

Evaluating the Performance of the ff99SB Force Field Based on NMR Scalar Coupling Data

Lauren Wickstrom,[†] Asim Okur,[‡] and Carlos Simmerling^{†*}

[†]Biochemistry and Structural Biology Program, and [‡]Department of Chemistry, State University of New York, Stony Brook, New York

ABSTRACT Force-field validation is essential for the identification of weaknesses in current models and the development of more accurate models of biomolecules. NMR coupling and relaxation methods have been used to effectively diagnose the strengths and weaknesses of many existing force fields. Studies using the ff99SB force field have shown excellent agreement between experimental and calculated order parameters and residual dipolar calculations. However, recent studies have suggested that ff99SB demonstrates poor agreement with J-coupling constants for short polyanalines. We performed extensive replica-exchange molecular-dynamics simulations on Ala₃ and Ala₅ in TIP3P and TIP4P-Ew solvent models. Our results suggest that the performance of ff99SB is among the best of currently available models. In addition, scalar coupling constants derived from simulations in the TIP4P-Ew model show a slight improvement over those obtained using the TIP3P model. Despite the overall excellent agreement, the data suggest areas for possible improvement.

A significant challenge in the use of computation to study complex biomolecular systems is force-field accuracy. Force fields are made up of a molecular-mechanics energy function with empirical parameters, which are typically obtained from fitting to experimental or high-level quantum mechanical (QM) data. These approximations may lead to inaccuracies in calculated kinetic and/or thermodynamic properties. One such force field is ff94 (1), which has a strong bias favoring helical content. Although it is not always apparent in short simulations, ff94 leads to overstabilization of helical systems and the adoption of stable helices for sequences that have nonhelical experimentally determined structures (2). Even in cases where the force field matches well to the QM data used in parameter development, errors can arise from inconsistencies in the model. For example, many nonpolarizable force fields employ partial charges that are intended to reproduce the enhanced dipoles found in aqueous solution, yet the dihedral potentials are fit to reproduce gas-phase QM energy profiles using these charges. These effects, combined with the relatively small size of the systems used for parameter development, indicate that validation against experimental data is vital.

ff99SB was developed to improve the secondary structure balance of the previous Amber protein force fields, as well as to improve the description of glycine residues (3). Although the parameters were fit with the use of QM data, the development relied on the validation of candidate parameters against experiment. Decoy sets of helical peptides, hairpins, and miniproteins demonstrated the correct energy minima. Calculated NMR order parameters for ubiquitin and lysozyme also showed better agreement with experiment than previous force

fields. Showalter and Bruschiweiler (4) demonstrated that ubiquitin dynamics as measured by residual dipolar couplings obtained from ff99SB simulation are “comparable to or better than the best static structural models and the NMR ensemble”. Other works have shown similarly good agreement between ff99SB simulations and NMR structural and relaxation data (5–8). Overall, these studies suggest that ff99SB is in at least reasonable agreement with experiment for a variety of proteins.

One disadvantage of performing such studies on complex systems is that it can be difficult to decompose inaccuracies into the specific force-field terms that need improvement. Short polyanalines have become useful, simple model systems for studying the conformational variability of unfolded states; however, in such states the structural preferences are weak and therefore the system is highly sensitive to small inaccuracies (9–12). A recent study of Ala_n ($n = 3–7$) by Graf et al. (13) showed that significant differences exist between the experimental and calculated J-coupling constants from unweighted simulation data. Building on this availability of extensive experimental data, Best et al. (14) performed a follow-up study on Ala₅ using variations of the Amber (15), CHARMM (16), GROMOS (17), and OPLS (18) force fields using various sets of Karplus parameters to calculate the scalar coupling constants. They evaluated the force fields using a χ^2 value that represented the sum of deviations of each calculated J-coupling constant compared with the experimental values, normalized by a factor related to the assumed systematic error in the coupling constant calculations. Of the parameter sets tested, ff99SB was ranked among the worst for this data set. Later, Best et al. (19) corrected key aspects of the data, with the result that ff99SB’s ranking improved significantly. We present here a more detailed analysis of ff99SB’s performance using two water models and peptides of different lengths.

We performed replica-exchange molecular-dynamics simulations (20,21) for 50 ns/replica of Ala₃ and Ala₅ in

Submitted March 30, 2009, and accepted for publication April 28, 2009.

*Correspondence: carlos.simmerling@stonybrook.edu

Asim Okur’s present address is Laboratory of Computational Biology, National Heart Lung and Blood Institute, National Institutes of Health, Bethesda, Maryland.

Editor: Kathleen B. Hall.

© 2009 by the Biophysical Society
0006-3495/09/08/0853/4 \$2.00

doi: 10.1016/j.bpj.2009.04.063

two explicit water models. Precision was quantified by means of fully independent simulations from different initial structures. Simulation details are included in the [Supporting Material](#). In this study we address the performance of the ff99SB force field, compare the effects of using different water models, and suggest improvements to ff99SB that may improve agreement with experiment.

We calculated the scalar coupling constants for each polyaniline simulation using the same three Karplus parameter sets employed in the study by Best et al. (14) (Table 1). We also employed the same χ^2 calculation as Best et al. to evaluate the deviation of the J-coupling constants from the experimental values. For Ala₃, the χ^2 values varied from 1.57 to 2.17 depending on the solvent model and the parameter set. The Ala₅ J-coupling constants were also quite sensitive to these different variables, but even with the larger peptide size, the χ^2 values remained < 2.0. Of importance, the χ^2 values for Ala₅ were at least as low as any of the other force fields evaluated by Best et al. (14).

In addition to the protein force field, the water model may also be expected to have a significant effect on accuracy for these short, solvent-exposed peptides. To compare solvent effects, we generated simulations using the TIP3P (22) and TIP4P-Ew (23) solvent models. Simulations using TIP4P-Ew have shown better agreement between calculated and experimental NMR observables (6,24) due to more realistic diffusion and tumbling in this water model. However, TIP3P has been shown to be better than TIP4P-Ew at reproducing solvation free energies of small molecules (25). Our data for both polyanilines indicate that the deviations from experiment are modestly smaller (3–16% reduction in the χ^2 value) when the TIP4P-Ew solvent model is used (Table 1). These data and those from the previous studies suggest that the combination of using the ff99SB force field with the TIP4P-Ew solvent model is one of the best combinations currently available, at least for short peptides.

To address the remaining issues with the force field, we compared the J-coupling constants across the sequence for the Ala peptides. We selected $^3J(\text{H}_\text{N}, \text{H}_\alpha)$ and the $^2J(\text{N}, \text{C}_\alpha)$ constant because of their sensitivity to the φ and ψ angles of the backbone (Fig. 1). For the middle residues of Ala₅, the calculated $^3J(\text{H}_\text{N}, \text{H}_\alpha)$ values show deviations ranging from 1.2 to 1.7 Hz from the experimental observables and ranging from 6.8 to 7.4 Hz depending on the parameter set (compared with the experimental values of 5.6–6.0 Hz). These trends are observed in the other polyaniline simula-

TABLE 1 Scalar coupling χ^2 values for Ala₃ and Ala₅ using three different Karplus equation parameter sets

| Peptide | Water | DFT1 (26) | DFT2 (26) | Original (13) |
|------------------|----------|-------------|-------------|---------------|
| ALA ₃ | TIP3P | 1.60 ± 0.04 | 1.89 ± 0.01 | 2.17 ± 0.01 |
| | TIP4P-Ew | 1.57 ± 0.01 | 1.75 ± 0.04 | 2.05 ± 0.04 |
| ALA ₅ | TIP3P | 1.44 ± 0.02 | 1.62 ± 0.03 | 1.81 ± 0.01 |
| | TIP4P-Ew | 1.36 ± 0.01 | 1.36 ± 0.01 | 1.55 ± 0.01 |

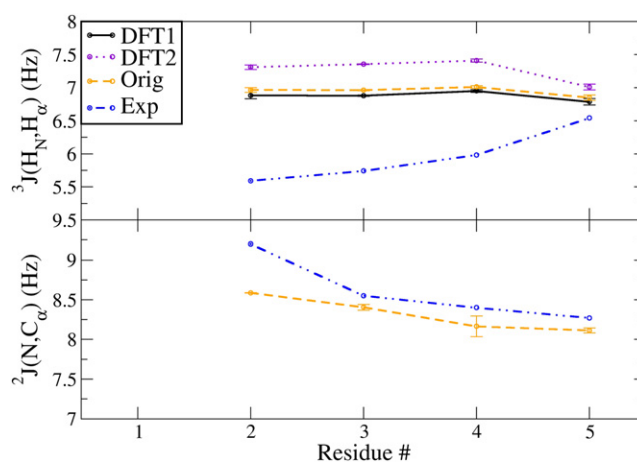


FIGURE 1 Average $^3J(\text{H}_\text{N}, \text{H}_\alpha)$ and $^2J(\text{N}, \text{C}_\alpha)$ scalar constants for simulations of ALA₅ in the TIP4P-Ew solvent model. These constants are calculated with the original DFT1 (black), DFT2 (purple), and original Karplus (orange) parameters, respectively. Experimental scalar values are plotted in each graph in blue. Error bars are calculated from the average difference between two simulations.

tions as well (see the [Supporting Material](#)); the coupling constants from the simulations are too large, indicating too much sampling of β -like local backbone conformations as compared with PP_{II}, although the latter remains the dominant conformation. In contrast, the calculated $^2J(\text{N}, \text{C}_\alpha)$ constants are in excellent agreement with the experimental values, with the largest deviation in residue 2, which shifts to values that suggest slightly too much α -helical conformation (which are generally ~6.50 Hz on the Karplus curve). Therefore, the most apparent issue for the local backbone conformations in these simulations is that the ensembles are shifted slightly away from favored PP_{II} conformations.

Our results show that scalar coupling calculations are sensitive to the implemented Karplus parameter sets (Table 1 and Fig. 1). Based on the calculated $^3J(\text{H}_\text{N}, \text{H}_\alpha)$ values, the DFT2 (26) parameters should have the worst χ^2 value; however, the Orig set (13) produces the worst results for Ala₅ in the TIP4P-Ew model. In Fig. S2, the average J-coupling scalar constants are shown for the other scalar constants involved in the χ^2 analysis. The parameter sets show similar trends for most of the scalar constants except for $^3J(\text{H}_\text{N}, \text{C}_\beta)$, where the DFT2 set performs the best, compensating for the errors in the $^3J(\text{H}_\text{N}, \text{H}_\alpha)$ and resulting in lower overall χ^2 . Thus the χ^2 data should be interpreted with caution, and the influence of Karplus parameters on individual errors must be considered. Furthermore, the Orig parameters implicitly include the effects of motional averaging, and one would therefore expect worse agreement with experiment when scalar couplings are back-calculated from the full ensembles using empirical parameters fit to experimental data (27). Since this is not apparent, it appears that the effect of force-field inaccuracies on the simulated ensembles may exceed the effects of including dynamic

fluctuations both implicitly in the Karplus parameters and explicitly in the molecular-dynamics ensembles.

Helical structural bias is a problem that has been associated with previous Amber force fields. We therefore focused on Ala₅ since its sequence is long enough to permit an α -helical hydrogen bond. We calculated the percentage of α -helical conformations sampled by the central residue of Ala₅ with the definition used by Best et al. (14) (Table S5). In the TIP3P and TIP4P-Ew solvent models, the α -basin populations are 20% and 15%, respectively, a significant improvement over the ff94 and ff99 force fields (90–95% of the ensembles sample an α -helical population depending on simulation conditions) (14). Nevertheless, one must use caution in interpreting these results in terms of helix formation. These calculations measure only the ϕ/ψ Ramachandran basin at the residue level; the structures may not sample an actual α -helical hydrogen bond. To test this, we repeated our calculations employing the dictionary for secondary structure prediction (DSSP) (28) definition for helicity, which resulted in populations of 0.4% and 0.0% of 3_{10} and α -helix in both solvent models. Hence, the ff99SB ensemble does not suffer from the ailments of the previous force fields because it samples local α -helical conformations only in the random coil state, and no measurable amount of helical conformations.

In conclusion, comparisons of the Ala₃ and Ala₅ ensembles in both water models exhibit an excellent agreement between experimental and calculated scalar couplings. We also find that these calculations are somewhat solvent model-dependent, indicating that the TIP4P-Ew water model is the better choice for comparisons with NMR scalar coupling data. The deviations of the calculated and experimental $^3J(H_N, H_\alpha)$ scalar constants with all of the parameter sets suggest that the deviations are largest in the ϕ -torsional potential, which could have effects on larger systems. Nevertheless, ff99SB does not present the helical bias issues of the previous force fields. In future work, we plan to use these experimental data as a reference to further improve our force-field parameters.

SUPPORTING MATERIAL

Simulation and analysis details, error analysis, and tables, figures, and references are available at [http://www.biophysj.org/biophysj/supplemental/S0006-3495\(09\)01031-5](http://www.biophysj.org/biophysj/supplemental/S0006-3495(09)01031-5).

We thank G. Hummer and R. Best for J-coupling scripts and feedback, and A. Roitberg for helpful discussions.

This work was supported by the National Institutes of Health (GM61678), National Center for Supercomputing Applications and National Institute for Computational Sciences (NPACI MCA02N028), and Department of Energy (DE-AC02-98CH10886).

REFERENCES

- Cornell, W. D., P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, et al. 1995. A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *J. Am. Chem. Soc.* 117:5179–5197.
- Okur, A., B. Strockbine, V. Hornak, and C. Simmerling. 2003. Using PC clusters to evaluate the transferability of molecular mechanics force fields for proteins. *J. Comput. Chem.* 24:21–31.
- Hornak, V., R. Abel, A. Okur, B. Strockbine, A. Roitberg, et al. 2006. Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins.* 65:712–725.
- Showalter, S. A., and R. Bruschweiler. 2007. Quantitative molecular ensemble interpretation of NMR dipolar couplings without restraints. *J. Am. Chem. Soc.* 129:4158–4159.
- Koller, A. N., H. Schwalbe, and H. Gohlke. 2008. Starting structure dependence of NMR order parameters derived from MD simulations: implications for judging force-field quality. *Biophys. J.* 95:L4–L6.
- Fawzi, N. L., A. H. Phillips, J. Z. Ruscio, M. Doucleff, D. E. Wemmer, et al. 2008. Structure and dynamics of the AB(21–30) peptide from the interplay of NMR experiments and molecular simulations. *J. Am. Chem. Soc.* 130:6145–6158.
- Wickstrom, L., A. Okur, K. Song, V. Hornak, D. P. Raleigh, et al. 2006. The unfolded state of the villin headpiece helical subdomain: computational studies of the role of locally stabilized structure. *J. Mol. Biol.* 360:1094–1107.
- Showalter, S. A., E. Johnson, M. Rance, and R. Bruschweiler. 2007. Toward quantitative interpretation of methyl side-chain dynamics from NMR by molecular dynamics simulations. *J. Am. Chem. Soc.* 129:14146–14147.
- Shi, Z. S., K. Chen, Z. G. Liu, and N. R. Kallenbach. 2006. Conformation of the backbone in unfolded proteins. *Chem. Rev.* 106:1877–1897.
- Zagrovic, B., J. Lipfert, E. J. Sorin, I. S. Millett, W. F. van Gunsteren, et al. 2005. Unusual compactness of a polyproline type II structure. *Proc. Natl. Acad. Sci. USA.* 102:11698–11703.
- Makowska, J., S. Rodziewicz-Motowidlo, K. Baginska, J. A. Vila, A. Liwo, et al. 2006. Polyproline II conformation is one of many local conformational states and is not an overall conformation of unfolded peptides and proteins. *Proc. Natl. Acad. Sci. USA.* 103:1744–1749.
- Mezei, M., P. J. Fleming, R. Srinivasan, and G. D. Rose. 2004. Polyproline II helix is the preferred conformation for unfolded polyalanine in water. *Proteins.* 55:502–507.
- Graf, J., P. H. Nguyen, G. Stock, and H. Schwalbe. 2007. Structure and dynamics of the homologous series of alanine peptides: a joint molecular dynamics/NMR study. *J. Am. Chem. Soc.* 129:1179–1189.
- Best, R. B., N. V. Buchete, and G. Hummer. 2008. Are current molecular dynamics force fields too helical? *Biophys. J.* 95:L7–L9.
- Case, D. A., T. E. Cheatham, T. Darden, H. Gohlke, R. Luo, et al. 2005. The Amber biomolecular simulation programs. *J. Comput. Chem.* 26:1668–1688.
- Brooks, B. R., R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan, et al. 1983. CHARMM—a program for macromolecular energy, minimization, and dynamics calculations. *J. Comput. Chem.* 4:187–217.
- van Gunsteren, W. F., S. R. Billeter, A. A. Eising, P. H. Hünenberger, P. Krüger, et al. 1996. Biomolecular Simulation: The GROMOS96 Manual and User Guide. Vdf Hochschulverlag, ETH Zurich, Switzerland.
- Jorgensen, W. L., D. S. Maxwell, and J. Tirado-Rives. 1996. Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *J. Am. Chem. Soc.* 118:11225–11236.
- Best, R. B., N. V. Buchete, and G. Hummer. 2008. Are current molecular dynamics force fields too helical? Erratum in *Biophys. J.* 95:4494.
- Hansmann, U. H. E. 1997. Parallel tempering algorithm for conformational studies of biological molecules. *Chem. Phys. Lett.* 281:140–150.
- Sugita, Y., and Y. Okamoto. 1999. Replica-exchange molecular dynamics method for protein folding. *Chem. Phys. Lett.* 314:141–151.
- Jorgensen, W. L., J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein. 1983. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* 79:926–935.

23. Horn, H. W., W. C. Swope, J. W. Pitera, J. D. Madura, T. J. Dick, et al. 2004. Development of an improved four-site water model for biomolecular simulations: TIP4P-Ew. *J. Chem. Phys.* 120:9665–9678.
24. Wong, V., and D. A. Case. 2008. Evaluating rotational diffusion from protein MD simulations. *J. Phys. Chem. B.* 112:6013–6024.
25. Shirts, M. R., and V. S. Pande. 2005. Solvation free energies of amino acid side chain analogs for common molecular mechanics water models. *J. Chem. Phys.* 122:134508.
26. Case, D. A., C. Scheurer, and R. Bruschweiler. 2000. Static and dynamic effects on vicinal scalar J couplings in proteins and peptides: a MD/DFT analysis. *J. Am. Chem. Soc.* 122:10390–10397.
27. Bruschweiler, R., and D. A. Case. 1994. Adding harmonic motion to the Karplus relation for spin-spin coupling. *J. Am. Chem. Soc.* 116:11199–11200.
28. Kabsch, W., and C. Sander. 1983. Dictionary of protein secondary structure—pattern-recognition of hydrogen-bonded and geometrical features. *Biopolymers.* 22:2577–2637.