solvates a functional group of similar size is blocked from interaction with other water molecules (Figure 9b). Each water molecule can make a total of four hydrogen bonds, but the fraction of solvent–solvent interactions that are populated is χ_S, as discussed above. Thus the number of water–water hydrogen bonds that are sterically blocked by the solute is $\phi=4\chi_S/12$. Using the α and β values for water and Equation (10) gives $\chi_S=0.93$, so $\phi=0.3$. This is the value of ϕ that was used to generate the interaction profiles presented here, but the shape of the plots is surprisingly insensitive to the exact value. It is worth noting that the value of χ_S allows us to estimate the number of nearest neighbors in the bulk solvent, $N=(0.93\times4)+(0.07\times12)=4.5$. This is consistent with experimental X-ray diffraction measurements on liquid water that give a value of 4.4 for N.

I would like to thank my research group and colleagues, particularly Nick Williams, Pablo Ballester, Jeremy Sanders, and David Leigh, who have contributed to this manuscript through stimulating discussions, as well as the Sheffield undergraduates for their constant encouragement to look for clearer explanations.

Received: December 30, 2003

[1] a) A. R. Fersht in *Enzyme Structure and Mechanism*, Freeman, New York, **1985**; b) W. P. Jencks in *Catalysis in Chemistry and Enzymology*, Dover, New York, **1987**; c) J. Rebek, *Acc. Chem. Res.* **1990**, *23*, 399–404; d) H. J. Schneider, *Angew. Chem.* **1991**, *103*, 1419–1439; *Angew. Chem. Int. Ed. Engl.* **1991**, *30*, 1417–1436; e) F. Vogtle in *Supramolecular Chemistry*, Wiley, New York, **1991**; f) J. M. Lehn in *Supramolecular Chemistry*, VCH, Weinheim, **1995**; g) J. C. Ma, D. A. Dougherty, *Chem. Rev.* **1997**, *97*, 1303–1324; h) C. G. Claessens, J. F. Stoddart, *J. Phys. Org. Chem.* **1997**, *10*, 254–272; i) W. W. Cleland, P. A. Frey, J. A. Gerlt, *J. Biol. Chem.* **1998**, *273*, 25529–25532; j) W. W. du Mont, F. Ruthe, *Coord. Chem. Rev.* **1999**, *189*, 101–133; k) A. M. Davis, S. J. Teague, *Angew. Chem.* **1999**, *111*, 778–792; *Angew. Chem. Int. Ed.* **1999**, *38*, 736–749; l) J. W. Steed, J. L. Atwood in *Supramolecular Chemistry*, Wiley, Chichester, **2000**; m) H. J. Schneider, A. Yatsimirsky in *Principles and Methods in Supramolecular Chemistry*, Wiley, Chichester, **2000**; n) C. A. Hunter, K. R. Lawson, J. Perkins, C. J. Urch, *J. Chem. Soc. Perkin Trans. 2* **2001**, 651–669; o) T. Lazaridis, *Acc. Chem. Res.* **2001**, *34*, 931–937; p) T. Steiner, *Angew. Chem.* **2002**, *114*, 50–80; *Angew. Chem. Int. Ed.* **2002**, *41*, 48–76; q) E. A. Meyer, R. K. Castellano, F Diederich, *Angew. Chem.* **2003**, *115*, 1244–1287; *Angew. Chem. Int. Ed.* **2003**, *42*, 1210–1250.

[2] a) M. J. Stone, *Acc. Chem. Res.* **2001**, *34*, 379–388; b) D. H. Williams, D. P. O'Brien, B. Bardsley, *J. Am. Chem. Soc.* **2001**, *123*, 737–738; c) C. T. Calderone, D. H. Williams, *J. Am. Chem. Soc.* **2001**, *123*, 6262–6267; d) F. P. Schmidtchen, *Chem. Eur. J.* **2002**, *8*, 3522–3529; e) K. N. Houk, A. G. Leach, S. P. Kim, X. Zhang, *Angew. Chem.* **2003**, *115*, 5020–5046; *Angew. Chem. Int. Ed.* **2003**, *42*, 4872–4897.

[3] a) C. Tanford in *The Hydrophobic Effect: Formation of Micelles and Biological Membranes*, Wiley, New York, **1973**; b) B. Lee, *Proc. Natl. Acad. Sci. USA* **1991**, *88*, 5154–5158; c) N. T. Southhall, K. A. Dill, A. D. J. Haymet, *J. Phys. Chem. B* **2002**, *106*, 521–533.

[4] M. S. Searle, D. H. Williams, *J. Am. Chem. Soc.* **1992**, *114*, 10690–10697.

[5] a) W. P. Jencks, *Proc. Natl. Acad. Sci. USA* **1981**, *78*, 4046–4050; b) P. R. Andrews, D. J. Craik, J. L. Martin, *J. Med. Chem.* **1984**, *27*, 1648–1657; c) A. Matouschek, A. R. Fersht, *Methods Enzymol.* **1991**, *202*, 82–112; d) H. J. Schneider, *Chem. Soc. Rev.* **1994**, *23*, 227–234; e) D. H. Williams, M. S. Westwell, *Chem. Soc. Rev.* **1998**, *27*, 57–63; f) C. A. Hunter, P. S. Jones, P. Tiger, S. Tomas, *Chem. Eur. J.* **2002**, *8*, 5435–5446.

[6] M. Rigby, E. B. Smith, W. A. Wakeham, G. C. Maitland in *The Forces Between Molecules*, Clarendon, Oxford, **1986**.
[7] Sometimes a fifth component, the charge transfer interaction, is considered, but for most organic molecules, the magnitudes of such effects are negligible in the ground state of a complex.
[8] A. P. Bisson, C. A. Hunter, J. C. Morales, K. Young, *Chem. Eur. J.* **1998**, *4*, 845 – 851.
[9] J. Jadzyn, M. Stockhausen, B. Zywucki, *J. Phys. Chem.* **1987**, *91*, 754 – 757.
[10] a) J. C. Slater, J. G. Kirkwood, *Phys. Rev.* **1931**, *37*, 682 – 697; b) *CRC handbook of Chemistry and Physics*, 72nd ed. (Ed.: D. R. Lide), CRC, Boca Raton, FL, **1991**.
[11] The values for hydrogen are not included in Table 2, due to uncertainties in defining the relevant surface area.
[12] There are some special cases where the geometry of a binding pocket prevents solvation, and here, dramatic enhancements in the association constant are observed, because the solvent is unable to compensate for the gain in solute–solute dispersion interactions. K. T. Chapman, W. C. Still, *J. Am. Chem. Soc.* **1989**, *111*, 3075 – 3077.
[13] S. Tsuzuki, K. Tanabe, *J. Phys. Chem.* **1991**, *95*, 2272 – 2278.
[14] M. H. Abraham, J. A. Platts, *J. Org. Chem.* **2001**, *66*, 3484 – 3491.
[15] L. Pauling, *Proc. Natl. Acad. Sci. USA* **1928**, *14*, 359.
[16] a) H. Hagelin, J. S. Murray, T. Brinck, M. Berthelot, P. Politzer, *Can. J. Chem.* **1995**, *73*, 483 – 488; b) J. C. Dearden, T. Ghafourian, *J. Chem. Inf. Comput. Sci.* **1999**, *39*, 231 – 235; c) J. A. Platts, *Phys. Chem. Chem. Phys.* **2000**, *2*, 973 – 980; d) J. A. Platts, *Phys. Chem. Chem. Phys.* **2000**, *2*, 3115 – 3120; e) A. M. Zissimos, M. H. Abraham, A. Klamt, F. Eckert, J. Wood, *J. Chem. Inf. Comput. Sci.* **2002**, *42*, 1320 – 1331.
[17] M. I. Page, W. P. Jencks, *Proc. Natl. Acad. Sci. USA* **1971**, *68*, 1678 – 1683.
[18] H.-J. Bohm, *J. Comput.-Aided Mol. Des.* **1994**, *8*, 243 – 256.
[19] a) F. H. Westheimer, L. L. Ingraham, *J. Phys. Chem.* **1956**, *60*, 1668; b) H. L. Anderson, *Inorg. Chem.* **1994**, *33*, 972 – 981; c) H. L. Anderson, S. Anderson, J. K. M. Sanders, *J. Chem. Soc. Perkin Trans. 1* **1995**, 2231 – 2245; d) X. Chi, A. J. Guerin, R. A. Haycock, C. A. Hunter, L. D Sarson, *J. Chem. Soc. Chem. Commun.* **1995**, 2563 – 2565.
[20] For the functional groups described in reference [14] the simplest molecule containing the relevant functional group was constructed by using Spartan'02 (Wavefuntion, Inc., Irvine CA), for example, methyl acetate was used to calculate the properties of the ester. The molecules were optimized by using AM1, and the maxima and minima in the molecular electrostatic potential surface were determined.
[21] Dimethyl ether was used to calculate the interaction profile, because it contains only one type of hydrogen-bond donor and one type of hydrogen-bond acceptor. However, the assumption is that the diethyl ether profile, which is more relevant to experiments at room temperature, is very similar.
[22] a) M. H. Abraham, P. L. Grellier, D. V. Prior, R. W. Taft, J. L. Morris, P. J. Taylor, C. Laurence, M. Berthelot, R. M. Doherty, M. J. Kamlet, J.-L. Abboud, K. Sraidi, G. Guiheneuf, *J. Am. Chem. Soc.* **1988**, *110*, 8534; b) J. Marco, J. M. Orza, R. Notario, J.-L. M. Abboud, *J. Am. Chem. Soc.* **1994**, *116*, 8841 – 8842; c) M. H. Abraham, M. Berthelot, C. Laurence, P. J. Taylor, *J. Chem. Soc. Perkin Trans. 2* **1998**, 187 – 191.
[23] a) J. Rubin, B. Z. Senkowski, G. S. Panson, *J. Phys. Chem.* **1964**, *68*, 1601 – 1602; b) J. Rubin, G. S. Panson, *J. Phys. Chem.* **1965**, *69*, 3089 – 3091; c) R. W. Taft, D. Gurka, L. Joris, P. von R. Schleyer, J. W. Rakshys, *J. Am. Chem. Soc.* **1969**, *91*, 4801 – 4808; d) M. J. Kamlet, J. F. Gal, P. C. Maria, R. W. Taft, *J. Chem. Soc. Perkin Trans. 2* **1985**, 1583 – 1589.
[24] F. Besseau, M. Lucon, C. Laurence, M. Berthelot, *J. Chem. Soc. Perkin Trans. 2* **1998**, 101 – 107.
[25] a) W. L. Jorgensen, J. Pranata, *J. Am. Chem. Soc.* **1990**, *112*, 2008 – 2010; b) T. J. Murray, S. C. Zimmerman, *J. Am. Chem. Soc.* **1992**, *114*, 4010 – 4011; c) J. Sartorius, H. J. Schneider, *Chem. Eur. J.* **1996**, *2*, 1446 – 1452.
[26] C. A. Hunter, C. M. R. Low, C. Rotger, J. G. Vinter, C. Zonta, *Chem. Commun.* **2003**, 834 – 835.
[27] M. W. Hosseini, *Coord. Chem. Rev.* **2003**, *240*, 157 – 166.
[28] a) H. J. Schneider, V. Rudiger, O. A. Raevsky, *J. Org. Chem.* **1993**, *58*, 3648 – 3653; b) H. Gohlke, G. Klebe, *Angew. Chem.* **2002**, *114*, 2764 – 2798; *Angew. Chem. Int. Ed.* **2002**, *41*, 2645 – 2676.

