ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/6918862

Cooperative Hydrogen Bonding Effects Are Key Determinants of Backbone Amide Proton Chemical Shifts in Proteins

ARTICLE *in* JOURNAL OF THE AMERICAN CHEMICAL SOCIETY · SEPTEMBER 2006

Impact Factor: 12.11 · DOI: 10.1021/ja0617901 · Source: PubMed

CITATIONS
51

READS
27

3 AUTHORS, INCLUDING:

Laura L Parker
University of Massachusetts Amherst
4 PUBLICATIONS **88** CITATIONS
SEE PROFILE

Jan Halborg Jensen
University of Copenhagen
116 PUBLICATIONS **6,472** CITATIONS
SEE PROFILE

Available from: Laura L Parker
Retrieved on: 16 January 2016

J|A|C|S
ARTICLES
Published on Web 07/08/2006

Cooperative Hydrogen Bonding Effects Are Key Determinants of Backbone Amide Proton Chemical Shifts in Proteins

Laura L. Parker, Andrew R. Houk,† and Jan H. Jensen*,‡

Contribution from the Department of Chemistry, University of Iowa, Iowa City, Iowa 52242

Received March 15, 2006; E-mail: Jan-Jensen@uiowa.edu

Abstract: A computational methodology for backbone amide proton chemical shift (δ_H) predictions based on ab initio quantum mechanical treatment of part of the protein is presented. The method is used to predict and interpret 13 δ_H values in protein G and ubiquitin. The predicted amide−amide δ_H values are within 0.6 ppm of experiment, with a root-mean-square deviation (RMSD) of 0.3 ppm. We show that while the hydrogen bond geometry is the most important δ_H-determinant, longer-range cooperative effects of extended hydrogen networks make significant contributions to δ_H. We present a simple model that accurately relates the protein structure to δ_H.

1. Introduction

The relationships between protein structure and NMR chemical shift values have been studied extensively for decades, and form the bases for software that can predict the NMR chemical shifs given the protein structure.[1−17] The NMR chemical shift predictors can significantly aid protein structure validation and refinement and, when combined with protein structure predictors, hold the promise for fully automated high-throughput determination of accurate protein structures.[18]

The chemical shifts of most main chain atoms can generally be predicted with correlation coefficients of 0.90 or better.[10,14,15,17,19] However, despite decades of work, the chemical shifts of backbone amide protons (δ_H) can only be predicted with a correlation coefficient of about 0.75.[1,3,9] This is particularly unfortunate since δ_H values are the most sensitive to the geometries[19] (and presumably the strengths) of the ubiquitous backbone−backbone and backbone−side chain hydrogen bonds critical to protein structure and stability. Thus, the relatively low accuracy of predicted δ_H values limits the accuracy with which these hydrogen bond geometries can be validated and refined.

The change in δ_H due to hydrogen bond formation is usually assumed to be due to a combination of magnetic anisotropy and electrostatic effects.[3,19,20] Both contributions scale as R^{-3} where R is the hydrogen bonding distance, and Wagner, Pardi and Wüthrich[21] demonstrated such a correlation for amide−amide hydrogen bonds in bovine pancreatic trypsin inhibitor in 1983. Though no correlation coefficient was given in this work, the plot of proton chemical shift vs R showed significant scatter, and subsequent work by Wishart, Sykes and Richards[3] indicated that a R^{-1} functional form correlates equally well with experimental data. However, Sitkoff and Case[9] clearly demonstrated a R^{-3} dependence, with a correlation coefficient of $r = 0.96$, for computed δ_H values for small structural models of amide hydrogen bonds. Two of the most popular chemical shift predictors, SHIFTS[14] and SHIFTX,[16] use the R^{-3} functional from for protein chemical shift predictions with parameters adjusted to reproduce experimental chemical shifts for proteins.

Recently, chemical shift values of protons in NH···O═C hydrogen bonds have been shown to correlate well ($r = 0.88$) with corresponding NC coupling constants ($^{3h}J_{NC'}$) measured for ubiquitin.[22] Coupling constants depend on molecular orbital overlap, which has an exponential distance dependence. Thus, Barfield[23] proposed such a functional form for δ_H values and obtained an excellent correlation ($r^2 = 0.98$, including angular terms) with ab initio data computed using model systems. However, the ab initio predictions did not correlate well with experiment ($r^2 = 0.51$). Poulsen and co-workers[24] obtained

† Present address: Department of Biochemistry, UCSF, San Francisco, CA 94143.

‡ Present address: Department of Chemistry, University of Copenhagen, Universitetsparken 5, 2100 Copenhagen, Denmark.

(1) Osapay, K.; Case, D. A. *J. Am. Chem. Soc.* **1991**, *113*, 9436.
(2) Spera, S.; Bax, A. *J. Am. Chem. Soc.* **1991**, *113*, 5490.
(3) Wishart, D. S.; Sykes, B. D.; Richards, F. M. *J. Mol. Biol.* **1991**, *222*, 311.
(4) Dedios, A. C.; Pearson, J. G.; Oldfield, E. *Science* **1993**, *260*, 1491.
(5) Williamson, M. P.; Asakura, T. *J. Magn. Reson. Series B* **1993**, *101*, 63.
(6) Szilagyi, L. *Prog. Nucl. Magn. Reson. Spectrosc.* **1995**, *27*, 325.
(7) deDios, A. C. *Prog. Nucl. Magn. Reson. Spectrosc.* **1996**, *29*, 229.
(8) Gronwald, W.; Boyko, R. F.; Sonnichsen, F. D.; Wishart, D. S.; Sykes, B. D. *J. Biomol. NMR* **1997**, *10*, 165.
(9) Sitkoff, D.; Case, D. A. *J. Am. Chem. Soc.* **1997**, *119*, 12262.
(10) Wishart, D. S.; Watson, M. S.; Boyko, R. F.; Sykes, B. D. *J. Biomol. NMR* **1997**, *10*, 329.
(11) Ando, I.; Kameda, T.; Asakawa, N.; Kuroki, S.; Kurosu, H. *J. Mol. Struct.* **1998**, *441*, 213.
(12) Cornilescu, G.; Delaglio, F.; Bax, A. *J. Biomol. NMR* **1999**, *13*, 289.
(13) Iwadate, M.; Asakura, T.; Williamson, M. P. *J. Biomol. NMR* **1999**, *13*, 199.
(14) Xu, X. P.; Case, D. A. *J. Biomol. NMR* **2001**, *21*, 321.
(15) Meiler, J. *J. Biomol. NMR* **2003**, *26*, 25.
(16) Neal, S.; Nip, A. M.; Zhang, H. Y.; Wishart, D. S. *J. Biomol. NMR* **2003**, *26*, 215.
(17) Wang, Y. J. *J. Biomol. NMR* **2004**, *30*, 233.
(18) Meiler, J.; Baker, D. *Proc. Nat'l. Acad. Sci., U.S.A.* **2003**, *100*, 15404.
(19) Wishart, D. S.; Case, D. A. Use of chemical shifts in macromolecular structure determination. In *Methods Enzymol.* **2001**, *338*, 1.
(20) Sitkoff, D.; Case, D. A. *Prog. Nucl. Magn. Reson. Spectrosc.* **1998**, *32*, 165.
(21) Wagner, G.; Pardi, A.; Wüthrich, K. *J. Am. Chem. Soc.* **1983**, *105*, 5948.
(22) Cordier, F.; Grzesiek, S. *J. Am. Chem. Soc.* **1999**, *121*, 1601.
(23) Barfield, M. *J. Am. Chem. Soc.* **2002**, *124*, 4158.
(24) Jensen, P. R.; Axelsen, J. B.; Lerche, M. H.; Poulsen, F. M. *J. Biomol. NMR* **2004**, *28*, 31.

significantly better correlation ($r = 0.84$) using a simpler exponential functional fitted to δ_H values (corrected for ring current effects) measured for protein G. However, significantly worse correlation was observed for other proteins for which high-resolution structures are not available.

In summary, the δ_H dependence on hydrogen bond geometry obtained by ab initio calculations on small model systems can be represented very well (r or $r^2 = 0.96-0.98$) by simple functional forms. However, when used to predict experimental δ_H values obtained for proteins the agreement is generally worse ($r \leq 0.75$). Possible causes may be that the protein structures used for the predictions are not sufficiently accurate or that there are important contributions to δ_H values that are currently not known.

We address both issues in the current study by constructing the simplest possible structural models of amide hydrogen bonds that lead to computed chemical shift values that consistently reproduce experimental values obtained for protein G and ubiqutin. The paper is organized as follows: First we discuss the computational methodology. Second we demonstrate the accuracy of our predictions for amide−amide hydrogen bonds, followed by an analysis of the structural determinants of the chemical shifts. Third, based on this analysis we propose a simple empirical formula relating the protein structure and the chemical shifts that accurately reproduces the experimental data using optimized geometries. Fourth, we compare our empirical approach to similar approaches proposed previously. Fifth, we compare our findings to cases where hydrogen bonding is to water molecules or amino acid side chains. Finally, we summarize our results and discuss future directions.

2. Computational Methodology

2.1. Ab Initio Methods. The accurate prediction of an experimentally measured NMR chemical shift presents a challenge to theory, due to the very high level of theory necessary for converged results and the effect of the molecular environment (e.g., solvent or protein). Chesnut[25] has proposed a scaling technique to address these effects and Rablen, Pearlman, and Finkbiner[26] have obtained the necessary parameters for proton chemical shifts relative to TMS in nonpolar solvents ($CDCl_3$ and CCl_4):

$$\delta_H(\mathrm{TMS}_{\mathrm{CDCl_3}}) = \sigma_{\mathrm{ref}} - 0.957\sigma_H \quad (1)$$

Here, σ_H is the isotropic chemical shielding calculated at the B3LYP/6-311++G(d,p)//B3LYP/6-31G+(d) level of theory, which in this study is approximated by B3LYP/6-311++G(d,p)// B3LYP/6-31G(d) and σ_{ref} is 30.60 ppm. Including the aqueous phase value of the chemical shielding of the proton[27] would increase σ_{ref} to 30.93 ppm. This correction is almost entirely due to solvent effects, since DSS and TMS have very similar chemical shifts in the same solvent.[28] This approach was used previously by Molina and Jensen[29] to successfully predict proton chemical shifts in the active sites of chymotrypsin and α-lytic protease. However, we found that a σ_{ref} of 30.30 ppm leads to a much better agreement with experiment in our case. The chemical shielding calculations are performed with the PQS program[30] on an 8-node Quantum Station, while the constrained geometry optimizations were performed with the GAMESS[31] program.

2.2. Structural Models. The structural models used in the NMR calculations are derived from the 1.10 Å X-ray structure of immunoglobulin binding domain of streptococcal protein G (1IGD;[32] hereafter referred to as "protein G") and the 1.32 Å X-ray structure of human ubiquitin (1OGW[33]) and protonated with PDB2PQR[34] or the WHATIF web interface. The model for residue Gln 40 was taken from an unpublished X-ray structure of K29Q ubiquitin, obtained by Andrew Robertson, S. Ramaswamy, and co-workers, for reasons explained below (the structure will be deposited in the PDB in the near future). The models are displayed in Figure 1 and generally contain (1) the amide group of interest flanked by the two neighboring amide groups, (2) any groups hydrogen bonded to the proton (the primary hydrogen bond) and oxygen (secondary hydrogen bond) of the amide group of interest and the amide group nearest to the primary hydrogen bond (the nonbonding amide), (3) any groups hydrogen bonded to the hydrogen acceptor in the primary hydrogen bond. Most models contain the two side chains on either side of the amide group of interest and the side chain of the hydrogen acceptor in the primary hydrogen bond. We find that these structural models are the simplest models needed to reproduce the observed proton chemical shifts.

For protein G residues only the positions of the OCNH···OCNH atoms (shown in bold in Figure 2) in the primary hydrogen bond are energy minimized while the remaining atoms are kept at their experimentally determined values. For ubiquitin, the positions of the (N)C_αH atoms are also energy minimized. The optimization of these extra atoms was found to have no effect on the calculated chemical shift. The side chains not involved in the hydrogen bonding network are represented as methyl groups, except for glycine residues, during the geometry optimization, but the side chains are added before the NMR chemical shielding calculations. The geometry optimizations are done in the gas phase, with the exception of Thr49, which was optimized with the conductor like polarizable continuum model[35] (C−PCM) as implemented in the GAMESS program, using the GEPOL-AS tessellation scheme,[36] 60 initial tesserae per atom and the united atom Hartree−Fock radii.[37]

The NMR chemical shift calculations are done both in the gas phase and in bulk solution as represented by the COSMO conductor like screening model[38] of solvation as implemented in the PQS program.

Smaller structural models such as those shown in Figures 4 and 6−8 are created to study how the various structural elements affect the amide protein chemical shifts, as described further below. Unless otherwise noted, the relevant atomic positions are reoptimized before NMR chemical shifts are computed.

3. Results and Discussion

3.1. Overall Accuracy and General Observations. The backbone amide proton chemical shifts (δ_H) computed using the computational methodology described above are listed in

(25) Chesnut, D. B. The Ab Initio Computation of Nuclear Magn. Reson. Chem.ical Shielding. In *Reviews in Computational Chemistry*; Boyd, D. B.; Ed.; 1996; Vol. 8; p 245.
(26) Rablen, P. R.; Pearlman, S. A.; Finkbiner, J. *J. Phys. Chem. A* **1999**, *103*, 7357.
(27) Porubcan, M. A.; Neves, D. E.; Rausch, S. K.; Markley, J. L. *Biochemistry* **1978**, *17*, 4640.
(28) Harris, R. K.; Becker, E. D.; Cabral de Menezes, S. M.; Goodfellow, R.; Granger, P. *Pure Appl. Chem.* **2002**, *73*, 1795.
(29) Molina, P. A.; Jensen, J. H. *J. Phys. Chem. B* **2003**, *107*, 6226.
(30) PQS version 2.4, P. Q. S.; Fayetteville, Arkansas, http://www.pqs-chem.com; sales@pqs-chem.com.
(31) Schmidt, M. W.; Baldridge, K. K.; Boatz, J. A.; Elbert, S. T.; Gordon, M. S.; Jensen, J. H.; Koseki, S.; Matsunaga, N.; Nguyen, K. A.; Su, S. J.; Windus, T. L.; Dupuis, M.; Montgomert, J. A. *J. Comput. Chem.* **1993**, *14*, 1347.
(32) Derrick, J. P.; Wigley, D. B. *J. Mol. Biol.* **1994**, *243*, 906.
(33) Alexeev, D.; Barlow, P. N.; Bury, S. M.; Charrier, J. D.; Cooper, A.; Hadfield, D.; Jamieson, C.; Kelly, S. M.; Layfield, R.; Mayer, R. J.; McSparron, H.; Price, N. C.; Ramage, R.; Sawyer, L.; Starkmann, B. A.; Uhrin, D.; Wilken, J.; Young, D. W. *Chembiochem* **2003**, *4*, 894.
(34) Dolinsky, T. J.; Nielsen, J. E.; McCammon, J. A.; Baker, N. A. *Nucleic Acids Res.* **2004**, *32*, W665.
(35) Barone, V.; Cossi, M. *J. Phys. Chem. A* **1998**, *102*, 1995.
(36) Li, H.; Jensen, J. H. *J. Comput. Chem.* **2004**, *25*, 1449.
(37) Barone, V.; Cossi, M.; Tomasi, J. *J. Chem. Phys.* **1997**, *107*, 3210.
(38) Klamt, A.; Schuurmann, G. *J. Chem. Soc.-Perkin Trans. 2* **1993**, 799.

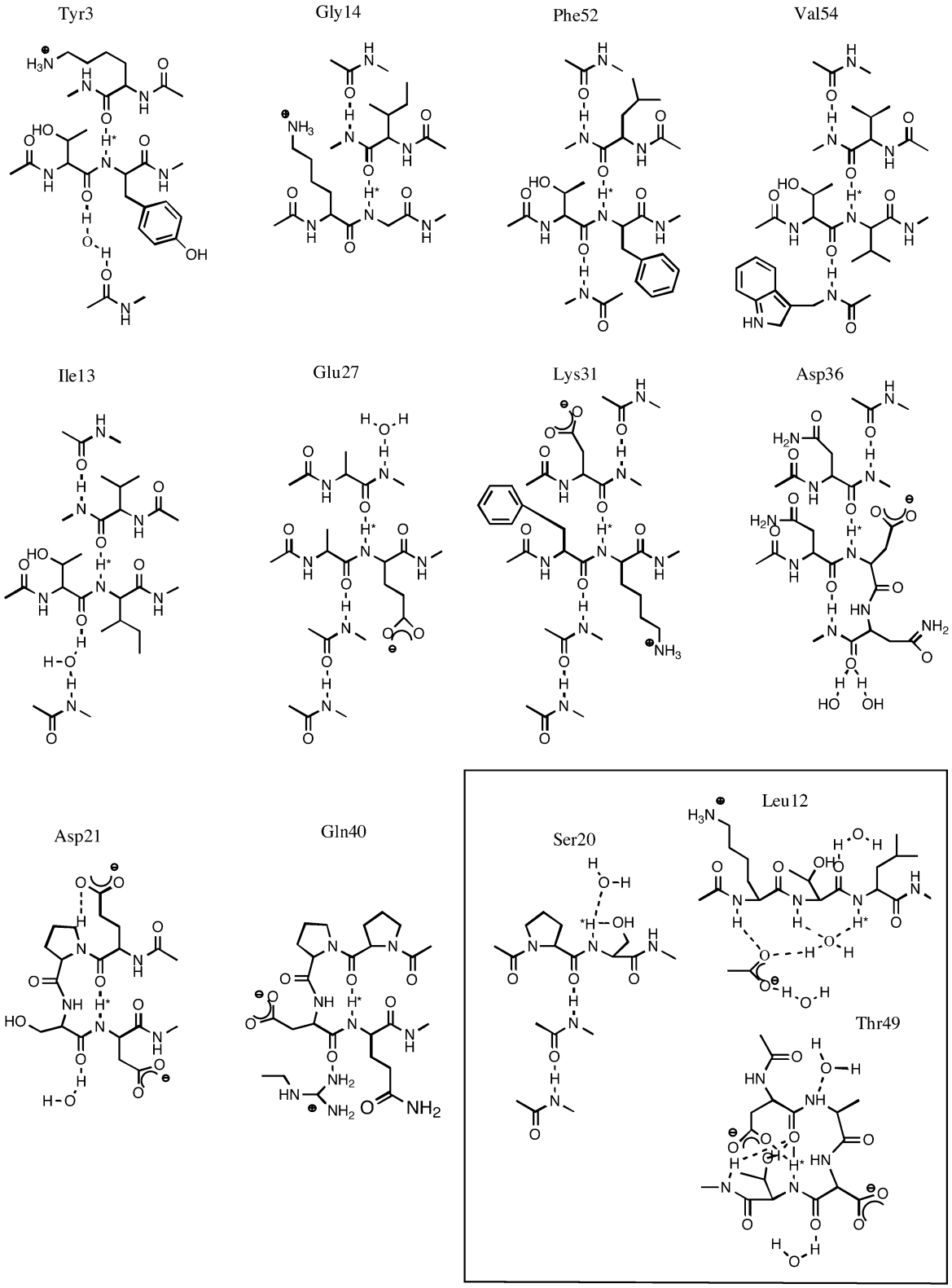


Figure 1. Sketch of the largest structural models (labeled "side-chain" in Table 1) used to predict amide proton chemical shift values.

Table 1 under the heading "solvation" together with the experimental values.[39,40] The experimental values are reproduced with a correlation coefficient of 0.95 and a root-mean-square deviation (RMSD) of 0.3 ppm, with the largest error (−0.6 ppm) observed for Gln40 in ubiquitin. [For Thr49 and Leu12 in protein G and Ser20 in ubiquitin the proton of interest is not hydrogen bonded to another amide. Data regarding these three residues are discussed separately in section 3.11 and not included when computing RMSD values, etc.] The effects of bulk solvation on the computed shifts are generally 0.1−0.2 ppm in the β-sheets and 0.2−0.5 ppm in the α-helixes and elsewhere. The largest effect was seen in Asp36 in protein G where

(39) Orban, J.; Alexander, P.; Bryan, P. *Biochemistry* **1992**, *31*, 3604.
(40) Weber, P. L.; Brown, S. C.; Mueller, L. *Biochemistry* **1987**, *26*, 7282.

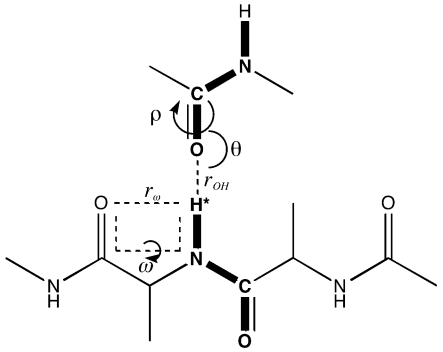


Figure 2. Sketch showing the structural parameters used in eqs 4 and 5 and the atoms whose position are energy minimized. For the residues in ubiquitin the position of the carbon atoms bonded to the amide nitrogen are also energy minimized.

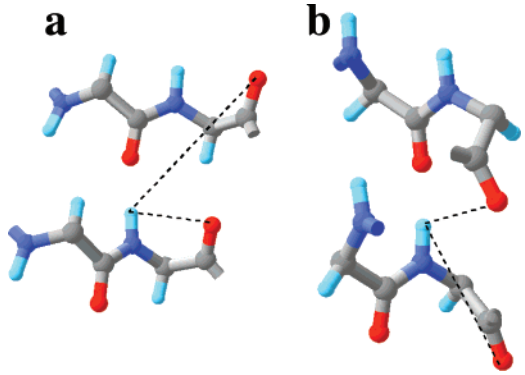


Figure 3. Local structures around the amide proton of interest in (a) a β-sheet (Phe54) and (b) an α-helix (Lys31). In β-sheets the "phi−psi" and "nonbonded" carbonyl group have large and small effect on δ_H, respectively; while the reverse is generally true in α-helices. See text for further discussion.

solvation increases the chemical shift by 0.5 ppm. We discuss the possible reasons for this in more detail below, but overall solvation decreases the RMSD relative to experiment by only 0.01 ppm. We focus on the gas-phase values throughout (unless otherwise noted) since they are easier to interpret as discussed next. The general approach is to start with the δ_H value of a di-alanine model of the amide group of interest, and then add successive structural elements until the models shown in Figure 1 are obtained. The process is exemplified for Lys31 in Figure 4 and the resulting chemical shifts are listed in Table 1.

3.2. Effect of Neighboring ("Phi-Psi") Interactions. The chemical shifts computed using the di-alanine (or "phi-psi") models of each amide group are listed in the second column of Table 1. The chemical shift values have a 1.6 ppm range (4.1−5.7 ppm), presumably due to the differences in backbone conformation. A thorough analysis of the data, as well as new data generated by constructing phi-psi models with other dihedral angles, revealed that the most important structural determinant of δ_H in these models is the proximity of the next amide group. We investigated several ways of quantifying this relationship and found the following functional form to work reasonably well.

$$f(\omega,r_\omega) = \frac{\cos^2(\omega)}{r_\omega^{\,3}} \qquad (2)$$

Here, ω and r_ω are the H−N−C═O dihedral angle and H−O

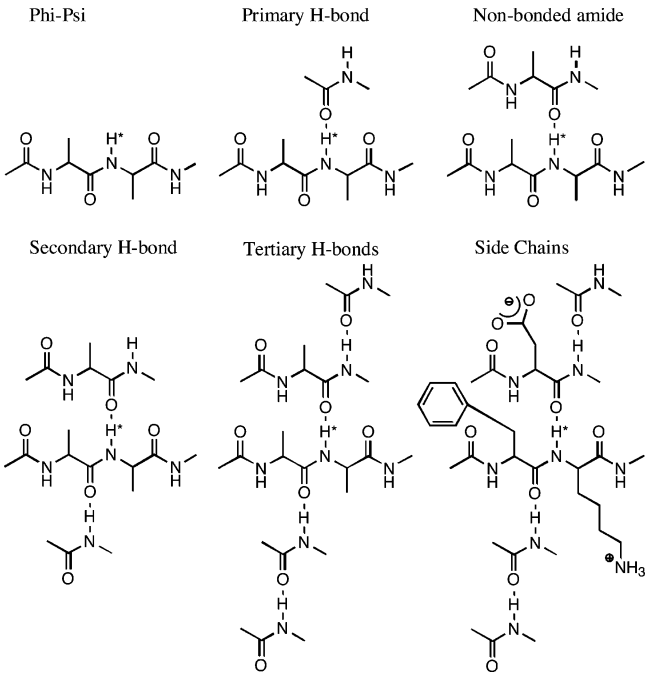


Figure 4. Sketch of all structural models used in the analysis of the structural determinants of the δ_H value of Lys31. The models used for the other residues can be found by comparison to Figure 1, with the exception of Leu12, Ser20, and Thr49 (see Figures 6−8).

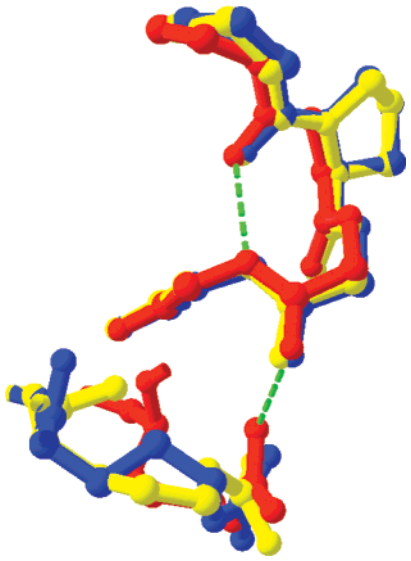


Figure 5. Overlay of local structures near Gln 40 in three X-ray structures of ubiquitin (red = unpublished structure, blue = 1OGW, and yellow = 1UBQ).

distance shown Figure 2. A plot of the chemical shifts computed using the phi-psi models (δ_ω) vs $f(\omega,r_\omega)$, is shown in Figure 9 and can be quantified as follows:

$$\delta_\omega = \begin{cases} 4.5 \text{ ppm if } f(\omega,r_\omega) \le 0.03 \\ 5.5 \text{ ppm if } 0.03 < f(\omega,r_\omega) \le 0.10 \\ 6.5 \text{ ppm if } 0.10 < f(\omega,r_\omega) \end{cases} \qquad (3)$$

This approach results in chemical shifts that have an RMSD of 0.17 ppm compared to the phi-psi values in Table 1, with the largest error (0.24 ppm) observed for Leu12 in protein G. Other denominators such as r^2 and r^4 were also tried but neither showed any advantage over r^3, so r^3 was used in analogy with

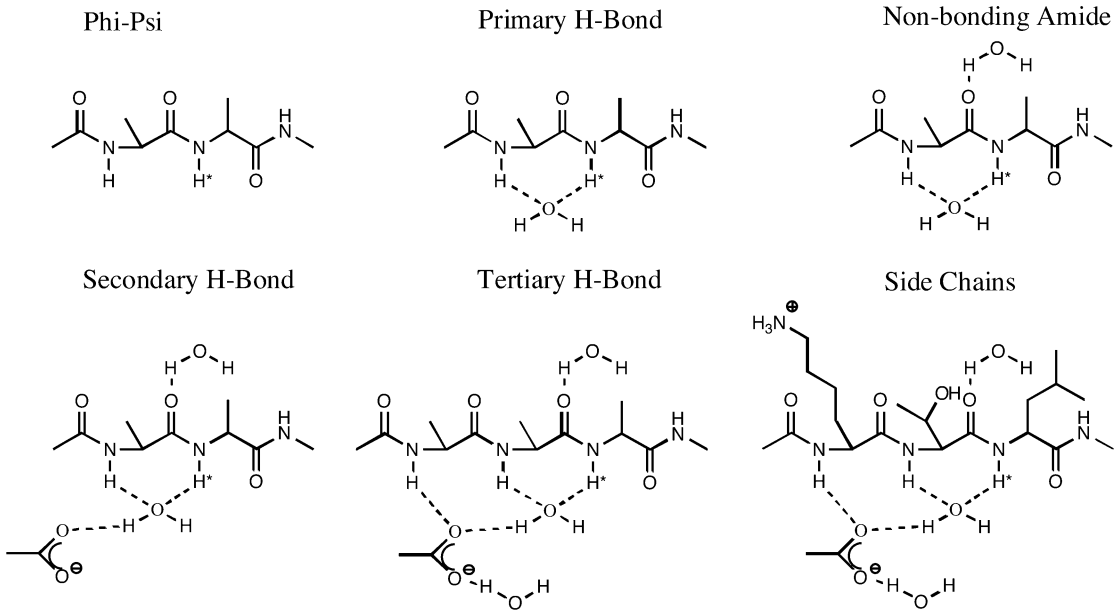


Figure 6. Sketch of all structural models used in the analysis of the structural determinants of the δ_H value of Leu12.

Table 1. Computed Chemical Shifts and Experimental Chemical Shifts Using Various Structural Models[a]

	phi–psi	primary HB	NB amide	second. HB	tertiary HB	side-chains	solvation	experiment
				β-sheet				
				Protein G				
Tyr3	5.2	9.2 (4.0)	8.9 (−0.3)	9.0 (0.1) w	9.2 (0.2)	9.3 (0.1)	9.5 (0.2)	9.1 (0.4)
Gly14	5.5	8.3 (2.8)	8.3 (0.0)		8.6 (0.3)	8.7 (0.1)	8.6 (−0.1)	8.3 (0.3)
Phe52	5.2	8.9 (3.7)	9.2 (0.3)	9.8 (0.6)	10.2 (0.4)	10.3 (0.1)	10.4 (0.1)	10.4 (0.0)
Val54	5.4	8.4 (3.0)	8.2 (−0.2)	8.6 (0.4)	9.1 (0.5)	8.5 (−0.6)[b]	8.5 (0.0)	8.3 (0.2)
				Ubiquitin				
Ile13	5.4	8.7 (3.3)	8.6 (−0.2)	8.9 (0.3) w	9.5 (0.6)	9.5 (0.0)	9.4 (−0.1)	9.4 (0.0)
				α-helix				
				Protein G				
Glu27	4.7	7.4 (2.7)	6.8 (−0.6)	7.9 (1.1)	8.2 (0.3)	8.3 (0.1)	8.7 (0.4)	8.3 (0.4)
Lys31	4.5	8.0 (3.5)	6.5 (−1.5)	7.7 (1.2)	8.6 (0.9)	8.9 (0.1)	9.2 (0.3)	9.0 (0.2)
Asp36	4.5	7.1 (2.6)	8.1 (1.0)	7.2 (−0.9)	8.3 (1.1)	8.6 (0.3)	9.1 (0.5)	8.7 (0.4)
				Other				
				Ubiquitin				
Asp21	5.2	6.8 (1.6)	7.1 (0.3)	7.4 (0.3) w	8.1 (0.7)	8.2 (0.1)	8.0 (−0.2)	8.0 (0.0)
Gln40	4.4	6.2 (1.8)	6.2 (0.0)	7.5 (1.3)		7.0 (−0.3)	7.2 (0.2)	7.8 (−0.6)
				Irregular				
				Protein G				
Leu12	5.7	5.7 (0.0) w		5.9 (0.2) w	6.6 (0.7)	6.8 (0.2)	7.1 (0.3)	7.6 (0.5)
Thr49[c]	4.1	5.0 (0.9)	5.7 (0.7)	6.4 (0.7) w	6.4 (0.0) w	6.4 (0.0)		7.0 (−0.6)
				Ubiquitin				
Ser20	4.1	6.3 (2.1) w, sc		6.9 (0.8)	6.8 (−0.1)		7.1 (0.3)	7.0 (0.1)
RMSD						0.28	0.28	

[a] The "side chain" models are shown in Figure 1. Examples of the remaining models are shown in Figure 4 for cases where the primary hydrogen bond it to an amide group and in Figures 6−8 for the other three cases. The column marked "solvation" indicates chemical shift calculations that include the effect of bulk solvation based on the "side-chain" structural models. Numbers in parentheses are changes in chemical shift compared to the previous column. w and sc indicates hydrogen bonding to water and side chains, respectively. [b] Trp43 side chain included. [c] Solvent included for all steps.

previous studies.[9] We also investigated the use of linear fits to the data in Figure 9, but this generally led to an increase in the RMSD compared to eq 3.

The relatively low chemical shift values observed for the amide groups in α-helices (4.5−4.7 ppm) compared to β-sheets (5.2−5.7 ppm; Table 1) are thus primarily due to comparatively larger r_ω values in the α-helix geometry (4.3−4.4 vs 2.1−2.7 Å; see also Figure 3). Additional calculations (data not shown) established that there are relatively small changes in the amide geometry due to changes in ω, which have a nonnegligible effect on the predicted δ_H values. Thus, the neighboring amide group affects δ_H in two ways: (1) by a polarization of the electron density and (2) by the resulting change in the internal geometry of the amide group (presumably mostly a change in the NH distance).

3.3. Effect of the Primary Hydrogen Bond. Addition of the primary hydrogen bonding partner (Figure 4) has a significant effect ($\Delta\delta_H = 1.6-4.0$ ppm) on, *and is usually the*

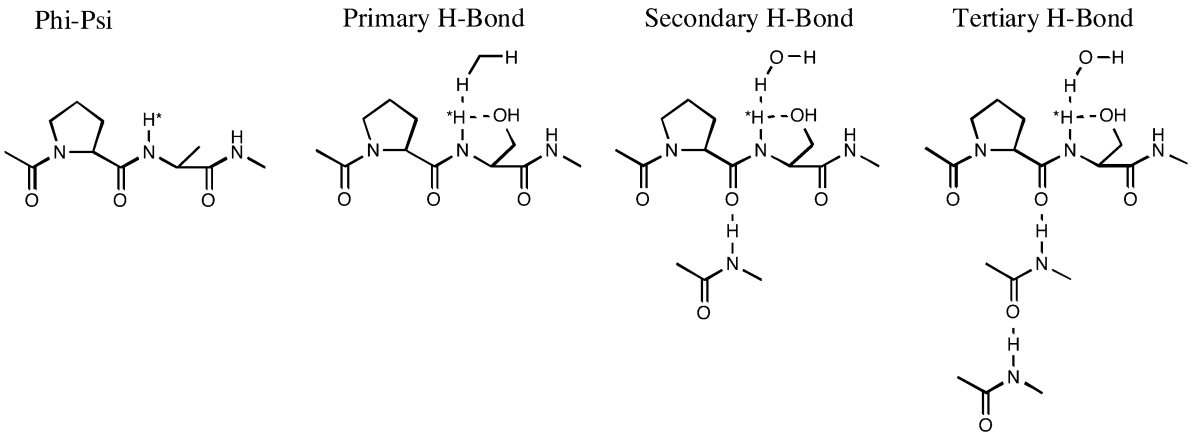


Figure 7. Sketch of all structural models used in the analysis of the structural determinants of the δ_H value of Ser20.

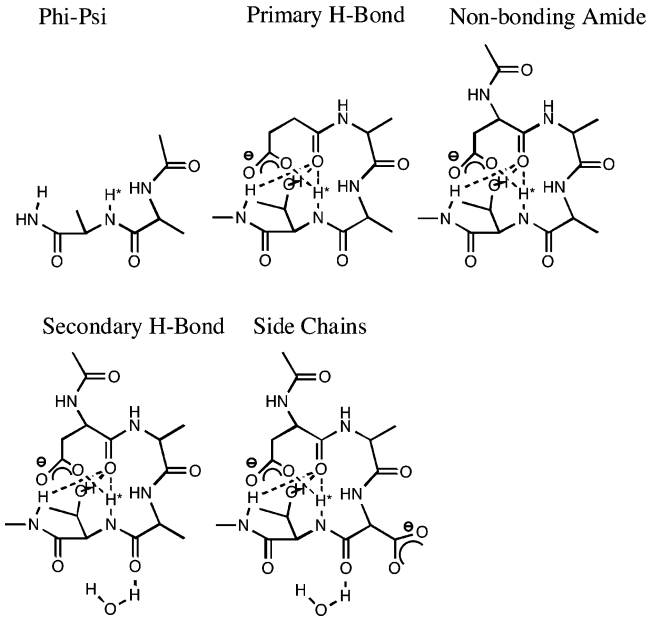


Figure 8. Sketch of all structural models used in the analysis of the structural determinants of the δ_H value of Thr49.

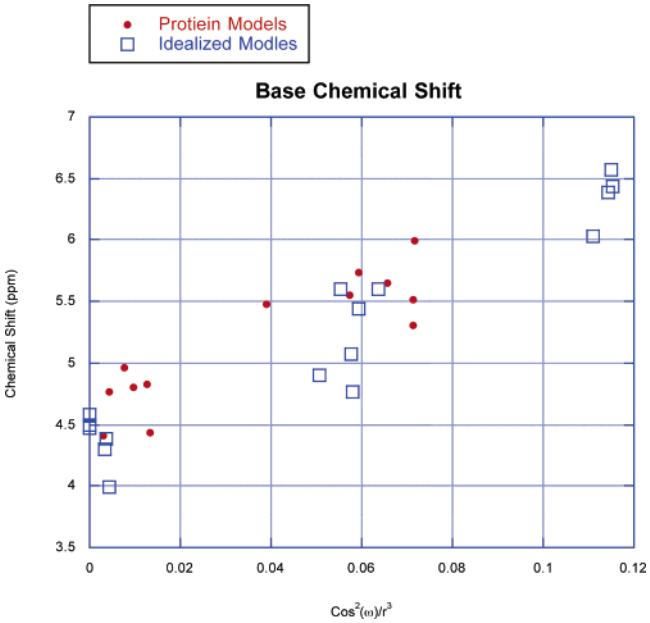


Figure 9. Chemical shifts calculated using various "phi−psi" models plotted as a function of eq 2. Filled circles denote results from structural models derived from X-ray studies are, while open squares refer to structures constructed by the authors.

single largest contributor to, δ_H as shown in Table 1. The size of the effect is primarily a function of the hydrogen bonding distance, but a careful study by Barfield[23] has shown that two other geometrical parameters have a nonnegligible effect:

$$\Delta\delta_{1°HB} = \{4.81\cos^2(\theta) + \sin^2(\theta)[3.01\cos^2(\rho) - 0.84\cos(\rho) + 1.75]\}e^{-2(r_{OH}-1.760)} \quad (4)$$

where r_{OH} is the hydrogen bond distance in Å, θ is the H···O=C hydrogen bonding angle, and ρ is the H···O=C−N dihedral angle (Figure 2). The values for these distances and angles are shown in Table 2. Though the functional form and constants were determined using different structural models and level of theory, eq 4 reproduces our calculated $\Delta\delta_H$ values in Table 2 (in the column labeled Primary HB) with an RMSD of 0.16 ppm.

3.5. Effect of the Non-Bonded Amide. Structural models such as that shown for Figure 2 have been used extensively to predict chemical shifts using ab initio quantum mechanics. However, we have found that the amide group next to the hydrogen bond acceptor ("nonbonded amide" in Figure 4) can have a significant (up to 1.5 ppm) effect on δ_H (Table 1). The effect is primarily seen for amide groups in α-helices, while in β-sheets the effect is only −0.2−0.3 ppm.

The change in hydrogen bond geometry due to the presence of this group is largely responsible for the effect on δ_H, and eq 4 is able to account for most of the effect. For example, for Lys31 in protein G, the change in chemical shift upon addition of the nonbonded amide −1.5 ppm compared to −1.0 ppm predicted by eq 4.

3.6. Effects of Secondary and Tertiary Hydrogen Bonds. Perhaps the most important finding in this study is that cooperative hydrogen bonding effects have significant effects on δ_H. Hydrogen bonding to the carbonyl group of the amide group of interest ("secondary" hydrogen bonding in Figure 4) and hydrogen bonding not directly to the amide group of interest ("tertiary" hydrogen bonding in Figure 4) can change δ_H by up to 1.3 and 1.1 ppm, respectively (Table 1). In Table 1, the column headed "Tertiary HB" lists the effect on δ_H of adding all tertiary hydrogen bonds, e.g., two amide groups in the case of Lys31 (Figure 4). The number and type of tertiary hydrogen bonds for each residue can be found by inspecting Figure 1. It

Table 2. Computed Structural Parameters and Corresponding Experimental Values Experiment (Using the Proton Position Determined by the WHATIF Program)[a]

	H···O (Å)			N···O (Å)			H···OC (deg)			H···O═CN (deg)		
	1° HB	3° HB	exp	1° HB	3° HB	exp	1° HB	3° HB	exp	1° HB	3° HB	exp
β-sheet												
Protein G												
Tyr3	1.82	1.85	1.79	2.84	2.86	2.77	161.3	156.4	174.8	−108.4	115.9	87.3
Gly14	2.06	2.02	2.07	3.06	3.01	3.03	138.8	141.7	146.6	155.2	169.6	180.0
Phe52	1.83	1.77	1.82	2.84	2.78	2.81	156.4	161.9	165.9	66.7	40.9	48.1
Val54	1.97	1.97	2.00	2.95	2.95	2.97	168.0	163.3	167.8	−6.5	−28.8	15.5
Ubiquitin												
Ile13	1.92	1.87	1.95	2.92	2.88	2.91	140.8	144.1	148.2	−130.8	−132.8	−133.2
α-helix												
Protein G												
Glu27	1.87	1.86	1.83	2.85	2.84	2.79	143.8	147.3	151.1	68.3	64.5	59.2
Lys31	1.83	1.82	1.80	2.80	2.80	2.77	153.7	150.3	154.3	67.7	71.7	66.7
Asp36	1.91	1.83	1.79	2.89	2.82	2.76	141.5	145.8	154.7	69.6	67.0	62.1
Other												
Ubiquitin												
Asp21	1.99	1.93	1.88	2.94	2.87	2.84	125.2	126.3	121.33	69.9	74.5	82.7
Gln40	1.98	1.96	1.96	2.95	2.92	2.88	123.3	125.5	121.84	73.9	71.8	80.0
Irregular												
Protein G												
Leu12[b]	2.86	2.64	2.11	3.41	3.39	3.05						
Thr49[c]	2.74	2.40	2.56	2.95	3.26	3.46	101.0	109.8	103.8	88.6	84.5	83.8
Ubiquitin												
Ser20[d]	2.35	2.35	2.35	2.75	2.75	2.75	80.8	80.3	79.5			
rmsd	0.22	0.16	0.16	0.18	0.12		7.3	7.1		13.7	16.9	

[a] Note that when the H···OC angle is near 180° small changes in atomic position can lead to large changes in the H···O═CN dihedral angle. [b] Water as primary hydrogen bond partner. [c] Amide and asp side chain both primary hydrogen bonding partners. [d] Ser side chain as primary hydrogen bonding partner.

is especially interesting to note that the hydrogen cooperativity extends through C═O···H−C_δN hydrogen bonds involving proline residues, as found for Asp21 in ubiquitin (Figure 1). Furthermore, for Gln 40 we were not able to reproduce the experimental value for δ_H using the 1OGW structure and similar attempts using the 1UBQ structure[41] also failed. Fortunately, an unpublished 1.05 Å structure of a K29Q mutant of ubiqutin obtained by Robertson, Ramaswamy, and co-workers was made available to us and this resulted in a predicted δ_H value (7.2 ppm) in reasonable agreement with experiment (7.8 ppm). The key difference between the new structure and 1OGW and 1UBQ is that the former exhibits a secondary hydrogen bond with the side-chain of Arg72 (Figure 5), while in the other two structures Arg42 is found near, but not in hydrogen bond range of, the amide of interest (Figure 5). If the position of the Arg72 side chain is energy minimized the predicted δ_H increases to 7.4 ppm.

The sign and magnitude of the δ_H change depend on two factors:

(1) A through-space effect and a through-bond effect. For Phe52, the introduction of the secondary and tertiary hydrogen bond without subsequent geometry optimization of the primary hydrogen bonding geometry increases δ_H by 0.1 and 0.3 ppm, compared to 0.6 and 0.4 ppm with geometry optimization. The source of this effect is presumably an inductive (or through-bond) effect mediated by the delocalization of the nitrogen lone pair. However, using the geometry obtained in the presence of the secondary and tertiary hydrogen bonding partners, and then removing them in the chemical shielding calculations, also underestimates the changes in chemical shift.

(2) The nature of the secondary and the nature and number of tertiary hydrogen bonding partners. Charged side chains tend to affect δ_H more than amide groups, which, in turn, tend to have a larger effect than water molecules. For example, the secondary amide hydrogen bonds in the structural models of Phe52 and Val54 increase δ_H by 0.4−0.6 ppm, whereas the corresponding water hydrogen bonds in the structural models of Tyr3 and Ile13 increase δ_H by 0.1 and 0.3 ppm (Table 2). Furthermore, more tertiary hydrogen bonding partners of the same type have a larger effect on δ_H. For example, the one and two tertiary amide hydrogen bonds increase δ_H by 0.4 and 0.9 ppm for Phe52 and Lys31, respectively.

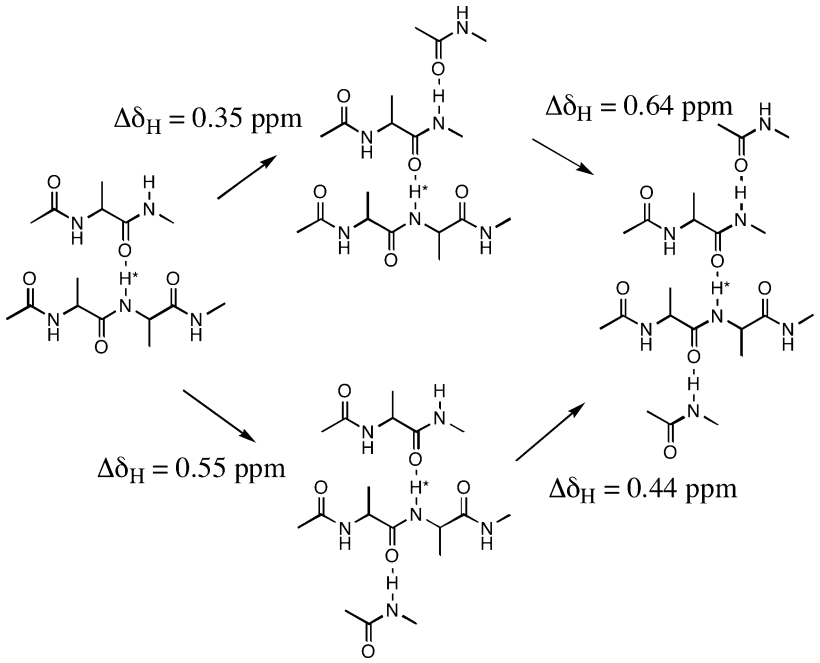
The order in which hydrogen bonds are added does not seem to matter greatly (i.e., the effect is roughly additive). For example, in the case of Phe52, the addition of the tertiary hydrogen bond in the absence of the secondary hydrogen bond increases δ_H by 0.35 ppm, compared to 0.44 ppm (Scheme 1). The secondary hydrogen bonds generally have a larger effect on the chemical shift compared to tertiary hydrogen bonds.

We discuss a quantitative model of these effects in section 3.9.

3.7. Side Chains Not Involved in Hydrogen Bonds. The largest structural models used here for chemical shift calculations are constructed by adding the two side chains to the di-alanine model and one side-chain to the primary hydrogen bonding partner as shown in Figure 4 and Figure 1. One exception is Val54 where the side chain of Trp43 is also included since it is very close to the amide proton of interest. As mentioned previously, the primary hydrogen bonding geometry is not reoptimized in the presence of these side chains. As shown in Table 2, side chains have only a modest (0.0−0.3 ppm) effect

(41) Vijaykumar, S.; Bugg, C. E.; Cook, W. J. *J. Mol. Biol.* **1987**, *194*, 531.

Scheme 1



on δ_H with the exception of Val54 where the ring current effect of Trp43 decreases δ_H by 0.6 ppm.

3.8. Effect of Solvent. Overall, the inclusion of bulk solvation effects as described by a continuum model has a modest effect on δ_H, with slightly larger changes for residues in the α-helix (0.3−0.5 ppm) compared to elsewhere in the protein (−0.2−0.5 ppm). This is most likely due to solvent screening of the interaction with the nonbonded amide, since the inclusion of solvent on the "nonbonded amide model" (Figure 4) has roughly the same effect on δ_H (data not shown). However, solvent effects can be very important for amide groups hydrogen bonded to charged groups as discussed in section 3.11.

3.9. Quantitative Interpretation. The final set of proton chemical shift values can be reproduced well (i.e., quantitatively rationalized) by

$$\delta_H = \delta_\omega + \Delta\delta_{1°HB} + \Delta\delta_{2°HB} + \Delta\delta_{3°HB} + \Delta\delta_{rc} \quad (5)$$

Here, δ_ω and $\Delta\delta_{1°HB}$ are given by eqs 3 and 4 and evaluated using the ab initio optimized primary hydrogen bonding geometry of the structural models shown in Figure 1 (where side chains are represented by Ala). $\Delta\delta_{rc}$ is modeled using the equation proposed by Case,[42] and is only applied to Val54 where the value (0.6 ppm) is computed using the SHIFTX web interface using the protein geometry.

$\Delta\delta_{2°HB} = 0.15$, 0.3, or 0.8 ppm depending on whether the secondary hydrogen bonding donor is water, an amide group, or a charged side-chain, respectively.

$\Delta\delta_{3°HB} = 0.1$, 0.2, or 0.8 ppm for any water, amide group, or charged side-chain hydrogen bonding partners at the C═O group of the secondary hydrogen bond donor plus 0.05, 0.1, or 0.8 ppm for any water, amide group, or charged side-chain hydrogen bonding partners at the NH group of the primary hydrogen bond acceptor, respectively. For example, in the case of Glu27 (Figure 1), $\Delta\delta_{3°HB}$ = 0.2 ppm (amide) + 0.05 ppm (water), whereas for Asp26, $\Delta\delta_{3°HB} = 2 \times 0.1$ ppm (2 waters) + 0.1 ppm (amide).

The proton chemical shifts predicted using eq 5 and the optimized geometries (Table 3) reproduce our quantum mechanical results with an RMSD value of 0.4 ppm ($r = 0.90$),

Table 3. Amide Proton Chemical Shift Predictions Using eq 5 and Three Other On-line Chemical Shift Predictors (See Text for Further Information)

	X-ray Geometries				optimized geometries				
	eq 5	ShiftS	ShiftX	Proshift	QM	eq 5	ShiftS	ShiftX	experiment
β-sheet									
Protein G									
Tyr3	10.3	9.2	9.4	8.9	9.5	9.5	9.3	9.6	9.1
Gly14	8.4	8.9	7.1	8.7	8.6	8.7	8.9	7.1	8.3
Phe52	10.1	9.5	9.8	9.0	10.4	10.5	9.5	9.8	10.4
Val54	8.3	8.4	8.3	8.9	8.5	8.5	8.3	8.2	8.3
Ubiquitin									
Ile13	9.0	8.9	9.2	8.8	9.4	9.5	8.7	8.9	9.6
α-helix									
Protein G									
Glu27	8.6	8.1	8.0	8.1	8.7	8.3	8.0	8.1	8.3
Lys31	8.9	8.3	8.4	8.0	9.2	8.7	8.3	8.4	9.0
Asp36	9.1	8.4	8.4	8.0	9.1	8.5	8.4	8.3	8.7
Other									
Ubiquitin									
Gln40[a]	7.2	7.9	8.1	−	8.0	7.4	7.9	8.1	7.8
Asp21	7.5	8.0	8.2	7.8	7.2	7.7	8.0	8.2	8.0
Irregular									
Protein G									
Leu12	8.6	7.7	7.7	7.8	7.1	7.3	7.7	7.8	7.6
Thr49	5.9	7.4	7.7	7.7	6.4	6.1	7.4	7.5	7.0
Ubiquitin									
Ser20	6.1	7.5	7.9	8.0	7.1	7.6	7.8	7.9	7.0
RMSD	0.7	0.4	0.6	0.7	0.3	0.3	0.5	0.6	
r	0.87	0.91	0.81	0.71	0.96	0.94	0.87	0.80	
RMSD[b]	0.5	0.5	0.5	0.7	0.3	0.3	0.5	0.6	
r^2[b]	0.83	0.83	0.82	0.54	0.94	0.95	0.79	0.78	

[a] Structure from new ubiquitin structure. [b] For comparison excluding irregular residues Leu12, Thr49, and Ser20.

(42) Case, D. A. *J. Biomol. NMR* **1995**, *6*, 341.

while it reproduces experiment with an RMSD of 0.3 ppm ($r = 0.94$).

A useful feature of eq 5 is that it is sensitive to the geometry of the primary hydrogen bond and may therefore be useful in validating, and ultimately refining, experimental geometries. Here, we explore this issue in a preliminary fashion by using eq 5 and the X-ray structures to compute the chemical shifts. As expected, the predicted chemical shifts do not agree as well with experiment, with RMSD values of 0.5 ppm ($r = 0.83$). Tyr3, Ile13, Asp21, and Gln 40 show predicted δ_H values that deviate from experiment by more than 0.3 ppm (the RMSD obtained using the optimized geometries). For example, using the X-ray structure the predicted value of Tyr3 is 10.3 ppm, significantly higher than the experimental value of 9.1 ppm and the corresponding prediction using the optimized geometry of 9.5 ppm. Interestingly, the largest deviation in the N···O distance obtained from the optimized and X-ray geometry (2.86 vs 2.77 Å) is observed for Tyr3. More generally, two of the four significant errors are observed for ubiquitin residues (Ile13, Asp21) calculated using an X-ray structure with a lower resolution than that of protein G (1.32 vs 1.10 Å).

3.10. Comparison to Other Methods. Several other empirical approaches exist for the prediction of NMR chemical shifts, and here we compare several of the results obtained with three of these methods to results obtained by eq 5 for the residues considered in this study. The three methods are SHIFTS developed by Case and co-workers,[14] SHIFTX developed by Wishart and co-workers,[10] and PROSHIFT developed by Meiler.[15] The first two methods are more similar to eq 5 in that they use physically motivated functional forms to relate the protein structure and proton chemical shifts, while PROSHIFT is based on a trained neural net. All three approaches contain adjustable parameters determined by fitting to experimental chemical shift values using X-ray structures.

Using X-ray structures, SHIFTS, SHIFTX, and PROSHIFT reproduce the experimentally determined δ_H with respective RMSD values of 0.4, 0.5, and 0.7 ppm. The first two values are similar to the RMSD of 0.5 ppm obtained using eq 5. Using protein structures in which the primary hydrogen bond geometries have been adjusted to reproduce our ab initio data has virtually no effect on the proton chemical shift values predicted by SHIFTS and SHIFTX (the PROSHIFT web interface did not allow user defined files to be uploaded) and the respective RMSD values are essentially unchanged: 0.4 and 0.6 ppm. These values are somewhat larger than the 0.3 ppm RMSD obtained with eq 5.

The better performance of the eq 5 in this case may be due to the fact that several of its terms are adjusted to best reproduce the experimental results using the optimized primary hydrogen bond geometries, and does not necessarily imply that the optimized geometries are more accurate than the X-ray geometries. However, it is instructive to consider a specific case. The amide proton chemical shift of Phe52 is 1.3 ppm larger than for Tyr3. This difference is relatively well reproduced ($\Delta\delta_H = 1.0$ ppm) by eq 5 using optimized geometries but not using experimental geometries ($\Delta\delta_H = -0.2$ ppm), and $\Delta\delta_H$ is underestimated by the other approaches using either optimized geometries ($\Delta\delta_H = 0.1-0.4$ ppm) or experimental geometries ($\Delta\delta_H = 0.2$ ppm). Furthermore, the optimized geometries are consistent with a stronger hydrogen bond for Val54 compared to Tyr3 (N···O = 2.78 vs 2.86 Å), whereas the X-ray geometry is not (N···O = 2.81 vs 2.77 Å).

Clearly, significantly more work is needed to establish if and how eq 5 can be used to refine protein structures. The first step, implementing eq 5 in a computer program similar to SHIFTS and SHIFTX, is ongoing.

3.11. Primary Hydrogen Bonds to Non-Amide Groups. So far, we have considered cases where the primary hydrogen bond is to another amide. Here, we consider three other residues (Thr49 and Leu12 in protein G and Ser20 in ubiquitin) where this is not the case.

Leu12 is located in a β-hairpin turn and the closest hydrogen bonding partner of the amide proton is a crystallographic water molecule. However, addition of this water molecule has a negligible effect on δ_H since the hydrogen bond is significantly weakened upon geometry optimization (Table 2). Since the hydrogen bond acceptor is water there is no nonbonded amide, and the secondary hydrogen bond is another water molecule, which increases δ_H slightly (0.2 ppm) as for the other residues (e.g., Tyr3 and Ile13). The tertiary hydrogen bonds involve the Glu61 side chain and two additional hydrogen bonds to it (Figure 6), which makes the single largest contribution to δ_H (0.7 ppm). This is primarily due to the fact the optimization of the water position in the presence of the Glu61 side-chain significantly shortens the primary hydrogen bond (Table 2). However, the $N-O_w$ distance is still 0.34 Å longer than in the X-ray structure and the chemical shift is 1.0 ppm lower than experiment. The additions of side chains and solvent effects increase δ_H by 0.5 ppm. The remaining 0.5 ppm error observed for this residue may be due to the difficulty in obtaining the correct position of the water. Interestingly, the predicted δ_H value based on eq 5 and the optimized geometry results is 7.8 ppm, in good agreement with the experimental value of 7.6 ppm, whereas use of the experimental structure results in a predicted δ_H value of 8.6 ppm.

Ser20 is located in the middle of a five-residue loop connecting a β-strand and an α-helix. Only a very weak interaction with the Ser20 side chain was observed in the 1OGW and 1UBQ X-ray structures and predictions based on these structures led to predicted chemical shift values that were too low compared to experiment. However, a primary hydrogen bond to a crystallographic water molecule was observed in the high-resolution structure of Robertson, Ramaswamy, and co-workers. Since the position of the water molecule is completely optimized, it was added to our structural models already constructed using the 1OGW structure (Figure 7). The water molecule is also in hydrogen bonding distance to the Ser20 side chain, which was included during the geometry optimization of the primary hydrogen bond geometry (based on our experience with Leu12). Computed in this way, the primary hydrogen bond increases the δ_H by 2.1 ppm (Table 1), and is thus somewhat lower but still comparable to the effect of a primary hydrogen bond to an amide group. The effects of secondary and tertiary hydrogen bonding as well as side chains and bulk solvation are all comparable to those obtained for the "regular" cases discussed above, and the final prediction of 7.1 ppm is in excellent agreement with the experimental value of 7.0 ppm. However, the predicted δ_H value based on eq 5 is overestimated by 0.6 ppm. This is pri-

marily due to an overestimation of the effect of the primary hydrogen bonds to the water and Ser20 side chain (2.7 ppm) by eq 4.

Thr49 is in a β-hairpin turn and has two primary hydrogen bonds to the side chain and backbone carbonyl group of Asp47 (Figure 8). Preliminary calculations quickly determined that the inclusion of bulk solvation in *both* the geometry optimization and chemical shift calculations were necessary for sensible results due to the presence of the charged Asp47 side chain, so these are included in all calculations involving this side chain. The combined effect of these two primary hydrogen bonds on δ_H is 0.9 ppm, a significantly smaller effect than observed for the other residues (1.6−4.0 ppm) with the exception of Leu12. The nonbonded amide increases δ_H further by 0.7 ppm as does the secondary hydrogen bond to a water molecule (Table 1). This latter effect is unusually large compared to our findings for Tyr3, Ile13, Leu12, and Asp21 where the corresponding hydrogen bond increases δ_H by 0.1−0.3 ppm. Tertiary hydrogen bonds and the additions of other side chains do not contribute to the δ_H of Thr49 and the resulting value is 6.4 ppm, somewhat lower than the experimental value of 7.0 ppm. The corresponding value predicted using eq 5 is even lower (6.1 ppm), primarily due to the unusually large contribution of the secondary hydrogen bond involving water of 0.7 ppm, which is approximated as 0.15 ppm in eq 5. It is possible that primary hydrogen bonds involving charged side chains are more affected by secondary hydrogen bonding. We will address this issue further in future studies.

In summary, the determination of δ_H for backbone amides hydrogen bonded to water molecules and side chains can present a challenge and it is not clear that eq 5 will offer useful predictions for such cases. Both issues will require further study, as will the role of dynamical motion of the protein in general and in loops in particular.

4. Summary and Future Directions

We have developed an electronic structure-based computational methodology for the prediction of backbone amide proton chemical shifts (δ_H) that reproduces experimental values with an RMSD of 0.3 ppm for 10 residues in protein G and ubiquitin. The largest deviation from experiment is −0.6 ppm (for Gln40 in ubiquitin), and this value can be reduced by including more atoms in the energy minimization (as discussed above). Detailed analyses of the results reveal the key determinants of δ_H. Consistent with similar studies,[19] the geometry of the hydrogen bond involving the proton of the interest (referred to here as the primary hydrogen bond) has the single largest effect on δ_H, which is thus a good "reporter" on the hydrogen bond geometry. Perhaps the most interesting finding in this study is that cooperative hydrogen bonding effects can have a significant effect on δ_H by affecting the primary hydrogen bond geometry and polarizing the electron density around the amide proton. Cooperative effects in hydrogen bonding networks are well established[43−54] but the effect on amide proton chemical shifts has not been successfully quantified before. Finally, amino acid side chains not directly involved in the hydrogen-bonding network have little effect on δ_H with the exception of very short-range ring current effects due to aromatic side chains.

We have found simple functional forms that relate the δ_H-determinants to the protein structure. In particular, we have found Barfield's equation[23] relating δ_H-changes to the primary hydrogen bond geometry very useful. When combined with the other functional forms developed here our empirical model [eq 5] predicts δ_H values with an RMSD of 0.3 ppm compared to experiment when energy minimized structures are used. Using X-ray structures the RMSD increases to 0.5 ppm, presumably due to inaccuracies in the hydrogen bond geometry. Other chemical shift predictors are significantly less sensitive to the protein geometry and predict the corresponding δ_H values with an RMSD of 0.4−0.7 ppm using either structure. This is not surprising since these methods have been parametrized using X-ray geometries. Future studies will address the use of eq 5 in protein structure validation and refinement. Finally, we plan to perform studies similar to this on other observables such as coupling constants, chemical shift anisotropies, and fractionation factors.

Acknowledgment. Support from the NSF (MCB 0209941) and the Iowa Research Experience for Undergraduates Program is gratefully acknowledged. The calculations were performed on IBM RS/6000 workstations and an eight-node QuantumStation obtained through NSF grants CHE 9974502 and MCB 0209941, respectively. We thank Andrew Robertson and S. Ramaswamy for making their unpublished X-ray structure available to us, and Jon Baker at PQS for making a the COSMO/NMR interface available to us ahead of its official release. Finally, we thank Andrew Robertson for a careful reading of the manuscript.

JA0617901

(43) Wieczorek, R.; Dannenberg, J. J. *J. Am. Chem. Soc.* **2005**, *127*, 17216.
(44) Koch, O.; Bocola, M.; Klebe, G. *Proteins−Struct. Funct. Bioinform.* **2005**, *61*, 310.
(45) Salvador, P.; Kobko, N.; Wieczorek, R.; Dannenberg, J. J. *J. Am. Chem. Soc.* **2004**, *126*, 14190.
(46) Viswanathan, R.; Asensio, A.; Dannenberg, J. J. *J. Phys. Chem. A* **2004**, *108*, 9205.
(47) Kobko, N.; Dannenberg, J. J. *J. Phys. Chem. A* **2003**, *107*, 6688.
(48) Ireta, J.; Neugebauer, J.; Scheffler, M.; Rojo, A.; Galvan, M. *J. Phys. Chem. B* **2003**, *107*, 1432.
(49) Lin, J. Q.; Luo, S. W.; Wu, Y. D. *J. Comput. Chem.* **2002**, *23*, 1551.
(50) Miller, J. S.; Kennedy, R. J.; Kemp, D. S. *J. Am. Chem. Soc.* **2002**, *124*, 945.
(51) Topol, I. A.; Burt, S. K.; Deretey, E.; Tang, T. H.; Perczel, A.; Rashin, A.; Csizmadia, I. G. *J. Am. Chem. Soc.* **2001**, *123*, 6054.
(52) Kobko, N.; Paraskevas, L.; del Rio, E.; Dannenberg, J. J. *J. Am. Chem. Soc.* **2001**, *123*, 4348.
(53) Jaravine, V. A.; Alexandrescu, A. T.; Grzesiek, S. *Protein Sci.* **2001**, *10*, 943.
(54) Guo, H.; Gresh, N.; Roques, B. P.; Salahub, D. R. *J. Phys. Chem. B* **2000**, *104*, 9746.