J|A|C|S ARTICLES

Published on Web 06/01/2009

Simulation of NMR Data Reveals That Proteins' Local Structures Are Stabilized by Electronic Polarization

Yan Tong,[†] Chang G. Ji,[†] Ye Mei,[‡] and John Z. H. Zhang*[,‡,§]

Institute of Theoretical and Computational Chemistry, Key Laboratory of Mesoscopic Chemistry of the Ministry of Education, School of Chemistry and Chemical Engineering, Nanjing University, Nanjing 210093, China, State Key Laboratory of Precision Spectroscopy, Department of Physics, East China Normal University, Shanghai 200062, China, and Department of Chemistry, New York University, New York, New York 10003

Received March 10, 2009; E-mail: john.zhang@nyu.edu

Abstract: Molecular dynamics simulations of NMR backbone relaxation order parameters have been carried out to investigate the polarization effect on the protein's local structure and dynamics for five benchmark proteins (bovine pancreatic trypsin inhibitor, immunoglobulin-binding domain (B1) of streptococcal protein G, bovine apo-calbindin D9K, human interleukin-4 R88Q mutant, and hen egg white lysozyme). In order to isolate the polarization effect from other interaction effects, our study employed both the standard AMBER force field (AMBER03) and polarized protein-specific charges (PPCs) in the MD simulations. The simulated order parameters, employing both the standard nonpolarizable and polarized force fields, are directly compared with experimental data. Our results show that residue-specific order parameters at some specific loop and turn regions are significantly underestimated by the MD simulations using the standard AMBER force field, indicating hyperflexibility of these local structures. Detailed analysis of the structures and dynamic motions of individual residues reveals that the hyperflexibility of these local structures is largely related to the breaking or weakening of relevant hydrogen bonds. In contrast, the agreement with the experimental results is significantly improved and more stable local structures are observed in the MD simulations using the polarized force field. The comparison between theory and experiment provides convincing evidence that intraprotein hydrogen bonds in these regions are stabilized by electronic polarization, which is critical to the dynamical stability of these local structures in proteins.

I. Introduction

Proteins in solution exhibit a variety of motions, ranging from local atomic fluctuations and bond oscillations to hinge bending motions, helix−coil transitions, and local and global unfolding processes. Both experimental and theoretical techniques are used extensively in attempts to characterize these dynamic fluctuations and understand their role in protein function and specificity. The internal dynamics plays an important role in the interplay between protein structure and function.[1] Molecular dynamics (MD) simulations readily provide detailed information on dynamical processes in proteins. Explicit and detailed comparison of the results of MD simulations with experimental data can provide much-needed insight into the internal motion and dynamics of proteins. NMR relaxation measurements have long been recognized as a benchmark crucial for establishing confidence in the accuracy of the simplified classical-mechanical description of intramolecular forces used in MD simulations,[2] which plays a central role in the analysis and interpretation of biomolecular NMR data.[3] Accurate and detailed descriptions of protein internal motions by MD simulations can provide unique insights into the time dependence of conformational fluctuations, especially on pico- to nanosecond time scales, which are directly accessible by MD simulations. Reliable detection of coupled motions can be used to quantify the conformational entropy of a protein from NMR spin relaxation. To this end, MD simulations of NMR order parameters provide an effective means for studying intraprotein dynamics by linking structural information from experiments with the mechanisms of the protein on the atomic scale. ^{15}N relaxation methods probe the motions of ^{1}H−^{15}N bonds in proteins (i.e., the backbone amide NH group and the side-chain NH groups of Asn, Gln, Trp, and Arg).

However, the reliability of simulation results depends on the accuracy of the force fields employed. The quality of the force field in an MD simulation is crucial for obtaining agreement between calculated and experimental order parameters.[4] Existing studies have shown evidence for some specific links between force-field deficiencies and disagreement between experimental and MD order parameters for some protein systems that cannot be reconciled simply by longer simulation times. Since electrostatic effects play a major role in determining the energetics and dynamics of biomolecules,[5−8] the ability to quantify

[†] Nanjing University.
[‡] East China Normal University.
[§] New York University.

(1) Ringe, D.; Petsko, G. A. *Prog. Biophys. Mol. Biol.* **1985**, *45*, 197.
(2) McCammon, J. A.; Gelin, B. R.; Karplus, M. *Nature* **1977**, *267*, 585.
(3) Jarymowycz, V. A.; Stone, M. J. *Chem. Rev.* **2006**, *106*, 1624.
(4) Showalter, S. A.; Brüschweiler, R. *J. Chem. Theory Comput.* **2007**, *3*, 961.
(5) Davis, M. E.; McCammon, J. A. *Chem. Rev.* **1990**, *90*, 509.
(6) Honig, B.; Nicholls, A. *Science* **1995**, *268*, 1144.

 10.1021/ja901650r CCC: $40.75

electrostatic interactions is essential for an accurate description of protein dynamics in solution and for structure−function correlation studies of proteins. Because conventional force fields lack the polarization effect, their accuracy and reliability are expected to have limitations. Although efforts have been made to develop polarizable force fields, their success has been quite limited because of practical difficulties.[9] Recently, we developed polarized protein-specific charges (PPCs) based on a fragmentation scheme (the MFCC scheme)[10−14] for electronic structure calculations on biomolecules and the continuum dielectric model for the solvent in a self-consistent manner.[15] On the basis of this approach, we developed PPCs that incorporate electrostatic polarization effects of proteins in their native structures into the atomic charge fitting. Its advantage over conventional nonpolarizable force fields has been demonstrated by a number of applications and explicit comparison of results with experimental data.[15−17] PPCs correctly describe the polarized electrostatic state of a protein and provide accurate electrostatic interactions in the vicinity of its native structure.

In this study, MD simulations of NMR order parameters were performed for five proteins [bovine pancreatic trypsin inhibitor (BPTI), immunoglobulin-binding domain (B1) of streptococcal protein G (GB1), bovine apo-calbindin D9K (calbindin), human interleukin-4 R88Q mutant (IL4), and hen egg white lysozyme (lysozyme)] using AMBER03 charges and PPCs. These proteins are benchmark systems that have received extensive experimental interest for various purposes.[4,18−26] Comparison between the current MD simulation results and measured experimental NMR order parameters was explicitly made, revealing new physical insight into the dynamical properties of intraprotein hydrogen bonding.

II. Computational Approach

A. Derivation of PPCs for Proteins. PPCs were fitted to electrostatic potentials (ESPs) by fragment quantum-mechanical calculations using an iterative approach, as described in ref 15. Specifically, each protein is first divided into capped amino-acid fragments using the MFCC approach[10] for quantum-chemical calculations in solvent. The continuum solvent model was employed, and solvent polarization was calculated with the Poisson−Boltzmann (PB) solver Delphi[27] and represented as induced charges on the solute−solvent interface with a probe radius of 1.4 Å.[28] The polarization of each protein fragment due to the rest of the protein fragments and the solvent was explicitly included as an external point charge in the fragment quantum calculation. The restrained electrostatic potential (RESP) method[29,30] was employed to fit fragment atomic charges to generate the ESP, and these newly fitted atomic charges were passed to the next round of the fragment quantum calculation. The process was iterated until the calculated dipole moment of the protein and the surface charges both converged and their variations were smaller than a certain criterion. In this way, accurate estimates of the electrostatic solvation free energies of proteins in water could be obtained. The solvent dielectric constant was set to 80, and a grid density of 4.0 grids/Å was used in numerically solving the PB equation. The quantum-chemical calculations on the protein fragments were performed using density functional theory at the B3LYP/6-31G* level.

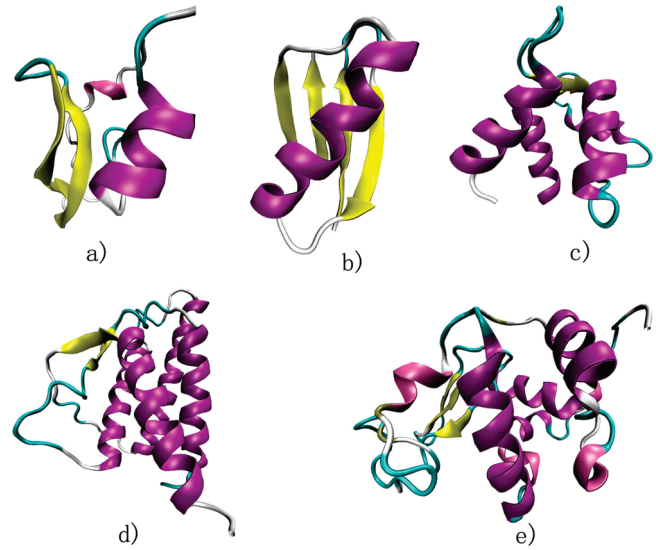

Figure 1. Model 1 from the ensemble of solution structures for (a) bovine pancreatic trypsin inhibitor (BPTI), (b) immunoglobulin-binding domain (B1) of streptococcal protein G (GB1), and (c) bovine apo-calbindin D9K (calbindin) and the crystal structures of (d) human interleukin-4 R88Q mutant (IL4) and (e) hen egg white lysozyme (lysozyme). α-Helices are shown in purple and β-sheets in yellow. Graphics were prepared using VMD.

B. Protein Systems and MD Simulations. We chose five benchmark protein systems (BPTI, GB1, calbindin, IL4, and lysozyme) on which MD simulations were performed for direct comparison of simulation results with experimental data. The native structures of these proteins are shown in Figure 1. The initial structures of these proteins used in the present calculation were taken from the Protein Data Bank (PDB entry codes 1pit, 2gb1, 1clb, 1hij, and 6lyt, respectively). Hydrogen atoms were added to the proteins using the Leap module in the AMBER 9 package.[30] The proteins were soaked in a periodic box of TIP3P water with the minimum distances from the protein atoms to the surfaces of the boxes set to 10 Å. Counterions were added to neutralize the

(7) Perutz, M. F. *Science* **1978**, *201*, 1187.
(8) Matthew, J. B. *Annu. Rev. Biophys. Chem.* **1985**, *14*, 387.
(9) Jorgensen, W. L. *J. Chem. Theory Comput.* **2007**, *3*, 1877.
(10) Zhang, D. W.; Zhang, J. Z. H. *J. Chem. Phys.* **2003**, *119*, 3599.
(11) Gao, A. M.; Zhang, D. W.; Zhang, J. Z. H.; Zhang, Y. *Chem. Phys. Lett.* **2004**, *394*, 293.
(12) Mei, Y.; Zhang, D. W.; Zhang, J. Z. H. *J. Phys. Chem. A* **2005**, *109*, 2.
(13) Mei, Y.; Wu, E. L.; Han, K. L.; Zhang, J. Z. H. *Int. J. Quantum Chem.* **2006**, *106*, 1267.
(14) Mei, Y.; Ji, C. G.; Zhang, J. Z. H. *J. Chem. Phys.* **2006**, *125*, 094906.
(15) Ji, C. G.; Mei, Y.; Zhang, J. Z. H. *Biophys. J.* **2008**, *95*, 1080.
(16) Ji, C. G.; Zhang, J. Z. H. *J. Am. Chem. Soc.* **2008**, *130*, 17129.
(17) Duan, L. L.; Mei, Y.; Zhang, Q. G.; Zhang, J. Z. H. *J. Chem. Phys.* **2009**, *130*, 115102.
(18) Sheinerman, F. B.; Brooks, C. L., III. *Proteins: Struct., Funct., Genet.* **1997**, *29*, 193.
(19) York, D. M.; Wlodawer, A.; Pedersen, L. G.; Darden, T. A. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 8715.
(20) Daggett, V.; Levitt, M. *Annu. Rev. Biophys. Biomol. Struct.* **1993**, *22*, 353.
(21) Harder, E.; Kim, B.; Friesner, R. A.; Berne, B. J. *J. Chem. Theory Comput.* **2005**, *1*, 169.
(22) Price, D. J.; Brooks, C. L., III. *J. Comput. Chem.* **2002**, *23*, 1045.
(23) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. *Proteins: Struct., Funct., Bioinf.* **2006**, *65*, 712.
(24) Buck, M.; Bouguet-Bonnet, S.; Pastor, R. W.; MacKerell, A. D. *Biophys. J.* **2006**, *90*, L36.
(25) Stocker, U.; van Gunsteren, W. F. *Proteins: Struct., Funct., Genet.* **2000**, *40*, 145.
(26) Soares, T. A.; Daura, X.; Oostenbrink, C.; Smith, L. J.; van Gunsteren, W. F. *J. Biomol. NMR* **2004**, *30*, 407.
(27) Rocchia, W.; Sridharan, S.; Nicholls, A.; Alexov, E.; Chiabrera, A.; Honig, B. *J. Comput. Chem.* **2002**, *23*, 128.
(28) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179.
(29) Bayly, C. I.; Cieplak, P.; Cornell, W. D.; Kollman, P. A. *J. Phys. Chem.* **1993**, *97*, 10269.
(30) Case, D. A.; et al. *AMBER 9*; University of California: San Francisco, 2006.

system. The MD simulations with PPCs were quite straightforward: we simply replaced the standard charges of the AMBER03 force field[31] with PPCs while keeping the rest of the AMBER parameters intact. This allowed us to cleanly examine the electrostatic effects on protein structure and dynamics by comparing the MD results obtained separately using the PPC and AMBER force fields. All of the MD simulations were performed using the AMBER 9 package, and the system was relaxed in a standard equilibration procedure. In the first step, only the solvent molecules were free to move, while protein atoms were constrained by an external force. In the second step, the whole system was energy-minimized until convergence was reached. After that, the system was heated to 300 K in 100 ps, and then a 5 ns NPT simulation with a time step of 2 fs was performed. The last 3.5 ns of each trajectory was used for analysis. The particle-mesh Ewald method[32] was used to treat long-range electrostatic interactions, while the van der Waals interactions were truncated at 12 Å. Langevin dynamics[33] was applied to regulate the temperature with a collision frequency of 1.0 ps^{-1}. The SHAKE algorithm was applied to all bonds involving hydrogen atoms during the MD simulations.

Backbone order parameters (S^2) are the plateau values of the autocorrelation function of the second-order Legendre polynomial of the N−H vector and are derived from NMR relaxation data. They reflect the mobility of the backbone at a given amino acid. Lower S^2 values reflect increased backbone flexibility. Order parameters were computed from the trajectories using the formula

$$S^2 = \frac{1}{2}\left[3\sum_{\alpha=1}^{3}\sum_{\beta=1}^{3}\langle\mu_\alpha\mu'_\beta\rangle^2 - 1\right] \quad (1)$$

where $\alpha,\beta = 1$, 2, and 3 denote x, y, and z, respectively, μ_α and μ'_β are the α and β components of the normalized interatomic vector $\boldsymbol{\mu}$ in the molecular frame at times t and t', respectively, and the average is performed over simulation time.[34,35]

III. Results and Discussion

The backbone order parameters S^2 obtained from MD simulations can be used for comparison of force fields in many systems.[22−24] The root-mean-square deviations (rmsd's) of the S^2 values calculated using PPCs and AMBER03 charges from the experimentally measured ones are given in Table 1. These numbers show that the order parameters calculated using PPCs are generally in better agreement with the experimental results than those using AMBER charges. Since the only difference in these two simulations was the charges (the other parameters in the AMBER force field remained the same), the improvement in the calculated order parameters is clearly due to the electrostatic or polarization effects embedded in the PPCs.

Since the rmsd measures only the average deviation of the order parameters for all of the backbone atoms, a more sensitive probe involves examination of the order parameters for individual residues. For this purpose, we show a comparison of residue-specific order parameters in Figure 2, in which order parameters calculated from the current simulations and from previous simulations by other groups[22−24] are plotted together for comparison with experimental values.[22,24,36,37] We give specific comparisons and analysis for each of the five proteins we studied.

Table 1. Root-Mean-Square Deviations of Backbone Order Parameters (S^2) Calculated Using PPCs and AMBER03 Charges from the Experimentally Derived S^2 Values

	BPTI	GB1	calbindin	IL4	lysozyme
PPC	0.056	0.107	0.101	0.101	0.072
AMBER03	0.093	0.099	0.131	0.132	0.080

A. BPTI. BPTI is a small disulfide-bonded protein with 58 residues that has been extensively studied by nearly all of the modern tools in protein science, thus making it an ideal system for investigating interrelationships among protein structure, dynamics, and stability.[36] Comparisons between simulation and experimentally derived main-chain S^2 values are shown in Figure 2a. Overall, the simulation using PPCs clearly yielded much improved agreement between S^2 and experiment. The rmsd's of the computed S^2 values from the experimental ones were 0.056 and 0.093 using PPCs and AMBER charges, respectively. More importantly, the residue-specific results show that the AMBER charges significantly underestimated the experimental order parameters at selected backbones, especially at Phe45, Cys14, and Cys38, as seen clearly in Figure 2a. The much smaller S^2 values at these residues simply mean that these local structures are too flexible. Analysis of protein structure shows that these local backbones are located in the isolated β-bridge and several of the coil regions.

In the native structure of BPTI, Phe45 forms two hydrogen bonds with Tyr21 and Cys14 forms a disulfide bridge with Cys38. Upon examination of the time evolution of the length of the $N^{45}-H^{45}\cdots O^{21}$ hydrogen bond (Figure 3a,b), a larger fluctuation is seen in the AMBER simulation, which results in overflexibility of the corresponding backbone structure. Further analysis shows that changes in local structure around the Cys14−Cys38 disulfide linkage also occur, as shown in Figure 4. Clearly, the local structure is better preserved in the PPC simulation, while large deviations of some local structures are seen in the simulation using AMBER03 charges. Thus, the underestimation of the order parameters at these regions in the AMBER03 simulation are associated with the instability of the hydrogen bonding and the partial denaturing of the local structure due to the lack of polarization.

In order to examine whether the simulation was converged with respect to simulation time, we plotted the order parameters calculated from the first 1.5 ns, the first 2.5 ns, and the full 3.5 ns of the simulation for BPTI using PPCs (Figure 2f). There is essentially no difference among these results, demonstrating that our simulation was converged with respect to simulation time.

B. GB1. The same analysis was performed for GB1, which is a small, well-characterized protein of 56 residues with a known NMR structure that contains no disulfide bonds or prosthetic groups. Topologically, GB1 consists of a four-stranded β-sheet and a long α-helix tightly packed against each other, as shown in Figure 1b. The rmsd's of the simulated order parameters from the NMR-measured value were 0.107 using PPCs and 0.099 using AMBER03 charges. Although there was not much difference in the overall rmsd's, there were some large

(31) Duan, Y.; Wu, C.; Chowdhury, S.; Lee, M. C.; Xiong, G. M.; Zhang, W.; Yang, R.; Cieplak, P.; Luo, R.; Lee, T.; Caldwell, J.; Wang, J. M.; Kollman, P. A. *J. Comput. Chem.* **2003**, *24*, 1999.
(32) Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993**, *98*, 10089.
(33) Pastor, R. W.; Brooks, B. R.; Szabo, A. *Mol. Phys.* **1988**, *65*, 1409.
(34) Smith, L. J.; Markand, A. E.; Dobson, C. M.; van Gunsteren, W. F. *Biochemistry* **1995**, *34*, 10918.
(35) Chandrasekhar, I.; Clore, G. M.; Szabo, A.; Gronenborn, A. M.; Brooks, B. R. *J. Mol. Biol.* **1992**, *226*, 239.
(36) Beeser, S. A.; Goldenberg, D. P.; Oas, T. G. *J. Mol. Biol.* **1997**, *269*, 154.
(37) Barchi, J. J., Jr.; Grasberger, B.; Gronenborn, A. M.; Clore, G. M. *Protein Sci.* **1994**, *3*, 15.

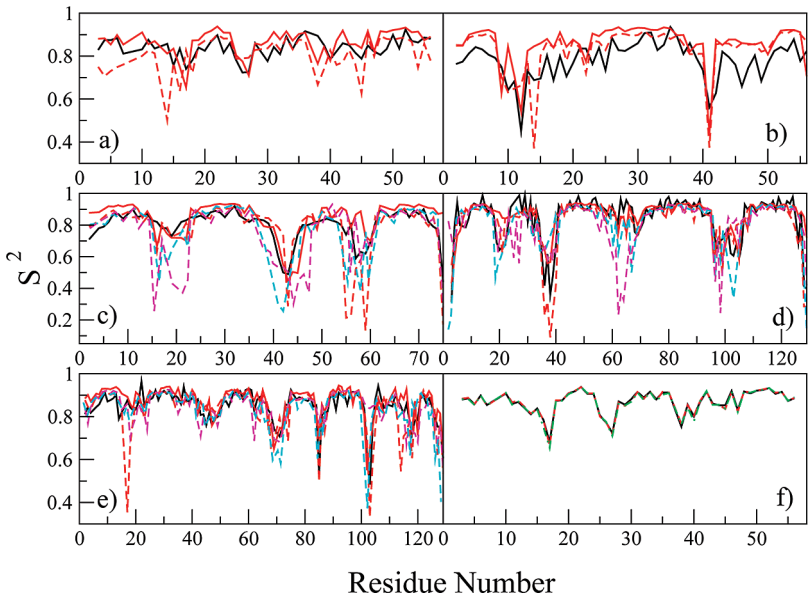

Figure 2. Order parameters for (a) BPTI, (b) GB1, (c) calbindin, (d) IL4, and (e) lysozyme. Experimental values are given by black lines, PPC results by red lines and AMBER03 results by dashed red lines. In (c) and (d), dashed lines in cyan and magenta denote simulation results from ref 22 calculated using the CHARMM22 and OPLS-AA force fields, respectively. In (e), dashed lines in cyan and magenta denote simulation results from refs 23 and 24, respectively, calculated using the ff99SB[23] and CHARMM22/CMAP[24] force fields. (f) Order parameters calculated using the first 1.5 ns (black), the first 2.5 ns (red), and the full 3.5 ns (green) of the trajectories in the simulation of BPTI using PPCs.

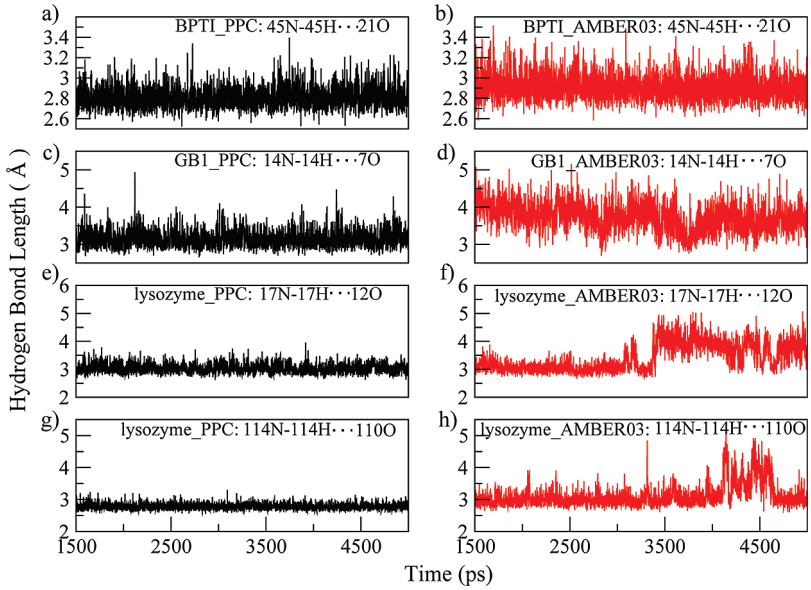

Figure 3. Time evolution of the lengths of selected hydrogen bonds during MD simulations of BPTI, GB1, and lysozyme: PPC (black, left column); AMBER (red, right column).

deviations in order parameters for specific residuals. For example, the calculated order parameters for GLY14 (which belongs to one strand of GB1) using AMBER03 charges were significantly smaller than the experimental value. Analysis shows that the corresponding backbone hydrogen bond $N^{14}-H^{14}\cdots O^{7}$ was overstretched with a much larger fluctuation in the simulation using AMBER03 charges relative to the PPC results, as shown in Figure 3c,d. This resulted in hyperflexibility of the local backbone structure. Similar hyperflexibility in the backbone motion was also seen at residue 41 in Figure 2b.

C. Calbindin. Calbindin is a 76-residue, right-handed, four-helix bundle with two small β-strands and is involved in calcium uptake from the stomach. The rmsd's of the simulated S^2 values of calbindin from the experimental value were 0.101 using PPCs and 0.131 using AMBER03 charges. Hyperflexibility in the backbone can be seen from the 55th to the 59th residues in the simulation using AMBER03 charges in Figure 2c. The S^2 values of five residues in the first loop (Asp55 to Gly59) and Gly43 in the second loop (Pro37 to Thr45) were significantly smaller than experimental values. The hyperflexibility of these backbone residues resulted in significant local structure changes. Figure 5A shows the final simulated structures of calbindin under the separate simulations using PPCs and AMBER03 charges overlaid onto the native structure; it is evident that the corresponding local structure changed significantly around residues Asp55−Gly59 using AMBER03 charges and that partial denaturing occurred there. In contrast, these local structures experienced much smaller fluctuations and were much

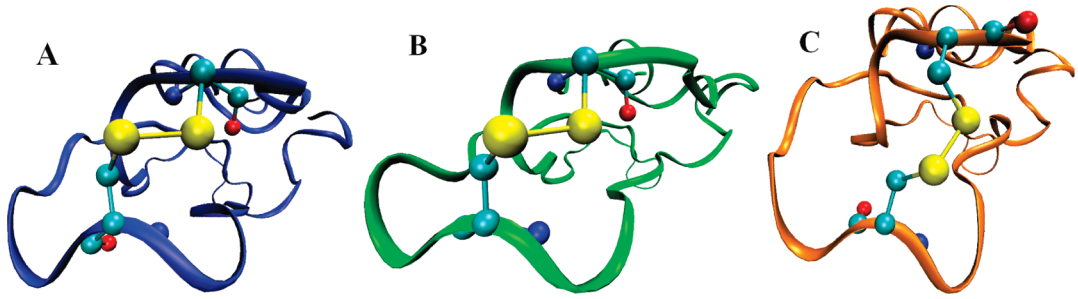

Figure 4. Local structures around the Cys14−Cys38 disulfide bond in BPTI: (A) native structure (blue) and final structures from MD simulations using (B) PPCs (green) and (C) AMBER03 charges (orange). Cys14 and Cys38 (excluding hydrogen atoms) are shown in the ball-and-stick representation.

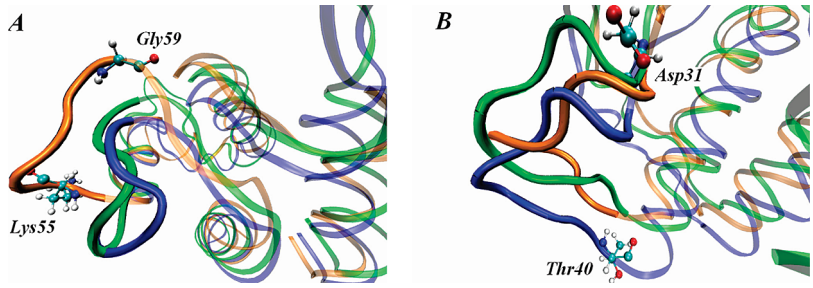

Figure 5. Overlay of the simulated structures using PPCs (green) and AMBER03 charges (orange) on the native structure (blue) for (A) calbindin and (B) IL4. The hyperflexibility and partly denatured local structure under the AMBER force field is seen in (A).

better preserved in the PPC simulation. The significantly improved agreement between the experimental values and the order parameters from the polarized force field simulation at these residues provides strong proof that these local structures are indeed preserved in experiment.

Figure 2c also shows simulated order parameters obtained using the CHARMM22 and OPLS-AA force fields.[22] It can be seen that incorrect hyperflexibility also exists in these simulations, and although some of these affected backbones are located in different parts of the protein, the overall result is generally similar to the case of simulation using AMBER.

D. IL4. IL4 is a four-helix bundle with two β-strands. The bundle is left-handed and contains two peculiar overhand connections that may be representative of the superfamily of cytokines to which it belongs. It has three disulfide bonds that are expected to limit mobility considerably. For IL4, the rmsd's of the calculated S^2 values from the experimental value were 0.101 for PPCs and 0.132 for AMBER03 charges. As shown in Figure 2d, hyperflexibility can be seen in one of the loops (Asp31−Thr40) in the simulation using AMBER03 charges. The S^2 values for residues Ala35−Asn38 in this region calculated using AMBER charges were smaller than those from the experimental NMR data. Local structural changes were seen in the simulations using PPCs and AMBER charges, as shown in Figure 5B, indicating inherent flexibility in this loop region, but the smaller S^2 values obtained using AMBER charges mean that loop hyperflexibility occurred when polarization was not included.

Figure 2d also plots the order parameters from MD simulations using the CHARMM22 and OPLS-AA force fields.[22] These simulations actually produced more significant hyperflexibility at several locations, as can be seen in Figure 2d.

E. Lysozyme. Hen egg white lysozyme is a compact protein containing 129 amino acids that folds into a compact globular structure. It has become a standard for evaluating the quality of force fields by comparing integral dynamical properties calculated from MD simulations to the results of NMR relaxation experiments.[23−26] The rmsd's from the experimental S^2 value were 0.072 for PPCs and 0.080 for AMBER03 charges. However, as shown in Figure 2e, the AMBER03-calculated S^2 values for residue Leu17 in the turn region and Arg114 in the α-helix region were significantly smaller than the experimentally measured ones. We note that Leu17 and Arg114 form the hydrogen bonds $N^{17}-H^{17}\cdots O^{12}$ and $N^{114}-H^{114}\cdots O^{110}$ with Met12 and Ala110, respectively. From Figure 3e−h, we observe that these two hydrogen bonds exhibited large fluctuations when AMBER charges were used, with the $N^{17}-H^{17}\cdots O^{12}$ hydrogen bond being partially broken. This clearly explains why the order parameter of the backbone Leu17 calculated using AMBER charges was much smaller than the experimental value. In contrast, these hydrogen bonds were much more stable in the simulation using PPCs, as shown in Figure 3e−h. The excellent agreement between the order parameters obtained from the PPC simulation and experimental data vindicated the correctness of the structure and dynamics described by PPCs. The S^2 values for lysozyme calculated using AMBER ff99SB and CHARMM22/CMAP[23,24] were also shown to be less accurate than those obtained using PPCs.

IV. Conclusions

Molecular dynamics simulations of NMR backbone relaxation order parameters have been carried out to investigate protein polarization effects on residue-specific protein dynamics for five benchmark proteins (BPTI, GB1, calbindin, IL4, and lysozyme). Our results clearly show that MD simulations based on the polarized force field yielded significantly improved agreement between the theoretical order parameters and the experimental values, especially in local regions such as loops, coils, and turns, where overly large fluctuations of local protein motions or hyperflexibility were seen in simulations based on standard nonpolarizable force fields. Detailed analysis shows that these local structures are critically stabilized by electronic polarization, especially the stability of hydrogen bonds. Without proper

inclusion of protein polarization, these local structures are less stable and exhibit hyperflexibility, as evidenced by comparison of residue-specific backbone NMR order parameters with experimental data. The present work shows that hyperflexibility of these local regions in proteins is due to the lack of electronic polarization, which artificially weakens the hydrogen bond in the standard force field.

Hydrogen bonding is the dominant interaction for secondary structures, and electronic polarization plays an important role in stabilizing the hydrogen bonds and the associated local structures. Without inclusion of polarization, the MD simulations failed to stabilize some intraprotein hydrogen bonds in BPTI, GB1, IL4, and lysozyme, causing hyperflexibility of the local structures in BPTI, calbindin, and IL4 as demonstrated by explicit comparison with experimentally measured backbone order parameters. In the five proteins we studied, the effect of protein polarization is important for stabilizing the secondary structures, such as an isolated β-bridge, a β-sheet, an α-helix, turn (loop) and coil regions, and disulfide bonds. As has been noted previously, existing molecular-mechanics force fields perform poorly for residues in the loop region, usually exaggerating their flexibility. This was confirmed by the current simulation studies, where significant disagreement between the experimental order parameters and the simulation results using standard AMBER charges mainly occurred in loops and turns where hydrogen bonds were destabilized because of the lack of polarization effects in the force field. Similar effects have been observed in protein−ligand binding, where the lack of protein polarization causes the breakage of the critical hydrogen bonds, resulting in the collapse of the binding complex structure.[16]

The explicit comparison between theory and NMR experiments in the current study provides convincing evidence that intraprotein hydrogen bonds and related local structures are dynamically stabilized by electronic polarization. It is encouraging to note that a recent MD study of BPTI[38] using an explicit polarization (X-Pol) approach,[39] in which partial atomic charges were derived from the semiempirical electronic wave function of the entire protein system, also found a significant effect of protein polarization on protein structure and dynamics. The X-Pol approach is similar in spirit to our PPC approach but differs in the specific techniques used to extract partial atomic charges from quantum-chemical calculations on proteins.

One question that may arise is whether the PPCs would overstabilize the protein's native structure. On the basis of the available computational results and their comparisons with experimental observations for a diverse set of protein systems,[15−17] including the current one for NMR study, it is clear that use of PPCs does not overstabilize the native structure of the protein. Of course, the accuracy of the PPCs is expected to degrade if the protein moves significantly away from its native structure. However, as long as the overall protein structure remains nativelike, PPCs should provide more accurate electrostatic interactions than typical mean-field-like force fields. An ideal approach would be to develop a fully polarizable force field that can adapt to conformational changes in the protein. However, the development of polarizable force fields has met with various difficulties, and its application to protein systems is thus significantly limited.[9]

Acknowledgment. This work was partially supported by the National Science Foundation of China (Grant 20773060) and the National Basic Research Program of China (Grant 2004CB719901). Y.M. is grateful for the financial support from the China Postdoctoral Science Foundation (Grant 20070420981) and the Postdoctoral Scientific Research Program of Jiangsu Province, China.

Supporting Information Available: Complete ref 30. This material is available free of charge via the Internet at http://pubs.acs.org.

JA901650R

(38) Xie, W.; Orozco, M.; Truhlar, D. G.; Gao, J. *J. Chem. Theory Comput.* **2009**, *5*, 459.

(39) Xie, W.; Gao, J. *J. Chem. Theory Comput.* **2007**, *3*, 1890.