JCTC Journal of Chemical Theory and Computation

Direct Comparison of Experimental and Calculated NMR Scalar Coupling Constants for Force Field Validation and Adaptation

Franziska F.-F. Schmid and Markus Meuwly*

Department of Chemistry, University of Basel, Klingelbergstrasse 80, 4056 Basel, Switzerland

Received June 24, 2008

Abstract: The ability to measure scalar coupling constants across hydrogen bonds ($^{3h}J_{NC'}$) from high-resolution NMR experiments allows the characterization of detailed structural properties of biomolecules. To analyze those, a parametrized model based on the linear combination of atomic orbitals relates H-bond geometries with the measured $^{3h}J_{NC'}$ coupling magnitude. In the present study the dependence of calculated $^{3h}J_{NC'}$ coupling constants on force field parameters is assessed. It is shown that increased polarity of the hydrogen bond improves the calculated $^{3h}J_{NC'}$ coupling constants and shifts the conformational ensemble sampled from the molecular dynamics (MD) simulations toward the experimentally measured one. Increased charges lead to more narrow distance and angle distributions and improve the agreement between calculated and measured $^{3h}J_{NC'}$ couplings. However, different secondary structures are better represented by different magnitudes of electrostatic interactions—different atomic partial charges in the present work—as indicated by root-mean square deviations (rsmds) between observed and calculated coupling constants $^{3h}J_{NC'}$. The parametrization of the empirical formula is found to be meaningful and robust, but the parameter values are not universal across different proteins and different secondary structural elements (α-helices, β-sheets and loops). Using standard and slightly increased CHARMM charges, predictions for the as-yet unknown scalar coupling constants for the V54A and I6A mutants of protein G are made.

1. Introduction

Hydrogen bonds (H-bonds) are important for the stabilization of biomolecular structures, formation of secondary and tertiary structures in proteins, ligand binding, and specificity. Therefore a detailed understanding of the energetics and dynamics of H-bonds is desirable. The ability to measure scalar coupling constants across H-bonds ($^{3h}J_{NC'}$) by using NMR techniques provided valuable information on the overall protein fold.[1] The donor and acceptor atom can now be identified unambiguously and the observed $^{3h}J_{NC'}$ coupling value is a measure of the orbital overlap. In general, scalar couplings are in the range of 0 to −1 Hz. On the basis of a linear combination of atomic orbitals, a purely geometrical model involving distances and angles to characterize an H-bond was proposed to calculate $^{3h}J_{NC'}$ values from structural data.[2] It has been shown that compared to coupling constants from more rigorous and computationally expensive density functional theory (DFT) calculations, deriving $^{3h}J_{NC'}$ values from the H-bond geometry alone is more economical and gives quantitatively correct results. This opens the possibility to calculate scalar $^{3h}J_{NC'}$ coupling constants from molecular dynamics (MD) simulations which has been recently done for different systems.[3,4]

Dynamics is likely to be important to accurately compute $^{3h}J_{NC'}$ couplings as the experimentally observed couplings are averages over the motions up to a time required for coherence transfer (typically 0.1 to 1 s). Compared to computations based on the experimental X-ray or NMR structures alone, significant improvement was found when the conformational motion is included.[3] Furthermore, the

* Corresponding author. Phone/Fax Number: +41 61 267 38 21/ +41 61 267 38 55. E-mail address: m.meuwly@unibas.ch.

10.1021/ct800241d CCC: $40.75
Published on Web 10/25/2008

analysis showed that the starting structure, in particular the resolution of the experimental structure, is important for the accuracy with which the coupling constants can be calculated.[4] Low-resolution structures can be improved by relaxing the structure, mostly the donor acceptor distance, where the values adapt to the ones found in high-resolution structures. In calculations on high-resolution structures of ubiquitin, protein G, and the Tudor domain for which extensive measurements are available, the best agreement between calculated and observed $^{3h}J_{NC'}$ couplings in a root-mean square (rms) sense is $\approx$0.14 Hz.[4] In addition, by comparing with previous simulations carried out in implicit solvent,[3] it was found that an atomistic representation of the solvent, which is water in the present case, is instrumental for meaningful simulations.

A different approach to model the structural features of H-bonds is to include constraints into the simulations by adding external biasing potentials, e.g. on $^{3h}J_{NC'}$, nuclear Overhauser effects (NOEs), or S^2 values, to the force field.[5] Such a procedure lowered the root-mean square diameter (rmsd) between calculated and measured $^{3h}J_{NC'}$ to 0.06 Hz and lead to a shift in the conformational ensemble, such that the distance between donor and acceptor, $r_{H\cdots O'}$, decreases and the directionality of the H-bond, $\theta_1(N-H\cdots O')$, was found to prefer a near-linear (180°) geometry. Such a procedure serves to better characterize the ensemble of protein structures represented by the measurements.

Given the relationship between geometry and scalar coupling constants, it is of interest to explore the influence of force field parameters (in particular nuclear charges) and the empirical parametrization of the model used to calculate $^{3h}J_{NC'}$ values.[2] Usually, force fields are parametrized in view of structural, (optical) spectroscopic, and thermodynamic measurements.[6] The former two provide information on bond lengths and force constants to which the force field can be fitted whereas the latter is rather more sensitive to the nonbonded interactions. However, scalar coupling constants across hydrogen bonds can be expected to be sensitive to parameters describing the interaction strength of a H-bond which is primarily reflected in the nuclear charges or the polarity of the hydrogen bond.

Furthermore, using atomistic simulations, it is possible to consider and separately analyze hydrogen bonds within different secondary structural elements (SSEs, α-helix, β-sheet, and loop) for which scalar coupling constants can be measured. To this end, three experimentally well-characterized proteins (ubiquitin, protein G, and the Tudor domain), for which an extensive set of $^{3h}J_{NC'}$ values has been measured, is used. Here, we address the following questions: What is the influence of the nuclear charges for atoms involved in the H-bond on the calculated coupling constants and the conformational ensemble sampled from MD simulations? How universal are the parameters in eq 2 (see below) for different proteins? And, are optimal parameter values identical for all secondary structural elements?

The present work is structured as follows. First, the proteins and methods used are described. Next, results for $^{3h}J_{NC'}$ from static structures and from MD simulations are presented and their relationship to the conformational space sampled is explored. Finally, the results are discussed, and predictions for particular mutants of protein G are made which can serve as a test to validate the computational approach used to analyze previously measured scalar coupling constants.

2. Materials and Methods

2.1. Setup of the Proteins. All molecular dynamics (MD) simulations were carried out with CHARMM[7] and the CHARMM22[6] force field. The starting structures for the three proteins considered, the Tudor domain, protein G, and human ubiquitin, were taken from the Brookhaven Protein Data Bank. All titratable side chains were in their standard protonation state for pH 7. For the Tudor domain, the 1.8 Å X-ray structure (code 1MHN)[8] and 10 NMR structures (code 1G5V)[9] are used. 56 residues are contained in both structures and correspond to residues 90−145. Protein G was set up from the 1.92 Å resolution X-ray structure (code 1PGB)[10] and from the 1.1 Å resolution X-ray structures (codes 1IGD[11] and 2IGD[12]). It consists of 56 residues, and the following mutations were introduced to match the sequence of the protein used in the experimental $^{3h}J_{NC'}$ measurement:[13] T1M, T2Q, V6I, I7L, K19E, E24A A29V, V42E, and, the *N*-terminal was deleted. The mutations were done by replacing the specific sidechains. The 76 residue protein human ubiquitin was taken from the 1.8 Å resolution X-ray structure (code 1UBQ)[14] and ten NMR structures (code 1D3Z).[15] After adding hydrogen atoms, the structures were relaxed with 3000 steps of steepest descent (SD) minimizations. All structures were solvated in a rectangular box of pre-equilibrated TIP3P water (box sizes are $50 \times 40 \times 40$, $56 \times 47 \times 40$, and $65 \times 50 \times 47$ Å^3, respectively, for Tudor domain, protein G, and ubiquitin). After removing solvent molecules which overlap with the protein, the systems contained between 8000 and 15 000 atoms.

2.2. Molecular Dynamics Simulations. Molecular dynamics simulations were carried out with periodic boundary conditions and images were updated every 10 time steps. A 12 Å cutoff was applied to the shifted electrostatic and switched van der Waals interactions. The systems have an overall charge of −3 for the Tudor domain, −4 for protein G, and 0 for ubiquitin. No counterions were added to neutralize the systems. The systems were heated and equilibrated at 300 K for 50 ps. All MD simulations were carried out using the leap Verlet algorithm with a time step of 1 fs and constraining hydrogen atoms with SHAKE,[16] which is consistent with the way in which force field parameters are optimized.[6] Several trajectories from different starting structures were run to sample the conformational ensemble. Overall, the analysis of the H-bond dynamics is based on 19 ns for the Tudor domain, 10 ns for protein G, and 15 ns for ubiquitin.

Atomic Partial Charges in the H-Bond Motif. To study the influence of varying atomic partial charges on the conformational ensemble, MD simulations were carried out using identical protocols. The atomic charges on the nitrogen, hydrogen, and oxygen atoms involved in the H-bonds (for the definition of the H-bonding motif, see Figure 1) were

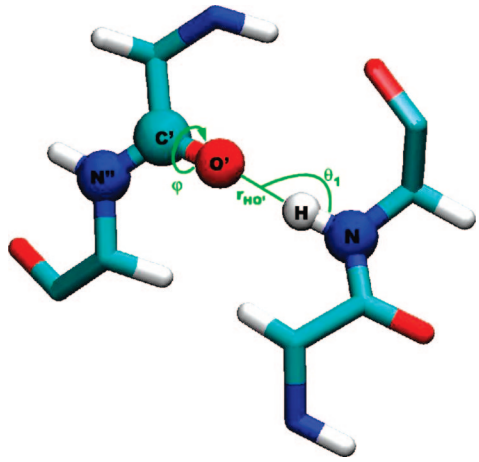

Figure 1. Ball and stick representation of the H-bond motif N−H···O′=C′−N″ together with the geometric properties incorporated in eq 1 to calculate $^{3h}J_{NC'}$ values: the donor−acceptor distance H···O′ ($r_{H\cdots O'}$), the directionality angle N−H···O′ (θ_1), and the dihedral angle H···O′=C′−N″ (ϕ).

Table 1. Atomic Partial Charges for H-Bond Atoms Applied in MD Simulations and ab initio Charges for Specific H-Bond Sites

	partial charges [e]				sum
	N−H···O=C				
$q_{-0.04}$	−0.45	0.27	−0.49	0.51	−0.16
q_0	−0.47	0.31	−0.51	0.51	−0.16
$q_{+0.04}$	−0.49	0.35	−0.53	0.51	−0.16
$q_{+0.1}$	−0.52	0.41	−0.56	0.51	−0.16
$q_{+0.2}$	−0.57	0.51	−0.61	0.51	−0.16
$q_{+0.4}$	−0.67	0.71	−0.71	0.51	−0.16
J12[a]	−0.64	0.43	−0.67	0.70	−0.18
J12[b]	−0.59	0.40	−0.62	0.65	−0.16
J19[a]	−0.69	0.42	−0.63	0.70	−0.20
J19[b]	−0.55	0.34	−0.51	0.56	−0.16

[a] ab initio NBO charges for coupling no. J12 and J19 of ubiquitin as obtained. [b] Scaled NBO charges to maintain the sum −0.16 e.

systematically increased and decreased. For all models, the total charge of the H-bond motif was preserved ($q_O + q_H + q_N = -0.16$ e). The different charge sets (see Table 1) are labeled q_0 for standard CHARMM charges and $q_{-0.04}$, $q_{+0.04}$, $q_{+0.1}$, $q_{+0.2}$, and $q_{+0.4}$ according to the charge increase/decrease on the hydrogen atom.

The variation of the partial charges is motivated from results of ab initio electronic structure calculations on two different H-bond motifs of ubiquitin (couplings J12 and J19) at the B3LYP/6-31G* level of theory. To assess possible electronic coupling between neighboring residues, the ab initio calculations included the preceding and succeeding amino acid of H-bond donor and acceptor (see Figures S1 and S2 in the Supporting Information). In the electronic structure calculations, coordinates for all except the H-bond hydrogen atom were kept fixed. These calculations were carried out with the Gaussian suite of programs.[17] Charges from a natural bond orbital (NBO) analysis are summarized in Table 1. The total charge of the H-bonding motif slightly differs from −0.16 e which is the value from the CHARMM force field used here. Therefore, appropriately scaled charges to make them directly comparable to the CHARMM force field are also reported.

2.3. Analysis of the Results. From the structures and the MD simulations, geometries describing the hydrogen bonds (the distances H···O′ ($r_{H\cdots O'}$) and N···O′ ($r_{N\cdots O'}$), the angles N−H···O′ (θ_1) and H···O′=C′ (θ_2), and the dihedral H···O′=C′−N″ (ϕ)) were extracted every 0.1 ps. This is done for all atoms for which scalar couplings across H-bonds are observed experimentally except for coupling no. 1 of ubiquitin, because for this coupling homotopic hydrogen atoms exist. A total of 15, 32, and 29 J-couplings were analyzed for the Tudor domain, protein G, and ubiquitin, respectively (see Figure 2). These H-bonds are located in different SSEs, namely the α-helix and β-sheet and loop regions (see Tables S6−S8 in the Supporting Information).

Using a model based on a linear combination of atomic orbitals, a formula which relates geometries characterizing a hydrogen bond to scalar $^{3h}J_{NC'}$ coupling constants was proposed, parametrized, and tested:[2]

$$^{3h}J_{NC'}(r_{HO'}, \theta_1, \phi) = [\alpha \cos^2 \theta_1 + f(\phi)]e^{\beta(r_{HO'} - r^0_{HO'})} - 0.10 \text{ Hz} \quad (1)$$

The coordinates involved are the H···O′ distance $r_{HO'}$, the N−H···O′ angle θ_1, and the H···O′=C′−N″ dihedral ϕ (see Figure 1 for a definition of coordinates). The contribution of the dihedral term $f(\phi)$ was found to be small and primarily removes systematic differences between α-helices and β-sheets.[2,4] Thus, eq 1 was simplified to

$$^{3h}J_{NC'}(r_{H\cdots O'}, \theta_1) = \alpha \cos^2(\theta_1)e^{\beta r_{H\cdots O'}} + \delta \text{ Hz} \quad (2)$$

which is the form used in the present work. The original parameters in eq 2 were fitted to 34 experimentally determined $^{15}N-^{13}C'$ coupling constants in protein G and yielded the following values: $\alpha_B = -357$ Hz, $\beta_B = -3.20$ Å^{-1}, and $\delta_B = 0$ Hz, where the index "B" refers to Barfield's work. The distance $r^0_{HO'}$ was 1.76 Å, the smallest H···O′ distance found in crystallographic structures.[2]

One of the aims of the present work is to study the universality of eq 2 for the Tudor domain, protein G, and ubiquitin. For this, the coupling constant $^{3h}J_{NC'}$ was calculated using eq 2 with the original parameters α_B and β_B for each snapshot from the MD simulations and then averaged over all structures from which the ensemble averaged coupling constant, $\langle ^{3h}J_{NC'}\rangle$ was obtained. Furthermore, optimal parameters α^{opt} and β^{opt} which minimize the sum of squares $\sum(^{3h}J^{calc}_{NC'} - {}^{3h}J^{obs}_{NC'})^2$ were determined where the sum extends over all couplings. As the parametrization of eq 2 suggests that the parameters α and β are correlated; α, which is only a scaling factor, is kept fixed, whereas β is optimized. The optimization is carried out for all $^{3h}J_{NC'}$ couplings and individually for couplings located in particular SSEs. Tests whereby the parameter α was varied were carried out after an optimal value for β had been found. However, the rmsd could usually not be improved any further. Thus, $\alpha_B = -357$ Hz was used throughout the present work.

3. Results

In the following, results for the Tudor domain, protein G, and ubiquitin are discussed. The Tudor domain contains

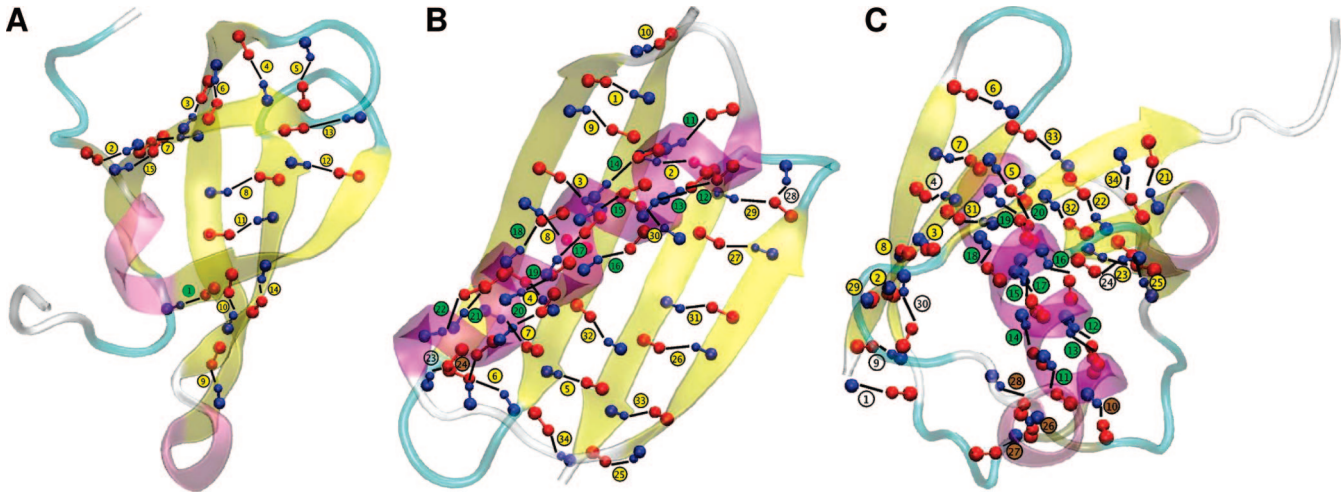

Figure 2. $^{3h}J_{NC'}$ coupling sites. The structure of (A) the Tudor domain, (B) protein G, and (C) ubiquitin in cartoon representation to highlight the different secondary structures. α-Helices are colored pink, β-strands are in yellow, and loops are in cyan. Hydrogen bonds corresponding to the experimentally measured scalar couplings are depicted in ball and stick representation for N—H (blue) and O═C (red). Each coupling is enumerated (J1−J*n*). Coupling constants used in this work are colored according to their secondary structure location in the protein: α-helix (green) and β-sheet (yellow) and loop (brown).

Table 2. Average rmsds between Calculated and Experimentally Measured Scalar Couplings $^{3h}J_{NC'}$ for Static Structures (X-ray and NMR)[a]

	all		α-helix		β-sheet		loop	
protein (coupl)	β [Å^{-1}]	rmsd [Hz]	β [Å^{-1}]	rmsd [Hz]	β [Å^{-1}]	rmsd [Hz]	β [Å^{-1}]	rmsd [Hz]
Tudor (15)	−3.20	0.25			−3.20	0.25		
fitted β	−3.05	0.20			−3.05	0.20		
protein G (32)	−3.20	0.14	−3.20	0.12	−3.20	0.15		
fitted β	−3.28	0.12	−3.26	0.09	−3.25	0.13		
ubiquitin (29)	−3.20	0.31	−3.20	0.20	−3.20	0.40	−3.20	0.14
fitted β	−3.46	0.23	−3.38	0.10	−3.42	0.28	−3.05	0.13

[a] Results are provided for both β_B and optimized β^{opt} to minimize the rmsd between calculated and observed couplings.

hydrogen bonds in β-sheets; for protein G, they are located in α-helices and β-sheets; ubiquitin has H-bonds in α-helices and β-sheets and loop regions. Results are reported and discussed separately for static structures and those from MD simulations.

3.1. Static Structures. First, eq 2 with parameters α_B and β_B was assessed by calculating rmsds between observed and calculated scalar couplings for the static X-ray and NMR structures of all three proteins. Including all experimentally observed couplings, the rmsds for the Tudor domain, protein G, and ubiquitin are 0.25, 0.14, and 0.31 Hz, respectively (see Table 2). The agreement between calculation and experiment is comparable to previous results.[2,4] Considering scalar coupling constants in particular SSEs, it is found that rmsds for α-helices (for ubiquitin 0.20 compared to 0.31 Hz) and loops (for ubiquitin 0.14 compared to 0.31 Hz) are smaller than by calculating rmsds including all couplings (see Table 2). In contrast, β-sheet couplings are less accurately estimated (for ubiquitin 0.40 compared to 0.31 Hz). This suggests that it is of interest to also consider individual parametrizations β_H (for α-helices), β_S (for β-sheets), and β_L (for loop regions) for different SSEs of eq 2.

Averages $\langle^{3h}J_{NC'}\rangle$ of the experimentally measured coupling constants for individual SSEs of each protein showed significant differences (up to 0.25 Hz in ubiquitin; see Table 3) in the magnitude of the scalar couplings. More negative values for the scalar couplings are found for β-sheets and less negative values for loop regions and α-helices (see Table 3), i.e. β-sheet hydrogen bonds are stronger than the ones in α-helix or loops. Thus, for all static protein structures, the parameter β was individually optimized for the three structure elements (see Table 2). It is found that these optimized β^{opt} values can differ considerably from $\beta_B = -3.20$ Å^{-1}. For protein G, the optimal β values for α-helix and β-sheet are similar, −3.26 and −3.25 Å^{-1}, respectively. However, for ubiquitin, β differs significantly for each secondary structural element: −3.38 Å^{-1} for α-helix, −3.42 Å^{-1} for β-sheet, and −3.05 Å^{-1} for loops. The parameter optimization decreased the rmsds for all secondary structure elements, although not equally well. Again, α-helices and loops are found to perform better than β-sheet (protein G 0.09 Hz rmsd (α-helix) vs 0.13 Hz (β-sheet), ubiquitin 0.10 Hz rmsd (α-helix) vs 0.28 Hz (β-sheet)).

Table 3. Average Experimentally Measured $^{3h}J_{NC'}$ Scalar Couplings for Different Secondary Structure Elements

	$^{3h}J_{NC'}$ [Hz]		
	ubiquitin	protein G	Tudor domain
overall	−0.50 ± 0.171	−0.43 ± 0.179	−0.55 ± 0.104
helix	−0.35 ± 0.149	−0.37 ± 0.211	
sheet	−0.61 ± 0.094	−0.47 ± 0.153	−0.55 ± 0.104
loop	−0.43 ± 0.161		

Table 4. β^{opt} Parameters [Å^{-1}] with Corresponding $^{3h}J_{NC'}$ rmsd [Hz] for Each Charge Set and Secondary Structure Element Calculated from Averages over MD Simulations of All Protein Systems Considered[a]

charge set	overall		helix		sheet		loop	
	rmsd	β	rmsd	β	rmsd	β	rmsd	β
$q_{-0.04}$	0.21	−3.09	0.21	−3.04	0.21	−3.09	0.21	−3.03
q_0	0.16	−3.19	0.15	−3.20	0.14	−3.19	0.20	−3.09
$q_{+0.04}$	**0.15**	−3.30	**0.14**	−3.32	0.14	−3.30	0.21	−3.20
$q_{+0.1}$	**0.15**	−3.45	0.15	−3.35	**0.13**	−3.45	0.17	−3.34
$q_{+0.2}$	**0.15**	−3.67	0.18	−3.82	0.17	−3.62	**0.11**	−3.54
q_{mixed}	0.22	−3.31	0.18	−3.32	0.25	−3.45	0.23	−3.34

[a] Optimal charge sets for specific secondary structures are emphasized.

3.2. Simulations: Influence of Modified Force Field Parameters. NMR experiments, from which scalar coupling constants $^{3h}J_{NC'}$ are determined, are by their very nature sensitive to the dynamics of this system. Therefore, MD simulations using explicit solvent were carried out to calculate ensemble-averaged $^{3h}J_{NC'}$ couplings. To better sample the conformational subspace, different starting structures were used. Snapshots, separated by 0.1 ps, along the trajectories with total lengths of 19, 15, and 10 ns for Tudor domain, ubiquitin, and protein G, respectively, were analyzed in the same way as the static structures described earlier, and an ensemble average $\langle^{3h}J_{NC'}\rangle$ was calculated. In addition to the static structures, different atomic partial charge sets were employed to investigate the influence of modified electrostatic interactions on the H-bond geometries and the conformationally averaged $^{3h}J_{NC'}$ couplings. The effect of the atomic partial charges is assessed by considering the rmsd between calculated and experimental $^{3h}J_{NC'}$ data and the ensemble distributions of the H···O′ distance, $r_{H\cdots O'}$, and the N−H···O′ angle, θ_1.

Quality of $^{3h}J_{NC'}$ Coupling Prediction. The relationship between partial charges and the rmsd between observed and calculated scalar coupling constants is tested by including all couplings ($\sum_{all}(^{3h}J_{NC'}^{calc} - {}^{3h}J_{NC'}^{obs})^2$) and distinguishing between SSEs ($\sum_{SSE}(^{3h}J_{NC'}^{calc} - {}^{3h}J_{NC'}^{obs})^2$). Thus for each charge set the optimal β, β^{opt} was determined. The optimized values β^{opt} and corresponding rmsds are summarized in Tables 4 and 5.

It is found that different charge sets affect the calculated coupling constants and also the value β^{opt} which minimizes $\sum(^{3h}J_{NC'}^{calc} - {}^{3h}J_{NC'}^{obs})^2$. Reducing the charge on the hydrogen atom by 0.04 e leads to a larger rmsd for all couplings and for couplings in particular SSEs compared to the standard charges q_0. Increasing the partial charge on the H-atom initially improves the agreement between calculated and experimentally measured $^{3h}J_{NC'}$ couplings. Depending on whether all couplings or couplings in particular SSEs are considered, the optimal charge set varies, as do the parameters β^{opt}. Considering all couplings, the lowest rmsd is found for charge set $q_{+0.1}$ (see the red line in Figure 3). Calculations for the three proteins individually yielded different optimal charge sets, namely $q_{+0.1}$ for the Tudor domain, q_0 for protein G, and $q_{+0.2}$ for ubiquitin (see Table 5). Thus, the ideal charge set depends on the particular protein. More generally, it is found that increased charges perform better. Comparison of rmsds between calculated and observed $\langle^{3h}J_{NC'}\rangle$ obtained from static structures with the ones from MD simulations showed that dynamics, i.e. conformational sampling in explicit solvent, leads to improved $\langle^{3h}J_{NC'}\rangle$ values (see Tables 2 and 5). For the Tudor domain and ubiquitin, an appreciable improvement is found. Including dynamics decreases the rmsd by 0.08 and 0.11 Hz, respectively. By contrast for protein G, $^{3h}J_{NC'}$ coupling predictions from static structures and from simulations are quite similar. This is consistent with previous simulations using a different protocol (see Discussion and Conclusions).[4]

It is also of interest to consider individual SSEs and to determine optimal charge sets and best achievable rmsds between experiment and simulations. Different SSEs favor distinct partial charges and also have significantly different parameter values (β_H^{opt} for α-helix, β_S^{opt} for β-sheet, and β_L^{opt} for loop structures). Analysis of the entire data from all three proteins in view of the SSEs yielded the following optimal charge sets: slightly enhanced charges $q_{+0.04}$ for α-helix and higher charges for β-sheet ($q_{+0.1}$) and loop regions ($q_{+0.2}$), respectively (see Figure 3 and Table 4). The corresponding β^{opt} values are $\beta_H^{opt} = -3.32$ Å^{-1}, $\beta_S^{opt} = -3.45$ Å^{-1}, and $\beta_L^{opt} = -3.54$Å^{-1}. Detailed analysis of SSEs for each protein showed that the lowering of the rmsd for the Tudor domain and ubiquitin observed from MD simulations is mainly due to improved couplings within the β-sheet structures.

On the basis of the result that different partial charges are favored by different SSEs, an additional 15 ns simulation using a mixed charge set, q_{mixed}, was run for ubiquitin. This charge set consisted of $q_{+0.04}$ for H-bonds located in α-helix structures and loops and $q_{+0.1}$ for β-sheet couplings. The use of q_{mixed} showed no obvious distortions of the protein structure as judged from the H-bond geometries ($r_{H\cdots O'}$ and θ_1) sampled during the simulation (not shown). In a first step, the $^{3h}J_{NC'}$ couplings using q_{mixed} were estimated by using optimal values for β determined above: $\beta_H = -3.32$ Å^{-1}, $\beta_S = -3.45$ Å^{-1}, and $\beta_L = -3.54$ Å^{-1}. It was found that the mixed charges, q_{mixed}, lead to inferior rmsds. Most notably, the β-sheet couplings were reproduced inaccurately, with a rmsd of 0.25 Hz as compared to 0.13 Hz from β-sheet couplings of all three proteins or even 0.07 Hz obtained with set $q_{+0.4}$ for ubiquitin. Couplings in α-helix and loop structures have comparable values for q_{mixed} and $q_{+0.04}$ charge sets. To determine whether the discrepancy of $\langle^{3h}J_{NC'}\rangle$ estimation is due to the use of the parameter β_S fitted for secondary structure and charge set with data from all three proteins, β_S was reoptimized for the current structure ensemble. With $\beta^{opt} = -3.22$ for β-sheets, an improved rmsd of 0.12 Hz was found which is in accord with 0.13 Hz rmsd from the entire data set (compare Tables 4 and 5). Finally, it is also found that increased partial charges across the H-bond not only improve the calculated $\langle^{3h}J_{NC'}\rangle$ values but they also decrease the standard deviations of the calculated couplings. Results for ubiquitin are shown in Figure 4, which are representative for all three proteins studied.

Geometric Properties of Ensemble. The influence of different atomic partial charges on the conformational ensemble can also be characterized from the MD simulations and compared with experiment. For all H-bond motifs, the

Table 5. Optimal Charge Set and β^{opt} Obtained from MD Simulations Given for Each Protein ($^{3h}J_{NC'}$ rmsd [Hz] and β^{opt} [Å^{-1}])

	Tudor			protein G			ubiquitin					
	charges	β^{opt}	rmsd	charges	β^{opt}	rmsd	charges	β^{opt}	rmsd	charges	β^{opt}	rmsd
overall	$q_{+0.1}$	−3.45	0.12	q_0	−3.18	0.14	$q_{+0.2}$	−3.45	0.12	q_{mixed}	−3.22	0.16
helix				q_0	−3.26	0.12	$q_{+0.2}$	−3.53	0.10	q_{mixed}	−3.24	0.18
sheet	$q_{+0.1}$	−3.45	0.12	q_0	−3.14	0.16	$q_{+0.2}$	−3.45	0.07	q_{mixed}	−3.22	0.12
loop							$q_{+0.2}$	−3.30	0.14	q_{mixed}	−3.10	0.22

donor−acceptor distance, $r_{H\cdots O'}$, and the H-bond directionality, θ_1, are considered separately for each protein and each charge set. Figure 5 summarizes the results for the Tudor domain, protein G, and ubiquitin. Experimental values ($r_{H\cdots O'}$, θ_1) from X-ray and NMR structures, where "experiment" refers to optimized H-positions with heavy atoms kept fixed, are indicated as blue squares and triangles, respectively, whereas structures from the simulation ensemble are individual dots (50 snapshots each coupling). The analysis was also carried out for N···O′ distances which are directly available from the experiment and the same effects, as described below, were found (see Figure S3 in the Supporting Information). Variation of the atomic partial charges influences the conformational ensemble, in particular the H-bond length $r_{H\cdots O'}$ and the H-bond angle θ_1 which are in turn used to calculate $^{3h}J_{NC'}$ couplings. Increased partial charges restrict the conformational space to H···O′ distances around 2 Å and favor linear H-bonds, i.e. $\theta_1 \approx 180°$. For increased charges, the geometry distributions from the simulations approach those from the experiments (see Figure 5). Thus, the conformational space sampled from the simulations better reflects the one observed and analyzed experimentally. The tighter geometries of both, H-bond distances and angles, is also reflected in decreased standard deviations for the $^{3h}J_{NC'}$ couplings from simulations with increased charges. This is related to increased Coulomb interactions. The improvement

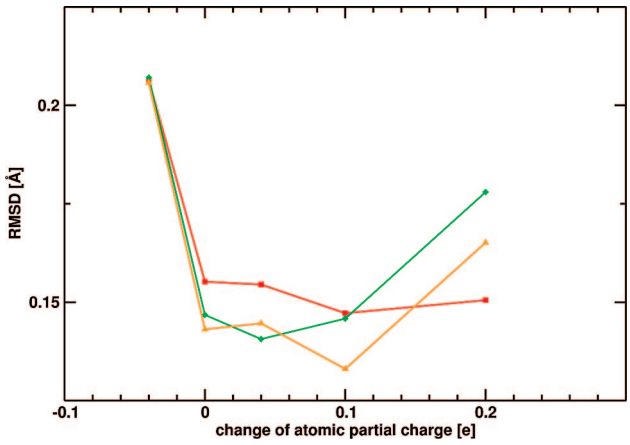

Figure 3. Performance of charge sets. The rmsd between experimental and calculated $\langle^{3h}J_{NC'}\rangle$ couplings is shown as a function of the charge set used for all couplings (red), for α-helix (green), and β-sheet (orange). Larger charge disparity leads to lower rmsd. Charge sets q_0 to $q_{+0.1}$ are suitable. In addition, the secondary structures have minor preferences for a specific charge set, i.e. $q_{+0.04}$ for α-helix and $q_{+0.1}$ for β-sheet.

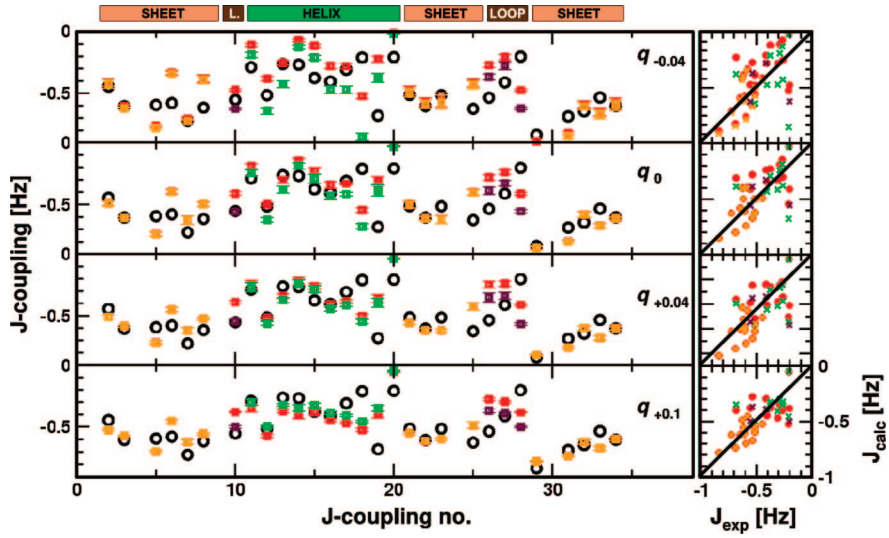

Figure 4. Calculated $^{3h}J_{NC'}$ values for ubiquitin with their respective standard errors from the different MD runs for the different charge sets ($q_{-0.04}$, q_0, $q_{+0.04}$, and $q_{+0.1}$) used in the present work. Secondary structure elements are indicated on top and predictions with optimized parameter β^{opt}_{SSE} are colored as follows: overall (red), α-helix (green), β-sheet (orange), and loop (brown). For comparison, experimental measurements are indicated as black circles. Correlation between J_{calc} and J_{exp} highlighting the quality of prediction depicted on the right. Note that for increased charge sets, the standard error decreases (left graph) and the agreement between calculated and experimentally measured $^{3h}J_{NC'}$ values is improved (right graph). In particular, parameter optimization for individual secondary structures (crosses) is superior to optimizations including all couplings (red circles). β^{opt}_{SSE} correspond to β values optimized for particular SSEs and charge sets; see Table 5.

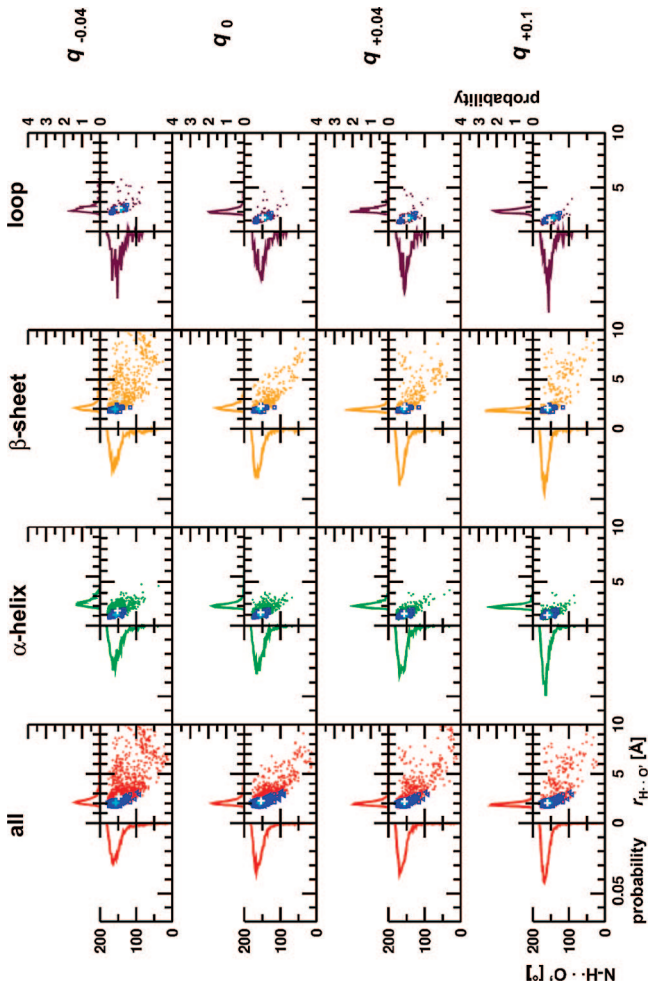

Figure 5. H-Bond geometry ensemble distributions. The geometric H-bond properties $r_{H\cdots O'}$ and θ_1 (angle N−H···O′) of 50 snapshots are shown as points for every coupling of the three proteins studied. Values found experimentally, i.e. H-position optimization while keeping heavy atom positions frozen, are indicated in blue (squares X-ray, triangles NMR). Experimental (cyan) and MD (gray) centers of gravity are shown as a cross (see Table 6). In addition, the probability distributions of the donor−acceptor distance and angle are given. Charge sets are arranged in rows and secondary structures are in columns.

of the H-bond geometries is not uniform across the different SSEs; i.e. they have different sensitivities with respect to atomic partial charges. For protein G, H-bonds located in the α-helix have a narrow $r_{H\cdots O'}$ and θ_1 distribution, close to the experimental values, which is largely independent of the charge set used. Disrupted H-bonds, characterized by a large H···O′ separation, mainly occur at the end of β-sheets. Even with an exaggerated charge difference ($q_{+0.4}$) H-bonds with very large donor−acceptor distances ($r_{H\cdots O'} > 5.0$ Å) are observed for coupling J5 of the Tudor domain and couplings J1, J6, J7, J10, J25, and J34 of protein G.

In summary, increased atomic partial charges shift the geometric ensemble toward the one found experimentally without affecting the overall dynamics (see discussion). Furthermore, the degree to which this occurs depends on the particular SSE considered. Finally, which charge set is suited to best reproduce on average (in a rmsd sense) experimentally measured $^{3h}J_{NC'}$ couplings depends on the protein studied: they are q_0 (i.e., CHARMM22 in the present case) for protein G, $q_{+0.1}$ for the Tudor domain, and $q_{+0.2}$ for ubiquitin.

3.3. Influence of H-Bond Location in Protein Structure. Although increased partial charges influence the conformational ensemble and improve the overall agreement with experiment, outliers occur. To identify outliers an arbitrary threshold of $r_{H\cdots O'} > 5.0$ Å is used. Such a value corresponds to an inexistent H-bond or a vanishing scalar coupling. As stated previously, hydrogen bonds corresponding to outliers are in general located at the end of β-sheet structures, e.g. adjacent to loop regions or chain termini. Throughout all charge sets used here no outlier (out of 29) for ubiquitin and only two (out of 15) for the Tudor domain (couplings 5 and 13) were identified. Contrary to that, for protein G, 9 couplings (out of 32) have donor−acceptor distances larger than 5.0 Å. Half of them disappear if charge sets $q_{\geq 0}$ are used. Persisting outliers in any of the increased charge sets are couplings J1, J6, J7, J10, J25, and J34. Identifying the optimal charge set for individual couplings showed that some $^{3h}J_{NC'}$ couplings are best captured with the $q_{-0.04}$ charge set, although in general stronger Coulombic interactions are superior. These are for the Tudor domain coupling no. J13, for protein G nos. J8, J13, J15, J19, J20, J27, and J33, and for ubiquitin nos. J2, J7, J23, and J33. However, their location includes α-helix and β-sheet structures where no systematic pattern is observed. In contrast, loop region H-bonds seem to require at least charge set q_0.

It is also of interest to consider whether the degree of solvent exposure of the residues involved in the H-bonds can be correlated with the quality of the calculated J-couplings. For this, the fraction of the solvent accessible surface area (SASA) to the total residue surface is calculated for each residue. A threshold of SASA >30% is typically used to define solvent exposure. Residues are considered deeply buried ("core" residues) if they have a SASA of <2%. The majority of experimentally measured H-bonds are located in the protein's interior (SASA < 30%). The few solvent-exposed couplings are in loop and α-helices. Couplings with appreciable errors (arbitrarily taken as $|^{3h}J_{NC'}^{exp} - {}^{3h}J_{NC'}^{calc}| > 0.1$ Hz) are found both in the "core" and in the surface area of the proteins. In addition, no apparent correlation between the polarity of the H-bond and the degree of solvent exposition of the residues involved in a H-bond was observed.

4. Discussion and Conclusions

In the present study, the applicability, parametrization, generality, transferability, and susceptibility to changes in the force field parameters of an empirical formula that relates geometric properties of a H-bond to measured scalar coupling constants from NMR experiments was investigated in detail. Analysis of optimized structures (starting from X-ray and NMR structures) and extensive MD simulations in explicit solvent showed that overall standard charges of the CHARMM22 force field (q_0) perform reasonably well for estimating $^{3h}J_{NC'}$ couplings with the original β_B previously derived from structures and DFT calculations alone. This is

Table 6. Geometric Centers of Gravity for Experiment and Simulations

	overall		helix		sheet		loop	
	$r_{H\cdot\cdot O'}$ [Å]	θ_1 [deg]	$r_{H\cdot\cdot O'}$ [Å]	θ_1 [deg]	$r_{H\cdot\cdot O'}$ [Å]	θ_1 [deg]	$r_{H\cdot\cdot O'}$ [Å]	θ_1 [deg]
$q_{-0.04}$	2.69	148.10	2.37	151.58	3.01	145.98	2.40	147.80
q_0	2.26	153.37	2.22	154.55	2.27	154.16	2.28	149.59
$q_{-0.04}$	2.21	155.60	2.11	156.96	2.23	156.17	2.22	151.23
$q_{+0.01}$	2.03	159.51	1.95	160.32	2.11	159.40	2.05	153.78
experiment	2.07	156.65	2.01	155.47	1.98	158.67	2.25	141.90

different if the conformational ensemble is considered where larger polarity of the H-bond shifts the distribution toward the experimentally measured one (see Table 6).

By allowing the parameter β to vary, the rmsd between calculated and experimentally determined $^{3h}J_{NC'}$ couplings can be reduced from 0.25 to 0.20 (Tudor domain), 0.14 to 0.12 (protein G), and 0.31 to 0.23 (ubiquitin), respectively. For couplings in particular SSEs, this can lead to improvements in rmsd of up to 50% (see Table 2). In all cases, β^{opt} differs appreciably from $\beta_B = -3.20$ Å^{-1}. It is also found that couplings in different SSEs are characterized by distinct values of β. Using independent multinanosecond MD simulations starting from various experimental X-ray and NMR structures, it was found that typically larger polarity of the H-bond (stronger interaction) leads to improved estimates for calculated $^{3h}J_{NC'}$ couplings. In all cases, the rmsd can be reduced to between 0.12 and 0.14 with the largest improvement for ubiquitin and the smallest one for protein G. Improvements for couplings located in particular SSEs can be even larger, such as for couplings in $\beta-$ sheets in ubiquitin which have a rmsd of 0.07 with increased polarity (from the $q_{+0.2}$ charge set). Another finding of the present work is that, considering all three proteins studied here, the lowest rmsd in α-helices and β-sheets are combinations (($q_{+0.04}$, $\beta = -3.32$ Å^{-1}) and ($q_{+0.10}$, $\beta = -3.45$ Å^{-1})), respectively. It is worthwhile to point out that the rmsds from using the "standard" parameters (q_0, $\beta = -3.20$ Å^{-1}) in both cases are only slightly worse. However, significant differences are found for the distribution of the conformational ensemble which hint toward a substantial improvement if larger charges are used (see Figure 5). It is instructive to compare "centers of gravity"—$r_{H\cdots O'}$ and θ_1 averages—for all hydrogen bonds in different SSEs from experiment and simulations (see Table 6). As can be expected, charge variation has a more pronounced effect on the H···O′ distance compared to the N−H···O′ angle. Increased polarity ($q_{+0.04}$ and $q_{+0.1}$) captures the experimentally measured average $r_{H\cdots O'}$ distance even though H-bond linearity is somewhat overestimated. Application of different atomic charges does not distort the protein structure which is confirmed by rmsd values below 2 Å compared to the starting structure (see Table S5 in the Supporting Information). In addition, calculated *B*-factors, which characterize flexible and rigid parts in the protein, agree quite well with experiment independent of the charges used (see Figure S4 in the Supporting Information). Together these results underline the applicability of modified atomic charges without altering the protein dynamics. As mentioned in the introduction, the expectation that modified partial charges across the H-bond lead to better scalar coupling constants rests on the observation that the geometry and energetics of a hydrogen bond is primarily described by electrostatic interactions in a force field. However, it might also be of interest to consider modified van der Waals parameters and their effect on calculated scalar coupling constants.

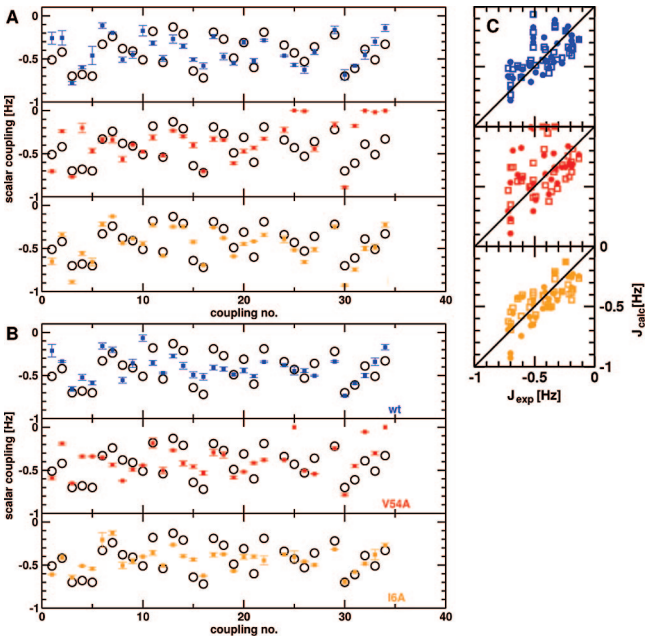

Figure 6. Coupling constants for protein G mutants V54A and I6A involved in coupling. For charge sets q_0 (A) and $q_{+0.1}$ (B), the calculated $^{3h}J_{NC'}$ values are shown for protein G wildtype (blue) and for the J32 mutants V54A (red) and I6A (orange). Experimental data for the wild type protein are shown as black circles. (C) Correlation between J_{calc} and J_{exp} for charge set q_0 (filled circle) and $q_{+0.1}$ (open square).

It is also of interest to compare the results from the present simulations with previous work. In a recent study, the same three proteins were studied with the same force field (CHARMM22) using NAMD and particle mesh Ewald (PME) for treating the long-range electrostatics.[4] This study was mainly concerned with assessing the differences between simulations in implicit and explicit solvent and with the question whether carrying out MD simulations (trajectory length between 0.5 and 1.4 ns) improves the conformational ensemble as reflected by the measured $^{3h}J_{NC'}$ coupling constants. For the static structures (see Table 1 and ref 4), rmsds were 0.15, 0.20, and 0.23 for protein G (structure 2IGD), ubiquitin (1UBQ), and the Tudor domain (1MHN), respectively. This compares with 0.14, 0.31, and 0.25 from the present work which, however, are averages over several structures (see section 3.1). Including dynamics, the earlier study found rmsds of 0.14, 0.15, and 0.30 (0.15 for rescaled

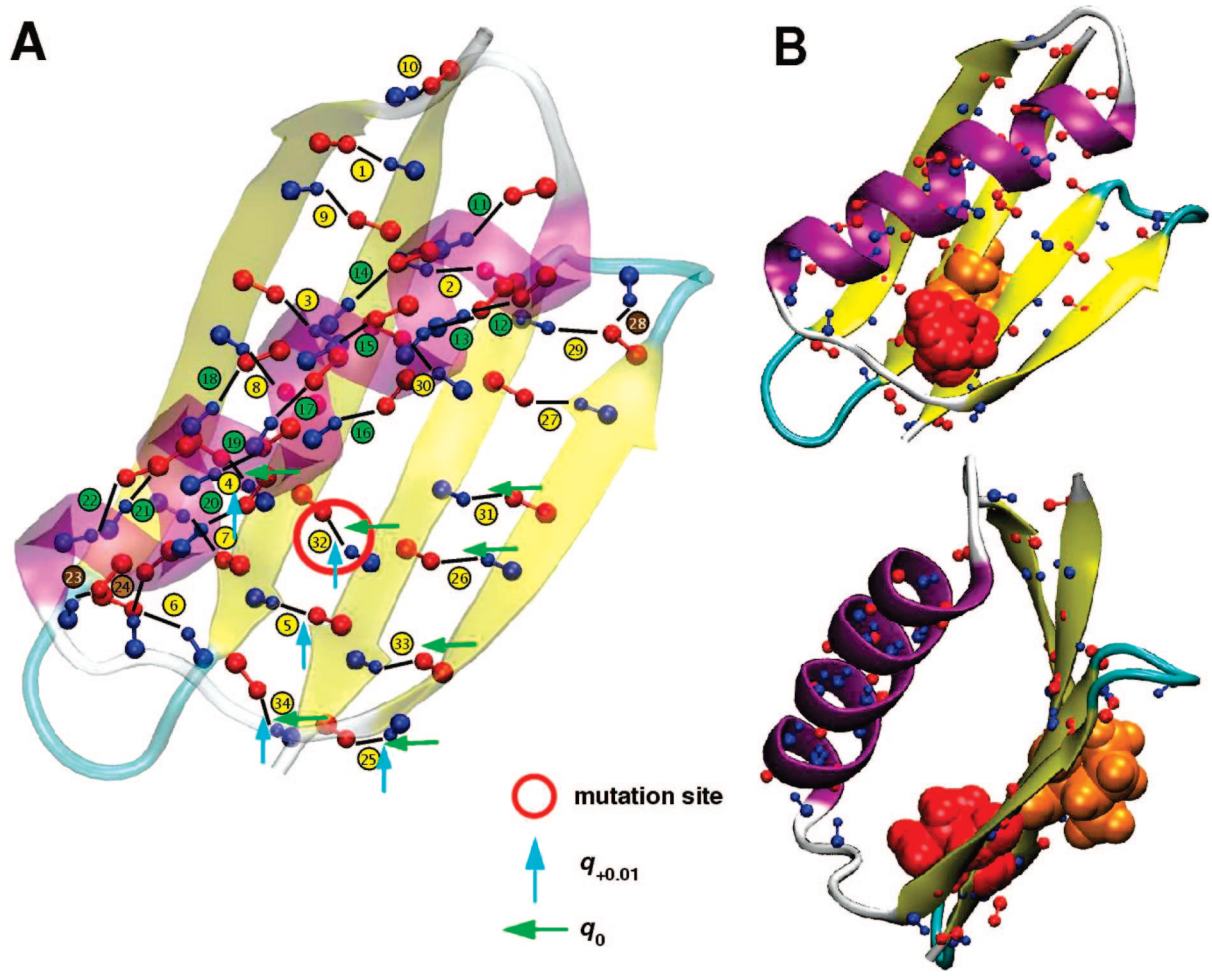

Figure 7. Differences for calculated coupling constants for protein G mutants V54A and I6A. (A) Coupling constants which differ by more than 0.3 Hz compared to experimental values for the wild type protein indicated as arrows (green for q_0 and blue for $q_{+0.1}$). As can be seen, couplings away from J32 are also more or less affected depending on the charge set used. The mutation site is shown as a red circle. (B) Spatial orientation of the mutated sidechains V54 (red) and I6 (orange).

$^{3h}J_{NC'}$ coupling constants)[4] Hz for protein G, ubiquitin, and the Tudor domain, respectively. This is in almost quantitative agreement with the present work which yields rmsds of 0.14, 0.15, and 0.25 (0.21) Hz using the same force field and the same parametrization to calculate $^{3h}J_{NC'}$ coupling constants ($\beta_B = 3.20$, q_0). This is quite remarkable as the simulation approaches differ. Here, electrostatic cutoffs, multiple trajectories and extended simulation times (between 10 and 19 ns) are used compared to PME with 1 trajectory (except for protein G) and simulation times between 0.5 and 1.4 ns in the earlier study.[4]

Using biased MD simulations, it was previously shown that by including NMR energy constraints, (e.g., $^{3h}J_{NC'}$, NOE, and S^2) to the force field, shorter donor−acceptor distances and a more pronounced directionality of the H-bond are favored and lead to an ensemble that better reflects the experimentally measured one.[5] In this approach, one is interested in whether experimental constraints move the conformational ensemble from MD simulations toward the experimentally observed one. However, as the constraints are sequence- and amino acid-specific, little can be inferred about the particular merits and weaknesses of the underlying force field and predicting expected scalar couplings upon (point) mutation is not possible.

On the basis of the validation simulation discussed above, $^{3h}J_{NC'}$ coupling constants for particular protein G mutants were calculated. The donor and acceptor residues for H-bond number 32 (V54 → I6) were computationally mutated to alanine. For both systems independent 5 ns simulations were carried out with charge sets q_0 and $q_{+0.1}$. This choice was motivated by the observation that q_0 was previously found to be optimal for protein G and for all three proteins investigated here slightly increased charges ($q_{+0.1}$) improve the correlation between experiment and simulations. The estimated scalar couplings for mutant I6A are similar to the wildtype, whereas mutant V54A shows significant differences (see Figure 6). For the latter, with either charge set, couplings number J4, J5, J25, J26, J31, J32, J33, and J34 differ by more than 0.3 Hz from the experiment of the wild type protein (see Figure 7A). These couplings are located in the vicinity of the mutation site, more precisely in the β-sheet toward the sheet end. Using q_0, scalar couplings vanish for couplings number J25, J26, J32, J33, and J34 due to loss of β-sheet character in the affected strands. Such effects should be observable experimentally.

In conclusion, larger polarity (i.e., stronger donor−acceptor interactions) of the H-bond improves the calculated $^{3h}J_{NC'}$ coupling constants and the conformational ensemble sampled

from the MD simulations. The parameter β_B is not universal across different proteins and different SSEs and optimized values moderately improve the calculated $^{3h}J_{NC'}$ coupling constants. The present work also suggests that scalar coupling constants can also be used to improve particular force field parameters. This will add a new aspect to force field development as NMR properties are intrinsically dynamical in nature. With additional measurements of scalar coupling constants in proteins (also on mutants of the ones discussed here), the available data could be considerably extended and a more exhaustive fit will be possible. Other dynamical quantities which are sometimes included are diffusion constants or, more recently, neutron structure factors[18] or NMR spin relaxation experiments.[19] In the future, it might be of interest to use force fields that take into account more details of the intermolecular interactions. One possibility is to use molecular mechanics with explicit proton transfer (MMPT) potentials which are 3-dimensional representations of model potentials for proton- or hydrogen-bonded motifs.[20,21] Another refinement that could be envisaged is to use higher multipole moments on the atoms involved in the H-bond. Such an extension has recently been found to improve the quality of simulations for infrared spectra of CO in myoglobin.[22]

Acknowledgment. We thank Prof. S. Grzesiek and Dr. I. Tubert-Brohman for valuable discussions and comments. This work was supported through SNF grant 200021-117810.

Supporting Information Available: Details on the structures for ab initio calculations, N···O′ distances, *B* factors, and rmsd derived from simulations and data on H-bonds studied in this work (PDF format). This material is available free of charge via the Internet at http://pubs.acs.org.

References

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

(2) Barfield, M. *J. Am. Chem. Soc.* **2002**, *124*, 4158.

(3) Markwick, P. R. L.; Sprangers, R.; Sattler, M. *J. Am. Chem. Soc.* **2003**, *125*, 644.

(4) Sass, H.-J.; Schmid, F.-F.; Grzesiek, S. *J. Am. Chem. Soc.* **2007**, *129*, 5898–5903.

(5) Gsponer, J.; Hopearuoho, H.; Cavalli, A.; Dobson, C. M.; Vendruscolo, M. *J. Am. Chem. Soc.* **2006**, *128*, 15127–15135.

(6) MacKerell, A. D., Jr; Bashford, D.; Bellott, M.; Dunbrack, R. L., Jr.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E., III; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J. E.; Watanabe, M.; Wiorkiewicz-Kuczera, J.; Yin, D.; Karplus, M. *J. Phys. Chem. B* **1998**, *102*, 3586.

(7) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187–217.

(8) Sprangers, R.; Groves, M. R.; Sinning, I.; Sattler, M. *J. Mol. Biol.* **2003**, *327*, 507.

(9) Selenko, P.; Sprangers, R.; Stier, G.; Buehler, D.; Fischer, U.; Sattler, M. *Nat. Struct. Biol.* **2001**, *8*, 27.

(10) Gallagher, T.; Alexander, P.; Bryan, P.; Gilliland, G. L. *Biochemistry* **1994**, *33*, 4721.

(11) Derrick, J. P.; Wigley, D. B. *J. Mol. Biol.* **1994**, *243*, 906.

(12) Butterworth, S.; Lamzin, V. S.; Wigley, D. B.; Derrick, J. P.; Wilson, K. S. Anisotropic Refinement of a Protein G Domain at 1.1 Angstrom Resolution, in press.

(13) Cornilescu, G.; Ramirez, B. E.; Frank, M. K.; Clore, G. M.; Gronenborn, a. M.; Bax, A. *J. Am. Chem. Soc.* **1999**, *121*, 6275.

(14) Vijay-Kumar, S.; Bugg, C. E.; Cook, W. J. *J. Mol. Biol.* **1987**, *194*, 531.

(15) Cornilescu, G.; Marquardt, J. L.; Ottiger, M.; Bax, A. *J. Am. Chem. Soc.* **1998**, *120*, 6836.

(16) Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. C. *J. Chem. Phys.* **1977**, *23*, 327–341.

(17) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A., Jr.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. *Gaussian 03*, revision B.01; Gaussian, Inc.: Wallingford, CT, 2004.

(18) Thomas, J. L.; Tobias, D. J.; MacKerell, A. D., Jr. *J. Phys. Chem. B Lett.* **2007**, *111*, 12941.

(19) Showalter, S. A.; Brüschweiler, R. *J. Chem. Theor. Comp.* **2007**, *3*, 961–975.

(20) Lammers, S.; Meuwly, M. *J. Phys. Chem. A* **2007**, *111*, 1638.

(21) Lammers, S.; Lutz, S.; Meuwly, M. *J. Comput. Chem.* **2008**, *29*, 1048.

(22) Plattner, N.; Meuwly, M. *Biophys. J.* **2008**, *94*, 2505.

CT800241D