JCTC Journal of Chemical Theory and Computation

Theoretical Calculation of Hydrogen-Bonding Strength for Drug Molecules

Ming-Hong Hao*

Department of Medicinal Chemistry, Boehringer Ingelheim Pharmaceuticals, Inc., Ridgefield, Connecticut 06877

Received January 18, 2006

Abstract: Hydrogen bond is an important type of interaction between drug molecules and their receptors. We present here a computational method for accurately predicting the hydrogen-bonding strength for different acceptors with respect to a given donor or vice versa. The method is based on quantum chemistry DFT calculation of the interaction energy between hydrogen bond donors and acceptors. An excellent linear correlation is observed between the calculated hydrogen-bonding energies and the experimentally measured hydrogen-bonding constants log K_β on a variety of known hydrogen bond acceptors and donors. These results not only indicate the predictive power of this method but also shed light on factors that determine the magnitude of experimentally measured hydrogen-bonding constants for different acceptors with respect to a given donor, suggesting a primarily enthalpic contribution from hydrogen-bonding energy. The method can be used for evaluating the effects of steric interference and inhibitor binding geometry on hydrogen-bonding strength in drug design.

Introduction

The importance of the hydrogen bond to drug design is well recognized. Hydrogen bonds are not only crucial in dictating the orientation of an inhibitor binding in the receptor but also contribute importantly to binding affinity. Hydrogen bond capacity is an essential factor in the strategy of bioisosterism for drug design and optimization. When a bioisostere is used to replace an existing moiety of a compound, the replacement has to match the hydrogen-bonding characteristics of the parent and would preferably further improve upon compound properties including binding potency. There is strong evidence indicating that the strength of different hydrogen bond donors and/or acceptors varies significantly. The hydrogen-bonding constants of commonly encountered hydrogen-bonding groups have been measured to vary over more than 3 orders of magnitudes.[1] Furthermore, in the drug optimization process, it is often observed that electron-withdrawing or donating substituents have opposite effects on the activity of the compounds. Some of the underlying causes of such structure−activity relationships (SAR) could be traced to a modulation of the hydrogen-bonding interaction of the inhibitor with the receptor.[2] Clearly, the variations in hydrogen bond strength could be utilized in drug design. A fundamental understanding of hydrogen bond interactions will greatly facilitate such efforts.

In a comprehensive monograph on the subject,[3] Jeffrey and Saenger summarized the known characteristics of hydrogen bonds, some of which are particularly instructive to a theoretical investigation: (1) hydrogen bonds are not properties of atom pairs but are dependent on the pair of atom groups that forms the extensive donor and acceptor subunits;[3,1] (2) the major component of hydrogen-bonding interaction is electrostatic;[3−5] and (3) hydrogen bonds are soft interactions, and hydrogen bond lengths and angles fluctuate according to local environments.[3,6,7] Because of the importance of hydrogen bonds to drug design, much work has been done in the past on the theoretical modeling of hydrogen bonds for QSAR studies. The approaches span from simple indicator methods[8,9] to parametrization approaches using theoretically calculated properties, such as atomic charges,[10] molecular electrostatic potential,[11] LUMO

* Corresponding author phone: (203)791-6033; e-mail: mhao@rdg.boehringer-ingelheim.com. Corresponding author address: Boehringer Ingelheim Pharmaceuticals, Inc., 900 Ridgebury Rd., P.O. Box 368, Ridgefield, CT 06877.

10.1021/ct0600262 CCC: $33.50
Published on Web 04/05/2006

and HOMO properties,[10,12] atom polarizability,[13] and superdelocalizability,[14,15] to model hydrogen bond strength. However, due to the character of being the property of a group of atoms and the susceptibility to local environments, hydrogen bonds could not be modeled accurately by a general semiempirical or rule-based method because there are many exceptions, such as steric factors, to be accommodated by a finite set of rules. To treat complicated systems such as hydrogen bonding, the ab initio quantum chemistry method is an ideal approach since all electronic and steric effects are fully taken into account in such a treatment. Over the past decade, much work has been published on the ab initio study of hydrogen bonds.[16−30] However, few ab initio calculations have been carried out on hydrogen bond systems that are directly related to medicinal chemistry in drug design. Hydrogen bond systems encountered in drug design involve relatively large organic molecules that interact with protein targets. We explore here the application of ab initio calculations to problems of hydrogen bonding relevant to drug design.

In this work, an ab initio procedure for calculating the hydrogen bond strengths for molecules of interest to medicinal chemistry was investigated. The fundamental quantum chemistry theory for calculation of the interaction energy between two molecular species has been well established.[20,31] However, there are practical issues as to the accuracy of the calculations on specific systems and the relevance of the calculated properties to a practical process. The ab initio calculation directly yields the binding energy for the hydrogen-bonding interaction, but the strength of the hydrogen bond is determined by its free energy change. Furthermore, hydrogen bond strength, as a function of free energy, will also be influenced by the interaction of the inhibitor with water solvent. These are clearly complicated systems. However, experimental observations help to simplify the perspective considerably. There had been historically the postulation of a linear relationship between hydrogen bond free energy and enthalpy.[32,33] Abraham el al. provided clear experimental evidence supporting such a relationship.[1] They observed a linear correlation between hydrogen-bonding energy, measured by spectroscopic wavelength shift, and hydrogen bond free energy measured by the equilibrium constant of hydrogen bond formation for many hydrogen bond donors and acceptors. On the basis of these observations and being concerned with only the relative hydrogen bond strength among different acceptors with respect to a given donor, it is possible that the ab initio calculated bonding energy can yield information, even though indirectly, for hydrogen-bonding free energy. In this paper, the term hydrogen-bonding strength will be used to refer to either the thermodynamic stability or the interaction energy of a hydrogen bond in a different context, even though the term is strictly speaking only a function of the free energy.

The paper is organized as follows: the basic computational method is presented first, followed by an investigation into the sensitivity of the calculated results on hydrogen bond geometry such as separation and orientation. The main results of this work are the calculations of the energy for a variety of hydrogen bond acceptors with respect to a given donor and the subsequent comparison of these calculated results with experimental determined hydrogen-bonding constants that provide a validation of the present approach. Additional issues such as secondary interactions and effects of different donors are further investigated. In the Discussion section, comparison of the accuracy of the present method with earlier approaches will be made. Insights gained from this study will also be elaborated on.

Theory and Method

All quantum chemical calculations in this work were carried out with the program package Jaguar from Schrodinger, Inc.[34] The standard energy difference method[20,31] was used to calculate the energy of the hydrogen-bonding interaction between a donor and acceptor

$$\Delta E_{HB}(R) = E_{AD}(R) - E_A - E_D \quad (1)$$

where ΔE_{HB} is the energy of bonding interaction, R is the set coordinates that define the structure of the hydrogen-bonding complex, $E_{AD}(R)$ is the total energy of the complex, and E_A and E_D are the individual energies of the donor and the acceptor, respectively. Because the calculations involve a complex system on one hand and two subunits on the other hand, the standard molecular orbital method will use different basis functions for the complex and for the subunits, respectively, thus incurring a basis set superimposition error (BSSE)[20,31] in calculating the energy difference. To correct for the BSSE, the counterpoise procedure[31] was applied in the calculation of E_A and E_D in which the virtual orbitals of the other subunit were included in the basis functions for the subunit that is to be calculated. Two quantum chemistry methods were used in this work. The majority of the calculations were carried out using density function theory (DFT) with the B3LYP procedure and a large 6-31++G** basis set. A subset of the calculations was carried out with both DFT and the MP2 procedures to evaluate the effects of the computational methods on the calculated hydrogen bond energy. In the MP2 calculations, the 6-31G** basis set was used.

One of the major technical issues encountered in this work was to find a proper procedure for geometry optimization. For simpler hydrogen bond systems, such as when one subunit is a water molecule, a single constraint on the hydrogen bond angle would be sufficient to keep the donor and the acceptor in a proper geometry through the geometry optimization process.[35] However, for calculations involving relatively large donors and acceptors in which only one or two internal variable constraints were applied, geometry optimization by energy minimization often results in structures that are substantially different from the desired hydrogen bond conformation. The reason for such an outcome is that, when two relatively large molecules are free to move, energy minimization always brings secondary interactions, in addition to hydrogen bonding, into the total energy. The secondary interactions not only perturb the hydrogen bond geometry but also cause difficulty in separating hydrogen-bonding energy from other nonbonded contributions. In a realistic and complete inhibitor−receptor complex, the hydrogen bond interactions between the donor

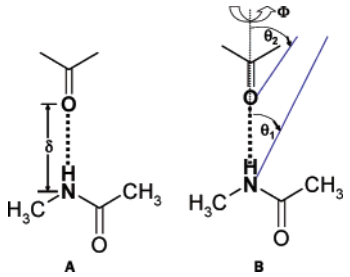

Figure 1. Variables used to define the hydrogen bond geometry between a donor and an acceptor. A: The separation between the donor and the acceptor. B: The variables defining the orientation of the acceptor with respect to the donor.

and the acceptor are not totally free to reach an absolute energy minimum but are subjected to numerous additional constraints at the binding site such as tethering from the protein structure. As a result, the hydrogen bond geometry in inhibitor−receptor complexes can be quite different from that produced from energy minimization of free subunits. We desired a procedure that computes hydrogen-bonding interaction at a specified geometry and can control the effects of secondary interaction contributions to hydrogen-bonding energy.

We devised a constrained energy minimization scheme for the calculation of hydrogen bond energy. The donor and the acceptor were first energy minimized separately with full geometry optimization. They were then brought together into a complex structure defined by one distance variable and three orientation variables. Figure 1 shows the definition of the separation variable, δ, and angular orientation variables, θ_1, θ_2, and Φ, used in this work. Without loss of generality, the donor is assumed to be fixed, and then the hydrogen bond geometry is specified by translation and rotation of the acceptor with respect to the donor. The separation between the donor and the acceptor is defined by the distance between the heavy atoms in the hydrogen bond. The orientation variable Φ describes the rotation of the acceptor around the N−O line (Figure 1) drawn between the heavy atoms of the donor and the acceptor; θ_1 defines the rotation of the acceptor away from the N−O line, with the rotation center being the heavy atom of the acceptor; θ_2 defines the rotation of the acceptor away from the N−O line, with the rotation center being the heavy atom of the donor. At a specified geometry, three atoms from the acceptor and three atoms from the donor are frozen during the geometry optimization process. These constrained atoms do not include the atoms that directly participate in the hydrogen bond. In this manner, geometry optimization can eliminate nonobvious steric problems in the system but will retain the specified hydrogen-bonding geometry. After the structure of the complex is geometrically optimized, the total energy of the complex system is first calculated, and then the energies of the donor and the acceptor are calculated at the optimized complex geometry, including the BSSE correction. To determine the energy minimum for a hydrogen bonding complex, we scanned over the separation and orientation variable space. The setup of different geometries for a given hydrogen bond system was done with a script program; the ab initio calculations on

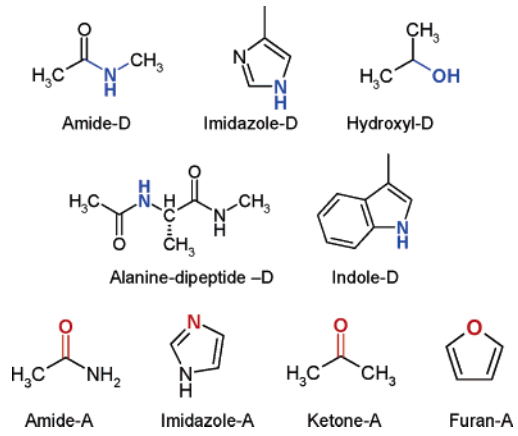

Figure 2. Chemical structures of the first group of donors and acceptors studied in this work. Donor atoms are labeled by blue color and acceptor atoms by red color.

multiple conformations were processed in parallel on a Linux cluster. A constrained geometry optimization job with the DFT method, run on one CPU of a Linux PC, took from 1 to 10 h, depending on the difficulty of the geometry optimization. A single point DFT calculation took less than 20 min on a Linux PC. MP2 calculations with geometry optimization would take a longer time.

Results

The most important hydrogen-bonding interactions to drug design are those between ligand and protein. We used an amide group to mimic the peptide hydrogen bond donor or acceptor. Hydroxyl, imidazole, and indole groups were used as surrogates for hydrogen-bonding donors from the protein side chains. Figure 2 shows a partial list of the chemical structures of the donors and acceptors studied in this work. The hydrogen bond donor and acceptor atoms are shown in blue and red colors, respectively. The labels for the different molecules will be used subsequently in describing the calculation results. We first focused on the dependency and sensitivity of the calculated hydrogen-bonding energy on the geometry of the hydrogen bond, using the amide-D (Figure 2) as donor and amide-A, imidazole-A, ketone-A, and furan-A, defined in Figure 2, as acceptors.

1. Hydrogen-Bonding Energy as a Function of Separation between Donor and Acceptor. Figure 3 shows the interaction energy, calculated with the DFT procedure, of four hydrogen bond complexes as a function of the separation. By definition, a larger negative energy indicates a stronger hydrogen-bonding interaction. Results presented in Figure 3 show that, for all four complexes, there is a sharp energy minimum at a separation of 3.0 Å. At larger separations, the interaction energy for all four complexes decays slowly, at a rate proportional to the inverse distance, resembling the behavior of electrostatic interactions. There are large differences in the interaction energy among the four hydrogen bond complexes: ranging from −6.53 kcal/mol for the imidazole acceptor to −2.13 kcal/mol for the furan acceptor. The minimum interaction energy for the amide and ketone acceptors is −5.44 and −4.06 kcal/mol, respectively. The order of hydrogen-bonding strength for these four acceptors by this calculation is imidazole > amide > ketone

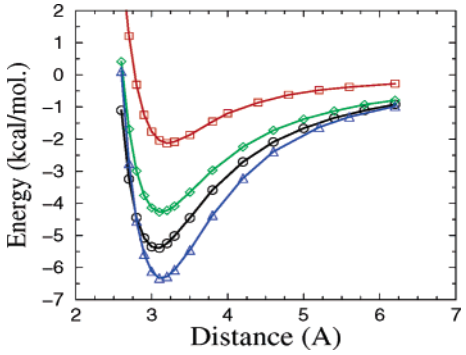

Figure 3. The interaction energy of a hydrogen bond as a function of the separation between the donor and the acceptor calculated with the DFT/B3LYP (6-31++G**) method. The donor is amide-D; the colored coded curves in blue, black, green, and red are for imidazole-A, amide-A, ketone-A, and furan-A as the acceptor, respectively. The molecular structures are given in Figure 2.

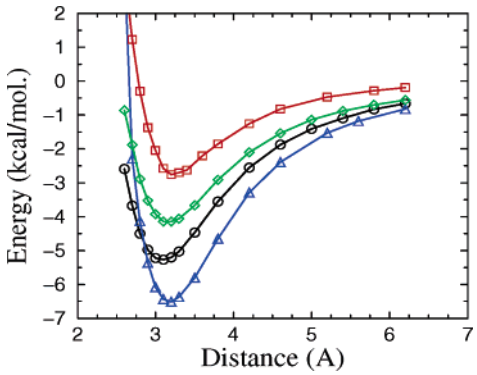

Figure 4. Hydrogen bond energy as a function of the separation between the donor and the acceptor calculated with the MP2 method. The labeling of the curves is identical to that of Figure 3.

> furan. Despite the large difference in the interaction energy, there is little difference in the energy-minimum distance between the four complexes. It was noticed that the energy-minimum distances of the hydrogen bonding calculated here appear slightly longer than the average hydrogen bond length observed from neutron diffraction measurements [Chapter 7 of ref 3]. However, they are within the range of calculated hydrogen bond lengths using different ab initio methods [Table 4.3 of ref 3, Chapter 2 of ref 20].

Quantitatively similar results for the interaction energy of the four hydrogen-bonding complexes were obtained from calculations with the MP2 method. Figure 4 shows the MP2 results. The minimum energy values calculated by the MP2 method are −6.51, −5.26, −4.15, and −2.76 kcal/mol for imidazole-A, amide-A, ketone-A, and furan-A, respectively. The order of hydrogen bond strengths for the four acceptors calculated by DFT and MP2 methods are identical. The calculated energy-minimum distances by the two methods are also similar. However, for furan-A, the weakest hydrogen bond acceptor of the four molecules, the minimum energy calculated by the MP2 method, is lower by 0.63 kcal/mol than that obtained by the DFT method. We observed that the MP2 method yielded, in our calculations, somewhat larger interaction energies for weaker hydrogen bond acceptors (such as the CH- - -O hydrogen bond) than the DFT method. However, MP2 calculations were considerably more time-consuming in geometry optimization on large complex systems. Since the calculation results with the two methods are similar for most common hydrogen bond acceptors, we chose to use the DFT method on all subsequent calculations in this work for convenience and consistency.

2. Directionality of Hydrogen-Bonding Interaction. Moving onto the directional properties of hydrogen bonds, we calculated the interaction energy for acceptors in different orientations with respect to the donor. Figure 5 shows sets of orientations between the imidazole acceptor and the amide donor generated by systematic variation of the variables of θ_1, θ_2, and Φ at a fixed separation of 3 Å between the two subunits. Figure 6 shows the hydrogen-bonding energy of these structures, where curve *a*, *b*, and *c* correspond to the geometric series shown in Figure 5, parts a−c, respectively. These results indicate that the directional property of a hydrogen-bonding interaction is anisotropic. Using the structure of the donor amide as the reference frame, i.e., fixing the position of amide donor, the orientation of the hydrogen bond of this system can vary in two orthogonal directions: i.e., the acceptor imidazole ring swings either within the plane or perpendicular to the plane of the amide structure. The ensemble of structures in Figure 5a is generated as a result of the imidazole ring swinging in the direction perpendicular to the structural plane of the amide donor. Curve *a* in Figure 6 shows the variation of the hydrogen-bonding energy as a function of such an orientational change. In this direction, the acceptor imidazole can rotate by ±45° from the normal direction, while the hydrogen-bonding energy changes by less than 0.5 kcal/mol. In contrast, when the imidazole ring swings within the structural plane of amide around angle θ_1, with the resulting conformations shown in Figure 5c, the hydrogen-bonding interaction weakens rapidly as shown by curve c of Figure 6. The directional dependency of the hydrogen bond energy as a function of angle θ_2, i.e., another orientational variation within the structural plane of the amide donor, and the changes of hydrogen bond energy are between the above two extremes, see the energy curve *b* of Figure 6 and in reference to the ensemble of structures shown in Figure 5b.

Figure 7 shows structural variations between the amide donor and the amide acceptor resulting from rotation of the acceptor around the angle Φ, with the angle θ_1 fixed at 0° and θ_2 at 15° and 30°, respectively. The separation between the donor and the acceptor was fixed at 3.0 Å. Figure 8 shows the calculated hydrogen-bonding energy for the two structural series of Figure 7. The peaks in the energy plots in Figure 8 correspond to structures in which the amide acceptor orients toward the right-hand side methyl group of the amide donor seen from the front of Figure 7. Significantly, as the acceptor tilts closer to the donor, i.e., when θ_2 changes from 15° to 30°, the energy peak increases quickly, as can be seen from a comparison of curves *a* and *b* of Figure 8. This result suggests that the primary reason for the weakening of the hydrogen-bonding interaction in these directions is steric interference between the donor and the acceptor.

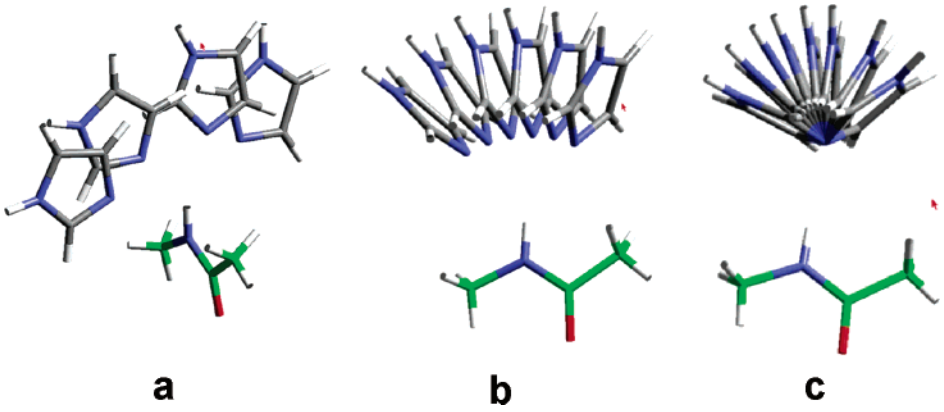

Figure 5. Different orientations between the imidazole acceptor and the amide donor at a fixed separation of 3 Å. a: Conformations generated from varying angle θ_2 with $\Phi = 90°$ and $\theta_1 = 0$. b: Conformations generated from varying angle θ_2 with $\Phi = 0$ and $\theta_1 = 0$. c: Conformations from a variation of angle θ_1 with $\Phi = 0$ and $\theta_2 = 0$.

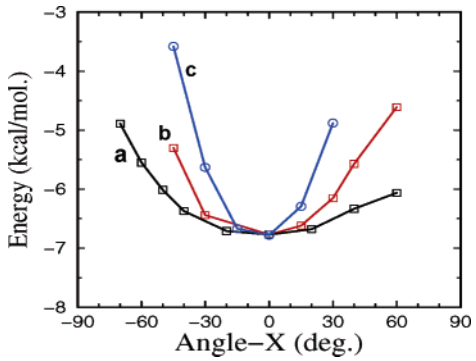

Figure 6. Hydrogen-bonding energy between the amide donor and the imidazole acceptor as a function of orientational variables. Curve *a* is for the conformational series of Figure 5a with the variable X = θ_2. Curve *b* is for the conformational series of Figure 5b with the variable X = θ_2. Curve *c* is for the conformational series of Figure 5c with the variable X = θ_1.

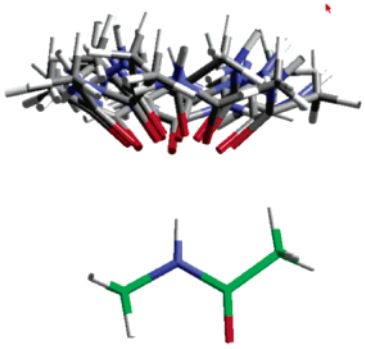

Figure 7. Orientation of the hydrogen bond between the amide donor and the amide acceptor at a constant separation of 3 Å. The conformations are generated by varying angle Φ with $\theta_1 = 0$ and $\theta_2 = 15°$ and 30°, respectively.

Combining the results from Figures 6 and 8, we observed that the hydrogen-bonding energy is rather insensitive to the orientation between the donor and the acceptor. There are rather broad ranges of directions in which acceptor (or donor) can tilt away from the normal direction of the hydrogen bond, by up to ±45°, while the energy changes less than 0.5 kcal/mol. However, the hydrogen-bonding energy could quickly deteriorate when steric interference arises between the donor and the acceptor. The directional anisotropy is determined by the specific chemical structure of the donor and the acceptor and has to be dealt with on a case-by-case basis. In contrast, hydrogen-bonding energy is much more sensitive to the separation between the donor and the acceptor: a 0.2 Å shift in separation can cause over a 1 kcal/mol change in energy. Based on these observations, in our subsequent calculations of hydrogen-bonding energies for different molecules, the initial geometry of the hydrogen bond was set up based on templates of analogous systems from X-ray crystal structures of inhibitor−protein complexes. The orientation between the donor and the acceptor was checked sparsely to ensure no steric interference. The separation was systematically scanned to locate the energy minimum of the hydrogen-bonding interaction.

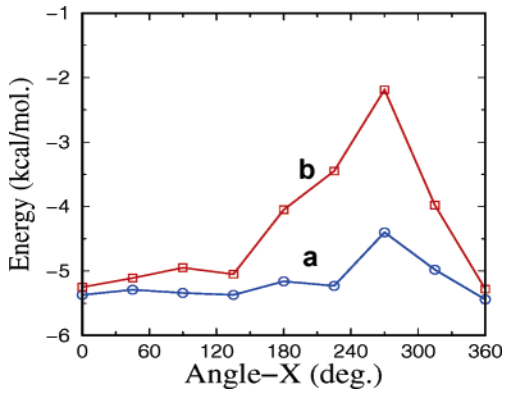

Figure 8. Interaction energy between the amide donor and the amide acceptor as a function of the hydrogen bond orientation at a constant separation of 3 Å. The hydrogen bond conformations are shown in Figure 7. The variable angle-X is Φ angle. Curves *a* and *b* correspond to angle $\theta_2 = 15°$ and 30°, respectively, while angle θ_1 is fixed at 0° and $\theta_2 = 30°$.

3. Comparison of the Hydrogen-Binding Energy of Different Acceptors. Having investigated the calculation procedure, we turned to the validation of the computational procedure by comparison of the calculation results with experimental data. Abraham et al.[1] determined the hydrogen-bonding constants for a large number of donors and acceptors. Hydrogen bond donors are generally confined only to a few types of NH or OH groups. In contrast, there is a large

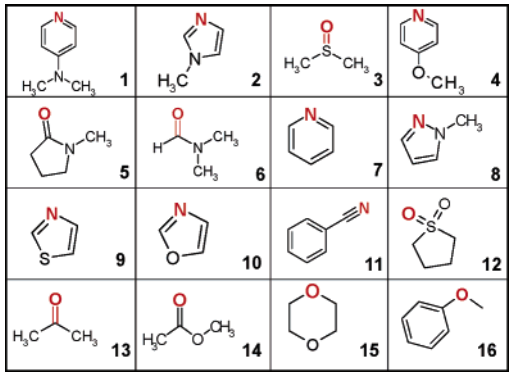

Figure 9. The structures of the hydrogen bond acceptors studied in this work. The atom in red color is the hydrogen bond acceptor.

variety of hydrogen bond acceptors, comprising sp^2 and sp^3 oxygen atoms and sp^1, sp^2, and some sp^3 nitrogen atoms from a large number of functional groups and heterocycles. It has been reported that theoretical modeling of hydrogen bond acceptors is more difficult than modeling hydrogen bond donors.[15] In this work, we selected 16 molecules that cover the major categories of acceptors in the list of Abraham et al.[1] In their experiment, the hydrogen-bonding constants for these acceptors were measured against a common hydrogen bond donor, 4-nitrophenol dissolved in 1,1,1-trichloroethane (TCE), which has a dielectric constant of 7.53. To avoid self-interference in the experiment, all hydrogen donors in an acceptor molecule have to be blocked. As a result, only a limited number of acceptors could be measured in such an experiment. The structures of our selected acceptors are shown in Figure 9. We calculated the interaction energies for these acceptors with the amide hydrogen bond donor (shown in Figure 2). In our calculations, the acceptors and donor were arranged in a fixed orientation based on the procedure described in the previous subsection. Energy calculations were carried out on each structure at 5 different donor−acceptor separations, from 2.8 to 3.2 Å at 0.1 Å intervals. Geometric optimization was done for each structure with constrained energy minimization previously described in the theory section. The lowest energy found from the 5 structures was used as the hydrogen-bonding energy for a given acceptor.

The computational results are collected in Table 1. Column 1 of the table is the label of the acceptor; column 2 contains the calculated hydrogen-bonding energy from this work; and column 3 shows the logarithm of experimental hydrogen-bonding constants, log K_β. The spreading scope of the calculated hydrogen bond energy and log K_β are not the same; and, physically, the first value is energy, while the second number corresponds to free energy. However, an excellent linear correlation exists between the two sets of values. Figure 10 shows the plot of the calculated ΔE_{HB} against experimental log K_β. The following relationship and statistics were found:

$$\log K_\beta = -0.892\Delta E_{HB} - 2.630 \qquad (2)$$

$$N = 16,\ R^2 = 0.94,\ F = 229$$

It is observed that the correlation obtained from this method

Table 1. Hydrogen-Bonding Energies of Different Acceptors with Respect to an Amide Donor

acceptor[a]	ΔE_{HB}[b]	log K_β [c]	acceptor[a]	ΔE_{HB}[b]	log K_β [c]
1	−6.78	3.54	9	−4.76	1.90
2	−6.73	3.68	10	−4.72	1.67
3	−6.20	3.06	11	−4.60	1.06
4	−5.95	2.87	12	−4.39	1.61
5	−5.84	3.12	13	−4.27	1.50
6	−5.37	2.81	14	−4.06	1.43
7	−5.36	2.52	15	−4.03	1.28
8	−5.10	2.22	16	−2.96	0.30

[a] The numerical labels of the acceptors correspond to those defined in Figure 9. [b] ΔE_{HB}, in units of kcal/mol, was defined by eq 1 and calculated with the DFT/B3LYP (6-31++G**) method including BSSE correction. [c] log K_β values adopted from Table 3 of ref 1, also listed in Figures 3 and 4 of ref 15.

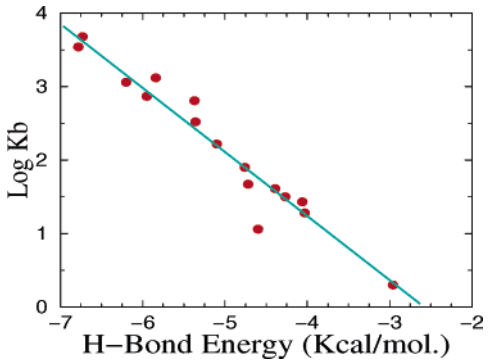

Figure 10. Correlation plot between theoretically calculated hydrogen-bonding energies (H-bond energy) and the logarithm of experimentally measured hydrogen-bonding constants (log K_β). The two sets of values fit the following relationship: log $K_\beta = -0.892E - 2.36$ with $N = 16$, $R^2 = 0.94$, $F = 229$.

is significantly better than that obtained by Gancia et al.[15] Those authors modeled hydrogen bond strength based on a parametrization of log K_β with respect to quantum-chemistry calculated properties of atomic electrophilic superdelocalizability (SE). The regression model between SE and log K_β has larger errors with R^2 in the range of 0.59−0.75. Moreover, in the parametrization approach, different types of acceptors, such as nitrogen and oxygen acceptors, have to be modeled with separated QSAR equations. In our present approach, all acceptors follow the same relationship between the calculated interaction energies and the logarithm of experimental hydrogen-bonding constants. This is consistent with a common physical process for the hydrogen bonding between the pairs of donors and acceptors.

The ab initio calculation results provide insights into the dominant factor for hydrogen-binding strength. While the hydrogen-bonding constant is determined by the binding free energy, the fact that there is an excellent linear correlation between the hydrogen-binding interaction energy and the binding constant suggests that the logarithm of binding constants for different acceptors with respect to given donors are largely determined by an enthalpic factor arising from the interaction energy between the donor and the acceptor. The entropic contribution to the hydrogen-bonding constant would then amount to a constant for all the acceptors. A constant entropic component is not a universal phenomenon;

very different entropic contributions have been observed in ITC experiments on ligands bound to a given receptor.[36] An alternative plausible explanation to the linear correlation between the hydrogen-binding energy and the logarithm of the hydrogen-binding constant could be that the entropy varies linearly with enthalpy. Further investigation will be needed to resolve these two possibilities.

There are interesting individual cases among the set of 16 acceptors. The oxazole molecule, #10 in Figure 9, has two hydrogen bond acceptors: a nitrogen and an oxygen atom. Because only one hydrogen-binding constant was reported for this molecule,[1] it was not clear whether the nitrogen or the oxygen is the actual acceptor in oxazole. Our calculation indicated that the hydrogen-bonding energy for the nitrogen and oxygen atoms in oxazole are −4.72 and −1.23 kcal/mol, respectively. Based on the strong statistic from eq 2, we can confidently predict that the nitrogen atom of the oxazole is the acceptor, whereas the oxygen atom has virtually no acceptor capacity in this particular molecule. A simple explanation of the difference between the hydrogen bond strength of nitrogen and oxygen atoms in oxazole is that the nitrogen atom is more basic than oxygen hence a better hydrogen bond acceptor. Another interesting case is sulfur containing acceptors. As shown in Table 1, sulfoxide is a stronger hydrogen bond acceptor than sulfone, and our calculation indicates that sulfonamide, with a calculated hydrogen-bonding energy of −3.87 kcal/mol, is an even weaker hydrogen bond acceptor than sulfone.

The hydrogen-bonding strength of an acceptor can be easily modified by chemical substitutions on the acceptor's structure. Attesting to this point are the three pyridine derivatives, #1, 4, and 7, shown in Figure 9 and Table 1. The calculated hydrogen-bonding energy of simple pyridine is −5.36 kcal/mol. Substitution on the para position with electron-donating groups such as methoxyl or dimethylamine gives molecules #1 and 4, which have larger hydrogen-bonding energies, −5.95 and −6.78 kcal/mol, respectively. Because dimethylamine is a stronger electron donating group than methoxyl, the former induces a larger increase in the hydrogen-bonding strength for pyridine. Following this line of reasoning, if electron-withdrawing groups are substituted on the pyridine, the hydrogen-bonding strength of the resulting molecules should be smaller than that of pyridine. Indeed, our calculations indicated that the hydrogen-bonding energy of pyridine para-substituted with nitro ($-NO_2$), nitrile ($-C{\equiv}N$), and −CF3 groups are −3.17, −4.27, and −4.26 kcal/mol, respectively. In general, substitution with an electron-withdrawing group will decrease the hydrogen bond strength of an acceptor but will increase the strength of a hydrogen-bonding donor. Electron-donating groups have opposite effects. The electronic effects can be transmitted effectively through conjugated systems. Modulating the hydrogen-bonding strength with proper substituents on a hydrogen donor and/or acceptor can be a very useful strategy in drug optimization and has been utilized in the drug optimization process.[2]

4. Effects of Hydrogen Bond Donors. One assumption in the previous subsection was that the *order* of the hydrogen-bonding energy for different acceptors should be generally the same regardless which donor was used in the calculation. The same assumption also holds for the experimentally measured hydrogen-bonding constants for acceptors,[1] which are transferable across different donors. We verified this assumption by calculating the interaction energy of the four acceptors shown in Figure 2 with the five donors shown in the same figure. Table 2 shows the hydrogen-bonding energies for these molecules: the columns correspond to the donors, while the rows correspond to acceptors. Crossing the columns, it is observed that the hydrogen-bonding strength of the donor is in the order of imidazole > indole > amide > hydroxyl. Crossing the rows, it is observed that the order of acceptor strength is as imidazole > amide > ketone > furan. As expected, the same order of hydrogen-bonding strength for the acceptor series holds across the different donors. In the other direction, the same order of hydrogen-bonding strength for the donor series holds across different acceptors.

Table 2. Hydrogen-Bonding Energy between Acceptors and Donors[a]

	amide-D	hydroxyl-D	indole-D	imidazole-D	alanine dipeptide-D
imidazole-A	−6.55	−6.24	−7.11	−7.96	−8.03
amide-A	−5.44	−4.51	−6.27	−6.92	−6.94
ketone-A	−4.06	−3.54	−5.12	−5.78	−5.52
furan-A	−2.13	−2.22	−2.47	−2.75	−3.01

[a] The hydrogen-bonding energy is in units of kcal/mol. The chemical structures of the donors (top row) and acceptors (first column) are shown in Figure 2. The energy values were calculated by the DFT/B3LYP(6-31++G**) method with BSSE correction.

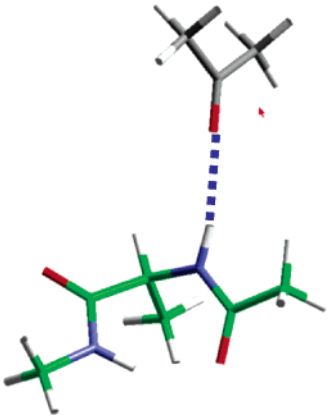

Figure 11. The conformation used in calculating the hydrogen-bonding energy for a ketone donor and an alanine dipeptide donor.

The case of alanine dipeptide as donor warrants further discussion. While we have been largely concerned so far with small donors such as amide and hydroxyl, the hydrogen-bonding groups of a receptor interacting with drug molecules are parts of the protein. It is of interest to know how a hydrogen bond donor from a larger structure interacts with a small acceptor or vice versa. Figure 11 shows the conformation of hydrogen-bonding interaction between a ketone acceptor and an alanine dipeptide donor. Compared to the amide donor shown in Figures 5 and 7, the alanine dipeptide clearly has more structural mass that is not directly involved in hydrogen bonding with the acceptor. Our

computational result has indicated that the amide donor is not the strongest in the set of five donors. Nonetheless, as shown in Table 2, the calculated interaction energy for all the acceptors with the alanine dipeptide donor is larger than with other donors. Clearly, there are secondary interactions other than hydrogen bonding that contributed to the calculated total interaction energy between the alanine dipeptide and the acceptors. This result suggests that, to obtain an accurate estimate for the hydrogen-bonding interaction in a given binding site of the receptor, secondary interactions have to be taken into account in the total energy calculation. It is likely that secondary interactions will vary in different binding sites and have to be dealt with on an individual basis. The constrained energy minimization procedure designed here could be used to calculate more complicated systems with a specified geometry. One difficulty in extending the present calculation to the protein binding site was that there is no experimental hydrogen bond parameter that can be compared with, thereby validate, theoretically calculated values.

Discussion

A key result to this work is the observation that ab initio calculated hydrogen-bonding energy linearly correlates with the logarithm of experimentally measured hydrogen-bonding constants. This provides a basis for theoretical prediction of hydrogen-bonding strength for drug design by ab initio calculations. The results presented here were obtained from DFT/B3LYP calculations with the 6-31++G** basis set. We had found that calculations at the MP2/6-31G** level gave quantitatively similar results. The MP2 calculations seemed to produce larger interaction energies for weaker hydrogen-bonding acceptors or donors, but the DFT/B3LYP calculation is faster and can provide consistent and satisfactory information.

It should be emphasized again that the calculated property here is the hydrogen-bonding energy. Since it is the free energy change from solvent to a receptor binding site that ultimately determines the hydrogen-bonding constant, it was not obvious that the interaction energy alone can explain the variations in hydrogen-binding strength from one acceptor to another. However, the excellent linear relationship between the calculated interaction energies at ab initio quantum level and logarithm of hydrogen-bonding constants for a variety of acceptors gives support to the theory that the enthalpic contribution from the hydrogen-bonding energy determines whether one hydrogen bond is stronger than another. Plausibly, entropic factors make only a constant contribution to the free energy change when different acceptors bind to a given donor. It is likely that a similar relationship also holds in the case of ligands binding to a receptor, at least for a series of compounds adopting the same binding mode. Another plausible explanation is that the entropic component varies proportionally with enthalpy in these systems.

Predicting the hydrogen-bonding strength by ab initio quantum chemical calculations has considerable advantages over semiempirical parametrization approaches.[8−15] First, the correlation between the ab initio calculated energies and experimental hydrogen-binding constants is much better than that from a parametrization approach.[15] Second, all different acceptors or donors can be evaluated by the same computational procedure in a unified correlation equation with the ab initio method. In contrast, with the parametrization approach, multiple equations would be needed for different types of donors or acceptors. For traditional QSAR models, the application range of a given model is even narrower. Third, the ab initio method can handle steric factors in a hydrogen bond system in a straightforward manner since ab initio calculation takes all electronic and steric factors of the system into account. Steric effects have been a major problem for the transferability of parametrization-based hydrogen bond models because, in such models, only one subunit of the donor−acceptor pair in a hydrogen bond is considered. A major drawback of the ab initio method is the high demand on computational resources. When fast Linux clusters become readily available, multiple ab initio calculations involved in distance scans can be carried out in parallel such that results for a handful of molecules typically encountered in an SAR study can be obtained overnight. Therefore, the computational demand of the ab initio method is not presently a major obstacle.

The directional properties of hydrogen bonds are important to both computational studies and practical applications. A survey of X-ray and neutron diffraction structures of organic and biological molecules[3] had found that the directional dependency of hydrogen bonds is weak: the distribution of hydrogen-bonding orientation almost evenly spreads over a broad band of ±50° from the normal direction for donor and/or acceptor. Our calculations also indicated that, at the normal hydrogen-bond distance, the interaction energy is not very sensitive to angular variations up to ±45° in certain directions. This property of the hydrogen bonds simplifies the task of calculating the minimum energy of hydrogen-bonding interaction. The orientation of the system needs only to be sparsely checked to ensure that there is no major steric problem. However, because of the sensitivity of the interaction energy to donor−acceptor separation, a scan over the distance variable would be required to identify the accurate energy minimum. Ireta et al.[26] reported that the accuracy of the DFT calculated hydrogen-bonding energy depends on the bond directionality. When the hydrogen bond deviates from linearity, the discrepancy between the DFT energy and the values calculated by the MP2 method[37,38] increases. The reason for the somewhat different observations between Ireta et al. and the present work is not clear.

Another interesting result of the present work is that the directionality of the hydrogen bond is not isotropic. While the statistics for a large number of hydrogen-bonding structures from X-ray data shows a nearly uniform distribution of hydrogen bond orientation over a broad range, individual molecules have different structures and steric characteristics. At the individual molecule level, hydrogen-bonding energy can deteriorate quickly in some directions. Using analogous X-ray structures as templates is a reasonable approach for setting up the geometry of a hydrogen bond system for energy calculation. When steric problems occur, the orientation of the hydrogen bond has to be adjusted to obtain a correct hydrogen-bonding energy.

Secondary interactions will be a complicating factor for an accurate estimate of the hydrogen-bonding interaction for ligand binding to a protein receptor. In an earlier study by Scheiner et al.,[35] it was reported that when the donor molecule increases in size from one to two amino acid units, the hydrogen-bonding energy of the water with the peptides changes only by ~10%. It is possible that a water molecule is a very small and simple system which would experience less secondary interactions with the donor molecules. For larger acceptors such as drug molecules interacting with protein receptors, secondary interactions accompanying the hydrogen-bonding interaction could be considerably large. To accurately estimate the strength for larger hydrogen bond donors or acceptors interacting with specific protein receptors, secondary interactions have to be dealt with adequately.

The usefulness of theoretically calculated hydrogen bond energies to drug design has been demonstrated by Bingham et al.[2] In their work, the SAR of a series of analogous molecules bound to the protein kinase IKK can be accounted for by the hydrogen bond strength of these compounds. That result was one of the motivations of the present study. With a more accurate and comprehensive calculation scheme, the present method can be used for problems involving hydrogen bond bioisostere in drug design; a comparison of the hydrogen-bonding energies of the bioisostere provides a more rational basis for the bioisostere approach. The calculated hydrogen-bonding energy by the present method can also be used in the refinement of force field parameters for molecular modeling simulation. Based on the author's experience, currently used force fields were not parameterized to reproduce the difference of the experimentally measured hydrogen bond strengths among different donors and acceptors.

Conclusion

We demonstrated in this work that ab initio calculated hydrogen-binding energy has an excellent linear correlation with logarithm of experimental hydrogen-bonding constants. This provides a basis for the theoretical prediction of hydrogen-bonding strength for a series of acceptors (or donors) with respect to a given donor (or acceptor). The ab initio approach advances the level of sophistication of theoretical modeling of hydrogen bonds and may provide deeper insights to the mechanism of hydrogen-bonding interactions. The method also provides a tool for handling steric effects, conformational properties, and secondary interactions of specific systems in predicting hydrogen bond strengths. These effects were traditionally not treated in empirical and parametrization approaches for hydrogen bonds. We believe that the method developed here is helpful in dealing with hydrogen bond related SAR problems in drug design.

Acknowledgment. The author thanks Dr. Ingo Mugge for his valuable comments and suggestions about this work and Drs. Sandy Farmer and James Stevenson for their thorough reading and editing of the paper.

References

(1) Abraham, M. H.; Duce, P. P.; Ptior, D. V.; Barratt, D. G.; Morris, J. J.; Taylor, P. J. *J. Chem. Soc., Perkin Trans*. **1989**, *11*, 1355−1374, 1989.

(2) Bingham, A. H.; Davenport, R. J.; Gowers, L.; Knight, R. L.; Lowe, C.; Owen, D. A.; Parry, D. M.; Pitt, W. R. *Bioorg. Med. Chem. Lett*. **2004**, *14*, 409−412.

(3) Jeffrey, G. A.; Saenger, W. *Hydrogen Bonding in Biological Structures*; Springer-Verlag: Berlin, 1989.

(4) Umeyama, H.; Morokuma, K. *J. Am. Chem. Soc*. **1977**, *99*, 1316−1332.

(5) Morokuma, K.; Kitaura, K. In *Chemical applications of atomic and molecular electrostatic* potentials; Politzer, P., Truhlar, D. G., Eds.; Plenum: New York, 1981; pp 215−242.

(6) Jeffrey, G. A.; Maluszynska, H. *Int. J. Biol. Macromol*. **1982**, *4*, 173−185.

(7) Taylor, R.; Kennard, O.; Versichel, W. *J. Am. Chem. Soc*. **1983**, *105*, 5761−5766.

(8) Fujita, T.; Nishioka, T.; Nakajima, M. *J. Med. Chem*. **1977**, *20*, 1071−1081.

(9) Charton, M.; Charton, B. I. *J. Theor. Biol*. **1982**, *99*, 629−644.

(10) Wilson, L. Y.; Famini, G. R. *J. Med. Chem*. **1991**, *34*, 1668−1674.

(11) Murray, J. S.; Politzer, P. *J. Org. Chem*. **1991**, *56*, 6715−6717.

(12) Dearden, J. C.; Cronin, M. T. D.; Wee, D. In *QSAR and molecular modeling: concepts, computational tools and biological applications*; Sanz, F., Giraldo, J., Manaut, F., Eds.; Proust Science Publishers: Barcelona, Spain, 1995; pp 117−119.

(13) Dearden, J. C.; Cronin, M. T. D.; Wee, D. *J. Pharm. Pharmacol*. **1997**, 49 Suppl. 4, 110.

(14) Dearden, J. C.; Ghafourian, T. *J. Chem. Inf. Comput. Sci*. **1999**, *39*, 231−235.

(15) Gancia, E.; Montana, J. G.; Manallack, D. T. *J. Mol. Graphics Modell*. **2001**, *19*, 349−362.

(16) Kollman, P. A.; Allen, L. C. *Chem. Rev*. **1972**, *72*, 283−303.

(17) *Modeling the hydrogen bond;* Smith, D. A., Ed.; American Chemistry Society: Washington, DC, 1994; Vol. 569.

(18) Scheiner, S.; Kar, T. *J. Am. Chem. Soc*. **1995**, *117*, 6970−6973.

(19) Pan, Y.; McAllister, M. A. *J. Org. Chem*. **1997**, *62*, 8171−8176.

(20) Scheiner, S. *Hydrogen Bonding. A Theoretical Perspective*; Oxford University Press: New York, 1997.

(21) Rablen, P. R.; Lockman, J. W.; Jorgensen, W. L. *J. Phys. Chem. A* **1998**, *102*, 3782−3797.

(22) Gu, J.; Leszczynski, J. *J. Phys. Chem. A* **2000**, *104*, 7353−7358

(23) Lukin, O.; Leszczynski, J. *J. Phys. Chem. A* **2002**, *106*, 6775−6782.

(24) Sukhanov, O. S.; Shishkin, O. V.; Gorb, L.; Podolyan, Y.; Leszczynski, J. *J. Phys. Chem. B* **2003**, *107*, 2846−2852.

(25) Domingo, L. R.; Andres, J. *J. Org. Chem.* **2003**, *68*, 8662−8668.

(26) Ireta, J.; Neugebauer, J.; Scheffler, M. *J. Phys. Chem. A* **2004**, *108*, 5692−5698.

(27) Guo, H.; Gresh, N.; Roques, B. P.; Salahub, D. R. *J. Phys. Chem. B* **2000**, *104*, 9746−9754.

(28) Alia, J. M.; Edwards, H. G. M. *J. Phys. Chem. A* **2005**, *109*, 7977−7987.

(29) Quinonero, D.; Garau, C.; Frontera, A.; Ballester, P.; Costa, A.; Deya, P. M. *J. Phys. Chem. A* **2005**, *109*, 4632−4637.

(30) Zhao, Y.; Truhlar, D. G. *J. Phys. Chem. A* **2005**, *109*, 4209−4212.

(31) Boys, S. F.; Bernardi, F. *Mol. Phys.* **1970**, *19*, 553−566.

(32) Jencks, W. P. *Adv. Enzymol*. **1975**, *43*, 219−410;

(33) Page, M. I. *Angew. Chem., Int. Ed. Engl*. **1977**, *16*, 449−459.

(34) *Jaguar v.5.5*; Schrodinger, Inc.: 2005.

(35) Scheiner, S.; Kar, T.; Gu, Y. *J. Biol. Chem*. **2001**, *276*, 9832−9837.

(36) Kroe, R. R.; Regan, J.; Proto, A.; Peet, G. W.; Roy, T.; Landro, L. D.; Fuschetto, N. G.; Pargellis, C. A.; Ingraham, R. H. *J. Med. Chem.* **2003**, *46*, 4669−4675.

(37) Vargas, R.; Garza, J.; Dixon, D. A.; Hay, B. P. *J. Am. Chem. Soc*. **2000**, *122*, 4750−4755.

(38) Vargas, R.; Garza, J.; Friesner, R. A.; Stern, H.; Hay, B. P.; Dixon, D. A. *J. Phys. Chem. A* **2001**, *105*, 4963−4968

CT0600262