Published in final edited form as:

J Phys Chem Lett. 2016 August 04; 7(15): 3032–3036. doi:10.1021/acs.jpclett.6b01229.

Performance of Protein-Ligand Force Fields for the Flavodoxin - Flavin Mononucleotide System

Michael J. Robertson, Julian Tirado-Rives, and William L. Jorgensen*

Department of Chemistry, Yale University, New Haven, Connecticut 06520-8107, United States

Abstract

The ability to accurately perform molecular dynamics and free energy perturbation calculations for protein-ligand systems is of broad interest to the biophysical and pharmaceutical sciences. In this work, several popular force fields are evaluated for reproducing experimental properties of the flavodoxin/flavin mononucleotide system. Calculated 3J couplings from molecular dynamics simulations probing ϕ and χ_1 dihedral angles are compared to over one thousand experimental measurements. Free energy perturbation calculations were also executed between different protein mutants for comparison with experimental data for relative free energies of binding. Newer versions of popular protein force fields reproduced 3J backbone and side chain couplings with good accuracy, with RMSD values near or below one hertz in most cases. OPLS-AA/M paired with CM5 charges for the ligand performed particularly well, both for the 3J couplings and FEP results, with a mean unsigned error for relative free energies of binding of 0.36 kcal/mol.

TOC GRAPHIC

*Corresponding Author: william.jorgensen@yale.edu. Phone: 203-432-6278. Fax: 203-432-6299.

Notes

The authors declare no competing financial interest.

Supporting Information

A document containing a full description of the simulation details.

Molecular modeling has emerged as a valuable partner to many experimental methods in biophysical chemistry. Particularly important is the ability to model protein-ligand systems, as most popular protein force fields now have small molecule complements. This permits easily accessible free energy perturbation calculations in performing lead optimization of potential pharmaceutical compounds.^{1,2} It has recently been demonstrated that, with a highly refined commercial protein and ligand force field, free energy perturbation calculations are able to provide useful estimates of relative free energies of binding across a broad range of systems.³ This and related studies have emphasized the quality of the calculated free energies of binding as a primary criterion for a simulation accurately reproducing the physics of the system. The focus of the present work is to explore a protein-ligand system where the protein dynamics can also be rigorously compared to experiment, and to do so for alternative force fields.

Several studies have systematically looked at the performance of different protein force fields in aqueous phase simulations, comparing the results to NMR and crystal structure data.^{4,5} While yielding important information, the systems that have been studied have often been proteins without cofactors or ligands. To explore a system with a ligand, the *Desulfovibrio vulgaris* flavodoxin/flavin mononucleotide system has been selected. Flavodoxin is a 148 residue electron transport protein found in bacteria that natively binds a flavin mononucleotide (FMN) ligand. The interaction between the wild type protein and the ligand is strong, with a K_d from a fluorometric assay of 240 ± 10 pM.⁶ The system has been studied extensively by NMR spectroscopy with the FMN ligand in its oxidized state, with over 1,300 3J couplings having been measured that probe the backbone ϕ and the side chain χ_1 angles.^{7,8} Binding free energies are also available in the literature for both the oxidized and reduced (FMNH₂) form of the ligand bound to wild-type and a few mutants of flavodoxin (Scheme 1).⁶ These two sets of experimental data, compared against results of molecular dynamics and free energy perturbation (FEP) simulations, can provide an idea of the interplay between accurate protein dynamics and reliable relative free energies of

binding. It is of particular interest to test the new OPLS-AA/M⁹ force field as well as a new program available in the BOSS package¹⁰ for generating CHARMM formatted parameter and topology files for a ligand with the OPLS-AA/CM1A¹¹ force field.

Triplicate 200 ns simulations for flavodoxin were performed using the NAMD program with the CHARMM22(C22)/CMAP¹²//CGenFF,¹³ CHARMM36(C36)¹⁴//CGenFF, OPLS-AA¹⁵//OPLS-AA/CM1A, OPLS-AA/M//OPLS-AA/CM1A, OPLS-AA//OPLS-AA/CM5,¹⁶ OPLS-AA/M//OPLS-AA/CM5, AMBER ff99¹⁷//GAFF¹⁸, AMBER ff99SB-ILDN¹⁹//GAFF, and AMBER ff14SB²⁰//GAFF force fields, where the notation protein force field/ligand force field has been employed. Three different sets of Karplus parameters were used to calculate ³J coupling constants from the ensemble of ϕ values, derived from work with the flavodoxin/flavin mononucleotide system,⁷ ubiquitin,²¹ and GB3.²² These are referred to as model 1, 2, and 3, respectively. Models 1 and 2 cover all of the experimental measurements of Schmidt *et al.*,⁷ while model 3 only includes parameters for ³J(H_NH _{α}), ³J(H_NC'), and ³J(H_NC β) couplings. For χ_1 couplings, only the residue specific Karplus parameters of Perez *et al.*⁸ that were developed from the flavodoxin/flavin mononucleotide system were used. Values reported in Figure 1 represent the average RMSD between experimental and simulated ³J couplings from the triplicate simulations with the error bars corresponding to the standard deviations. Results are reported both for all residues and for the subset of residues that are within 7.5 Å of the ligand.

All three of the more recent protein force fields (AMBER ff14SB, CHARMM 36, and OPLS-AA/M) performed well, with a RMSD of ~0.6–0.8 hertz for all backbone couplings and ~1–1.1 hertz for the side chain χ_1 s. While AMBER ff99SB-ILDN, ff14SB, and OPLS-AA/M performed the best for protein residues, CHARMM 36 or CHARMM 22/CMAP are competitive. In addition, while the more recent AMBER and OPLS-AA force fields represent significant improvements over their previous incarnations (ff99 and OPLS-AA), the performance of CHARMM 36 is mixed compared to CHARMM 22/CMAP. It may be that the CMAP, which creates agreement between the quantum mechanical and molecular mechanical ϕ - ψ potential energy surfaces with a cubic spline, already improved the force field to near the limit of accuracy for a point-charge model. Also worth noting is that some of the improvement of CHARMM 36 focused on the unfolded state and transitions to the folded state. The performance of AMBER ff14SB compared to ff99SB-ILDN is also mixed. Interestingly, both the CHARMM and AMBER force fields tended to have markedly higher RMSD values for the ³J couplings within 7.5 Å of the ligand compared to the full protein with Karplus parameter models 1 and 2 and the side-chain Karplus parameters. This is not the case for OPLS-AA/M, and to some extent AMBER ff14SB. This may be a consequence of the OPLS family of force fields originally being parameterized for small molecules, with the protein component of the force field developed based upon analogy to the small molecule parameters. Alternatively, while much of the protein has alpha helical or beta sheet structure, the region immediately around the ligand contains more loops. This secondary structure motif may be better captured in OPLS-AA/M and ff14SB than the other force fields.

Free energy perturbation calculations were performed between the wild type, G61A, G61L, and G61V variants of flavodoxin in the presence of oxidized FMN. The mutated residue,

Gly-61, is located in the binding loop surrounding the core of FMN (Figure 2). As the current implementation of FEP calculations in NAMD is based on CHARMM-formatted input, and not Amber inpcrd and prmtop files, only the CHARMM and OPLS-AA force fields were examined. While these mutations can be made with the AMBER software, it was the goal of this study to have highly consistent simulations in terms of software, integrator, thermostat/barostat, and method of performing the mutations. The results are provided in Table 1. All perturbations were performed forwards and backwards in triplicate, with the standard deviation reported as the uncertainties.

The latest OPLS and CHARMM force fields primarily updated the parameters for the torsional energetics, with each generally adopting the non-bonded (charge and Lennard-Jones) parameters of its predecessor. The alterations of torsional parameters are expected to change the protein dynamics and yield differences in the calculated relative free energies of binding. This is confirmed in Table 1 with differences of 2 – 3 kcal/mol for the OPLS results.

The recent force fields did well for the perturbations, with the MUE in calculated relative free energies of binding ranging from 0.36 – 1.12 kcal/mol. Calculated values typically overlap the experimental results within the error bars. It is also noted that the original OPLS-AA force field, which performed significantly less well in reproducing the 3J couplings than the other force fields, also provides poorer binding results, with a MUE of 2.38 kcal/mol, and the incorrect prediction that the wild-type protein is a weaker binder of FMN than the three variants. The new OPLS-AA/M improves significantly upon this with a MUE of 0.36 kcal/mol and with wild-type flavodoxin correctly predicted as the best binder.

For the mutations involving leucine and valine side chains, it is worth noting differences in χ_1 rotamers between different force fields. Presented in Table 2 are the rotamer populations averaged over all six simulations for the different force fields taken from the lambda window at the appropriate endpoint for that residue. It is clear that there are substantial differences, which undoubtedly contribute to the differences in the relative free energies of binding for the G61L and G61V mutations.

While the torsional energetics and protein dynamics are clearly important, the charge model for the ligand also has effects on the calculated relative free energies of binding. Most notable is the flavin nitrogen closest to Gly-61, which has significantly different charge with the three force fields (–0.7 with CGenFF, –0.3 with CM5, and –0.1 with CM1A) and thus yields different strengths of electrostatic interactions, especially with the Asp-62 NH. The CM5 charge model paired with OPLS-AA type force fields appears to yield a lower mean unsigned error compared to the older CM1A model, 0.36 vs. 0.63 kcal/mol, though with the statistical uncertainties, the improvement is uncertain.

Overall, the results are encouraging, with the more recent versions of each family of force fields succeeding in reproducing the experimental 3J couplings to a high degree of accuracy. While OPLS-AA/M//OPLS-AA/CM5 and the two most recent AMBER force fields performed the best for the NMR data, both versions of CHARMM that employed a CMAP also performed well. The AMBER and CHARMM force fields performed less well for the

residues within 7.5 Å of the ligand compared to all residues, while the difference was less significant for the OPLS-AA/M force field. The CHARMM36 and C22 CMAP force fields with CGenFF and OPLS-AA/M//OPLS-AA/CM5 or CM1A were also found to perform well for computing relative free energies of binding for flavodoxin and the variants of Gly61. The importance of accurate protein dynamics in calculating relative free energies of binding was reflected in that the worst performing protein force field for the 3J couplings also delivered the least accurate results for relative free energies of binding. Overall, the present results show that there has been significant improvement in protein force fields for both dynamics and energetics.

Computational Methods

All systems were prepared from the low temperature crystal structure 3FX2²³ with an A2P mutation to correspond to the wild type *desulfovibrio vulgaris* enzyme used in both the NMR and binding studies. For the CHARMM and OPLS based force fields, input files for NAMD²⁴ were prepared from CHARMM formatted parameter files. Parameters for the oxidized flavin mononucleotide ligand for use with the CHARMM force fields were prepared with CGenFF typed with the MATCH webserver.²⁵ For use with the OPLS-AA protein force fields, a newly developed program for use with BOSS was used to generate CHARMM formatted OPLS-AA/CM1A parameters. CM5 charges were calculated with Gaussian09²⁶ and CM5Pac²⁷ and substituted for the CM1A charges with the BOSS/Gaussian interface²⁸ to produce OPLS-AA/CM5 parameters for the ligand. Input files for NAMD with the AMBER type force fields were prepared with the Ambertools²⁹ Antechamber³⁰ program utilizing the GAFF force field with HF/6–31G(d) RESP derived charges for the ligand. All systems were solvated in cubic boxes with lengths of 63 Å with TIP3P water and enough sodium and chlorine atoms to achieve a charge neutral system with an approximately 150 mM concentration of salt. Full details on the molecular dynamics simulations are provided in the Supporting Information.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Gratitude is expressed to the National Institutes of Health (GM032136) for support, to the National Science Foundation for a Graduate Research Fellowship under Grant No. DGE-1122492 (M.J.R.), and to the Yale Center of Research Computing for computer access.

References

1. Chodera JD, Mobley DL, Shirts MR, Dixon RW, Branson K, Pande VS. Alchemical free energy methods for drug discovery: progress and challenges. *Curr Opin Struct Biol.* 2011; 21:150–160. [PubMed: 21349700]
2. Jorgensen WL. Efficient drug lead discovery and optimization. *Acc Chem Res.* 2009; 42:724–733. [PubMed: 19317443]
3. Wang L, Wu Y, Deng Y, Kim B, Pierce L, Krilov G, Lupyan D, Robinson S, Dahlgren MK, Greenwood J, et al. Accurate and reliable prediction of relative ligand binding potency in

- prospective drug discovery by way of a modern free-energy calculation protocol and force field. *J Am Chem Soc.* 2015; 136:2695–2703.
4. Beaucamp KA, Lin Y-S, Das R, Pande VS. Are protein force fields getting better? A systematic benchmark on 524 diverse NMR measurements. *J Chem Theory Comput.* 2012; 8:1409–1414. [PubMed: 22754404]
 5. Lindorff-Larson K, Maragakis P, Piana S, Eastwood MP, Dror RO, Shaw DE. Systematic validation of protein force fields against experimental data. *PLoS One.* 2012;7.
 6. O'Farrell PA, Walsh MA, McCarthy AA, Higgins TM, Voordouw G, Mayhew SG. Modulation of the redox potentials of FMN in *Desulfovibrio vulgaris* Flavodoxin: thermodynamic properties and crystal structures of glycine-61 mutants. *Biochemistry.* 1998; 37:8405–8416. [PubMed: 9622492]
 7. Schmidt JM, Blümel M, Löhr F, Rüterjans H. Self-consistent 3J coupling analysis for the joint calibration of Karplus coefficients and evaluation of torsion angles. *J Biomolecular NMR.* 1999; 14:1–12.
 8. Pérez C, Löhr F, Rüterjans H, Schmidt JM. Self-Consistent Karplus Parameterization of 3J couplings depending on the polypeptide side-chain torsion χ_1 . *J Am Chem Soc.* 2001; 123:7081–7093. [PubMed: 11459487]
 9. Robertson MJ, Tirado-Rives J, Jorgensen WL. Improved peptide and protein torsional energetics with the OPLS-AA force field. *J Chem Theory Comput.* 2015; 11:3499–3509. [PubMed: 26190950]
 10. Jorgensen WL, Tirado-Rives J. Molecular modeling of organic and biomolecular systems using BOSS and MCPRO. *J Comput Chem.* 2005; 26:1689–1700. [PubMed: 16200637]
 11. Udier-Blagovi M, Morales de Tirado P, Pearlman SA, Jorgensen WL. Accuracy of free energies of hydration using CM1 and CM3 atomic charges. *J Comput Chem.* 2004; 25:1322–1332. [PubMed: 15185325]
 12. MacKerell AD Jr, Feig M, Brooks CL. III Extending the treatment of backbone energetics in protein force fields: limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations. *J Comput Chem.* 2004; 25:1400–1415. [PubMed: 15185334]
 13. Vanommeslaeghe K, Hatcher E, Acharya C, Kundu S, Zhong S, Shim J, Darian E, Guvench O, Lopes P, Vorobyov I, et al. CHARMM General Force Field (CGenFF): A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. *J Comput Chem.* 2010; 31:671–690. [PubMed: 19575467]
 14. Best RB, Xiao Zhu X, Shim J, Lopes P, Mittal J, Feig M, MacKerell AD Jr. Optimization of the additive CHARMM all-atom protein force field targeting improved sampling of the backbone phi, psi, and side chain chi1 and chi2 dihedral angles. *J Chem Theory Comput.* 2012; 8:3257–3273. [PubMed: 23341755]
 15. Jorgensen WL, Maxwell DS, Tirado-Rives J. Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *J Am Chem Soc.* 1996; 118:11225–11236.
 16. Marenich AV, Jerome SV, Cramer CJ, Truhlar DG. Charge Model 5: An extension of Hirshfeld population analysis for the accurate description of molecular interactions in gaseous and condensed phases. *J Chem Theory Comput.* 2012; 8:527–541. [PubMed: 26596602]
 17. Wang J, Cieplak P, Kollman PA. How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules? *J Comput Chem.* 2000; 21:1049–1074.
 18. Wang J, Wolf RM, Caldwell JW, Kollman PA, Case DA. Development and testing of a general AMBER force field. *J Comput Chem.* 2004; 25:1157–1174. [PubMed: 15116359]
 19. Lindorff-Larsen K, Piana S, Palmo K, Maragakis P, Klepeis JL, Dror RO, Shaw DE. Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins.* 2010; 78:1950–1958. [PubMed: 20408171]
 20. Maier JA, Martinez C, Kasavajhala K, Wickstrom L, Hauser KE, Simmerling C. ff14SB: Improving the accuracy of protein side chain and backbone parameters from ff99SB. *J Chem Theory Comput.* 2015; 11:3696–3713. [PubMed: 26574453]

21. Hu J, Bax A. Determination of ϕ and χ_1 Angles in Proteins from ^{13}C - ^{13}C Three-bond J couplings measured by three-dimensional heteronuclear NMR. How planar is the peptide bond? *J Am Chem Soc.* 1997; 119:6360–6368.
22. Vögeli B, Ying J, Grishaev A, Bax A. Limits on variations in protein backbone dynamics from precise measurements of scalar couplings. *J Am Chem Soc.* 2007; 129:9377–9385. [PubMed: 17608477]
23. Watt W, Tulinsky A, Swenson RP, Watenpaugh KD. Comparison of the crystal structures of a flavodoxin in its three oxidation states at cryogenic temperatures. *J Mol Biol.* 1991; 218:195–208. [PubMed: 2002503]
24. Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, Chipot C, Skeel RD, Kale L, Schulten K. Scalable molecular dynamics with NAMD. *J Comput Chem.* 2005; 26:1781–1802. [PubMed: 16222654]
25. Yesselman JD, Price DJ, Knight JL, Brooks CL. III MATCH: An atom-typing toolset for molecular mechanics force fields. *J Comput Chem.* 2012; 15:189–202.
26. Frisch, MJ., Trucks, GW., Schlegel, HB., Scuseria, GE., Robb, MA., Cheeseman, JR., Scalmani, G., Barone, V., Mennucci, B., Petersson, GA., et al. Gaussian 09, revision D.01. Gaussian, Inc; Wallingford, CT: 2009.
27. Marenich, AV., Cramer, CJ., Truhlar, DG. CM5PAC. University of Minnesota; Minneapolis: 2011.
28. Vilseck JZ, Kostal J, Tirado-Rives J, Jorgensen WL. Application of a BOSS-Gaussian interface for QM/MM simulations of Henry and methyl transfer reactions. *J Comput Chem.* 2015; 36:2064–2074. [PubMed: 26311531]
29. Case, DA., Babin, JT., Berryman, RM., Betz, RM., Cai, Q., Cerutti, DS., Cheatham, TE., III, Darden, TA., Duke, RE., Gohlke, H., et al. AMBER 14, 2014. University of California; San Francisco:
30. Wang J, Wang W, Kollman PA, Case DA. Automatic atom type and bond type perception in molecular mechanical calculations. *J Mol Graph Model.* 2006; 25:247–260. [PubMed: 16458552]

Figure 1.

Plots of the RMSD between calculated and experimental 3J couplings for the flavodoxin/ flavin mononucleotide system with different force fields. The plots in the left hand column represent results for all residues in the protein (A,B,C,D), while the right hand column plots are values reported for all residues within 7.5 Å of the flavin mononucleotide ligand. The first three rows of plots provide the RMSD in backbone 3J values with Karplus parameters developed from work with flavodoxin (A,E), ubiquitin (B,F), and GB3 (C,G). The final row of plots (D,H) are for side chain χ_1 3J couplings.

Figure 2. Crystal structure 3FX2 for FMN bound to flavodoxin. Gly-61, which is mutated, is represented explicitly along with the two adjacent residues, Asp-62 and Trp-60.

Scheme 1.

Table 1*G* Results from FEP Calculations for Flavodoxin-FMN (kcal/mol)

Force Field	WT → G61A	G61A → G61L	G61A → G61V	MUE
Experiment	0.8 ± 0.4	0.6 ± 0.2	0.6 ± 0.4	0.0
C36//CGenFF	2.26 ± 0.75	1.00 ± 0.62	-0.89 ± 0.94	1.12
C22 CMAP//CGenFF	1.53 ± 0.98	0.30 ± 0.77	0.52 ± 0.81	0.37
OPLS-AA/M//OPLS/CM5	0.55 ± 0.76	0.27 ± 0.77	0.11 ± 1.32	0.36
OPLS-AA/M//OPLS/CM1A	-0.04 ± 0.66	1.01 ± 0.85	-0.02 ± 1.09	0.63
OPLS/AA//OPLS/CM1A	-2.81 ± 0.61	-1.03 ± 1.52	2.49 ± 1.12	2.38

Table 2

Rotamer Populations from FEP Calculations for Flavodoxin-FMN

Force Field	%m	Valine %p	%t	%m	Leucine %p	%t
C36//CGenFF	0%	0%	100%	46%	0%	54%
C22 CMAP//CGenFF	0%	14%	86%	0%	80%	20%
OPLS-AA/M//OPLS/CM5	0%	7%	93%	24%	0%	75%
OPLS-AA/M//OPLS/CM1A	0%	4%	96%	14%	0%	86%
OPLS/AA//OPLS/CM1A	9%	25%	66%	11%	0%	88%