Protein−Solvent Hydrogen Bonding Studied by NMR $^1J_{NC'}$ Coupling Constant Determination and Molecular Dynamics Simulations

Nenad Juranić,* Vladimir A. Likić,
Franklyn G. Prendergast, and Slobodan Macura

Department of Biochemistry and Molecular Biology
Mayo Graduate School, Mayo Clinic and Foundation
Rochester, Minnesota 55905

Received December 15, 1995
Revised Manuscript Received April 18, 1996

At present, there is no experimental technique for direct measurement of individual H-bonds between protein and solvent. It is important to obtain specific information on such H-bonds because they play a central role in the stability and function of biological macromolecules. For example, protein stability in solution and the folding pattern of the peptide backbone strongly depend on the balance between intramolecular H-bonds and intermolecular H-bonds with water (solvent).[1] Formation of H-bonds at protein−DNA recognition sites in many instances is mediated by hydration water.[2] Recently, we have shown that protein main-chain H-bonding in solution can be monitored by the protein main-chain amide group $^1J_{NC'}$ coupling constant.[3] This coupling constant increases from predominantly NH to predominantly CO H-bonding of amide groups. Thus, the $^1J_{NC'}$ coupling constant reveals characteristic patterns for secondary structure elements of proteins. We have demonstrated that the same coupling constant is sensitive to H-bonding of the protein amide group to water molecules of the solvent.[3−5] For further analysis of the $^1J_{NC'}$ coupling constant dependence on H-bonding an independent assessment of the protein−solvent H-bonds is needed. Crystal-structure X-ray analysis provides H-bond data for tightly bound water molecules.[6] However, this is only a small fraction of total protein−water H-bonds and thus is insufficient for the present analysis.

A more detailed picture of the hydrogen-bonding patterns can be obtained from the molecular dynamics (MD) simulations with explicit water.[7−15] Here we report the relationship of the $^1J_{NC'}$ coupling constant to protein−solvent H-bonds evaluated by the use of MD simulations of the solvated protein.

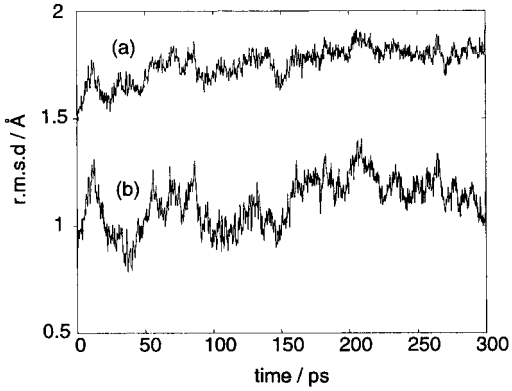

Figure 1. Root mean square deviations (rmsd) of the instantaneous structure of ubiquitin in water from the crystal structure: (a) all atoms; (b) protein backbone.

We investigated human ubiquitin, a small globular protein, MW 8600, with known X-ray structure[16] and $^1J_{NC'}$ coupling constants.[3] Initial heavy atom coordinates for ubiquitin were obtained from the crystal structure.[16] For water we have chosen a simple and well-established three-point-charge model, TIP3P.[17] MD simulations were carried out with the program CHARMm,[18] using the Molecular Simulations all-atom parameter set, version 22.[19]

During the simulation, the ubiquitin structure remained in close vicinity of the crystal structure as is shown in Figure 1. Only the C-terminal tail, residues 73−76, exhibited somewhat larger fluctuations. The dependence of the rms deviations between the instantaneous structure in water and the crystal structure suggested that equilibrium was reached after 50−100 ps (Figure 1). The rms deviations between the average molecular dynamics structure and the crystal structure was 1.46 Å for all atoms, and 0.96 Å for backbone atoms. In the course of simulation the coordinates of ubiquitin and water were saved every 200 time steps (0.2 ps), and 1250 coordinate frames of solvated ubiquitin (from 50 to 300 ps) were analyzed for both intra- and intermolecular hydrogen bonds. The criteria for evaluating the presence of a H-bond were solely geometric.[11] A H-bond is recorded if the H···O distance and the H-bond angle, H···O═C or N−H···O, were less than selected cutoff values. The H-bonds are reported by a score number, a ratio of the events in which the H-bond was present and the total number of sampling events. The H-bond scores were sorted into four classes: (hbCOw) protein backbone carbonyl oxygen to water hydrogen; (hbCOi) backbone carbonyl oxygen to

* To whom correspondence should be addressed. Phone: (507) 284-0114. Fax: (507) 284-8433. E-mail: juranic@mayo.edu.

(1) Eberhardt, E.; Raines, R. *J. Am. Chem. Soc.* **1994**, *116*, 2149−2150.
(2) Wüthrich, K. *NMR in Structural Biology*; Anonymous World Scientific: London, 1995; pp 647−655.
(3) Juranić, N.; Ilich, P. K.; Macura, S. *J. Am. Chem. Soc.* **1995**, *117*, 405−410.
(4) Juranić, N.; Moncrieffe, M.; Macura, S.; Prendergast, F. G. *Book of Abstracts*, 36th Experimental Nuclear Magnetic Conference, Boston, 1995; 1995; p 152.
(5) Juranić, N.; Moncrieffe, M.; Kurian, E.; Prendergast, F. G.; Macura, S. International Conference on Molecular Structural Biology, Vienna, 1995, Proceedings, 1995; p 21.
(6) Billeter, M. *Prog. Nucl. Magn. Reson. Spectrosc.* **1995**, *27*, 635−645.
(7) Levitt, M.; Sharon, M. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 7557−7561.
(8) Brooks, C.; Karplus, M. *J. Mol. Biol.* **1989**, *208*, 159−181.
(9) Komeiji, Y.; Uebayasi, M.; Someya, J.; Yamato, I. *Proteins: Struct., Funct., Genet.* **1996**, *16*, 268−277.
(10) Tirado-Rives, J.; Jorgensen, W. *J. Am. Chem. Soc.* **1990**, *112*, 2773−2781.
(11) Hartsough, D. S.; Mertz, K. M. J. *J. Am. Chem. Soc.* **1993**, *115*, 6529−6537.
(12) Vijayakumar, S.; Vishveshwara, S.; Ravishanker, G.; Beveridge, D. L. In *Modeling the Hydrogen Bond*; Smith, D., Ed.; American Chemical Society: Washington, DC, 1994; pp 175−193.
(13) Daggett, V.; Levitt, M. *Annu. Rev. Biophys. Biomol. Struct.* **1993**, *22*, 353−380.
(14) Cheatham, T.; Miller, J.; Fox, T.; Darden, T.; Kollman, P. *J. Am. Chem. Soc.* **1995**, *117*, 4193−4194.
(15) Wodak, S.; Belle, D.; Prevost, M. In *Computer Modelling in Molecular Biology*; Goodfellow, J., Ed.; VCH Publishers: New York, 1995; pp 62−99.
(16) Vijay-Kumar, S.; Bugg, C.; Cook, W. *J. Mol. Biol.* **1987**, *194*, 531−544.
(17) Jorgensen, W.; Chandrasekhar, J.; Madura, D. *J. Chem. Phys.* **1983**, 926−931.
(18) Brooks, B.; Bruccoleri, R.; Olafson, B.; States, D.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187−217.
(19) The protein initial structure was immersed in a 56.0 Å water cube, and water molecules closer than 2.6 Å to the protein were deleted. The resultant structure, which contained ubiquitin and 5396 water molecules, was treated with periodic boundary conditions. The system was extensively minimized with the combined steepest descents and adopted basis Newton−Raphson method. In the first minimization stage the protein was fixed, allowing only water molecules to move; subsequently both protein and water were allowed to relax. In the molecular dynamics simulations we used the leapfrog method with an integration time step of 1 fs. The bonds containing hydrogens were constrained with the SHAKE algorithm.[22] The nonbonded energies and forces were smoothly truncated at 10.0 Å, and the nonbonded interaction lists[23] were updated every 10 time steps. The minimized ubiquitin−water structure was heated to 293.0 K within 5 ps, and then equilibrated for 25 ps with velocity scaling every 5 ps. Starting from the structure produced by the equilibration, we carried out a 300 ps molecular dynamics simulation which was analyzed for hydrogen bonds. The average temperature for a 300 ps simulation was 295.6 K.

S0002-7863(95)04210-7 CCC: $12.00

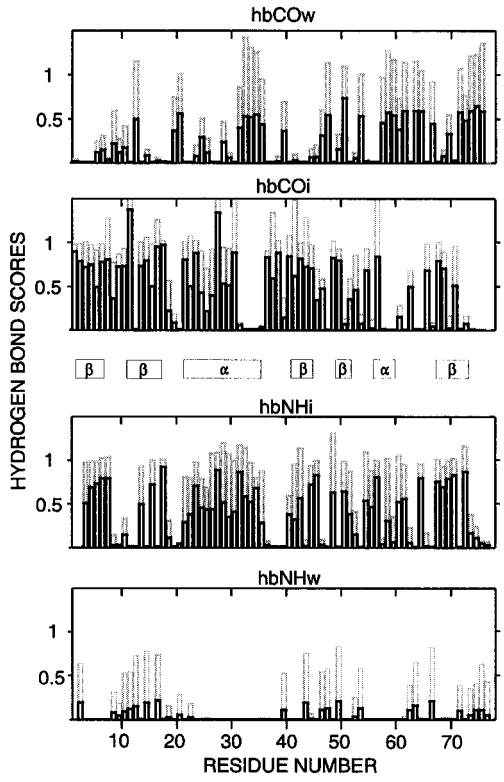

Figure 2. H-bond scores of the backbone amide groups as obtained in MD simulation of solvated human ubiquitin at two cutoffs, 2.1 (full lines) and 2.5 Å (dashed lines). Presented are protein−solvent H-bonds formed between carbonyl oxygen and water hydrogen (hbCOw) and amide protons and water oxygens (hbNHw), and protein−protein H-bonds formed between carbonyl oxygen and protein polar hydrogens (hbCOi) and amide protons and protein polar oxygens (hbNHi).

intraprotein polar hydrogen; (hbNHw) protein backbone amide hydrogen to water oxygen; (hbNHi) backbone amide hydrogen to intraprotein polar oxygen.

Figure 2 shows the protein backbone H-bond scores determined from the same MD simulations data set using cutoffs of 2.1 Å (full lines) and 2.5 Å (dashed lines) and an angle cutoff of 120°. The scores of intramolecular H-bonds (hbCOi and hbNHi) at a 2.5 Å cutoff indicate that most of the intraprotein H-bonding network, as known from crystal structure x-ray analysis and solution NMR,[16,20] is well reproduced. Protein−solvent H-bonds have high scores at carbonyl oxygens (hbCOw) and rather low scores at amide hydrogens (hbNHw), in agreement with previous observation in statistical surveys of protein H-bonding. For the scores with a 2.1 Å cutoff the overall H-bond profile is similar to that obtained with the 2.5 Å cutoff, but the scores are generally lower and more variable. We found that scores with a 2.1 Å cutoff correlate better with measured $^1J_{NC'}$ coupling constants, indicating the dominant role of the short distance H-bonds on the coupling constant. We have also correlated (electrostatic) H-bonding energy[7] with the coupling constant, but correlation obtained with the geometric criteria was better.

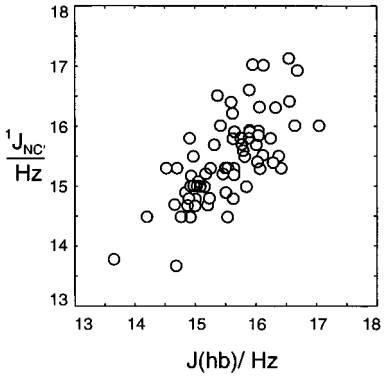

Figure 3. Empirical relationships between the $^1J_{NC'}$ coupling constant and the H-bonding populations obtained from MD simulations. A correlation coefficient of 0.72 and a standard deviation of 0.65 Hz are obtained by applying the equation J(hb)/Hz = 14.3 + 4.1(hbCOs) + 1.6(Hb(COi) − 2.1(hbNHs) − 0.8(hbNHi) to data for 72 peptide groups in human ubiquitin.

For correlation of the coupling constants with H-bond scores in the absence of a suitable theoretical predictor, the constant was expressed as a linear combination of the individual H-bond contributions: J(hb) = J_0 + a(hbCOw) + b(hbCOi) + c-(hbNHw) + d(hbNHi). A least-squares fitting gave a correlation coefficient of 0.7 with $a = 4.1 \pm 0.5$ Hz, $b = 1.6 \pm 0.3$ Hz, $c = -2.1 \pm 1.0$ Hz, and $d = -0.8 \pm 0.3$ Hz (Figure 3). The result shows that H-bonding to water (hbCOw and hbNHw scores) has about a 2-fold larger impact on the $^1J_{NC'}$ coupling constant than intraprotein H-bonding (hbCOi and hbNHi scores). It also shows that carbonyl oxygen H-bonding (hbCOw and hbCOi scores) increases the $^1J_{NC'}$ coupling constant about twice as much as H-bonding of amide protons (hbNHw and hbNHi scores) decreases it. Taking into account that hbNHw scores are generally low, *the protein−solvent* C═O···HOH H-bond has the largest impact on the $^1J_{NC'}$ coupling constant. The extreme values of the coupling, $^1J_{NC'} > 17$ Hz and $^1J_{NC'} < 14$ Hz, are gained only for hbCOw > 0.5 and hbCOw < 0.05, respectively. Therefore, the extremely high values of the coupling constant, $^1J_{NC'} > 17$ Hz, corresponds to a high population of C═O···HOH hydrogen bonds, and extremely low values, $^1J_{NC'} < 14$ Hz, to the absence of H-bonds at the respective carbonyl sites.

Presently observed modest correlation between the $^1J_{NC'}$ coupling constants of the backbone amide groups and H-bond population of the protein main chain may be improved by improving methods of H-bond calculation from MD (geometric scores vs electrostatic energy, better force field in MD) and by including the dependence of the coupling constant on other parameters (local geometry, nature of the residues). Nevertheless, the correlation shows a dominant role of C═O···HOH and a negligible role of N−H···OH_2 protein−solvent H-bonding on the coupling constant. Consequently, the extreme values of the coupling constant are unambiguous indicators of the high population ($^1J_{NC'} > 17$ Hz) or absence ($^1J_{NC'} < 14$ Hz) of the short H-bonds between water hydrogen and the protein backbone amide oxygen. Overall, this suggests that the $^1J_{NC'}$ coupling constant is a useful experimental parameter in the study of protein−solvent H-bonding.

Acknowledgment. We thank Dr. Zsolt Zolnai for help in the mathematical analysis of the correlations. We also thank the reviewers for their valuable suggestion for revision of the paper.

JA9542101

(20) Weber, P.; Brown, S.; Mueller, L. *Biochemistry* **1987**, *26*, 7282−7290.
(21) Baker, E.; Hubbard, R. *Prog. Biophys. Mol. Biol.* **1984**, *44*, 97−179.
(22) Ryckaert, J.; Ciccotti, G.; Berendsen, H. J. *J. Comput. Phys.* **1977**, *23*, 327−341.
(23) Allen, M.; Tildesley, D.*Computer Simulation of Liquids*; Claredon Press: Oxford, 1989.