J|A|C|S
ARTICLES
Published on Web 03/30/2004

Investigation of the NMR Spin−Spin Coupling Constants across the Hydrogen Bonds in Ubiquitin: The Nature of the Hydrogen Bond as Reflected by the Coupling Mechanism

Tell Tuttle, Elfi Kraka, Anan Wu, and Dieter Cremer*

Contribution from the Department of Theoretical Chemistry, Göteborg University, Reutersgatan 2, S-41320 Göteborg, Sweden

Received April 21, 2003; E-mail: cremer@theoc.gu.se

Abstract: The indirect scalar NMR spin−spin coupling constants across the H-bonds of the protein ubiquitin were calculated, including the Fermi contact, the diamagnetic spin−orbit, the paramagnetic spin−orbit, and the spin dipole term, employing coupled perturbed density functional theory in combination with the B3LYP functional and different basis sets: (9s,5p,1d/5s,1p)[6s,4p,1d/3s,1p] and (11s,7p,2d/5s,1p)[7s,6p,-2d/4s,2p]. Four different models based on either the crystal or the aqueous solution structure of ubiquitin were used to describe H-bonding for selected residue pairs of ubiquitin. Calculated and measured ^{3h}J(NC′) coupling constants differ depending on the model used, which is due to the fact that the geometry of ubiquitin is different in the solid state and in aqueous solution. Also, conformational averaging leads to a decrease of the magnitude of the measured ^{3h}J(NC′) constants, which varies locally (larger for β-sheets, smaller for α-helix). Two different spin−spin coupling mechanisms were identified. While mechanism I transmits spin polarization via an electric field effect, mechanism II involves also electron delocalization from the lone pair of the carbonyl oxygen to the antibonding orbital of the N−H bond. Mechanism I is more important in the crystal structure of ubiquitin, while in aqueous solution, mechanism II plays a larger role. It is possible to set up simple relationships between the spin−spin coupling constants associated with the H bond in proteins and the geometrical features of these bonds. The importance of the ^{3h}J(NC′) and ^{1}J(N−H) constants as descriptors for the H-bond is emphasized.

1. Introduction

The experimental observation of spin−spin coupling constants (SSCCs) across hydrogen bonds (HBs), $^{n,h}J$,[1,2] has created many exciting possibilities for the application of these parameters to biologically important compounds such as proteins[3], DNA,[4] etc. It has long been established that hydrogen bonding is the primary interaction that determines the secondary structure of these molecules.[5] However, while crystallographic studies have been able to provide the relative positions of heavy atoms in biomolecules, the position of the hydrogen atoms have proven to be far more elusive. Therefore, the position of the hydrogen atoms, the directionality of the HB (linear or bent), and other geometrical features of the HB have previously been inferred from the resulting spatial proximity of the donor and acceptor moiety.[5,6]

Magnetic properties of molecules such as NMR chemical shifts sensitively depend on the geometry of a molecule. It has been shown that critical geometrical parameters, which change depending on the environment (gas phase, solution phase, etc.), can be determined by a careful comparison of calculated and measured NMR chemical shifts.[7,8] SSCCs have also been found to be sensitive to geometrical features such as bond length and bond angles.[9−11] The revelation of the relationship between SSCCs across a HB and the geometrical features of the bond affords us the opportunity to accurately determine the position of the hydrogen atom and, in turn, the geometry of the HB. This knowledge will lead to a more accurate description of, e.g., interresidue hydrogen bonding in polypeptides and proteins, and consequently a better description of the molecule as a whole.

The first measurements of trans-HB SSCCs by Dingley and Grezeseik[1] triggered a flood of follow-up studies, including both experimental[2,12−17] and theoretical investigations.[18−23] Special

(1) Dingley, A. J.; Grezeseik, S. *J. Am. Chem. Soc.* **1998**, *120*, 8293.
(2) Pervushin, K.; Ono, A.; Fernández, C.; Szyperski, T.; Kainosho, M.; Wüthrich, K. *Proc. Natl. Acad. Sci.* **1999**, *121*, 6019.
(3) (a) Wang, Y.-X.; Jacob, J.; Cordier, F.; Wingfield, P.; Stahl, S. J.; Lee Huang, S.; Torchia, D.; Grezeseik, S.; Bax, A. *J. Biomolecular NMR* **1999**, *14*, 181. (b) Lohr, F.; Mayhew, S. G.; Ruterjans, H. *J. Am. Chem. Soc.* **2000**, *122*, 9289.
(4) Dingley, A. J.; Masse, J. E.; Feigon, J.; Grezeseik, S. *J. Biomol. NMR* **2000**, *16*, 279.
(5) See, e.g.: Jeffrey, G. A.; Saenger, W. *Hydrogen Bonding in Biological Structures*; Springer, Berlin, 1991.
(6) Bernstein, J.; Etter, M. C.; Leiserowitz, L. In *Structure Correlation*; Bürgi, H.-B., Dunitz, J. D., Eds.; VCH: Weinheim, 1994.
(7) (a) Cremer, D.; Reichel, F.; Kraka E. *J. Am. Chem. Soc.* **1991**, *113*, 9459. (b) Cremer, D.; Svensson, P.; Kraka, E.; Konkoli, Z.; Ahlberg, P. *J. Am. Chem. Soc* **1993**, *115*, 7457. (c) Arshadi, M.; Johnels, D.; Edlund, U.; Ottosson, C.-H.; Cremer, D. *J. Am. Chem. Soc.* **1996**, *118*, 5120.
(8) (a) Ottosson, C.-H.; Kraka, E.; Cremer, D. In *Theoretical and Computational Chemistry*; Maksic, Z., Ed.; Elsevier: Amsterdam, 1999; Vol. 6, pp 231−301. (b) Cremer, D.; Olsson, L.; Reichel, F.; Kraka E. *Israel J. Chem.* **1993**, *33*, 369.
(9) (a) Kowalewski, J. *Prog. NMR Spectrosc.* **1977**, *11*, 1. (b) Kowalewski, J. *Annu. Rep. NMR Spectrosc.* **1982**, *12*, 81.
(10) (a) Contreras, R. H.; Facelli, J. C. *J. Annu. Rep. NMR Spectrosc.* **1993**, *27*, 255. (b) Contreras, R. H.; Peralta, J. *E. Prog. NMR Spectrosc.* **2000**, *27*, 321.
(11) Grant, D. M., Harris, R. K., Eds. *Encyclopedia of Nuclear Magnetic Resonance*; Wiley: Chichester, UK, 1996; Vols. 1−8.
(12) Cordier, F.; Grezeseik, S. *J. Am. Chem. Soc.* **1999**, *121*, 1601.

efforts were made to determine the $^{3h}J(NC')$ constants across the C═O···H−N units in the protein ubiquitin.[12,13] Ubiquitin plays a central role in protein degradation[14−26] and is presently discussed in connection with neurodegenerative diseases.[27,28] Furthermore, due to its relatively small size (just 76 residues), it is one of the most well described proteins. Its structure was investigated by X-ray diffraction[29] and by NMR spectroscopy,[12,13,15,16,30−33] and as such it is an excellent target for exploring the nature of the HB in proteins (covalent or electrostatic) utilizing the measured trans-HB SSCCs and comparing them with the corresponding calculated SSCCs. In this way, the geometrical dependence of the trans-HB SSCCs can be determined and the transmission mechanism of spin coupling across the HB investigated. Important steps in this direction were made by Bagno[22] and Barfield[23] who determined largely the geometrical dependence of $^{3h}J(NC')$ in ubiquitin by utilizing the X-ray structure and describing the immediate environment of a given HB in ubiquitin with suitable models (formamide dimers, residue pairs, etc.). These authors could reproduce the measured $^{3h}J(NC')$ with surprising accuracy, which is astonishing due to the fact that the approach used suffered from various calculational restrictions.

Any quantum chemical description of the SSCCs of ubiquitin has to consider the following points:

(1) Previous calculations of the SSCCs of ubiquitin were exclusively done using finite perturbation theory (FPT)[34] in connection with density functional theory (DFT)[35] to calculate the Fermi contact (FC) term of the SSCC only, which was expected to dominate the value of $^{3h}J(NC')$. FPT-DFT is a numerical method and is not able to provide a full account of the indirect scalar SSCC, which according to Ramsey[36] is the result of four rather than just one coupling mechanism, namely, besides the FC mechanism, also the diamagnetic spin−orbit (DSO), paramagnetic spin−orbit (PSO), and spin dipole (SD) transferring mechanism. Therefore, it is preferable to replace FPT-DFT by an analytical DFT method that covers all Ramsey terms.[37,38]

(2) Previous investigations[22,23] focused primarily on the measured $^{3h}J(NC')$ of ubiquitin,[12,13] although other trans-HB SSCCs, even if they have not been measured, are needed to complete the description of HBs in ubiquitin. In addition, it is important to investigate those SSCCs associated with a HB system that may provide important information on the nature and the geometry of HBs in ubiquitin. For example, most of the $^{1}J(NH)$ SSCCs were measured[30] and their comparison with calculated values may lead also to important insights into H-bonding in ubiquitin.

(3) Both the $^{3h}J(NC')$ and $^{1}J(NH)$ SSCCs of ubiquitin were measured in aqueous solution, while the previous quantum chemical investigations of the $^{3h}J(NC')$ constants were based on the solid-state structure of ubiquitin.[29] Although it was shown that the secondary structure of the protein does not change from solution to the crystal phase, changes in the geometry of the backbone and the interresidue hydrogen bridges are indeed likely. Recent refinement work on the structure of ubiquitin carried out for both the crystal state and the solution phase by Linge and co-workers[39,40] clearly demonstrates this point. Therefore, any quantum chemical investigation should start from their structures to explicitly consider the solvent effects.

(4) Vibrational effects can have a strong influence on SSCCs,[41] especially if large amplitude vibrations associated with conformational changes of a protein backbone are involved. Markwick, Sprangers, and Sattler[42] demonstrated that the simulation of dynamic effects corresponding to conformational movements of the protein backbone and the calculation of conformationally averaged $^{3h}J(NC')$ values leads to characteristic lowering of their magnitude and a better agreement between theory and experiment in the case of the SMN Tudor domain.

(5) Recent work by Juranic and co-workers[15,16] indicates that trans-HB SSCCs are sensitive to the extended environment of an HB system, including several peptide groups. Therefore, quantum chemical investigations simulating only the immediate HB environment with the help of a formamide dimer as done in one of the previous investigations[23] miss this influence and therefore may lead to erroneous trans-HB SSCCs.

In view of points 1−5, it seems that the agreement between calculated and measured $^{3h}J(NC')$ values as obtained indepen-

(13) Cornilescu, G.; Hu, J.-S.; Bax, A. *J. Am. Chem. Soc.* **1999**, *121*, 2949.
(14) Cornilescu, G.; Ramirez, B. E.; Frank, M. K.; Clore, G. M.; Gronenborn, A. M.; Bax, A. *J. Am. Chem. Soc.* **1999**, *121*, 6275.
(15) Juranic, N.; Macura, S. *J. Am. Chem. Soc.* **2001**, *123*, 40099.
(16) Juranic, N.; Moncrieffe, M. C.; Likic, V. A.; Prendergast, F. G.; Macura, S. *J. Am. Chem. Soc.* **2002** *124*, 14221.
(17) Dingley, A. J.; Cordier, F.; Grezeseik, S. *Concepts Magn. Res.* **2001**, *13*, 103.
(18) (a) Scheurer, C.; Brüschweiler, R. *J. Am. Chem. Soc.* **1999**, *121*, 8661. (b) Case, A.; Scheurer, C.; Brüschweiler, R. *J. Am. Chem. Soc.* **2000**, *122*, 10390.
(19) Benedict, H.; Shenderovich, I. G.; Malkina, O. L.; Malkin, V. G.; Denisov, G. S.; Golubev, N. S.; Limbach, H. H. *J. Am. Chem. Soc.* **2000**, *122*, 1979.
(20) Arnold, W. D.; Oldfield, E. *J. Am. Chem. Soc.* **2000**, *122*, 12835.
(21) Wilkens, S. J.; Westler, W. M.; Weinhold, F.; Markley, J. L. *J. Am. Chem. Soc.* **2002**, *121*, 1601.
(22) Bagno, A. *Chem. Eur. J.* **2000**, *6*, 2925.
(23) Barfield, M. *J. Am. Chem. Soc.* **2002**, *124*, 4158.
(24) Peters, J.-M. *Ubiquitin and the Biology of the Cell*; Plenum Publishing Co.: New York, 1998.
(25) Zwickl, P.; Baumeister, W. *Proteasome-Ubiquitin Degradation Pathway*; Springer: New York, 2002.
(26) Ciechanover, A. J., Masucci, M. G., Eds. *Recent Advances in Human Biology*; World Scientific: New York, 2003.
(27) Klimascheweski, L. *News Physiol. Sci.* **2003**, *18*, 29.
(28) Ehlers, M. D. *Trends Neurosci.* **2003**, *26*, 4.
(29) Vijay-Kumar, S.; Bugg, C. E.; Cook, W. J. *J. Mol. Biol.* **1987**, *194*, 531.
(30) Tjandra, N.; Grzesiek, S.; Bax, A. *J. Am. Chem. Soc.* **1996**, *118*, 6264.
(31) (a) Ottiger, M.; Bax, A. *J. Am. Chem. Soc.* **1997**, *119*, 8070. (b) Cornilescu, G.; Marquardt, J. L.; Ottinger, M.; Bax, A. *J. Am. Chem. Soc.* **1998**, *120*, 6836.
(32) (a) Hennig, M.; Bermel, W.; Schwalbe, H.; Griesinger, C. *J. Am. Chem. Soc.* **2000**, *122*, 6268. (b) Peti, W.; Hennig, M.; Smith, L. J.; Schwalbe, H. *J. Am. Chem. Soc.* **2000**, *122*, 12017.
(33) Kloiber, K.; Konrat, R. *J. Am. Chem. Soc.* **2000**, *122*, 12033.
(34) (a) Pople, J. A.; McIver, J. W.; Ostlund, N. S. *J. Chem. Phys.* **1968** *49*, 2960. (b) Kowalewski, J. *Ann. Rept. NMR Spectrosc.* **1982**, *12*, 81. (c) Fukui, H. *J. Chem. Phys.* **1976**, *65*, 844. (d) Cuevas, G.; Juaristi, E., Vela, A. *J. Phys. Chem. A* **1999**, *103*, 932.
(35) (a) Kohn, W.; Sham, L. *Phys. Rev. A* **1965**, *140*, 1133. For reviews on DFT, see for example: (b) Parr, R. G.; Yang, W. *International Series of Monographs on Chemistry 16: Density-Functional Theory of Atoms and Molecules*; Oxford University Press: New York, 1989. (c) Labanowski, J. K., Andzelm, J. W., Eds. *Density Functional Methods in Chemistry*; Springer: Heidelberg, 1990. (d) Seminario, J. M., Politzer, P., Eds. *Theoretical and Computational Chemistry*;; Elsevier: Amsterdam, 1995. (e) Laird, B. B., Ross, R. B., Ziegler, T., Eds., *Chemical Applications of Density Functional Theory*; ACS Symposium Series 629; American Chemical Society: Washington, DC, 1996. (f) Chong, D. P., Ed. *Recent Advances in Computational Chemistry*; World Scientific: Singapore, 1997; Vol. 1, Part II.
(36) Ramsey, N. F. *Phys. Rev.* **1953**, *91*, 303.
(37) Sychrovský, V.; Gräfenstein, J.; Cremer, D. *J. Chem. Phys.* **2000**, *113*, 3530.
(38) Helgaker, T., Watson, M., Handy, N. C. *J. Chem. Phys.* **2000**, *113*, 9402.
(39) Linge, J. P.; Williams, M. A.; Spronk, C. A. E. M.; Bonvin, A. M. J. J.; Nilges, M. *Proteins: Struct., Funct., Genet.* **2003**, *50*, 496.
(40) Spronk, C. A. E. M.; Linge, J. P.; Hilbers, C. W.; Vuister, G. W. *J. Biomol. NMR* **2002**, *22*, 281.
(41) See, e.g.: (a) Jordan, M. J. T., Toh, J. S., DelBene, J. *Chem. Phys. Lett.* **2001**, *346*, 288. (b) Ruden, A. T.; Lutnas, O. B.; Helgaker, T.; Ruud, K. *J. Chem. Phys.* **2003** *118*, 9572 and references therein.
(42) Markwick, P. R. L.; Sprangers, R.; Sattler, M. *J. Am. Chem. Soc.* **2003**, *125*, 644.

dently by Bagno[22] and by Barfield[23] for ubiquitin is due to a fortunate cancellation of errors. Better theory is now available to calculate the SSCCs associated with a HB system in such a way that all SSCCs associated with a HB system can be reliably obtained. This current work critically reviews the choice of theory and model on the evaluation of the SSCCs of ubiquitin as well as the effect of the aforementioned environmental factors. The coupling mechanism for the SSCCs of the HB system is analyzed, and two different coupling mechanisms are discussed in view of possible covalent or electrostatic H-bonding. Diagrams and three-dimensional representations are given that make an interpretation of measured SSCCs in terms of the stereochemistry of the HB possible. A relationship between the value of the NC′, and other SSCCs, and the geometry of the system is established. This is then utilized for the determination of the position of the H atom in the HB system.

2. Computational Methods

Geometry optimizations were carried out with DFT[35] employing the hybrid functional B3LYP[43] and Pople's 6-31G(d,p) basis set.[44] Additional geometry optimizations were performed with second-order many-body perturbation theory with the Møller−Plesset perturbation operator (MBPT2)[45] using Dunning's aug-cc-pVDZ basis set.[44] This was necessary to verify the strength of the HB. In some cases, optimizations were performed for both the gas phase and in aqueous solution. To examine the effect of a water solvent on hydrogen bonding, the PISA continuum model (PCM)[47] was used employing the dielectric constant $\epsilon(H_2O) = 78.2$ at 298 K.[48] B3LYP and MBPT2 binding energies were determined correcting in all cases for basis set superposition errors (BSSE) with the help of the counterpoise method.[49]

The calculation of the indirect scalar SSCC J values involving the nuclei of the HB system were computed as the sum of the four Ramsey terms[36] (FC, DSO, PSO, SD) that comprise the coupling constant. This was done using coupled perturbed DFT (CP-DFT) as described by Cremer and co-workers,[37] utilizing the B3LYP functional[43] and two basis sets, (9s,5p,1d/5s,1p)[6s,4p,1d/3s,1p] and (11s,7p,2d/5s,1p)[7s,-6p,2d/4s,2p], which have been developed for the calculation of magnetic properties.[50] SSCC calculations were done for both gas-phase geometries and PCM geometries obtained for the aqueous solution. In this way, the geometry effect on the SSCC caused by the solvent was determined. In addition, the PCM solution of the wave function was used to calculate the SSCC, which made it possible to determine the change in the SSCC caused by the different electronic structure of the target molecule in aqueous solution (as compared to the gas phase).

For the investigation of hydrogen bonding in ubiquitin, both the crystallographic structure[29] and the water-refined structure[39] of this protein were utilized. The residue pairs investigated in this report were generated by using the heavy atom coordinates and then determining the unknown H position by B3LYP/6-31G(d,p) optimizations. Since the whole protein cannot be calculated, which is also not necessary, the extended chemical environment of a HB system and its impact on the SSCC $^{n,h}J$ was modeled in several different ways. The most elaborate model (model 1) considered all residues directly or indirectly involved into a particular C═O···H−N interaction as shown in Figure 1a for the HB between donor residue 32 and acceptor residue 28 (32−28). In case of the crystal structure,[29] the conformation of the backbone was taken from experiment with the positions of the H atoms optimized at B3LYP/6-31G(d,p). The interacting residues were further simplified by removal of the side chains, while the terminal positions were saturated by conversion into a NH_2 or CH_3 group. In connection with model 1, it was also investigated whether a zwitterionic structure with a terminal $-NH_3^+$ and a terminal $-COO^-$ group has an impact on the $^{n,h}J$ values. It was found that the $^{n,h}J$ values change considerably, leading to a relatively large deviation from measured values. Therefore, a zwitterionic model was no longer considered.

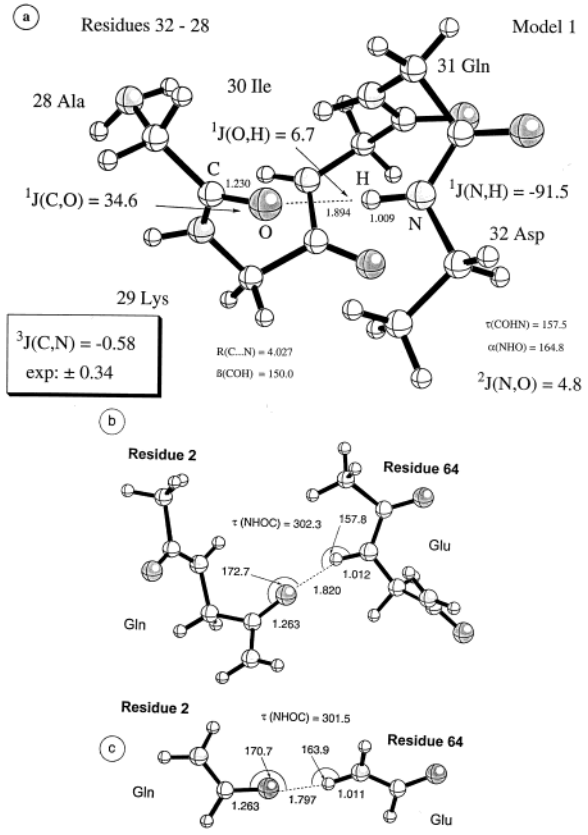

Figure 1. (a) H-bonding between residues 32 (donor) and 28 (acceptor) of the protein ubiquitin (model 1). (b) H-bonding between residues 64 and 2 according to model 2 made up of glycine residues. (c) H-bonding between residues 64 and 2 according to model 3 made up of two formamide molecules. Distances in Å, angles in degrees, and spin−spin coupling constants in Hz. Experimental value was taken from ref 12.

Model 1 led to calculations for 19 heavy atoms and about 323 basis functions (in the case of the SSCCs: 551 basis functions), which will be feasible, however, not advisable if done on a routine basis. Also, model 1 becomes more and more problematic when the interacting residues are separated by a larger number of intermediate residues. In this case, model 1 is simplified by keeping just the two H-bonded residues, again in the conformation taken from the ubiquitin crystal structure, but deleting all connecting intermediate residues. In this way it is possible to model the interaction between residues 2 and 64 as shown in Figure 1b (model 2). It was investigated whether a variation of the terminal groups (either as NH_2/C(═O)CH_3 or as $NHCH_3$/

(43) (a) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648. (b) Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098. (c) Lee, C.; Yang, W.; Parr, R. P. *Phys. Rev. B* **1988**, *37*, 785.
(44) Hariharan, P. C.; Pople, J. A. *Theor. Chim. Acta* **1973**, *28*, 213.
(45) For a recent review see: Cremer, D. *Encyclopedia of Computational Chemistry*; Schleyer, P.v. R., Allinger, N. L., Clark, T., Gasteiger, J., Kollman, P. A., Schaefer, H. F., III, Schreiner, P. R., Eds.; Wiley: Chichester, 1998; Vol. 3, p 1706. See also: Møller, C.; Plesset, M. S. *Phys. Rev.* **1934**, *46*, 618.
(46) Dunning, T. H. Jr. *J. Chem. Phys.* **1989**, *90* 1007.
(47) (a) Miertus, S.; Scrocco, E.; Tomasi, *J. Chem. Phys.* **1981**, *55*, 117. (b) Barone, V.; Cossi, M.; Tomasi, J. *J. Chem. Phys.* **1997**, *107*. 3210. (c) Cammi, R.; Cossi, M.; Tomasi, J. *J. Chem. Phys.* **1996**, *104*, 4611. (d) Mennucci, B.; Tomasi J. *J. Chem. Phys.* **1997**, *106*, 5151. (e) Tomasi J.; Mennucci, B. In *Encyclopedia of Computational Chemistry*; Schleyer, P.v. R., Allinger, N. L., Clark, T., Gasteiger, J., Kollman, P. A., Schaefer, H. F., III, Schreiner, P. R.; Wiley: Chichester, 1998: Vol. 1, p 2547.
(48) *CRC Handbook of Chemistry and Physics on CD-ROM* 2000 Version. Lide, D. R. Ed. CRC Press LLC, 2000.
(49) Boys, F.; Bernardi, F. *Mol. Phys.* **1970**, *19*, 553.
(50) Kutzelnigg, W.; Fleischer, U.; Schindler, M. In *NMR−Basics, Principles and Progress*; Springer: Heidelberg, 1990; Vol. 23, p 165.

C(═O)CH_3) plays a role; however, effects are rather small so that, in all cases considered, the simpler model can be taken.

The strongest simplification is chosen for the third model (model 3), in which the two residues of model 2 are simulated by two formamide molecules (Figure 1c) yielding a formamide dimer with a HB geometry typical of the residue pair of ubiquitin to be simulated. As shown in Figures 1b and 1c, further simplification of the model leads to a small change in the H position (as expressed by the H···O distance and the NHO and the COH angle), although the heavy atoms are exactly kept in the positions they would have in ubiquitin according to its crystallograhpic structure.[29] These small changes are a result of the changes in the environment of the HB accompanying the simplification of the model. We note in this connection that in the previous investigations of the trans-HB SSCC, model 3 was exclusively used[23] or a slightly modified form of model 2.[22]

The effect of the aqueous environment on the structure of ubiquitin was examined using the explicitly solvated MDS (molecular dynamics simulation) generated structure of Linge and co-workers.[39] From this structure the residues were modulated using the same simplifications as chosen in model 3.

The effects of the three methodological variables (method, XC functional, basis set) were investigated by their systematic variation. (1) The effect of the type of theory, CP-DFT or FPT-DFT, was evaluated by comparison of the SSCC that results from each method. The FPT calculations were performed by the application of a perturbation to each of the coupling nuclei of a particular H-bond in the residue pairs. The B3PW91[43,51] functional and the 6-311G(d,p) basis[52] were applied with a tight SCF convergence criteria (changes in the density matrix $<10^{-8}$). To determine the effect of the magnitude of the perturbation constant, this calculation was performed twice, once with the perturbation constant, λ, set to 0.0025 and second with λ set to 0.025.[53] The calculation was then carried out using CP-DFT with the same basis set and XC functional applied in the FPT case. (2) The effect of the XC functional was then examined by comparing the SSCCs using CP-DFT first with the B3PW91 functional and subsequently with the B3LYP functional. Both calculations operated with the 6-311G-(d,p) basis. (3) Finally, the effect of the basis set was investigated by comparing results of the CP-DFT/B3LYP method in combination with basis sets [7s,6p,2d/4s,2p], [6s,4p,1d/3s,1p],[50] and 6-311G(d,p).[52]

Independent formamide dimers were used to determine the structural dependence of all SSCCs connected with a HB system on the HB-geometry. The calculation of SSCCs for this system were carried out using CP-DFT, the B3LYP functional, and the [6s,4p,1d/3s,1p] basis. The dependence of the SSCC on the COH and NHO angles, the NC′ distance, and the COHN dihedral angle was investigated by variation of the HB geometry and evaluation of the consequent change in the calculated value of the HB−SSCCs.

All SSCC calculations were performed with COLOGNE 2002,[54] while for the geometry optimizations, Gaussian 98[55] was used.

(51) Perdew, J. P.; Wang, Y. *Phys. Rev. B.* **1992**, *45*, 13244.
(52) Krishnan, R.; Binkley, J. S.; Seeger, R.; Pople, J. A. *J. Chem. Phys.* **1980**, *72*, 650.
(53) Peralta, J. E.; Ruiz de Azua, M. C.; Contreras, R. H. *Theor. Chem. Acc.* **2000**, *105*, 165.
(54) Kraka. E.; Gräfenstein; J.; Filatov, M.; He, Y.; Gauss, J.; Wu, A.; Polo, V.; Olsson, L.; Konkoli, Z.; He. Z.; Cremer D. *COLOGNE 2003*; Göteborg University: Göteborg, 2003.
(55) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Zakrzewski, V. G.; Montgomery, J. A., Jr.; Stratmann, R. E.; Burant, J. C.; Dapprich, S.; Millam, J. M.; Daniels, A. D.; Kudin, K. N.; Strain, M. C.; Farkas, O.; Tomasi, J.; Barone, V.; Cossi, M.; Cammi, R.; Mennucci, B.; Pomelli, C.; Adamo, C.; Clifford, S.; Ochterski, J.; Petersson, G. A.; Ayala, P. Y.; Cui, Q.; Morokuma, K.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Cioslowski, J.; Ortiz, J. V.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Gonzalez, C.; Challacombe, M.; Gill, P. M. W.; Johnson, B. G.; Chen, W.; Wong, M. W.; Andres, J. L.; Head-Gordon, M.; Replogle, E. S.; Pople, J. A. *Gaussian 98*, revision A.9; Gaussian, Inc.: Pittsburgh, PA, 1998.

Table 1. Comparison of SSCCs Calculated with CP-DFT and FPT-DFT for Residue Pair 4-65 from the Ubiquitin Protein[a]

SSCC	CP-FC	CP-total	FP-$\lambda = 0.0025$	FP-$\lambda = 0.025$
N···C′	−0.79	−0.84	−0.96	−0.88
N·H	−83.40	−85.35	−81.84	−85.52
N···O	5.86	5.85	5.52	5.76
O···H	7.29	7.57	3.76	6.38
C···H	−1.14	−0.86	6.80	−0.13
C·O	19.71	30.96	37.20	20.45

[a] B3PW91/6-311G(d,p) calculations. All SSCCs are given in Hz. CP-DFT: coupled perturbed DFT. FPT-DFT: finite perturbation theory DFT.

3. Methodological Considerations

In the following, only an extract of the results of this work are given to illustrate the most important results. As far as methodological questions are considered, these are discussed for residue-pair 4-65, which is a typical example from a β-sheet of ubiquitin. In other cases, we consider just three typical examples from the α part, two from the β-sheets, and two from the irregular (random coil) parts to test the flexibility of ubiquitin folding and its impact on the trans-HB SSCCs.

The ability to calculate all of the terms that contribute to the SSCC provides the most compelling argument for the choice of CP-DFT over FPT-DFT in the calculation of SSCCs. However, the results obtained by Barfield,[23] using FPT at B3PW91/6-311G(d,p) and calculating just the FC term under the assumption that the DSO, PSO, and SD terms are negligible provide ^{3h}J(NC′) estimates that are closer to measured ^{3h}J(NC′) than the CP-DFT results to be presented in this work. Because of the fact that the performance of CP-DFT is superior to that of FPT-DFT,[37,38] this result suggests some fortuitous error cancellation in the case of the FPT SSCCs, which may disguise important features of the HB systems in ubiquitin. We will show that this is indeed the case. For this purpose, a comparison of the relative performance of the two methods is carried out to determine the effect of the choice of theory on the calculated value. The results of this comparison are shown for residue 4-65 in Table 1.

These results highlight two important points of consideration in the choice of theory. First, there is indeed only little variation between the FC term and the total SSCC ^{3h}J(NC′). However, for the other SSCC of the HB system, this assumption no longer holds as the magnitude of the DSO, PSO, and SD terms depends on the type of nucleus, the availability of π and/or lone pair electrons, etc. Hence, the variation between the CP and FPT results is not consistent, and as such, there cannot be a systematic error or cancellation of errors that consistently alters the calculated values. The randomness in the errors clearly suggests that the choice of the more complete CP method is the most appropriate choice. Second, the variation of the SSCC with the value of λ in FPT is further disconcerting, as this appears to, once again, cause an inconsistent variation in the derived results. Pople and co-workers[34a] suggested that the value of λ should be in the area of 10^{-3} to produce the most accurate FC values, which was confirmed by Contreras and co-workers.[53] However, the above results show that the larger value of λ produces results that are more consistent in their agreement with the FC term obtained at the CP-DFT level of theory. It has been shown[53] that the FPT result for FC depends not only on the choice of λ but also on the choice of the nucleus perturbed. In addition, an adjustment of the convergence criteria of the SCF procedure in

Table 2. Comparison of SSCCs Obtained with CP-DFT/B3LYP and the Three Basis Sets (11s,7p,2d/5s,1p)[7s,6p,2d/4s,2p], (9s,5p,1d/5s,1p)[6s,4p,1d/3s,1p], and 6-311G(d,p) for Residue Pair 4-65 from the Ubiquitin Protein[a]

SSCC	(11s,7p,2d/5s,1p)[7s,6p,2d/4s,2p]	(9s,5p,1d/5s,1p)[6s,4p,1d/3s,1p]	6-311G(d,p)
N···C′	−0.99	−1.01	−0.92
N·H	−95.14	−94.20	−90.45
N···O	6.61	6.99	6.31
O···H	8.10	8.25	7.97
C···H	−0.85	−0.96	−0.93
C·O	27.44	30.15	28.95

[a] All SSCCs are given in Hz.

DFT to the actual λ value is absolutely necessary (the smaller λ requires the tighter convergence criteria).

The use of FPT in the calculation of SSCCs will be a potentially misleading practice if not carried out after determining the appropriate λ value in test calculations for each SSCC. The inability to predict the DSO, PSO, and SD terms can produce inaccurate results in any case where the FC term is not dominant. In addition, the empirical nature of choosing the correct λ is problematic. Barfield, when correctly using $\lambda = 0.0025$ (in view of refs 34a and 52), obtained less accurate values than could have been calculated with $\lambda = 0.025$. Although this point is partially disguised by the small magnitude of ^{3h}J(NC′) (exclusively calculated in ref 23), it represents a serious source of errors for FPT-DFT SSCCs. In many publications based on FPT, this is not considered at all, and dubious (if not erroneous) FC values have been published, which are used as estimates for the actual SSCC. Even when these values may agree with measured *J*-values, this agreement occurs for the wrong reason and can disguise structural or electronic effects on the SSCC. Therefore, the prediction of SSCCs should be carried out using the more complete and analytical CP-DFT method in order to decrease the number of unknown variables present in the problem.

Effect of the XC Functional. Previous investigations[37] into the effect of the XC functional on the calculation of the SSCC have shown that in the gas phase, the B3LYP hybrid functional performs better than the B3PW91 functional. However, in the calculation of ^{3h}J(NC′) for ubiquitin, results closer to the experimentally observed values were obtained[23] using the B3PW91 functional. Thus, to determine how the chosen functional was affecting the SSCCs, a sample calculation was performed on residue 4-65, with CP-DFT, comparing the two XC functionals (see Tables 1 and 2).

B3PW91 gives a lowered value of the absolute magnitude for all SSCCs of the HB system except the C′O coupling constant. This lowering leads to a deterioration in the agreement between theory and experiment in the case of ^{1}J(NH). Typical values for this parameter in ubiquitin are in the range −93.5 to −96.5 Hz[30] (SSCCs for the NH bond anti to the C═O bond are larger in magnitude than those syn to the C═O group[56]). Hence, the B3LYP result is clearly in better agreement with experiment than the B3PW91 value. Also, the few data known for ^{1}J(C═O) SSCCs (e.g., acetone: 22 ± 4 Hz[57]) favor the B3LYP rather than the B3PW91 value of ^{1}J(C═O) (see Table 1). These observations are in line with previous findings, namely, that B3LYP provides more accurate results.[37] Thus, despite the fact that for the ^{3h}J(NC′) SSCC the B3PW91 values are closer to experiment than the B3LYP value, this agreement arises from a fortunate error cancellation in a singular case, rather than a more accurate description of the coupling interaction by the functional. As such, a consistent description of SSCCs requires the use of B3LYP rather than B3PW91.

Effect of the Basis Set. SSCCs are second-order properties, and therefore their calculation implies much higher demands on the basis set than in the case of first-order properties. In particular, basis sets for SSCCs require an improved description of the core region (to obtain reliable FC values) and decontracted pd-basis sets (to obtain more accurate PSO and SD terms).[58] Standard basis sets used for energy and geometry calculations are not sufficient for reliable SSCC calculations. This is the reason basis sets [6s,4p,1d/3s,1p] and [7s,6p,2d/4s,2p] designed for the calculation of magnetic properties[49] were used in this work. However, the previous calculations on ^{3h}J(NC′) SSCCs in ubiquitin[22,23] utilized the 6-311G(d,p) basis set, i.e., a standard basis for energy calculations, which is not necessarily suited for SSCC calculations. Thus, a comparison of the performance of the various basis sets was carried out to determine what effect the choice of the basis set has on the calculation of the SSCC of a HB system (see Table 2).

This comparison for residue 4-65 shows that the effect of using the less accurate 6-311G(d,p) basis set is to lower the absolute magnitude, in all cases, of the SSCC, which is particularly problematic for the SSCC ^{1}J(NH) (absolute value becomes too small, Table 2). Consequently, the ^{3h}J(NC′) SSCCs produced by the smaller basis set artificially appear closer to the experimental value as a result of a systematic error in the performance of the basis set rather than a more accurate description of the coupling interaction. Reliable and consistent descriptions of the SSCCs can only be obtained by larger basis sets designed for the calculation of magnetic properties rather than standard basis sets.

Determination of SSCCs in Ubiquitin Residue Pairs. Through analysis of the results presented above, we have concluded that the appropriate choice of the method required for the calculation of the SSCCs in the interresidue HB system is the combination of CP-DFT using a B3LYP hybrid functional and at least the [6s,4p,1d/3s,1p] basis set. As such these specifications were used to predict the spin−spin coupling interactions between the four atoms, which comprise a HB system between two residues of ubiquitin. CP-DFT enables the accurate calculation of the four Ramsey terms of the total SSCC. Each of these terms was found to contribute to varying degrees in each of the six SSCCs that were calculated. The percentage

(56) Berger, S.; Braun, S.; Kalinowski, H. O. *NMR—Spectroskopie von Nichtmetallen*; Thieme: New York, 1992; Band 2.
(57) Berger, S.; Braun, S.; Kalinowski, H.-O. *NMR Spectroscopy of the Non-Metallic Elements*; Wiley: Chichester, 1997.
(58) See, e.g.: (a) Helgaker, T.; Jaszunski, M.; Ruud, K. *Chem. Rev.* **1998**, *99*, 293. (b) Fukui, H. *Progress in NMR Spectroscopy* **1999**, *35*, 267.

Table 3. Relative Contribution of Each Ramsey Term to the SSCCs for Residue Pair 4-65 from the Ubiquitin Protein[a]

SSCC	%DSO	%PSO	%FC	%SD
N···C′	−1.63	5.29	90.28	2.80
N·H	0.30	1.31	98.11	0.28
N···O	0.24	−1.22	98.07	0.47
O···H	−6.51	12.25	78.85	−2.38
C···H	10.04	−21.97	63.03	−4.96
C·O	−0.31	39.31	56.45	−3.93

[a] CP-DFT/B3LYP/(9s,5p,1d/5s,1p)[6s,4p,1d/3s,1p] calculations. The percentage is the percentage contribution of the individual term to the sum of the absolute values of each term.

contribution of each term, for the representative residue 4-65, is given in Table 3.

The above discussion reveals that the agreement previously obtained[22,23] for the value of ^{3h}J(NC′) resulted from a cancellation of errors due to the choice of a numerical rather than an analytical method, the choice of an inappropriate functional and a basis set that was too small, the dominance of the FC term in the NC′ coupling interaction, and the neglect of solvent and vibrational effects on the SSCC. However, all the other SSCCs of the HB system are much more sensitive to these effects and therefore reveal the severe shortcomings of the theory used in previous work on the SSCCs of ubiquitin. For example, the ^{1h}J(OH), ^{2h}J(C′H), and ^{1}J(C′O) values all contain substantial contributions from all SSCC terms. Notably, the PSO term comprises ca. 40% of the total magnitude of the C′O coupling interaction. Accordingly, the mere calculation of the FC term and the neglect of the other three terms will lead to an inaccurate description of these coupling interactions. The determination of the relative contribution of the four terms to the SSCCs allows both the accurate determination of the magnitude as well as insight into the mechanism by which the coupling occurs.

4. Choice of a Suitable Model

Currently, restrictions resulting from computational expense dictate the requirement of models to represent the interacting residues in the ubiquitin protein. Thus, a model for the individual HB residues in ubiquitin must be identified that will recreate the chemical environment experienced in the protein while at the same time remain computationally inexpensive. A previous investigation[18] has suggested that the chemical environment of the coupled nuclei may affect the value of the resulting coupling constant.

Model 1 was found to lead to reasonable SSCCs across the HB (Figure 1a), although the absolute magnitude of ^{3h}J(NC′) is twice as large as the measured value. Since the applicability of model 1 is strongly limited, we concentrate on models 2 and 3 in the following. In Table 4, some energetic and geometrical parameters of models 2 and 3 are listed (including the effects of the water refinement). There is a clear difference in the complex binding energies for the two models. Model 2, which gives a more accurate description of the environment of the HB leads to a larger (absolute) value of the binding energy (6.5−8.9 kcal/mol) than model 3 (5.2−6.2 kcal/mol, Table 4). This is simply a result of the fact that two amino acid residues have more nonbonded (attractive) interactions than two formamide molecules. However, the difference in the models causes only a small difference in the H position of the HB because in both models the heavy-atom framework remains the same. This is confirmed by all geometrical parameters involving the H atom of the HB bridge (Table 4).

There is a stronger change when considering the influence of the solvent water. All R(C···N) distances become shorter (up to 0.26 Å, Table 4), most of the β angles decrease significantly, and there are substantial changes in the τ values. The force field used in the MDS calculations lead to a shortening of most of the N−H and C═O bond lengths so that most of the R(O···H) values (apart from 64-2 and 23-54, Table 4) become longer. These geometrical changes lead to a significant decrease in the HB binding energies, which in addition vary now between 2.2 and 5.2 kcal/mol (Table 4).

With regard to the SSCCs obtained for models 2 and 3, the

Table 4. Comparison of the Complexation Energies and the Geometrical Parameters of the HB System Using Model 2, Model 3, and the Water Structure[a]

residues	ΔE	R(C···N′)	R(C═O)	R(O···H)	R(N–H)	β(COH)	α(NHO)	τ(COHN)	Sec[b]
				model 3					
64-2	−5.18 (−6.62)	4.022	1.263	1.797	1.011	170.7	163.9	−58.5	β
72-40	−5.24 (−7.07)	3.813	1.226	1.751	1.015	139.8	158.7	−178.3	rc
4-65	−5.96 (−7.56)	4.047	1.226	1.847	1.014	162.8	173.1	−114.3	β
27-23	−6.17 (−7.49)	4.170	1.242	1.986	1.015	156.6	174.5	101.6	α
33-29	−5.59 (−7.13)	4.064	1.255	1.911	1.015	149.8	167.2	129.4	α
23-54	−5.98 (−7.59)	4.093	1.251	1.892	1.013	170.6	159.4	−86.7	rc
				model 2					
64-2	−6.90 (−8.34)	4.022	1.263	1.820	1.012	172.7	157.8	−57.6	β
72-40	−7.09 (−8.92)	3.813	1.226	1.739	1.019	140.6	160.7	−172.4	rc
4-65	−7.45 (−9.05)	4.047	1.226	1.868	1.015	160.2	163.6	−133.7	β
27-23	−8.87 (−10.19)	4.170	1.242	1.984	1.015	156.3	177.6	141.6	α
33-29	−8.75 (−10.29)	4.064	1.255	1.904	1.015	150.2	170.3	133.8	α
23-54	−6.50 (−8.11)	4.093	1.251	1.876	1.013	169.6	163.8	−85.6	rc
				water					
64-2	−4.27 (−6.05)	3.891	1.229	1.751	0.985	157.8	173.7	−66.5	β
72-40	−2.57 (−4.38)	3.557	1.229	1.750	0.984	125.2	159.7	97.7	rc
4-65	−5.23 (−6.69)	4.022	1.231	1.873	0.983	157.3	167.0	92.0	β
27-23	−4.09 (−5.89)	4.055	1.232	2.066	0.983	141.8	145.0	161.6	α
33-29	−2.23 (−3.69)	4.025	1.238	2.009	1.041	135.5	170.9	−102.0	α
23-54	−4.92 (−6.63)	3.999	1.228	1.835	0.985	173.7	162.6	−9.2	rc

[a] Energy given in kcal/mol includes BSSE corrections; numbers in parentheses are energies without BSSE correction. Distances are given in Å, and angles are given in degrees. [b] Secondary structure (Sec) is given in the following way: β = β-sheet; α = α-helix; rc = random coil. The aqueous solution structure of ubiquitin ("water") was taken from ref 39.

Table 5. Comparison of SSCCs Obtained with Model 2 and Model 3 with Measured Values[a]

residue pair	^{1}J(NH)	^{1}J(CO)	^{1h}J(OH)	^{2h}J(NO)	^{2h}J(CH)	^{3h}J(NC′)	^{3h}J(NC′) water	^{1}J(NH)[b] exp	^{3h}J(NC′)[c] exp
64-2	−90.91	37.90	9.59	7.76	−1.25	−1.35		(±) 93.518	(±) 0.8
	(−92.84)	(39.55)	(9.83)	(8.83)	(−1.28)	(−1.49)	−1.09		
72-40	−88.89	29.54	9.80	8.21	−0.97	−0.91			(±) 0.3
	(−88.37)	(30.84)	(9.49)	(7.77)	(−0.81)	(−0.77)	−0.34		
4-65	−103.88	28.38	7.36	7.49	−0.82	−1.07		(±) 93.934	(±) 0.6
	(−94.20)	(30.15)	(8.25)	(6.99)	(−0.96)	(−1.01)	−0.88		
27-23	−90.98	34.00	5.67	4.37	−0.61	−0.61		(±) 94.388	(±) 0.5
	(−90.29)	(34.89)	(5.47)	(4.35)	(−0.57)	(−0.61)	−0.31		
33-29	−91.34	38.99	7.04	5.32	−0.85	−0.68		(±) 93.478	(±) ≤ 0.25
	(−87.42)	(39.19)	(6.72)	(4.93)	(−0.73)	(−0.59)	-0.00_2		
23-54	−93.88	33.58	7.90	6.32	−0.97	−1.05		(±) 94.164	(±) 0.5
	(−92.85)	(36.41)	(7.60)	(5.94)	(−0.87)	(−1.04)	−1.12		

[a] Model 3 values are given in parentheses. All SSCCs are in Hz. For the definition of model 2 and model 3, see text. [b] Experimental values are from ref 30. [c] Experimental values from are refs 12 and 13.

one-bond SSCCs show the largest dependence on the model used (see Table 5). We find deviations of 0.5−4 Hz and, in the case of ^{1}J(NH) of residue pair 4-65, even 9.7 Hz caused by a stronger change in the relative orientation of the two residues, respectively, of the molecules modeling them. For the three-bond SSCC, the changes are marginal.

The conclusion can be drawn that in all of these cases, where a full account of the SSCC of a HB system has to be given, a more extended geometrical model such as model 2 is clearly preferable. The best account of the SSCCs seems to be provided with model 1 (Figure 1a) despite the fact that in this case only a relatively small basis set can be used. The influence of the model becomes even more obvious when considering environmental factors. Utilizing the ubiquitin structure derived for an aqueous solution[39] changes all of the calculated ^{3h}J(NC′) SSCCs to smaller magnitudes in line with the fact that the β-angles become smaller. The water-refined structure, however, has the disadvantage that because of the force field used the NH (and CO) bond are too short, thus leading in the case of the ^{1}J(NH) SSCCs to a significant reduction of their magnitudes (all values between −87.8 and −88.9 Hz), as is typical of an increase in the NH bond strength. Similar observations can be made for the ^{1}J(CO) SSCCs. Clearly, the water-refined structure of ubiquitin has to be improved by using for the modeling of the HB systems DFT geometries while keeping the backbone geometries.

Effect of Vibrational and Conformational Averaging. A direct evaluation of vibrational and conformational averaging effects on J is beyond the scope of this work as it requires a further development of the current methodology. Vibrational averaging can have a strong effect on the measured SSCC, which then deviates from the calculated one.[41] However, in the case of proteins, conformational averaging has an even more important impact on measured trans-HB SSCCs, as was recently shown by Markwick, Sprangers, and Sattler.[42] The conformational changes of the protein backbone change the geometry of the HB system and lead to much smaller values of calculated ^{3h}J(NC′) SSCCs, which agree perfectly with the corresponding measured SSCCs for the protein investigated (SMN Tudor domain). In the following chapter, where we will discuss the geometrical dependence of the SSCCs associated with a HB, we will consider this effect in detail.

As for the role of vibrational averaging, we take a qualitative approach to describe this effect. Large changes in J due to vibrational averaging can be expected in the case of large amplitude vibrations guided by a strongly asymmetric HB potential. For example, if the NH stretching vibration is influenced by a flat potential on the side to a longer NH bond (as it could be in the case of a strong H bonding interaction), the r_z value of the NH bond length could be larger than the r_e value. Accordingly the vibrationally averaged SSCC ^{1}J(NH) and other SSCCs involving the bridging H atom could significantly differ from their equilibrium values.

For the purpose of determining if this averaging effect is present in the HB systems of ubiquitin, the SSCC and energy of the system was calculated at varying lengths of the NH bond, thus simulating a N−H stretching vibration in the formamide dimer (model 3). This simulation also provides an indication as to the possibility of the formation of a transition structure, in which the H atom (or proton) is shared between the N and O atom, forming a H-shared HB system. The results of these calculations are shown in Figure 2.

A stretching of the NH bond (similarly as a compression of this bond) leads to a strong increase in the energy. This has to do with the fact that the stretching vibration initiates a migration of a proton, which leads to a zwitterionic structure of much higher energy. As the H migrates across the HB, the absolute value of ^{3h}J(NC′) increases to |−8| Hz (Figure 2c), which is opposite to the trend in the measured ^{3h}J(NC′) values that are smaller than 1 Hz. Strong changes are also found for J(O,H), J(C,H) (Figure 2a), J(NH), J(CO) (Figure 2b), and J(NO) (Figure 2c). The changes in the J values provide suitable references for the discussion of the SSCCs of a HB system. The increase of $|^{3h}J(NC')|$, coupled with the high energy barrier that exists for the migration, suggests that vibrational effects will play a negligible role not only for ^{3h}J(NC′) but also for all other SSCCs associated with the HB system.

There is, however, the possibility that in aqueous solution, the situation changes because strong solvation effects might stabilize the zwitterionic form. Clearly, the formamide dimer is no longer a good model for describing this situation. One has to consider the folding of the protein ubiquitin and whether it provides sufficient space for the intercalation of water molecules. The X-ray diffraction structure[29] reveals that in the vicinity of a HB, there is no space left for additional water molecules that may increase the local dielectric constant ϵ to 78.2. It is more realistic to consider an ϵ value typical of proteins ($\epsilon \approx 4$) and to calculate the potential energy curve for N−H stretching under these conditions. As shown in Figure 2 (see dashed curve), there is little change in the potential. Thus, we regard the effects of vibrational averaging in general as

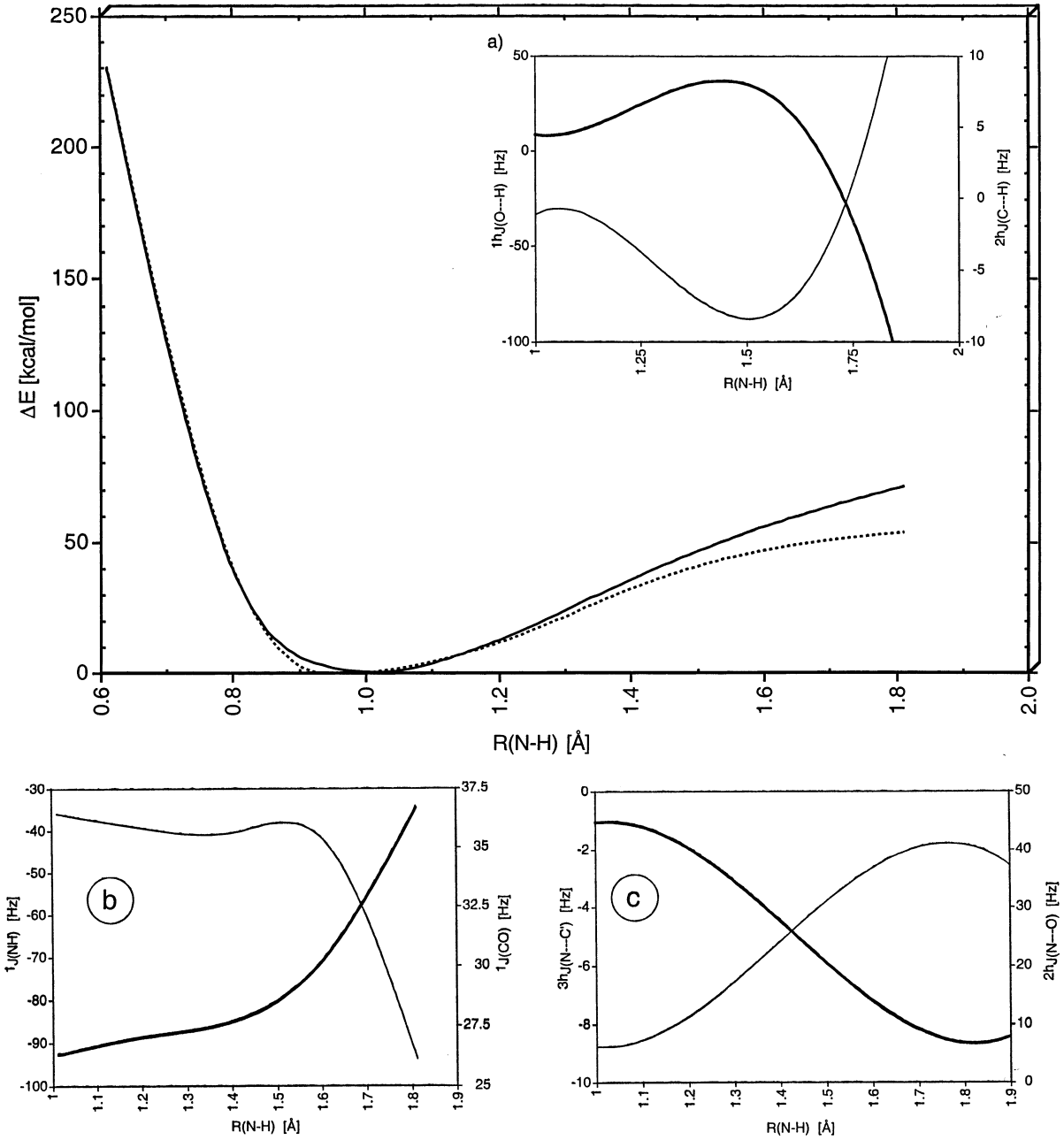

Figure 2. Energy change for a transfer of H from N to O in the formamide dimer. B3LYP/6-31G(d,p) calculations. The dashed curve gives the energy change in solution for a dielectric constant of 4. PISA/B3LYP/6-31G(d,p) calculations. The inserts give the corresponding changes in the values of (a) ^{1h}J(O···H) (bold line) and ^{2h}J(C···H), (b) ^{1}J(NH) (bold line) and ^{1}J(CO), (c) ^{3h}J(N···C′) (bold line) and ^{2h}J(N···O).

negligible and contend that conformational averaging is much more important.

5. Geometrical Dependence of the SSCCs of the Hydrogen Bond. The folding of a protein leads to different types of H-bonds: (a) σ-HBs: H-bonding that involves the σ-type lone-pair orbital, which requires that N−H and C═O bond are located in a common plane. (b) π-HBs: H-bonding that involves the π-type bond orbital of the C═O group, which implies a perpendicular approach of the N−H bond to the plane of the second O═C−NH group. (c) All intermediate arrangements can be considered as $\sigma-\pi$ HBs. These structures have been investigated before using MBPT2[45] with a cc-pVDZ basis set[46] to guarantee a reliable description of H-bonding.[59] In this work, these calculations were repeated to form a suitable model 3.

In Figure 3, the σ-HB and the π-HB dimers are shown. The former HB system leads to greater stability (BSSE-corrected complexation energy: 5.8 kcal/mol) than the latter (3.4 kcal/mol). As well as the optimized σ-HB dimer the planar σ-HB system (Figure 3b) was also investigated and found to have a binding energy of 4.3 kcal/mol. The π-HB system occupies a saddle point on the potential energy surface and is characterized by a cis structure (τ(NHOC) = 9.2°, Figure 3c). For the purpose of obtaining also a more open π-structure, the dihedral angle τ(NHOC) was increased to 180°, leading to the form shown in Figure 3d. In this case, the binding energy is just 3.1 kcal/mol.

The geometrical dependence of the SSCCs was quantified for the formamide dimer (model 3) by systematically varying parameters R(N···C′) between 3.5 and 4.4 Å and β between

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

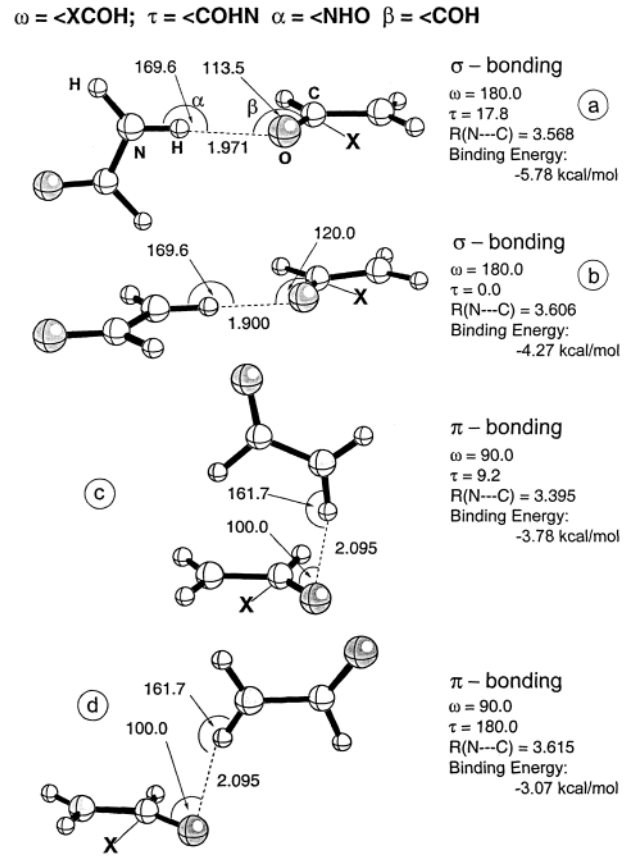

Figure 3. MBPT2/aug-cc-pVDZ geometries and binding energies for σ- and π-type H bonding in the formamide dimer. Distances in Å, angles in degree. (a) Minimum: σ-H bonding. (b) Enforced planarity: σ-type H bonding. (c) Enforced π-type H bonding: β fixed to 100° and ω to 90°; the rest are optimized. (d) Enforced π-type H bonding: β fixed to 100°, ω to 90°, and τ to 180°; the rest are optimized.

180 and 110° while keeping τ at its equilibrium value (see Figure 3). The angle ω helps to distinguish between a σ-HB (ω = 180°) and a π-HB (ω = 90°, Figure 3) and was varied between these extremes. In all geometry variations of the formamide dimer, the angle α (see Figure 3) was kept at its optimized value of 169.6°, which guarantees an almost linear approach of the NH bond toward the oxygen atom and, in turn, an optimal coupling mechanism.

In Figure 4, calculated SSCCs ^1J(NH) of the formamide dimer are given for different values of β and τ as a function of R(N···C′) and in the insert as a function of R(H···O), which of course is related to R(N···C′).

NH Spin−Spin Coupling Constant. The value of ^1J(NH) shows a different dependence on the distance R(N···C′) (or R(H···O)) depending which β and τ (ω) values are chosen. It varies by not more than 1.5 Hz in the range −91.0 to −92.4 Hz. Although these variations are small, it is worth discussing them because ^1J(NH) has been measured with modern NMR techniques.[30]

(I) The value of ^{1}J(NH) increases, i.e., adopts smaller negative values, with increasing R(N···C′) when β is close to 180°. This holds for both σ, $\sigma-\pi$, or π-type H-bonding (lines J, K, and I in Figure 4). (II) In all other cases, i.e., β between 110 and 150° and σ, $\sigma-\pi$, or π-type H-bonding (Figure 4), ^1J(NH) decreases to more negative values with increasing R(N···C′). This effect is strongest for β values between 110 and 120°. In general a larger value of β leads to more negative ^1J(NH) values; also π-type H-bonding implies more negative ^1J(NH) values than σ-type H-bonding (compare lines D with C and I with H in Figure 4).

The different behavior of ^1J(NH) = $f(R$(N···C′)) indicates that two different contributions to the spin−spin coupling mechanisms are active. For a linear approach of the NH bond toward the C═O bond (β close to 180°), the NH bond experiences the electric field of the negative charge at the O atom and of the C═O bond dipole moment. Negative charge is pushed from the H atom to the N atom; the N−H bond becomes shorter, and spin polarization is increased, thus leading to a larger magnitude of $|^1J(NH)|$ (situation I). Clearly, the electric field effect increases as the distance R(N···C′) decreases. This is demonstrated in Figure 5, where the SSCC ^{1}J(NH) for the monomer is given as a function of increasing point charges located at the atomic positions of the second formamide molecule in the formamide dimer, thus mimicking in this way the electric field effect. The point charges (corresponding to a charge coefficient of 1 in Figure 5) were obtained by a natural bond orbital (NBO) analysis[60] of the dimer. Since the atomic charges exaggerate the electric field effect relative to the continuous charge distribution of the second formamide molecule of the dimer, the charge coefficient was determined that leads to the SSCCs calculated for the dimer (0.86 to 0.88; Figure 5).

For a β-value of 180° (linear arrangement) and an R(O,H) value of 1.9 Å, the electric field effect of the second formamide molecule increases the magnitude of ^1J(NH) by 3 Hz (^1J(NH) decreases from −90.2 to 93.2 Hz, Figure 5). A lengthening of R(O,H) from 1.9 to 2.2 Å changes ^1J(NH) from −93.2 to −92.7 Hz, thus indicating a reduction of the electric field effect with an increase of R(O,H), thus explaining the trend in the ^1J(NH) values observed in region I (Figure 4). In the bent form (β = 120°), the electric field effect is smaller (−92.2 to −92.0 Hz) because the C═O dipole moment and the NH bond are no longer linearly arranged. If the NH bond is in the plane perpendicular to the plane of the second formamide molecule, the electric field effect increases substantially because σ-moment and π-moment add to yield a large effective dipole moment influencing the NH bond.

In Table 6, important orbital contributions to the FC term of the SSCC ^{1}J(NH) are listed. The FC term is dominated by the one-orbital contribution σ(NH), which describes spin polarization effects caused by repolarization or delocalization of charge[61,62] where the latter effect plays a small role. FC(σ(NH)) changes by 1.6 Hz from −139.0 to −137.4 Hz (Table 6) for β, decreasing from 180 to 120°. This is a direct reflection of a weakening of the electric field effect. Hence, in region I, changes in ^1J(NH) with R(O,H) are dominated by the electric field effect as described by Figure 5.

There must be a different spin−spin coupling mechanism in region II because the observed changes (see Figure 4) cannot be explained by the electric field effect. An increase (rather than decrease) in the magnitude of ^1J(NH) with R(O,H) or

(60) (a) Carpenter, J. E.; Weinhold, F. *J. Mol. Struct. (THEOCHEM)* **1988**, *46*, 41. (b) Reed, A. E.; Weinstock, R. B.; Weinhold, F. *J. Chem. Phys.* **1985**, *83*, 735. (c) Reed, A. E.; Curtiss, L. A.; Weinhold, F. *Chem. Rev.* **1988**, *88*, 899.
(61) Tuttle, T.; Grafenstein, J.; Wu, A.; Kraka, E.; Cremer, D. *J. Phys. Chem. B*, **2004**, *108*, 1115.
(62) (a) Wu, A.; Gräfenstein, J.; Cremer, D. *J. Phys. Chem. A.* **2003**, *107*, 7043. (b) Wu, A.; Cremer, D. *Phys. Chem. Chem. Phys.* **2003**, *5*, 4541.

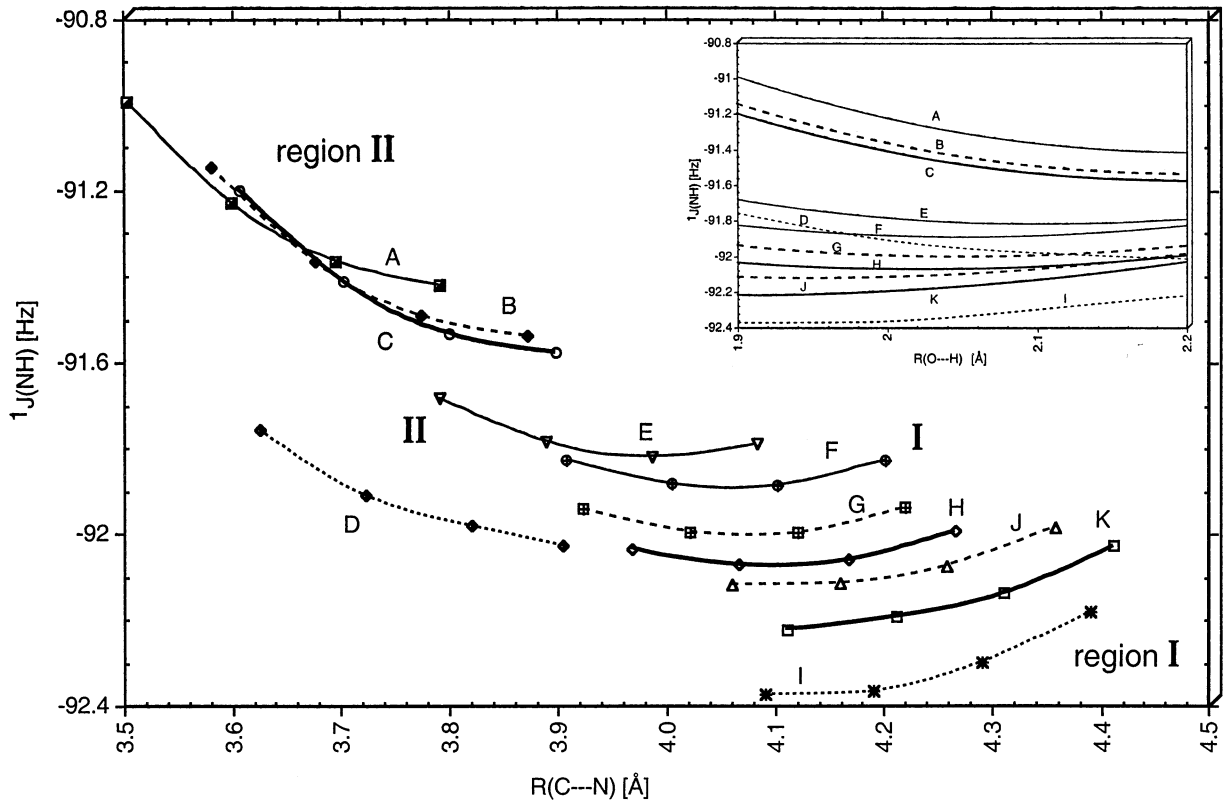

Figure 4. Model 3: Relationship between ^1J(NH) and distance R(N···C′) for different β, τ, and ω values, keeping α at 169.6°. The inset gives the dependence of ^1J(NH) on the distance R(O···H). Line A: $\beta = 112.5°$; $\tau = -44.1°$; $\omega = 140.0°$. Line B: $\beta = 118.0°$; $\tau = -22.8°$; $\omega = 160.0°$. Line C: $\beta = 120.0°$; $\tau = 0.0°$; $\omega = 180.0°$. Line D: $\beta = 120.0°$; $\tau = 0.0°$; $\omega = 90.0°$. Line E: $\beta = 131.6°$; $\tau = -59.2°$; $\omega = 140.0°$. Line F: $\beta = 140.0°$; $\tau = -90.0°$; $\omega = 140.0°$. Line G: $\beta = 144.5°$; $\tau = -36.1°$; $\omega = 160.0°$. Line H: $\beta = 150.0°$; $\tau = 0.0°$; $\omega = 180.0°$. Line I: $\beta = 150.0°$; $\tau = 0.0°$; $\omega = 90.0°$. Line J: $\beta = 160.0°$; $\tau = -90.0°$; $\omega = 160.0°$. Line K: $\beta = 180.0°$; τ is undefined; $\omega = 180.0°$.

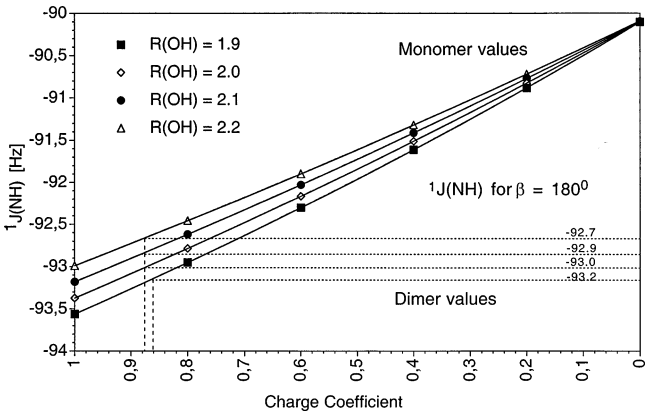

Figure 5. Dependence of SSCC ^1J(NH) of the formamide molecule on the NBO atomic charges of a second formamide molecule fixed at the positions of the atoms of the latter in the formamide dimer (see Figure 3b). Four different distances R(O,H) are considered: R(O,H) = 1.9, 2.0, 2.1, and 2.2 Å. The NBO charges were determined for the dimer and switched on with a charge coefficient increasing from 0 to 1 in the monomer calculations. Horizontal dotted lines give the values of SSCC ^1J(NH) at the four distances R(O,H) in the formamide dimer. All values are given in Hz. CP-DFT/B3LYP/(11s,7p,2d/5s,1p) [7s,6p,2d/4s,2p] calculations.

R(C··· N) (^1J(NH) values become more negative; Figure 4) is due to changes in the σ*(NH) population. At shorter distances and an alignment of lone pair orbital 2 (lp2) at the carbonyl O atom, there is a transfer of charge from lp2 into σ*(NH), which leads to a weakening of the NH bond and a repolarization of the bond (negative charge is shifted from N to H because σ*(NH) has a larger amplitude at H). Both effects lead to similar changes in the spin ploarization and to a reduction of the magnitude of ^1J(NH), which becomes less negative. With increasing R(O,H), the delocalization effect descreases, thus causing an increase (rather than decrease) in the magnitude of ^1J(NH).

The charge-transfer effect (delocalization) effect is responsible for a covalent nature of spin−spin coupling and can be assessed from the oxygen lone pair (lp)-orbital contributions to FC(NH), especially those of lp2, which is oriented in the direction of the NH bond. For an R(O,H) value of 1.6 Å, the total lp contribution is 1.5 Hz; however, this decreases to 0.6 Hz for R(O,H) = 1.9 Å (Table 6). Hence, the change of ^1J(NH) from −91.0 to −91.5 Hz (Figure 4, line B) is dominated by the lp2-orbital contributions and its involvement in delocalization, causing a covalent contribution to the spin−spin coupling mechanism. We note that the distinction between electric field (region I) and covalent contribution (region II) to the coupling mechanism *concerns only the changes* in ^1J(NH) caused by changes in β and R(O,H). The coupling mechanism of ^1J(NH) in a monomer is covalent anyway, as reflected by the dominating contribution FC(σ(NH))[61] (Table 6).

The covalent coupling mechanism (mechanism II) is enforced when H migrates from N to O atom (see Figure 2a). It is weakened when β increases to 180°. The covalent effect changes more than the electric field effect with R(C···N) as reflected by the steeper function ^1J(NH) = f(R(N···C′)) for $\beta = 120°$ (Figure 4). For π-type H-bonding, the electric field effect becomes stronger and the covalent effect weaker, which is

Table 6. Decomposition of the NMR Spin−Spin Coupling Constants across and at the Hydrogen Bond in the Formamide Dimer into Orbital Contributions[a]

terms	^{3h}J(CN)		^{1h}J(OH)		^{2h}J(ON)		^{1}J(NH)		^{1}J(CO)	
FC(σ(NH))	−0.5		*2.6*	*2.0*	3.1	*2.4*	−139.0	*−137.4*	0.3	
FC(σ(CO))	−0.3		0.2	*0.2*	−0.7	*−0.3*	0.1		−130.7	*−127.6*
FC(lp1(O))	−0.1		−0.8	*0.5*	2.5	*−1.1*	0.1		27.4	*22.6*
FC(lp2(O))	−0.1		−1.0	*−2.0*	2.2	*4.3*	0.1	*0.6*	23.9	*27.1*
FC(ob)	0.3				−1.2	*−0.9*	17.1	*16.8*	21.7	*20.7*
FC(σ(NH),σ(CO))	−0.8	0.1	−1.8	*−0.7*	−2.0	*−0.8*			0.2	*0.1*
FC(σ(NH),lp1(O))	−0.1		3.1	*−1.6*	3.5	*−2.2*			0.2	
FC(σ(NH),lp2(O))	−0.1		2.9	*5.7*	3.2	*6.7*		*−0.1*	0.2	*0.2*
FC(lp1(O),lp2(O))	0.1		0.7	*0.5*	−1.4	*−0.9*		*0.1*	−14.3	*−13.5*
FC(σ(CO),lp1(O))	−0.1			*−0.2*		*0.4*			31.7	*32.6*
FC(σ(CO),lp2(O))	−0.1	*0.1*		*0.2*	0.1	*0.3*			30.8	*30.4*
FC(ob,ob)	0.6	*−0.1*	0.5	*0.3*	−2.3	*−1.8*	24.1	*23.6*	4.6	*4.1*
FC(total,one)	−0.6		1.0	*0.7*	5.9	*4.3*	−121.6	*−119.9*	−57.4	*−57.1*
FC(total,two)	−0.6	*−0.2*	5.4	*4.2*	1.2	*1.2*	24.0	*23.5*	53.4	*53.9*
FC(total,all)	−1.0	*0.1*	6.7	*4.9*	6.0	*4.3*	−91.2	*−90.1*	0.2	*16.4*
PSO(total,one)			1.0	*0.8*	−0.5		−0.6	*−0.6*	5.9	*6.0*
PSO(total,two)			0.3	*0.2*			−0.1	*−0.1*	6.0	*6.0*
PSO(total,all)	−0.1		1.3	*1.0*	−0.1	*−0.1*	−1.5	*−1.3*	13.1	*13.7*
J(total)	−1.1	*0.1*	7.2	*4.9*	5.9	*4.2*	−93.2	*−92.1*	27.4	*28.4*

[a] All contributions and SSCCs are given in Hz. Values in normal font correspond to $\beta = 180$; those in italics correspond to $\beta = 120°$. Orbital contributions smaller than |0.05| Hz have been omitted. Also, the ^{2h}K(CH) values, which are all 0.1 SI units or smaller, have been dropped. The symbol ob (other bonds) denotes contributions from bonds not explicitly considered. Abbreviations FC(total,one) and FC(total,two) denote the sum of all one-orbital and two-orbital contributions listed in the table. FC(total,all) gives the true value for the formamide dimer, including the orbital contributions of remote bonds not listed here. The latter is also true for PSO(total,all) and J(total).

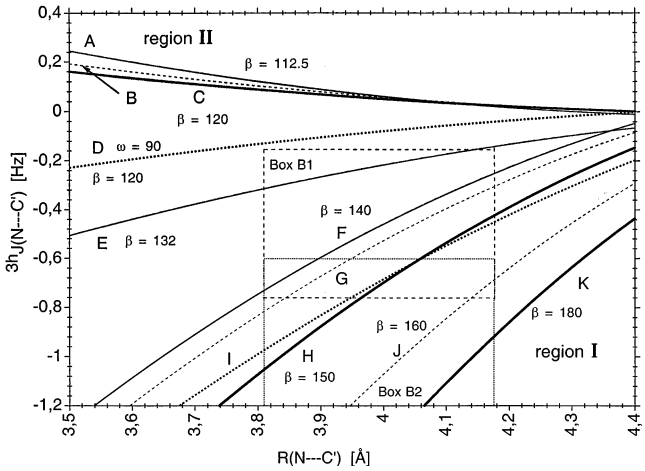

Figure 6. Model 3: Relationship between ^{3h}J(N···C′) and distance R(N···C′) for different β, τ, and ω values keeping α at 169.6°. The dashed boxes indicate the regions of calculated and measured SSCCs. Regions I and II indicate geometrical situations in which spin−spin coupling mechanisms I and II, respectively, are dominant. For explanations of lines, see Figure 4.

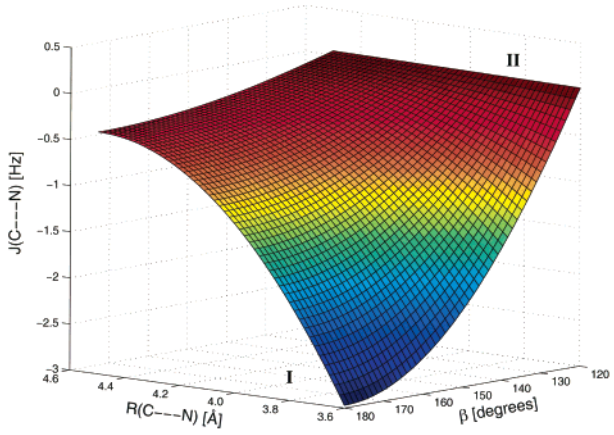

Figure 7. Model 3: Perspective drawing of the SSCC ^{3h}J(N···C′) calculated as a function of the distance R(N···C′) and the angle β. Regions I and II indicate geometrical situations in which spin−spin coupling mechanisms I and II, respectively, are important.

reflected by the shifting of curve D (Figure 4) to more negative values and its flattening relative to curve C (Figure 4).

There are some curves such as E, F, or G (β = 132, 140, 145°, respectively, Figure 4) that resemble very flat parabola. They represent the transition from a more covalent to a more electrostatic contribution to the coupling mechanism. The minima of these curves can actually be taken as an indicator for the change in the type of the N−H,O═C interactions.

We conclude that accurate measurement of the SSCC ^{1}J(NH) of a protein leads to valuable information about the type of the N−H,O═C interactions. Using diagrams such as the ones shown in Figure 4 and combining these with geometrical information (obtained, e.g., from X-ray diffraction studies), the stereochemistry of these interactions can be analyzed and a general insight into the conformational features of a protein can be gained.

NC′ Spin−Spin Coupling Constant. In Figures 6 and 7, the calculated SSCC ^{3h}J(NC′) values for the formamide dimer is given as a function of R(N···C′) at fixed values of β and ω as well as a function of both R(N···C′) and β in the form of a perspective drawing.

The reduced SSCC ^{3h}K(NC′) is positive, as one would expect for a three-bond SSCC according to the Dirac vector model.[60] The negative value of ^{3h}J(NC′) is a consequence of the negative gyromagnetic ratio of the ^{15}N nucleus ($\gamma = -2.7126 \times 10^7$ rad T^{-1} s^{-1} [45]). Figure 6 reveals that again there are two different types of curves ^{3h}J(NC′) = f(R(N···C′)). For β values larger than 130° (situation I), the SSCC ^{3h}J(NC′) decreases in magnitude (i.e., adopts less negative values) with increasing R(N···C′) values. The SSCC decreases rapidly in magnitude (^{3h}J(NC′) becomes less negative), whereas for β < 180°, the decrease in magnitude becomes gradually slower (Figure 6). The strongest geometrical dependence of ^{3h}J(NC′) is found for β, while the dependence on ω is moderate (compare lines H and I in Figure 6).

In the linear form (line K in Figure 6), spin−spin coupling is transmitted by mutual spin polarization of the C═O and N−H

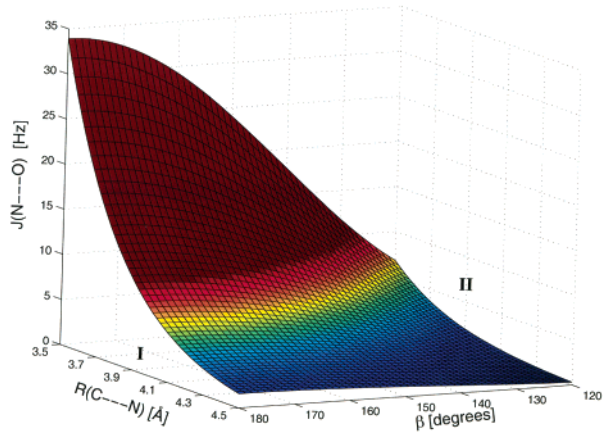

Figure 8. Model 3: Perspective drawing of the SSCC ^{2h}J(N···O) calculated as a function of the distance R(N···C′) and the angle β. Regions I and II indicate geometrical situations in which spin−spin coupling mechanisms I and II, respectively, are important.

bond density as reflected by the steric exchange term[61] FC(σ(NH),σ(CO)) (−0.8 Hz, Table 6) and the electric field effect between C═O and N−H, leading to orbital contributions FC(σ(NH)) and FC(σ(CO)) of −0.5 and −0.3 Hz (Table 6). These three orbital contributions, which are all associated with the electric field effect, determine the magnitude of ^{3h}J(NC′) in the linear form. They reduce to a small positive value of 0.1 Hz upon decreasing β to 120° (Table 6; see lines A−D in Figure 6). There are other small orbital contributions of different signs, which cancel each other. One can say that the remaining small positive orbital contribution results from the lp-orbitals. Again, lp2 dominates the sum of the lp contributions, thus indicating that in the bent form, the covalent spin−spin coupling mechanism involving charge transfer from lp2 to σ*(NH) is responsible for a small but finite ^{3h}J(NC′) value.

In the perspective drawing of Figure 7, the two regions of different coupling mechanism can be clearly distinguished: In front (for β values close to 180°), the most negative values of ^{3h}J(NC′) are found because of a strong electric field effect in the linear arrangement (region I). There is little difference for ^{3h}J(NC′) whether there is σ or π-type H-bonding involved between the N−H and the C═O group. This matters, however, for the second coupling mechanism mediated by a charge transfer from C═O to N−H (region II). For π-type H-bonding, a charge transfer is more difficult because the π(C═O) orbital is lower in energy than the n(C═O) orbital. Hence, π-type H-bonding is weaker (see above) and adds to the coupling mechanism predominantly via the electric field mechanism rather than the charge-transfer mechanism (compare line D with line C in Figure 6). Figure 7 reflects the situation of σ-type H-bonding, for which at $\beta = 120°$ the charge-transfer mechanism dominates, thus leading to ^{3h}J(NC′) values slightly larger than zero (region II).

N···O Spin−Spin Coupling Constant. The SSCC ^{2h}J(NO) reveals a strong dependence on both R and β (Figure 8). It changes in the range $3.5 < R(N\cdots C') < 4.5$ Å by more than 30 Hz and similarly for the range $120° < \beta < 180°$ (Figure 8). Both ^{2h}J(NO) and ^{2h}K(NO) should have the same sign considering that both ^{15}N and ^{17}O possess a negative gyromagnetic ratio ($\gamma(^{17}O) = -3.6266 \times 10^7$ rad T$_1$ s^{-1} [56]). The SSCC ^{2h}J(NO) is positive for a linear approach of the two groups (β close to 180°; region I), which seems to be a result of a strong contribution by the electric field effect. With increasing distance R(N···C′), the value of ^{2h}J(NO) decreases exponentially from 35 to 2 Hz as one would expect such that, at infinite separation, the nuclei are uncoupled (^{2h}J(NO) = 0).

For a decrease of β to 120° (region II), ^{2h}J(NO) decreases strongly for short R (Figure 8). Again, the coupling mechanism changes in the way that a charge transfer from the O electron lone pair to the N−H antibonding MO is involved, which leads to a negative contribution to ^{2h}J(NO). Consequently, the distance dependence resulting from the electric field effect is less severe in this region.

The analysis is supported by the orbital contributions listed in Table 6. The magnitude of SSCC ^{2h}J(NO) (5.9 Hz) is dominated by the sum of the one-orbital contributions to the FC term (5.9 Hz, Table 6), which in turn is the result of FC(σ(NH)) = 3.1, FC(lp1(O)) = 2.5, FC(lp2(O)) = 2.2 Hz and some smaller negative contributions. Again, the lp and the σ(NH) contribution can be related to the electric field effect. When the latter decreases upon decreasing β to 120°, the one-orbital contributions also decrease with the exception of FC(lp2(O)), which increases to 4.3 Hz as a result of lone pair delocalization. Lone pair delocalization is also reflected by an increase of the charge transfer from 0.014 to 0.025 e (BSSE-corrected NBO charges obtained with the (11s,7p,2d/5s,1p)-[7s,6p,2d/4s,2p] basis set[50]). When at the same time the distance R(O,H) is reduced to 1.6 Å, the charge transfer from the first formamide molecule (acting with its C═O group) to the second (acting with its NH group) increases from 0.025 to 0.054 e. A similarly strong increase is calculated for the lp2(O) orbital contribution, confirming that charge transfer and lp2(O) contribution are useful indicators for the covalent character of the HB. Clearly, a change in the spin−spin coupling from an electrostatic to a more covalent mechanism is obtained upon bending of the HB.

Although there is almost no experimental evidence on ^nJ(NO) SSCCs (one of the exceptions: ^1J(NO) of N_2O is 50.8 Hz[57]), enrichment with ^{15}N and ^{17}O could lead to more information on ^nJ(NO). However, this implies that the relaxation rate ($1/T_1$) of the ^{17}O nucleus is sufficiently small (the resonance signal sufficiently narrow) compared with the coupling constant. Apart from this, SSCC ^{2h}J(NO) is a sensitive tool to determine the nature of the coupling mechanism from calculations.

C═O Spin−Spin Coupling Constant. The value of the SSCC ^1J(CO) in acetone is known to be 22.4 Hz.[57] Also, it is known that the one-bond C═C SSCC increases by 11 Hz in the presence of an electronegative α-substituent such as NH_2.[63] Hence, the ^1J(CO) values of 30 to 32 Hz found in this work are in line with expectations.

In Figure 9, ^1J(CO) is given as a function of R(N···C′) for different β and ω values. The value of ^1J(CO) always decreases for decreasing R(N···C′) independent of the values of β and τ, which is actually in line with the geometrical dependence of the other SSCCs of the HB-system of the formamide dimer. The electric field effect leads to a change opposite for ^1J(CO) than for ^1J(NH), i.e., it must lead to a decrease rather than an increase in the magnitude of the SSCC (density is drawn out of the C═O bond; region I in Figure 9). At $\beta = 120°$, the negative charge of the O is transferred to the N−H bond, which also

(63) Kalinowski, H. O.; Berger, S.; Braun, S. *^{13}C NMR-Spektroskopie*; Georg Thieme Verlag: Stuttgart, 1984; see also references therein.

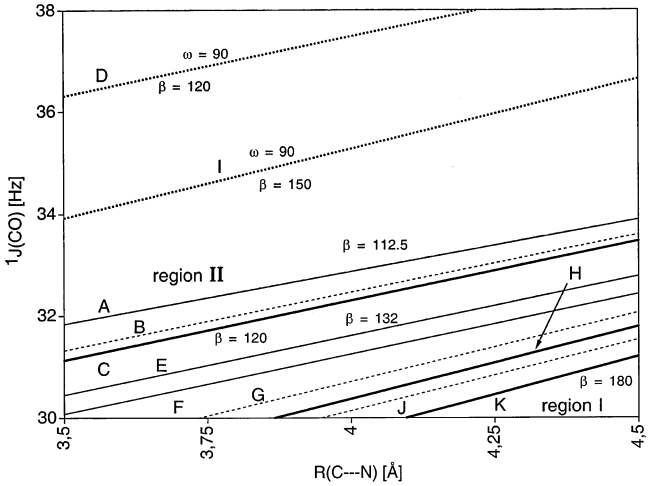

Figure 9. Model 3: Relationship between ^1J(CO) and distance R(N···C′) for different β, τ, and ω values, keeping α at 169.6°. Regions I and II indicate geometrical situations in which spin−spin coupling mechanisms I and II, respectively, are important. For explanations of lines, see Figure 4.

leads to a stronger weakening of the C,O coupling mechanism the closer the interacting groups are (region II). Hence, both coupling mechanisms have a similar dependence on the distance R, which implies that the lines in Figure 9 are almost parallel.

There is a significant increase in ^1J(CO) for π-type H-bonding (lines D and I in Figure 9). In this situation, π-type density, if at all, is drawn from the CO bond, which effects only the PSO (39% of total SSCC, Table 3) and not the FC term (56%), thus yielding a larger SSCC, which also decreases with decreasing R(N···C′) for reasons similar to those discussed for σ-type HB.

Influence of Noncontact Terms. The noncontact terms PSO, DSO, and SD have varying influence on the different SSCCs of a HB-system. The actual orbital contributions to the noncontact terms can be significant; however, there is considerable cancellation both between different orbital terms and between different noncontact terms. This in turn depends also on the geometry of the HB system, and therefore, it is difficult to predict the role of the noncontact terms. Clearly, the noncontact terms have to be calculated in each case because there is no basis for assuming the role of the FC term to always be dominant. This can be demonstrated for the SSCC ^1J(NH), which gives (from the experimental side) an easy access to the HB system. ^1J(NH) will vary by 2 Hz if the HB changes from a linear to a bent arrangement (Figure 2 and Table 6). The PSO term contributes −1.4 Hz to the value of ^1J(NH) and 0.4 Hz to the variation of the SSCC, i.e., an investigation based only on the calculation of the FC term cannot provide a reliable description of the SSCC ^1J(NH) and its dependence on the type of H-bonding. In the case of the SSCC ^{3h}J(NC′), the noncontact terms do not contribute more than 0.3 Hz to the value of the SSCC, which may be considered as being negligible. Considering, however, that this SSCC can vary between 0.2 and −2 Hz and that it is measured with an accuracy better than 0.1 Hz, it is quite important to include the noncontact terms into the calculation of the SSCC again to obtain a reliable description of the dependence of the SSCC on the type of H-bonding.

The calculation of the noncontact terms, especially PSO and SD, is absolutely necessary in the case of SSCCs ^1J(CO) and ^{1h}J(OH) (see Tables 3 and 6) but not needed for the SSCC ^{2h}J(NO). The importance of PSO and SD term increases when the geometry changes from σ to π H-bonding because, in the latter case, H-bonding is supported by the π-orbital of the C═O group, which leads to nonnegligible PSO and SD contributions.

Table 7. Comparison of Calculated and Predicted COH Angles (β) Observed for Selected Residue Pairs of the Protein Ubiquitin[a]

residues	β(DFT)	$\beta(J_{theor})$	$\beta(J_{exp})$	β(water)
64-2	172.4	≈180[b]	150	157.8
72-40	140.6	136	131	125.2
4-65	160.9	171	143	157.3
27-23	155.3	157	147	141.8
33-29	148.7	144	130	135.5
23-54	170.7	∼180[b]	146	173.7

[a] Optimized β values obtained at the B3LYP/6-31G(d,p) level of theory with model 3. Angle $\beta(J_{theor})$ was determined using the value of ^{3h}J(NC′) calculated at CP-DFT/B3LYP. Angle $\beta(J_{exp})$ was determined using the value of ^{3h}J(NC′) obtained experimentally.[12,13] Values β(water) were obtained from the aqueous solution structure of ubiquitin.[39] [b] Approximate value due to region extension beyond the domain of the calculated ^{3h}J(NC′) hypersurface $J(\beta,\alpha,\omega,\tau)$.

Application of Geometrical Relationships. The derivation of the geometrical relationships described above enables us to determine the position of the hydrogen atom in the HB without the necessity of optimizing the structure. To test the reliability of the relationship, the angle β(COH) was predicted in each of the interresidue structures using the NC′ separation obtained from the crystal structure[29] and the calculated ^{3h}J(NC′) constants. The same angle was also calculated in each structure using the separation as before, although with the experimentally determined SSCCs. The predicted results are compared to those obtained through the geometry optimization and those obtained from the water-refined structure[39] in Table 7.

The β angles, which are determined for the calculated SSCCs ($\beta(J_{theor})$ in Table 7), deviate by less than 10° from the optimized angles β (β(DFT) in Table 7), whereas these deviations reflect simplifications introduced by model 3 relative to model 2. A much larger deviation is found when the angle β is determined with the help of the measured ^{3h}J(NC′) constants ($\beta(J_{exp})$ in Table 7), reflecting the fact that the measured values are up to 0.6 Hz smaller than the absolute magnitude of the calculated ones. This is shown in Figure 6, where box B1 describes the area of the measured values. These values would imply, according to Figure 6, relatively small β values, which are supported by the geometry of ubiquitin determined in the water-refined structure (β(water) in Table 7). The calculated ^{3h}J(NC′) values for the residue pairs investigated are in the area of box B2 (note that box B1 and box B2 overlap somewhat), indicating that the crystal structure used in the calculations leads systematically to magnitudes of the ^{3h}J(NC′) SSCCs that are too large.

We have demonstrated that the discrepancy between measured and calculated ^{3h}J(NC′) values is not an artifact of the theory used and the simplifications made. On the contrary, it results from the different geometries of ubiquitin in aqueous solution and in the solid state. In the crystalline phase, the N−H···O═C HBs are more linearly arranged than in the aqueous solution structure, where these effects are larger in the β-sheets than the random coil or α-helix parts of ubiquitin. This becomes obvious when comparing calculated and measured ^{3h}J(NC′) values: the deviations are on the average 0.3 Hz for the α-helix but increase to 0.5−0.6 Hz in other parts of ubiquitin, which correspond to the difference between box B1 and box B2 in Figure 6.

A decrease of the angle β is best accomplished by a gliding movement of the backbones in neighboring parts with regard to each other so that the linear N−H···O═C arrangement is just a transient point, while those arrangements with small positive or negative β values (close to ±120°) are the turning points of the movement. Alternatively, there could be fluctuations in the backbone torsions leading to similar effects. For example, Case and co-workers[18b] calculated for ubiquitin an average ϕ-angle fluctuation amplitude of ±24°. Clearly, these movements, either gliding motions, torsional fluctuations, or any combination of both, are smaller for the α-helix than the β-sheets or the random coils. In any case, the smaller β values dominate in the conformational averaging and lead to ^{3h}J(NC′) values substantially smaller (Figure 6, lines A−C) than those obtained with the crystal or even the static (rather than conformationally averaged) water-refined structure of ubiquitin. In previous investigations, error cancellations have disguised this point.

5. Conclusions

From this study the following conclusions can be drawn:

(1) The use of FP-DFT in the calculation of SSCCs in HB systems is a potentially misleading practice due to the neglect of the PSO, DSO, and SD terms. Furthermore, the results of such calculations are highly dependent on the choice of the perturbation constant, λ, with no systematic variation between the selected values. These factors suggest that the use of the more comprehensive CP-DFT is the most appropriate choice for the calculation of SSCCs across HB systems. Previous investigations of the SSCCs in ubiquitin[22,23] were carried out with inferior methodological approaches that disguised the dependence of these SSCCs on the environment and conformational averaging (points 3 and 4).

(2) The XC functional and basis set effects were also considered, and it was found that B3PW91 functional lowers the calculated SSCC values relative to B3LYP. Similarly, the 6-311G(d,p) basis set produces, in most cases, lower magnitude SSCCs relative to the [6s,4p,1d/3s,1p] or [7s,6p,2d/4s,2p] basis sets. Since CP-DFT/B3LYP/VQZ+P or better descriptions of SSCCs have been found to be most reliable, it is misleading and disguises important effects when an XC/basis set combination is used, which artificially brings the calculated SSCCs of ubiquitin in better agreement with the measured ones.

(3) The crystal and aqueous solution structures of ubiquitin differ in a typical way, which can be best assessed by the angles β and their influence on the ^{3h}J(NC′) SSCCs (see Figure 6). In the crystal structure, the HBs are more linearly arranged (larger distance R(N···C′), while in the aqueous solution structure the NH bonds are oriented more in the direction of one of the electron lone pairs at the carbonyl oxygen (β closer to 120°; smaller distance R(N···C′). This leads to a general decrease of SSCCs ^{3h}J(NC′).

(4) The analysis of calculated and measured ^{3h}J(NC′) values indicates typical deviations speaking for considerable flexibility both in the backbones and the HB geometry, which can be described (apart from torsional fluctuations of the backbone) as a back-and-forth gliding movement between backbone parts connected by HBs relative to each other. The more linear HB arrangements of the crystal structure of ubiquitin are just transient points of these movements, which lead to conformationally averaged β and ^{3h}J(NC′) values with significantly smaller magnitudes than found for either the solid-state or solution-phase structure of ubiquitin. The conformational flexibility of the backbones is greater in the β-sheets and random coils than in the α-helix of ubiquitin.

(5) The modeling of the intramolecular HB environment requires the inclusion of several residues to obtain reliable trans-HB SSCCs. The most extended but also most costly model, model 1, leads to the best representation of SSCCs. This is in line with observations made by Juranic and co-workers[15,16] who got a better description of measured ^{3h}J(NC′) SSCCs when considering geometrical parameters of the peptide bond and the backbone in the vicinity of the HB. It is, however, not clear whether these effects are just an artifact caused by relating conformationally averaged trans-HB SSCCs measured in aqueous solution to geometrical parameters obtained from the crystal structure of the protein.

(6) Geometrical relationships for all six SSCCs of the HB system of the formamide dimer (model 3) were derived and the most important ones graphically displayed. Most interesting is the dependence of the SSCCs on R(N···C′), β, and ω. The quantitative relationship for ^{3h}J(NC′) was tested to predict the position of the hydrogen in the residue pairs investigated, and in each case the predicted position agreed well with the optimized geometry, suggesting that for this type of coupling model 3 (formamide dimer) serves well its purpose. We note that the exponential dependence of ^{3h}J(NC′) (or related SSCCs) on the distance between the coupling centers suggested by several authors[14,64,65] is a simplification, as shown in Figures 4, 6, 7, and 8 of this work.

(7) Two different mechanisms for spin−spin coupling were identified, one for which the spin polarization is transmitted via the HB predominantly by an electric field effect (mechanism I), and one for which spin polarization is transferred more via electron delocalization from the O electron lone pair orbital into the σ*(N−H) orbital (mechanism II). Mechanisms I and II can lead to a different geometrical dependence as demonstrated for the SSCCs $^{1}J(NH)$, $^{2h}J(N\cdots O)$ or ^{3h}J(NC′).

(8) Mechanisms I and II are active at the same time so that it is not possible to say that one of them is exclusively operative. It is also not possible to say that mechanism I indicates electrostatic H-bonding and mechanism II covalent-bonding. If mechanism I is operative (linear arrangement of N−H and O═C bonds), there are still tail interactions between the bond orbitals (transfer of NH bonding density into the σ*(CO) orbital and vice versa), which lead to a finite density in the intermolecular space and a covalent contribution to the SSCCs across the HB. Measured SSCCs in combination with appropriate calculations can provide insight into the geometry of the HB system and, in turn, into the ease of covalent H-bonding.

(9) In the crystal structure of ubiquitin, the electric field effect plays an important role for spin−spin coupling across a HB (mechanism I). This drastically simplifies the relationship between SSCCs and the geometrical parameters of the HB. A distinction between σ- and π-HBs is only possible in special cases, e.g., in the case of the SSCC ^{1}J(CO), which significantly increases when π-bonding is involved.

(64) (a) Del Bene, J. E.; Perera, S. A.; Bartlett, R. J. *J. Am. Chem. Soc.* **2000**, *122*, 3560. (b) Del Bene, J. E.; Bartlett, R. J. *J. Am. Chem. Soc.* **2000**, *122*, 10480. (c) Del Bene, J. E.; Perera, S. A.; Bartlett, R. J. *J. Phys. Chem. A* **2001**, *105*, 930.
(65) Pecul, M.; Sadlej, J. *Chem. Phys. Lett.* **1999**, *308*, 486.

The question whether a HB in proteins is covalent or electrostatic is an academic question because hydrogen bonding involves both electrostatic and electron delocalization (covalent) interactions. Clearly, the latter will increase if β decreases to values close to 120°. The backbone motions of ubiquitin will lead to changes in the HB geometry, which are synchronized in the α-helix and the β-sheets and most likely do not lead to a strong collective change in the overall binding energy. The HBs to the water molecules will also play a (probably) stabilizing role in these collective changes, which guarantee that, despite the conformational flexibility reflected by the difference between calculated and measured (conformationally averaged) SSCCs $^{3h}J(\mathrm{NC}')$, the secondary structure of ubiquitin is retained.

In general, it is problematic to conclude from the magnitude of trans-HB SSCCs the nature of hydrogen bonding. A SSCC reflects first of all the degree of spin polarization between the coupling nuclei. Spin polarization is associated with the first-order density rather than the zeroth order density that is responsible for bonding. However, the SSCCs depend on the geometry of the HB and, provided this relationship is decoded (as has been done in this work), the SSCCs lead to some indirect information on the bonding situation.

Acknowledgment. We thank Professor Alexandre Bonvin (Utrecht University), Professor Michael Nilges (Institut Pasteur), and Professor Jens Linge (Institut Pasteur) for providing the refined structures of ubiquitin both in the crystal state and in the solution phase. This work was supported by the Swedish Research Council (Vetenskapsrådet). An allotment of computer time at the National Supercomputer Center (NSC) at Linköping is gratefully acknowledged.

JA030246E