J|A|C|S
ARTICLES
Published on Web 08/08/2002

Hydrogen Bonding in High-Resolution Protein Structures: A New Method To Assess NMR Protein Geometry

Rebecca S. Lipsitz, Yugal Sharma, Bernard R. Brooks, and Nico Tjandra*

Contribution from the Laboratory of Biophysical Chemistry, Building 50, Room 3503, National Heart, Lung, and Blood Institute, National Institutes of Health, Bethesda, Maryland 20892-8013

Received May 10, 2002. Revised Manuscript Received July 16, 2002

Abstract: An analysis of backbone hydrogen bonds has been performed on nine high-resolution protein X-ray crystal structures. Backbone hydrogen-bond geometry is compared in the context of X-ray crystal structure resolution. A strong correlation between the hydrogen-bond distance, R_{HO}, and the hydrogen-bond angle, θ_{NHO}, is observed when the X-ray crystal structure resolution is <1.00 Å. Ab initio calculations were performed to substantiate these results. The angle and distance limits found in our correlation for the backbone hydrogen-bond geometry can be used to evaluate the quality of protein structures and for further NMR structure refinement.

Introduction

Ever since Pauling described the properties of hydrogen bonds and their role in forming the α-helix and β-sheet,[1,2] protein chemists have viewed hydrogen bonding as an important context in which to understand protein stability,[3] enzyme catalysis,[4] protein folding,[5,6] and formation of secondary structure.[7] No description of protein structure is complete without including hydrogen bonding, and most secondary structure algorithms look for specific patterns of hydrogen bonds to determine regions and elements of secondary structure.[8] In the past 50 years, there has been an extensive amount of work on hydrogen bonding using both theoretical and empirical data. Surveys of crystal structures from the Cambridge Structural Database and the Protein Data Bank have been carried out in an attempt to accurately characterize properties of the hydrogen-bond geometry such as the proton donor−oxygen acceptor bond length and the N−H···O angle. The majority of the data in these studies is comprised of molecules whose structures have been determined using X-ray crystallography. Therefore, the precision and, in turn, the accuracy of these studies are limited by the crystal structure resolution.

Despite the vast number of papers in the literature dealing with the topic of hydrogen bonding in biological settings, only a handful of comprehensive studies have been published in the past 20 years. Two of these papers by Baker and Hubbard[9] and Taylor and Kennard[10] are still cited extensively in the current literature on hydrogen bonding. In Baker and Hubbard's survey of 16 proteins, the resolution ranges from 1.40 to 1.80 Å, which is sufficient to obtain many useful data concerning hydrogen-bonding properties in proteins. However, the number of protein structures with a resolution better than 1.00 Å has increased steadily over the last 20 years, allowing for some additional observations to be made.

The definition of a hydrogen bond varies widely. In part, this stems from the choice of criteria used to define a hydrogen bond, that is, distance and angle ranges or an energy cutoff from an electrostatic potential function with distances and angles as variables. Another issue that arises when trying to define the hydrogen bond is that there are no precise distance and angle limits above which the hydrogen-bond interaction ceases to exist. As a result, most definitions are not definitive. The energy cutoff is somewhat arbitrary because "there is no discontinuity in energy as a function of distance or alignment that governs the interaction".[11]

It is assumed from small molecule studies that ideal hydrogen bonds have a linear orientation between the donor proton and acceptor oxygen.[12] However, the manner in which the hydrogen-bond angle compensates for deviations from linearity has not been clearly detailed. Hydrogen-bond geometry is often described in terms of separate distance limits and angle ranges and not in terms of correlated distances and angles, that is, a restricted distance for a given angle or, vice versa, a restricted angle for a given distance. An empirically derived correlation, albeit not a strong one, between R_{HO} and the hydrogen-bond angle, θ_{NHO}, has been shown previously.[9,10] Because the average

* To whom correspondence should be addressed. Phone: (301) 402−3029. Fax: (301) 402-3404. E-mail: nico@helix.nih.gov.

(1) Pauling, L.; Corey, R. B.; Branson, H. R. *Proc. Natl. Acad. Sci. U.S.A.* **1951**, *37*, 205−211.
(2) Pauling, L.; Corey, R. B. *Proc. Natl. Acad. Sci. U.S.A.* **1951**, *37*, 729−740.
(3) Shi, Z.; Krantz, B. A.; Kallenbach, N.; Sosnick, T. R. *Biochemistry* **2002**, *41*, 2120−2129.
(4) Frey, P. A.; Whitt, S. A.; Tobin, J. B. *Science* **1994**, *264*, 1927−1930.
(5) Creighton, T. E. *Curr. Opin. Struct. Biol.* **1991**, *1*, 5−16.
(6) Hunt, N. G.; Gregoret, L.; Cohen, F. E. *J. Mol. Biol.* **1994**, *241*, 214−225.
(7) Stickle, D. F.; Presta, L. G.; Dill, K. A.; Rose, G. D. *J. Mol. Biol.* **1992**, *226*, 1143−1159.
(8) Kabsch, W.; Sander, C. *Biopolymers* **1983**, *22*, 2577−2637.
(9) Baker, E. N.; Hubbard, R. E. *Prog. Biophys. Mol. Biol.* **1984**, *44*, 97−179.
(10) Taylor, R.; Kennard, O. *Acc. Chem. Res.* **1984**, *17*, 320−326.
(11) Berndt, K. D. *Protein Secondary Structure*; 1996.
(12) Taylor, R.; Kennard, O.; Versichel, W. *J. Am. Chem. Soc.* **1983**, *105*, 5761−5766.

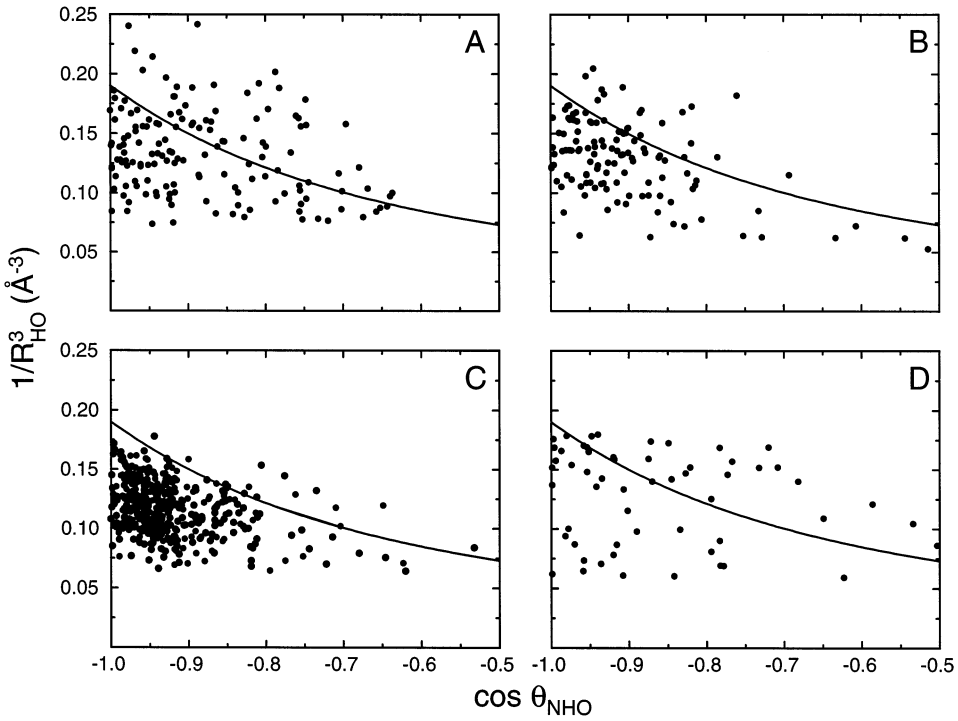

Figure 1. Correlation between R_{HO} and θ_{NHO}. Plots of $1/R^3_{HO}$ and cosine θ_{NHO} are shown for protein structures solved by X-ray crystallography and NMR: (A) thymidine kinase − 3.00 Å, (B) trypsin − 1.75 Å, and (C) all backbone hydrogen bonds in regions of secondary structure from nine high-resolution X-ray crystal structures (resolution ≤ 0.95 Å) (D). For comparison, a solution structure of cytochrome b_5 solved by NMR is also shown. In all of the plots, a function is drawn that defines "good" hydrogen-bond geometry. This function corresponds to a 5 kcal/mol·bohr force on the heavy atoms (see text and Table 1 for details).

backbone hydrogen-bond angle in proteins does deviate from linearity, 155° (α-helices) and 160° (β-strands),[9] an inquiry into the relationship between hydrogen-bond length and hydrogen-bond angle would provide a better understanding of hydrogen-bond geometry in proteins. Here, we examine the correlation between hydrogen-bond length, R_{HO}, and hydrogen-bond angle, θ_{NHO}, for backbone hydrogen bonds using newly available high-resolution protein crystal structures. We propose that this correlation can be used to assess the quality of a protein structure and, in the case of protein structures solved by NMR, used to refine the structure.

Methods

Protein Structure Data. X-ray crystal structures were retrieved from the Protein Data Bank by screening for protein structures with a reported resolution better than 1.00 Å.[13] Structures of oligonucleotides, hormones, or peptides were not used. The nine protein structures that were used for the geometric analysis are listed in Table 1. The crystal structures do not explicitly include hydrogen atoms. The program MOLMOL was used to place hydrogen atoms.[14] A comparison of hydrogen-bond distances in protein structures after using other programs (XPLOR[15] and REDUCE[16]) to place hydrogen atoms showed that average variations in R_{HO} and θ_{NHO} were no more than 0.02 Å and 3.5°, respectively.

Hydrogen-Bond Geometry Analysis and Protein Structure Refinement. The criteria used to select hydrogen bonds were the hydrogen-bond distance, $R_{HO} \leq 2.50$ Å, and the hydrogen-bond angle, $120° \leq \theta_{NHO} \leq 180°$. The HBDA (hydrogen-bond distance angle) module was incorporated into XPLOR 3.84 and was used to refine the structure of Bax which was determined using solution NMR.[17] For α-helix residues, input restraints were of the type i, $i-4$, where i refers to the residue of the NH donor, and $i-4$ refers to the one with the O acceptor atom. Input restraints contain the identities of hydrogen-bonded atoms. The empirical potential used was set such that only those hydrogen bonds whose calculated HBDA values are greater than zero (above the upper boundary limit in Figure 1) are penalized by energies proportional to the square of their deviation from zero. The HBDA force constants were increased from 2 to 500 kcal/Å^6 during the simulated annealing period of the refinement. This force constant is relatively weak as compared to those used for other terms in the refinement. The average final HBDA energy was 0.50 kcal mol^{-1}, which is at least 1 order of magnitude smaller than other energy terms such as the NOE energy term. Thirty protein structures were calculated, and the 10 lowest energy structures were used to evaluate the success of the HBDA refinement.

Table 1. High-Resolution Protein Crystal Structures

protein	PDB ID	resolution (Å)	ref
acetylxylan esterase	1G66	0.90	34
ribonuclease A,	1DY5	0.87	35
endoglucanase Cel5A catalytic core	7A3H	0.95	21
high-potential iron protein, H42Q	1B0Y	0.93	36
lysozyme	4LZT	0.95	37
lysozyme	1IEE	0.94	38
penicillopepsin + inhibitor	1BXO	0.95	39
parvalbumin	2PVB	0.91	40
rubredoxin	1BRF	0.95	41

(13) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. *Nucleic Acids Res.* **2000**, *28*, 235−242.
(14) Koradi, R.; Billeter, M.; Wuthrich, K. *J. Mol. Graphics* **1996**, *14*, 29−32.
(15) Brunger, A. T. *X-PLOR Version 3.1*; Yale University: New Haven, CT, 1992.
(16) Word, J. M.; Lovell, S. C.; Richardson, J. S.; Richardson, D. C. *J. Mol. Biol.* **1999**, *285*, 1733−1747.
(17) Suzuki, M.; Youle, R. J.; Tjandra, N. *Cell* **2000**, *103*, 645−654.

Ab initio Calculations. Ab initio calculations were performed on a system containing Ala28 of ubiquitin and acetamide (AcAm) to represent its hydrogen-bond partner using Gaussian 98 (revision A.6). The initial geometry of the system is taken out of the ubiquitin crystal structure.[18] The procedure for the ab initio calculation follows the previously reported protocol.[19] In short, the structure was highly optimized using 6-311**G basis set at the Hartree−Fock level of theory. The hydrogen-bond distance and angle of this optimized geometry were varied, without further optimization, to create the two-dimensional map of the energy of the system as a function of these two variables. Because of the large range of distances and angles included in the calculations, the structure was not optimized for each individual combination of distance and angle. However, a few distance/angle combinations were checked to assess the differences in angle and distance between optimized and nonoptimized structures. The difference between the nonoptimized and optimized structures was no greater than 0.07 Å (R_{HO}) and 6.0° (θ_{NHO}).

Initially, the stability of hydrogen-bond geometry was assessed by the overall energy of the system. This energy, however, is a measure of the global quantity of the system; thus it does not isolate instability of the system due to unfavorable hydrogen-bond geometry. We have chosen instead to calculate the forces on the specific atoms involved in the hydrogen bond. Any residual forces on these atoms are a direct measure of the acceptability of the hydrogen-bond geometry. All of the energy and force calculations were carried out using 6-311++G-(2d,2p) basis set at the DFT level of theory. Reported forces are relative to the values calculated for the starting optimized geometry.

Results

An analysis of protein structures from the Protein Data Bank[13] solved by both X-ray crystallography and NMR indicates that there is a fundamental geometric relationship between hydrogen-bond distance, R_{HO}, and hydrogen-bond angle, θ_{NHO}, for backbone hydrogen bonds, that is, those hydrogen bonds where the proton donor and carbonyl oxygen are both along the backbone. This relationship is most clearly seen in terms of a $1/R_{HO}^3$ and cosine θ_{NHO} correlation. Figure 1 shows this correlation for several protein structures. Each hydrogen bond is represented as a data point. The backbone hydrogen bonds from thymidine kinase, 3VTK,[20] 3.0 Å; endoglucanase, 7A3H, 0.95 Å;[21] and cytochrome b5, 1BFX,[22] are typical examples of hydrogen bonds from X-ray and NMR protein structures.

The data in Figure 1 demonstrate several general points concerning the geometry of hydrogen bonds. First, for X-ray crystal structures, the correlation between hydrogen-bond distance and angle depends strongly on the protein structure resolution. Backbone hydrogen bonds from X-ray crystal structures have a tendency to be confined to specific angles and distances, but a high-resolution structure is needed to see this clearly. The backbone hydrogen-bond geometry of most X-ray crystal structures is limited to an area that can be described by a polynomial function. Figure 1D shows the backbone hydrogen-bond geometry for a protein solved by solution NMR, cytochrome b5. Although the term "resolution" does not apply to NMR protein structures in the manner that it does to X-ray crystal structures, cytochrome b5 is a high-resolution structure by virtue of the quantity and type of data used; this includes a large number of distance restraints (NOEs) as well as geometric restraints on the amide bond vector in the form of pseudocontact shift data. Yet, despite the amount of data used in the structure determination, the data points in Figure 1D extend beyond the upper distance/angle threshold values seen for X-ray crystal structures. In particular, the range of hydrogen-bond angle is much broader in NMR protein structures than it is in X-ray crystal structures.

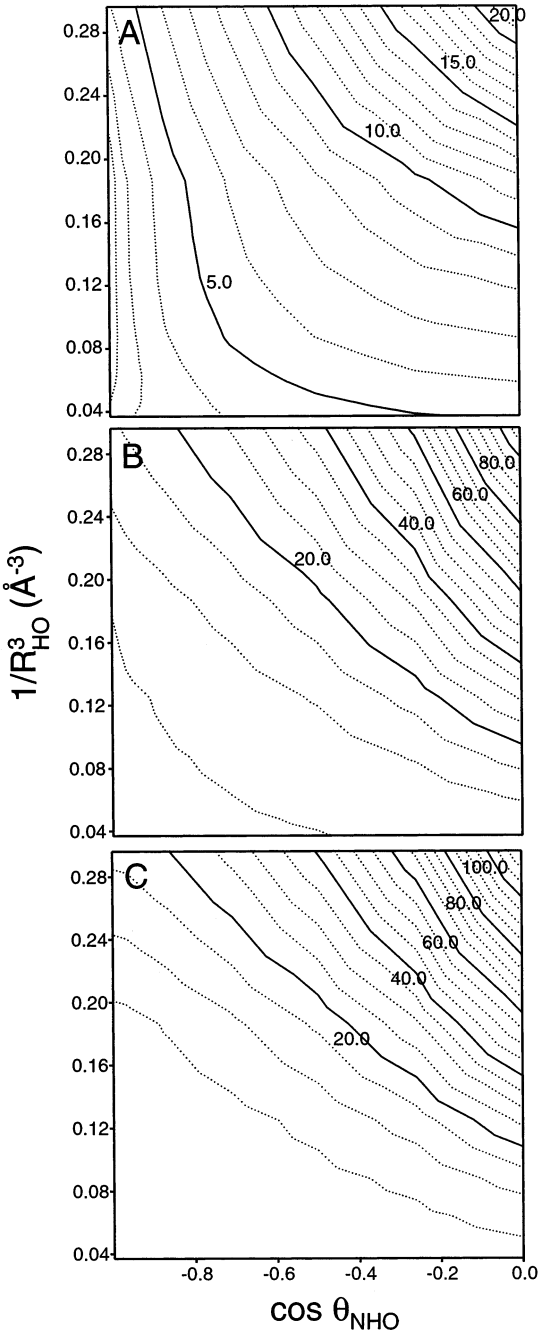

Figure 2. Ab initio calculations demonstrating the magnitude of the force (kcal/mol·bohr) on an AcAm hydrogen bond for the H^N (A), N (B), and O (C) hydrogen-bond atoms, respectively.

Figure 2 shows ab initio calculations on the hydrogen bonds in the model compound Ala28 of ubiquitin and acetamide (AcAm). It is apparent that the magnitude of the forces on the

(18) Vijay-Kumar, S.; Bugg, C. E.; Cook, W. J. *J. Mol. Biol.* **1987**, *194*, 531−544.
(19) Sharma, Y.; Kwon, O. Y.; Brooks, B.; Tjandra, N. *J. Am. Chem. Soc.* **2002**, *124*, 327−335.
(20) Wild, K.; Bohner, T.; Folkers, G.; Schulz, G. E. *Protein Sci.* **1997**, *6*, 2097−2106.
(21) Davies, G. J.; Mackenzie, L.; Varrot, A.; Dauter, M.; Brzozowski, A. M.; Schulein, M.; Withers, S. G. *Biochemistry* **1998**, *37*, 11707−11713.
(22) Arnesano, F.; Banci, L.; Bertini, I.; Felli, I. C.; Koulougliotis, D. *Eur. J. Biochem.* **1999**, *260*, 347−354.

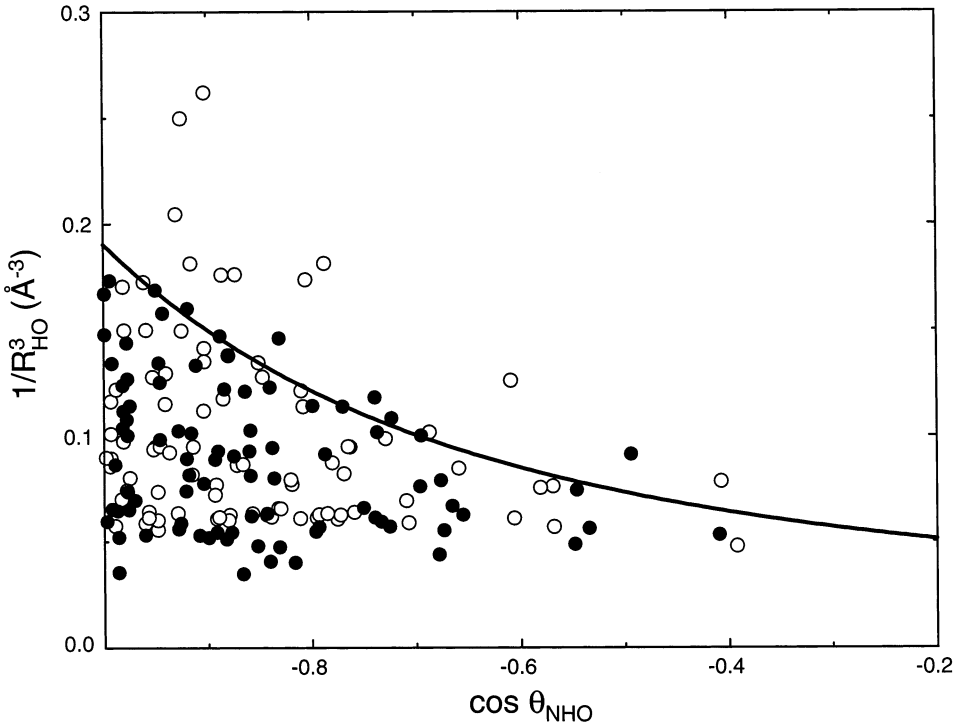

Figure 3. Refinement of Bax with respect to hydrogen-bond geometry. Sixty-six backbone hydrogen bonds were used. They are shown with respect to the hydrogen-bond geometry function before refinement (○) and after (●).

atoms H, N, and O forming the hydrogen bond follows the trend of hydrogen-bond distance dependence on the angle that is found in the empirical data, thus substantiating the observation. A similar behavior was also seen in the change of the total energy of the system (data not shown) as a function of hydrogen-bond distance and angle. The calculations show that unfavorable hydrogen-bond geometry will dominate the energy. The magnitudes of the forces on nitrogen and oxygen are larger than the proton. Furthermore, the magnitude of the force increases for short hydrogen-bond distance as the angle decreases or as a decreasing distance for a fixed angle.

To use backbone hydrogen-bond geometry as a tool to assess and/or improve the quality of a NMR solution structure, a function was derived using data from nine high-resolution X-ray crystal protein structures listed in Table 1. Although definitions vary, by convention hydrogen bonds in proteins are defined with an R_{HO} distance $\leq$ 2.40 Å and $\theta_{HNO} \leq 35°$,[23] or R_{HO} distance $\leq$ 2.40 Å and $\theta_{COH} \geq 90°$.[12] In our studies, an upper distance limit of 2.50 Å and an N−H···O angle $\geq 120°$ were used. This definition is somewhat stringent because definitions based on hydrogen-bond energy show that even at larger R_{HO} distances there is still a potentially significant hydrogen-bond energy.[8] The backbone hydrogen-bond data from the protein structures listed in Table 1 are shown in Figure 1C. An empirical function was derived which reflects the distance/angle boundary corresponding to the ab initio calculated force on the heavy atoms of 5 kcal/mol·bohr. This function is given as

$$\frac{1}{R^3} = A + \left[\frac{B}{(2.07 + \cos\theta_{NHO})^3}\right] \qquad (1)$$

where A and B are constants and equal to 0.019 and 0.21 Å^{-3}, respectively. Figure 1C shows that when this curve is applied to a set of hydrogen bonds from nine high-resolution protein crystal structures, the data points are distributed almost entirely below the curve. There are a few data points above the curve which stem from the probability distribution describing the likelihood of a hydrogen-bond heavy atom having a force greater than 5 kcal/mol·bohr. The quality of an NMR protein structure can be assessed by examining where the backbone hydrogen bonds occur with respect to this curve.

A "good" NMR structure which has been refined using a large number of NOE distance restraints and geometric residual dipolar coupling restraints may still contain many data points above this curve. Therefore, another application of the hydrogen-bond distance geometry correlation is to refine NMR structures using hydrogen-bonded atoms identified from hydrogen−deuterium exchange or $^{3h}J_{NC'}$ scalar coupling data as input restraints.[24] A module termed HBDA (hydrogen-bond distance angle) was developed for use with XPLOR 3.84[15] to carry out the refinement. The HBDA refinement seeks to optimize the hydrogen-bond geometry so that the data points are closer to the curve without disrupting the structure.

The HBDA refinement was demonstrated on the protein Bax, a pro-apoptotic member of the Bcl-2 family of proteins. Bax is a 192 residue protein whose secondary structure elements are all α-helices with the exception of an 18 residue unstructured loop. Figure 3 shows the hydrogen-bond distance/angle geometry correlation for Bax before and after applying the HBDA refinement. In the original structure determination of Bax, 87 hydrogen-bond restraints were identified. In Figure 3, it can be seen that 12 of these hydrogen bonds are above the curve. Bax has five i, $i-4$/i, $i-3$ backbone bifurcated hydrogen bonds (identified according to the criteria of Kabsch and Sander), of

(23) Berndt, K. D.; Guntert, P.; Wuthrich, K. *J. Mol. Biol.* **1993**, *234*, 735−750.

(24) Cordier, F.; Grzesiek, S. *J. Am. Chem. Soc.* **1999**, *121*, 1601−1602.

Table 2. Data for Bax Structure Calculations (10 Lowest Energy Structures Out of 30)

	without HBDA	with HBDA
energies (kcal mol^{-1})		
overall	298.5 ± 7.2	309.6 ± 6.4
NOE	7.2 ± 1.7	8.0 ± 3.1
dipolar	21.7 ± 3.2	21.8 ± 2.3
dihedral	0.32 ± 0.52	0.21 ± 0.30
HBDA	N/A	0. 28 ± 0.16
HBDA RMSD (Å^{-3})	0.048 ± 0.011	0.007 ± 0.002
pairwise RMSDs (residues 16−188) (Å)		
backbone	1.70 ± 0.30	1.51 ± 0.14
heavy atom	2.45 ± 0.23	2.24 ± 0.13
number of NOE violations	4.0 ± 2.1	3.6 ± 1.8
number of angle violations	1.7 ± 1.6	3.3 ± 1.6
Ramachandran plot analysis		
residues in core regions (%)	76.7 ± 2.1	77.2 ± 2.1
residues in additional allowed regions (%)	17.9 ± 2.1	17.4 ± 1.9
residues in generously allowed regions (%)	4.3 ± 1.8	3.8 ± 1.2
residues in disallowed regions (%)	1.1 ± 0.6	1.7 ± 0.9

which four are within helices rather than at the helix termini as is most common. The fifth bifurcated backbone hydrogen bond is located at a helix C-terminus and was therefore not used as a restraint. Of the four bifurcated hydrogen bonds that were used in the refinement, the *i*, *i* − 4 hydrogen bond is both shorter and more linear than the *i*, *i* − 3 hydrogen bond. All HBDA refinement restraints were of the *i*, *i* − 4 type; so, in the case of bifurcated hydrogen bonds, no *i*, *i* − 3 hydrogen-bond restraints were used. (In many cases, the lengths and angles of the *i*, *i* − 3 backbone hydrogen were such that they met the definition of a hydrogen bond even though traditionally a repetition of *i*, *i* − 4 hydrogen bonds is necessary to define an α-helix.)

In the HBDA refinement, only those hydrogen bonds that were used as restraints in the original structure determination were included in the refinement. Therefore, a number of hydrogen bonds with "bad" geometry, which typically are found outside the regions of well-defined secondary structure, were not refined. In addition, loop residues and helix termini were not included in the refinement. The 87 data points in Figure 3 correspond only to those hydrogen bonds within well-defined regions of secondary structure. The HBDA refinement was performed using a harmonic potential with a force constant that was increased slowly from the initial force constant of 2 kcal/Å^6 to a final force constant of 500 kcal/Å^6 during the simulated annealing period. Following HBDA refinement, the number of residues on average with hydrogen-bond distance/angle geometry above the cutoff curve decreases by 50% to six residues. The residues above the upper boundary limit lie just slightly above the curve. The improvement corresponds to a change of HBDA RMSD from 0.055 Å^{-3} before refinement to 0.007 ± 0.002 Å^{-3}. The pairwise backbone RMSDs for the 10 lowest energy structures calculated without and with HBDA refinement are 1.70 ± 0.30 and 1.51 ± 0.14 Å, respectively. The overall energies for the 10 lowest energy structures refined with or without HBDA are comparable. Table 2 shows the calculation statistics for the two sets of structures. The experimental energy terms such as NOE, dihedral, and residual dipolar couplings are essentially the same. In addition, the HBDA refinement also helps to improve the overall convergence of the lowest energy structures irrespective of the initial simulated annealing temperature, which is another indication that the hydrogen-bond data are consistent with the rest of the experimental data used in the structure calculation. The Ramachandran plots of the refined structures are better than the plot of the starting structure. Thus, the statistics on the refined structures demonstrate that the HBDA refinement improves the precision of the structures without distorting any structural elements.

Discussion

Many studies based on both empirical data and ab initio calculations have sought to characterize hydrogen bonds.[25−27] From these studies, it is clear that an accurate description of hydrogen bonds in proteins must simultaneously take into account a series of parameters including R_{HO} bond length, associated bond angles, electrostatic terms, and steric interactions. Indeed, part of the reason there is such a great range in the definition of hydrogen bond in the literature is due to the inclusion of only a subset of these terms. Most of the hydrogen-bond studies based on experimental data have been performed on either small molecules or proteins with medium resolution. These studies have resulted in several well-accepted generalizations concerning hydrogen-bond geometry. Perhaps one of the most well-established generalizations is that as the hydrogen-bond distance decreases, there becomes a more pronounced tendency for the hydrogen-bond geometry to adopt a linear conformation.[10] Here we have attempted to extend these studies by examining high-resolution protein structures with resolution below 1.00 Å and ab initio studies to illustrate interdependence between backbone hydrogen-bond distance and the hydrogen-bond angle, θ_{NHO}.

Figure 1 demonstrates a key distinction between backbone hydrogen bonds from X-ray crystal structures and from NMR structures. In X-ray crystal structures, θ_{NHO} increases as the hydrogen-bond distance decreases, whereas in NMR structures, there is no noticeable shift in hydrogen-bond distance in response to N−H···O angle increases. The discrepancy in hydrogen-bond geometry between X-ray and NMR structures arises in part due to the nature of the data. In X-ray crystallography, atoms are defined according to the electron density, and protons are not directly observed, whereas in NMR, protons are observed directly, and hydrogen bonds are determined in a semiquantitative manner from hydrogen−deuterium exchange and/or $^{3h}J_{NC'}$ scalar coupling data; typically, all hydrogen bonds are used in the structure calculations as distance restraints of the same magnitude. In addition, most of those restraints are implemented in the calculation so as to achieve idealized, linear hydrogen-bond geometry. In recent years, ^{1}H−^{15}N residual dipolar couplings have been used in conjunction with NOEs in NMR structure determination to impose geometric restraints on individual bond vectors relative to an alignment frame.[28,29] The structural information provided by hydrogen-bond geometry and residual dipolar couplings is not redundant and arises from

(25) Grabowski, S. J. *Chem. Phys. Lett.* **2001**, *338*, 361−366.
(26) Ippolito, J. A.; Alexander, R. S.; Christianso, D. W. *J. Mol. Biol.* **1990**, *215*, 457−471.
(27) McDonald, I. K.; Thornton, J. M. *J. Mol. Biol.* **1994**, *238*, 777−793.
(28) Tjandra, N.; Bax, A. *Science* **1997**, *278*, 1111−1114.
(29) Tolman, J. R.; Flanagan, J. M.; Kennedy, M. A.; Prestegard, J. H. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 9279−9283.

different underlying physical processes. Interestingly, protein structures determined with residual dipolar couplings generally have a better HBDA distribution than structures determined without residual dipolar couplings, which suggests that residual dipolar couplings have the benefit of improving the quality of the hydrogen-bond geometry.

The hydrogen bond is usually described as an electrostatic interaction due to Coulombic attractions between the proton donor and acceptor oxygen,[10] yet debate remains concerning whether the hydrogen-bond geometry is defined and or limited by electrostatic interactions, that is, the orientation that maximizes contact between the proton donor and the carbonyl oxygen sp^2 hybridized lone pair electrons or steric interactions, that is, the distance and orientation that limits the amount of overlap between the donor nitrogen and the acceptor oxygen atoms. A recent study by Buck and Karplus[30] aims to shed light on this controversy by using molecular dynamics to examine the acceptable range of R_{HO} and θ_{NHO} under the upper boundary limit described here with lysozyme as an example protein. They indicate that the R_{HO} minimum is principally governed by van der Waals interactions with electrostatic interactions having a significant influence on the hydrogen-bond orientation at longer distances. Incidently, the hydrogen bonds in that particular X-ray crystal structure of lysozyme[31] fall almost entirely under the boundary curve described here. It is not clear whether a molecular dynamics study carried out on a structure that does not have sufficiently good hydrogen-bond distribution will lead to the same outcome.

The results of ab initio calculations shown in Figure 2 would suggest that when the hydrogen-bond angle is linear there is no restriction on the hydrogen-bond length, but, at smaller hydrogen-bond angles, there is a sharp upper hydrogen-bond distance limit for a hydrogen bond to maintain optimal geometry. Our ab initio result also shows that forces on the donor nitrogen and acceptor oxygen atoms are at least 4 times larger than the protons. This is in keeping with the argument that the hydrogen-bond geometry is strongly governed by steric interactions at the boundary defined by eq 1. Within the region below the boundary curve, there is no preference for a hydrogen bond to adopt a particular geometry.[32] This is in agreement with our ab initio result. In contrast, Buck and Karplus found a preferred population within the same region.[30]

Because hydrogen bonding dictates protein stability and structure, it is appropriate to use hydrogen-bonding restraints to refine the protein structure. The only similar type of method to date that we are aware of uses a method where the hydrogen-bond energy is minimized to position polar hydrogen atoms in protein crystal structures.[33] This method seeks to optimize particular donor/acceptor configurations to reduce the ambiguity of polar side-chain conformations. The hydrogen-bond force field is derived from a statistical analysis of energetically favorable donor/acceptor pairs in the Cambridge Structural Database. In contrast to what has been shown here, the authors maintain that only at very large R_{HO} values is there any correlation between hydrogen-bond distance and angle and the hydrogen-bond force field they develop uses the hydrogen-bond distance and angle as separate, unrelated parameters.

The HBDA refinement of the protein Bax, the results of which are shown in Figure 3, confirms that the backbone hydrogen-bonding geometry provides a useful criteria to be used either as a diagnostic tool to identify irregularities in localized regions and/or for improving the quality of the structure. The HBDA refinement has general application to protein structures determined using solution NMR methods. It is also expected that similar distance/angle relationships govern the geometry of other types of hydrogen bonds such as those between the backbone amides and water molecules or side chains. Even though many of these types of hydrogen bonds are dynamic or exist transiently, examining them in terms of their distance/angle relationship may reveal individual hydrogen bonds with unique roles, that is, stabilization of active sites. Changes and/or aberrations in these hydrogen bonds may be associated with a protein's functional state. Furthermore, because the ab initio calculations were performed without assumptions on the system, the results are expected to be applicable to molecules other than proteins. In addition, the application of this type of refinement is not limited to NMR structures; it can also be applied to other structure determination procedures. The role of hydrogen bonding is important in all biological systems, and the methods presented here can easily be extended to study hydrogen-bond geometry in other macromolecules.

JA020676P

(30) Buck, M.; Karplus, M. *J. Phys. Chem. B* **2001**, *105*, 11000−11015.
(31) Artymiuk, P. J.; Blake, C. C. F. *J. Mol. Biol.* **1981**, *152*, 737−762.
(32) Adalsteinsson, H.; Maulitz, A. H.; Bruice, T. C. *J. Am. Chem. Soc.* **1996**, *118*, 7689−7693.
(33) Hooft, R. W. W.; Sander, C.; Vriend, G. *Proteins: Struct., Funct., Genet.* **1996**, *26*, 363−376.
(34) Ghosh, D.; Sawicki, M.; Lala, P.; Erman, M.; Pangborn, W.; Eyzaguirre, J.; Gutierrez, R.; Jornvall, H.; Thiel, D. J. *J. Biol. Chem.* **2001**, *276*, 11159−11166.
(35) Esposito, L.; Vitagliano, L.; Sica, F.; Sorrentino, G.; Zagari, A.; Mazzarella, L. *J. Mol. Biol.* **2000**, *297*, 713−732.
(36) Parisini, E.; Capozzi, F.; Lubini, P.; Lamzin, V.; Luchinat, C.; Sheldrick, G. M. *Acta Crystallogr., Sect. D* **1999**, *55*, 1773−1784.
(37) Walsh, M. A.; Schneider, T. R.; Sieker, L. C.; Dauter, Z.; Lamzin, V. S.; Wilson, K. S. *Acta Crystallogr., Sect. D* **1998**, *54*, 522−546.
(38) Sauter, C.; Otalora, F.; Gavira, J. A.; Vidal, O.; Giege, R.; Garcia-Ruiz, J. M. *Acta Crystallogr., Sect. D* **2001**, *57*, 1119−1126.
(39) Khan, A. R.; Parrish, J. C.; Fraser, M. E.; Smith, W. W.; Bartlett, P. A.; James, M. N. *Biochemistry* **1998**, *37*, 16839−16845.
(40) Declercq, J. P.; Evrard, C.; Lamzin, V.; Parello, J. *Protein Sci.* **1999**, *10*, 2194−2204.
(41) Bau, R.; Rees, D. C.; Kurtz, D. M.; Scott, R. A.; Huang, H. S.; Adams, M. W. W.; Eidsness, M. K. *J. Biol. Inorg. Chem.* **1998**, *3*, 484−493.