How well do simulations reflect reality?

BCH 581
Travis Hughes
Associate professor
Department of Biomedical and Pharmaceutical Sciences.
University of Montana

Overall simulation of a crystal (@295K) has similar RMSD to the crystal (@295K).

- A crystal composed of 12 copies of lysozyme were simulated for 1μs using several different force fields (including amber ff14SB and charm C36)
- RMSD of simulation compared to crystal structure is similar to that between crystal structures

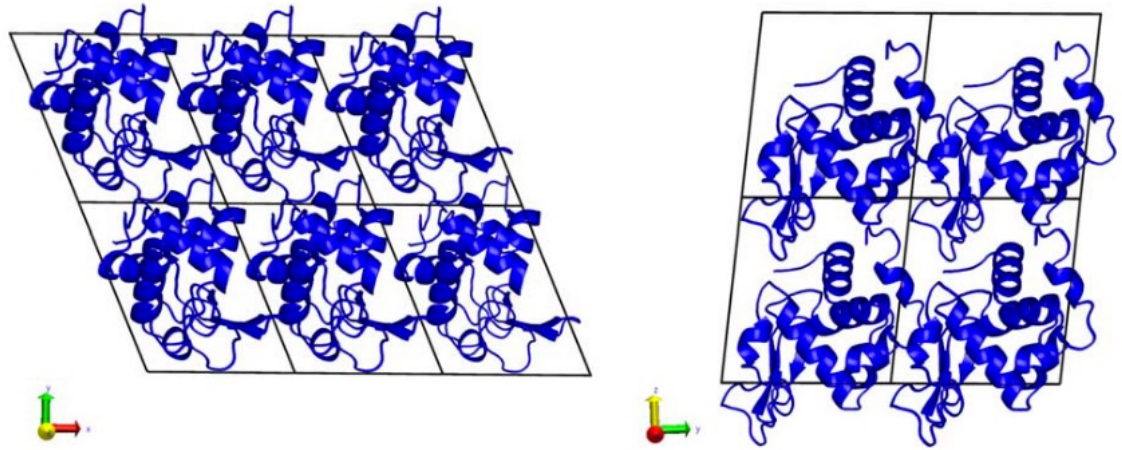


Figure 1. Simulation setup of the HEWL supercell. The P1 space group unit cell was extended three times along the crystallographic *a* axis and two times each along the *b* and *c* axes. Addition of solvent is described in Table I.

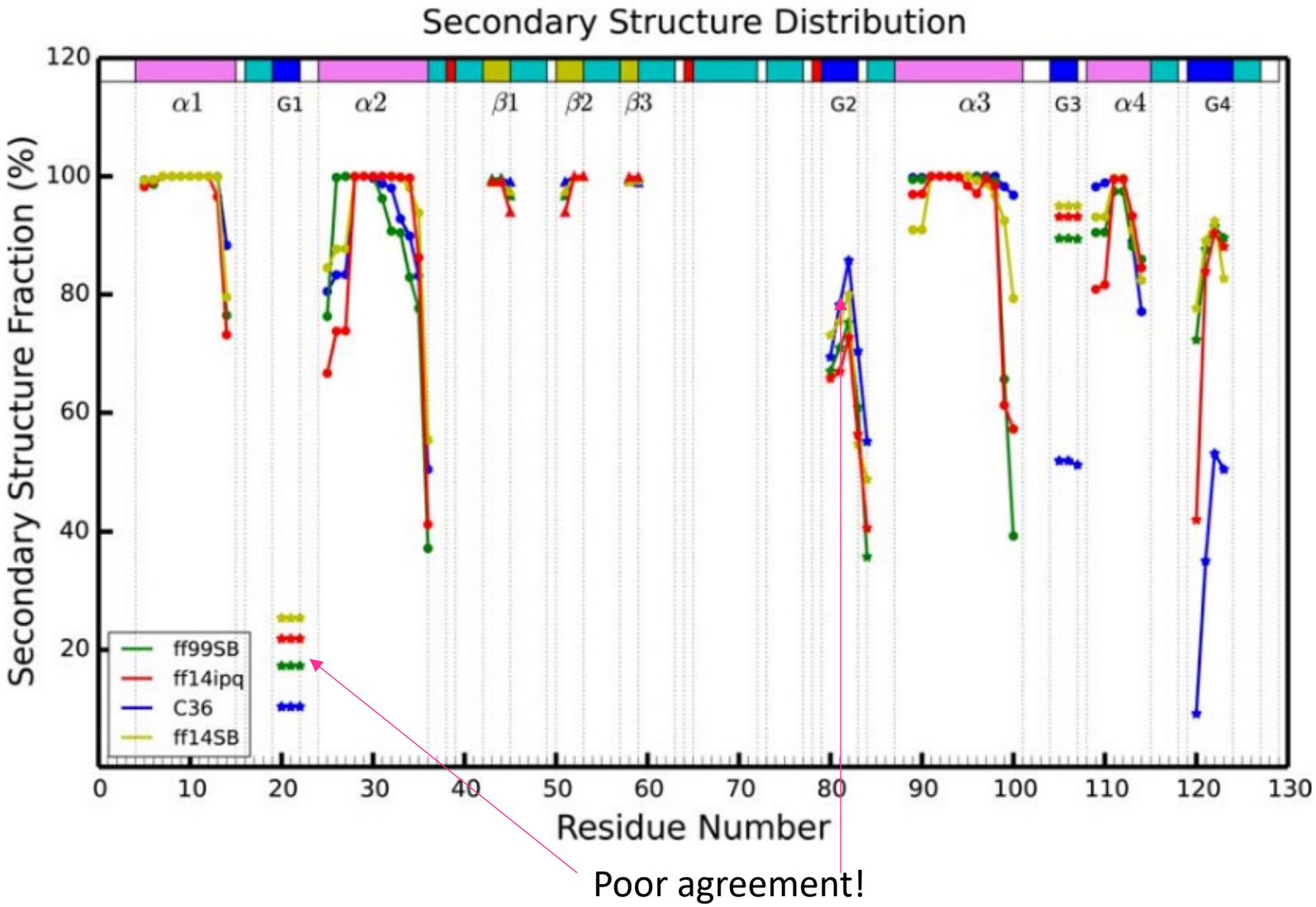
	Backbone RMS[a]	Heavy atom RMSD[b]
ff99SB	0.41	0.84
ff14ipq	0.40	0.77
C36	0.47	1.00
ff14SB	0.37	0.79

Backbone RMSD between 4LZT (simulated structure) and 3LZT is 0.28 Angstroms and between two other lysozyme structures is 0.37 Angstroms

http://doi.wiley.com/10.1002/pro.2713 PROTEIN SCIENCE 2016 | VOL 25:87–102

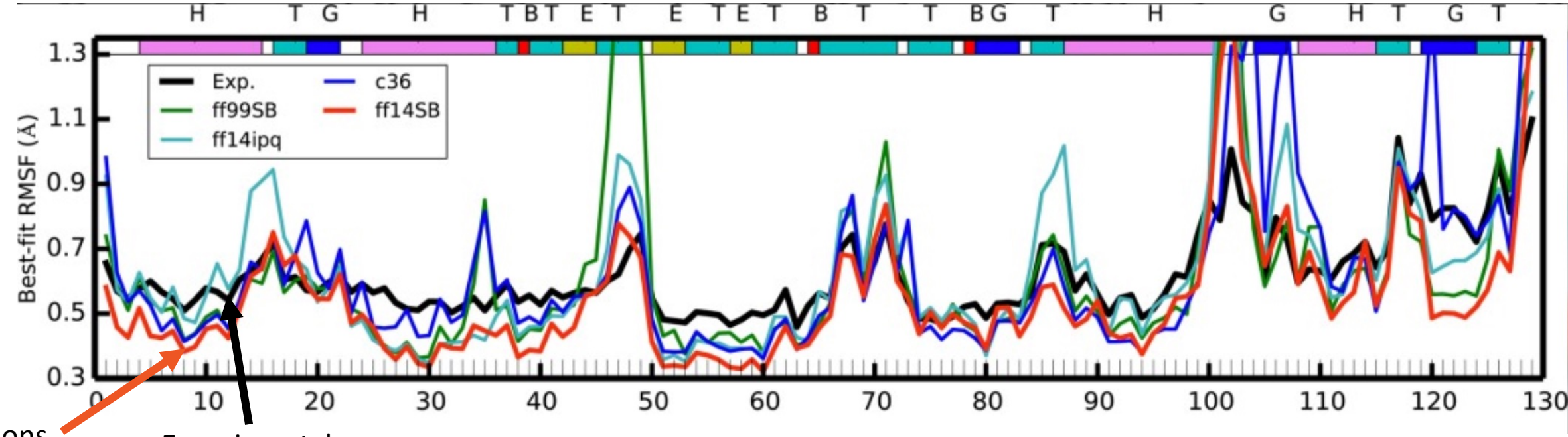
There are some problems with helix termini and 3_{10} α helix

G=3_{10} α helix
α=α helix
β=β sheet



http://doi.wiley.com/10.1002/pro.2713 PROTEIN SCIENCE 2016 | VOL 25:87–102

Simulations with ff14SB reproduce experimental fluctuations well



Simulations

Experimental
(B values from
room temperature crystal)

http://doi.wiley.com/10.1002/pro.2713 PROTEIN SCIENCE 2016 | VOL 25:87–102

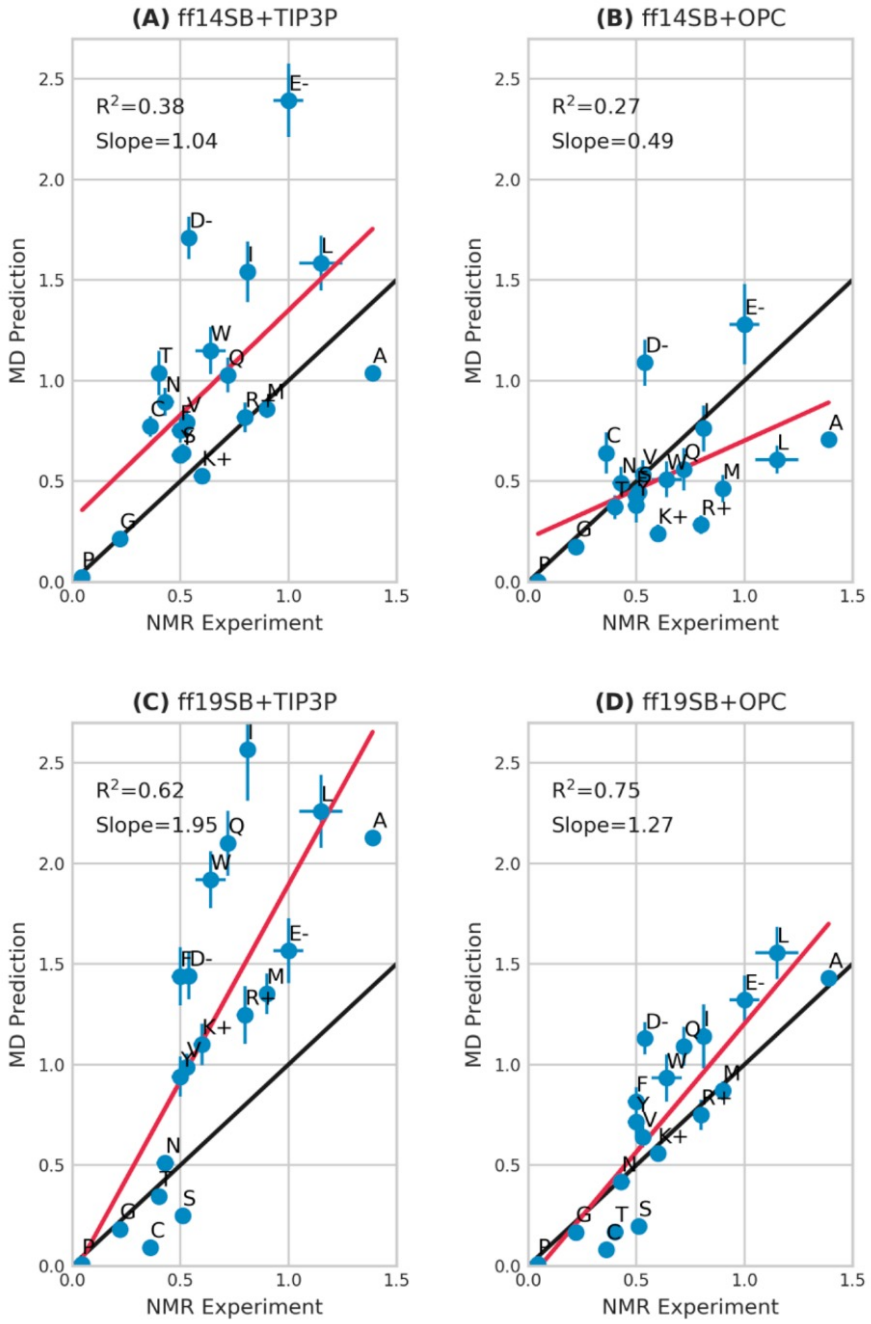
What about hydrogen bonds?

- Hydrogen bonds in the simulation were defined as <3.2 A between nitrogen or oxygen with a covalent hydrogen on one of the atoms (any angle).
  - Using this definition, the hydrogen bonding observed in the simulation reflected that observed in the crystal structure:
    - 3 “strong” H-bonds identified in the crystal are maintained >75% of the time in the simulation
    - One “weak” H-bond (as defined by crystal structure) is, as expected, less prevalent in the simulations (<30%).

http://doi.wiley.com/10.1002/pro.2713 PROTEIN SCIENCE 2016 | VOL 25:87–102

The water model-molecule forcefield combination used affects accuracy, some are not very accurate.

Helical propensity of amino acids

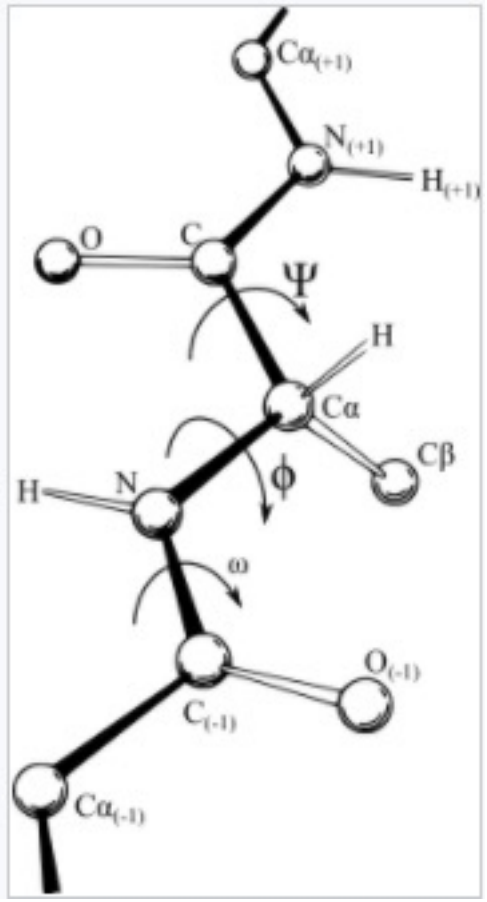


DOI: 10.1021/acs.jctc.9b00591
J. Chem. Theory Comput. 2020, 16, 528−552

Tom Cheatham (U of U) discusses convergence, reproducibility and accuracy in simulations

- https://www.youtube.com/watch?v=C3ptGf22Xiw
- 2:55-6:55 is the main part about accuracy.

Ramachandran plots AMBER ff vs. QM



Journal of Chemical Theory and Computation Article

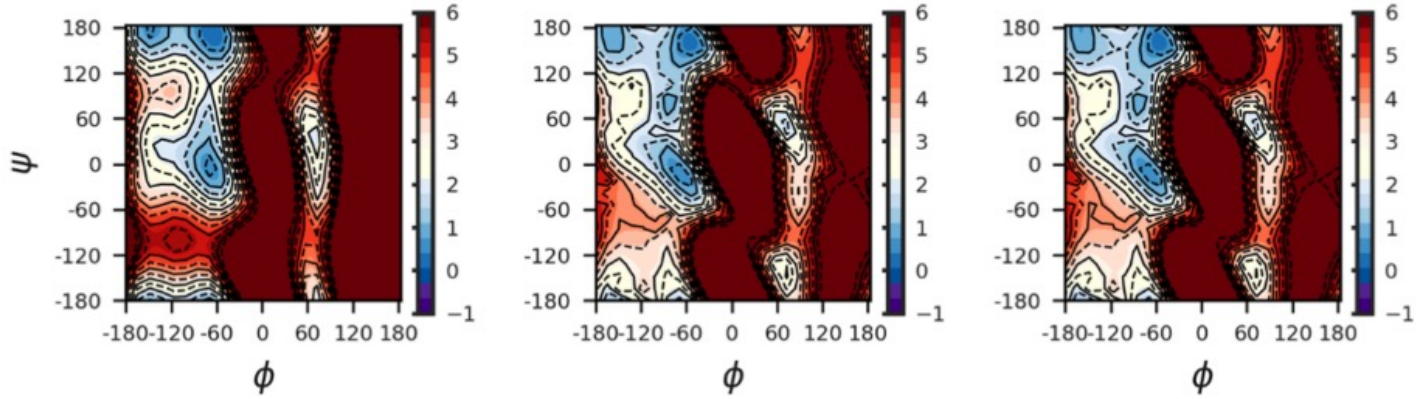


Figure 3. Ala dipeptide Ramachandran potential energy surfaces (kcal/mol) calculated in (left) ff14SB+GBSA, (middle) QM+SMD, and (right) ff19SB+GBSA. All energies were zeroed relative to the lowest energy in the ppII region (defined in Table S6). The values above the color bar range are depicted in dark red. Solid, labeled contours indicate integer energy values (kcal/mol), whereas dashed contours indicate half-integer energies. The bicubic spline interpolation implemented in Python was used to calculate values between grid points.

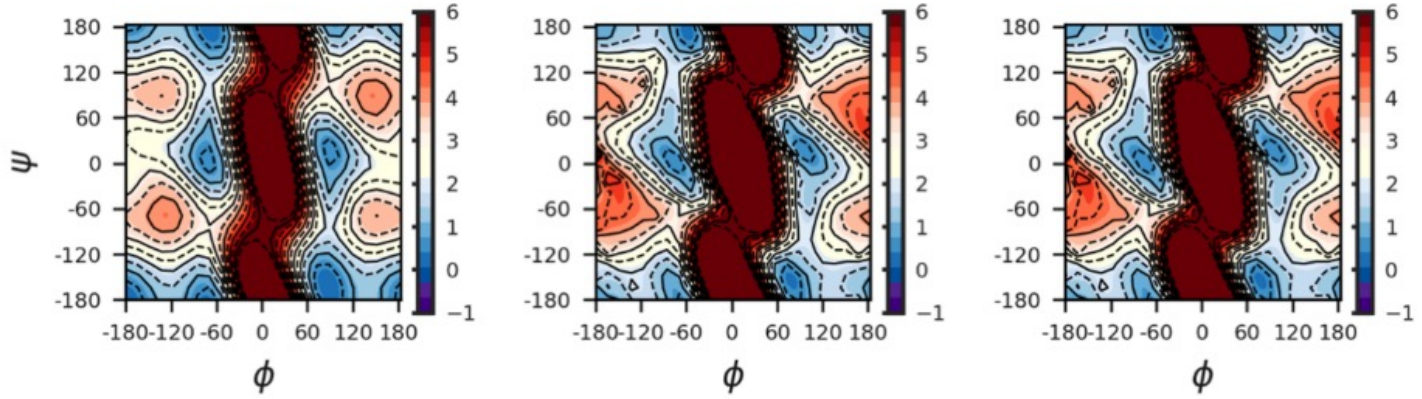
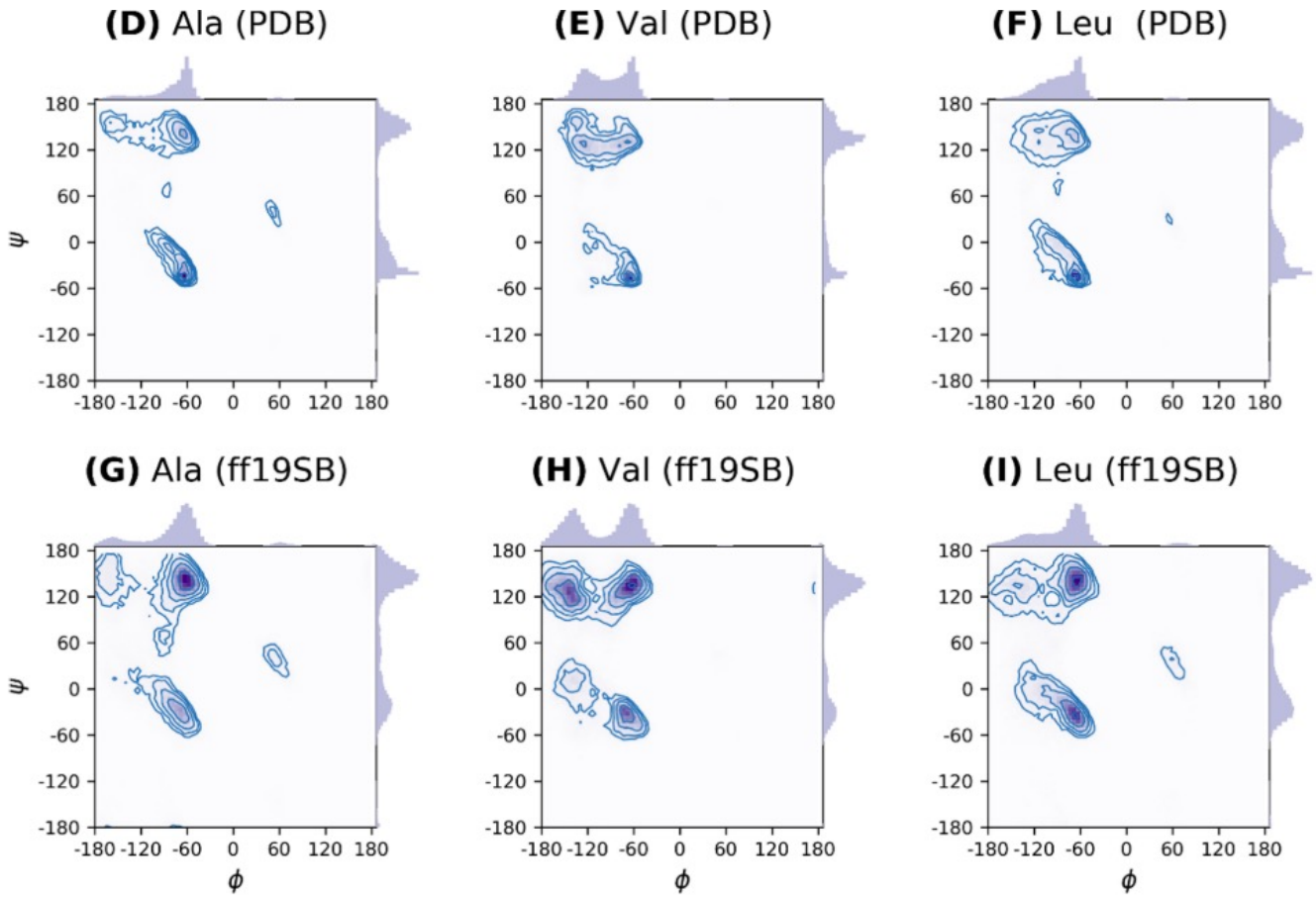


Figure 4. Gly dipeptide Ramachandran potential energy surfaces (kcal/mol) calculated in (left) ff14SB+GBSA, (middle) QM+SMD, and (right) ff19SB+GBSA. All energies were zeroed relative to the lowest energy at ppII region (defined in Table S6). The values above the color bar range are depicted in dark red. Solid, labeled contours indicate integer energy values (kcal/mol), whereas dashed contours indicate half-integer energies. The bicubic spline interpolation implemented in Python was used to calculate values between grid points.

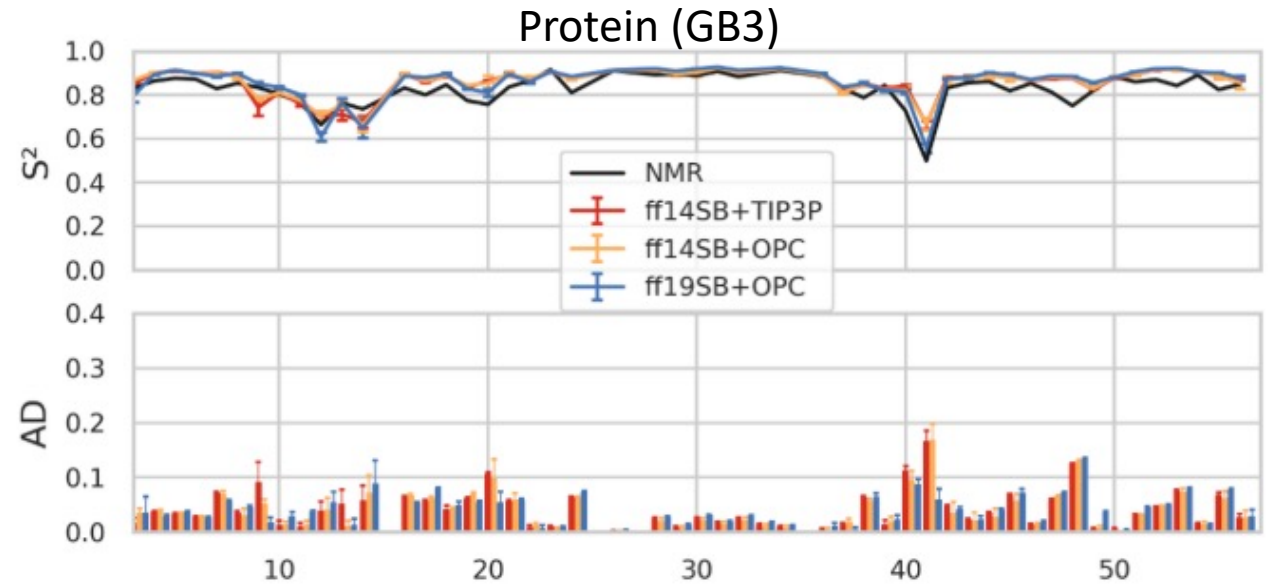
DOI: 10.1021/acs.jctc.9b00591
J. Chem. Theory Comput. 2020, 16, 528−552

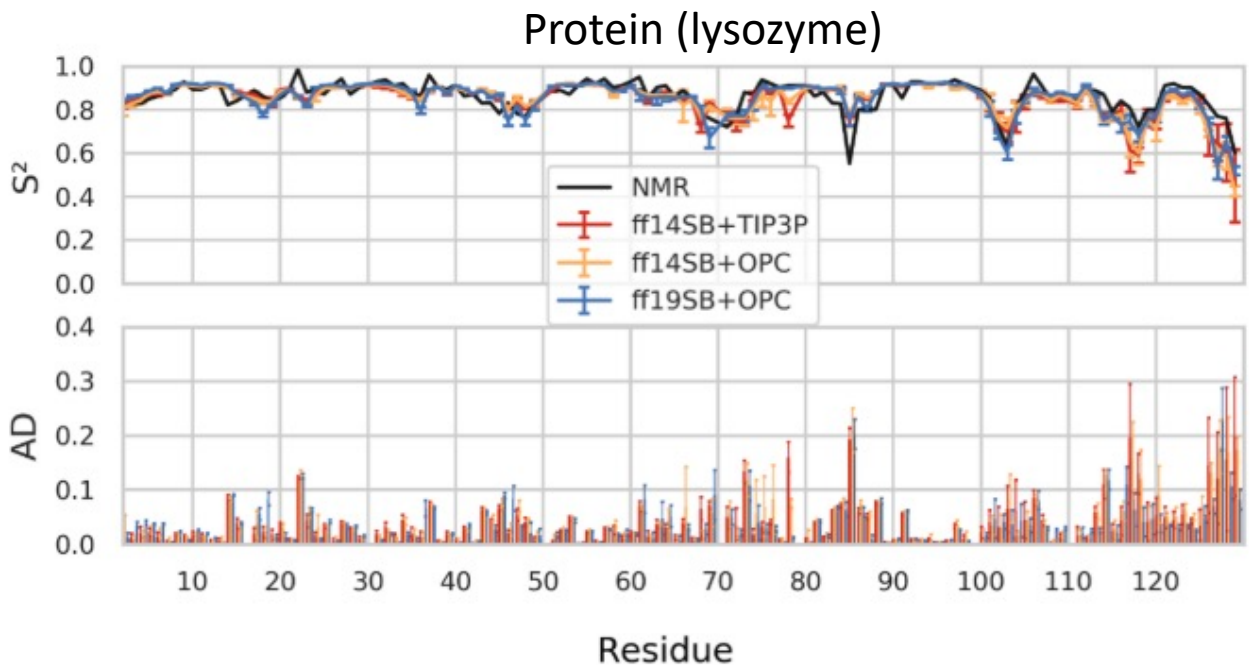
Ramachandran plots AMBER ff vs. PDB



DOI: 10.1021/acs.jctc.9b00591
J. Chem. Theory Comput. 2020, 16, 528−552

Short time scale dynamics (NMR order parameters)

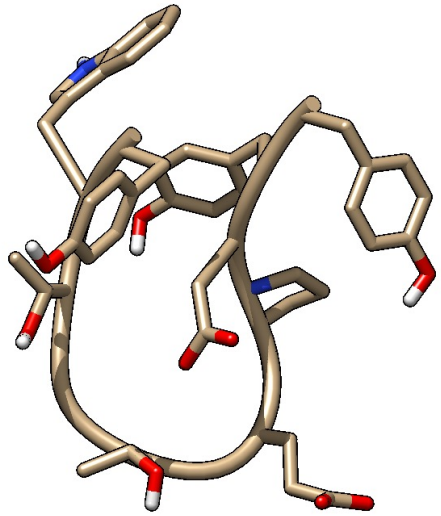


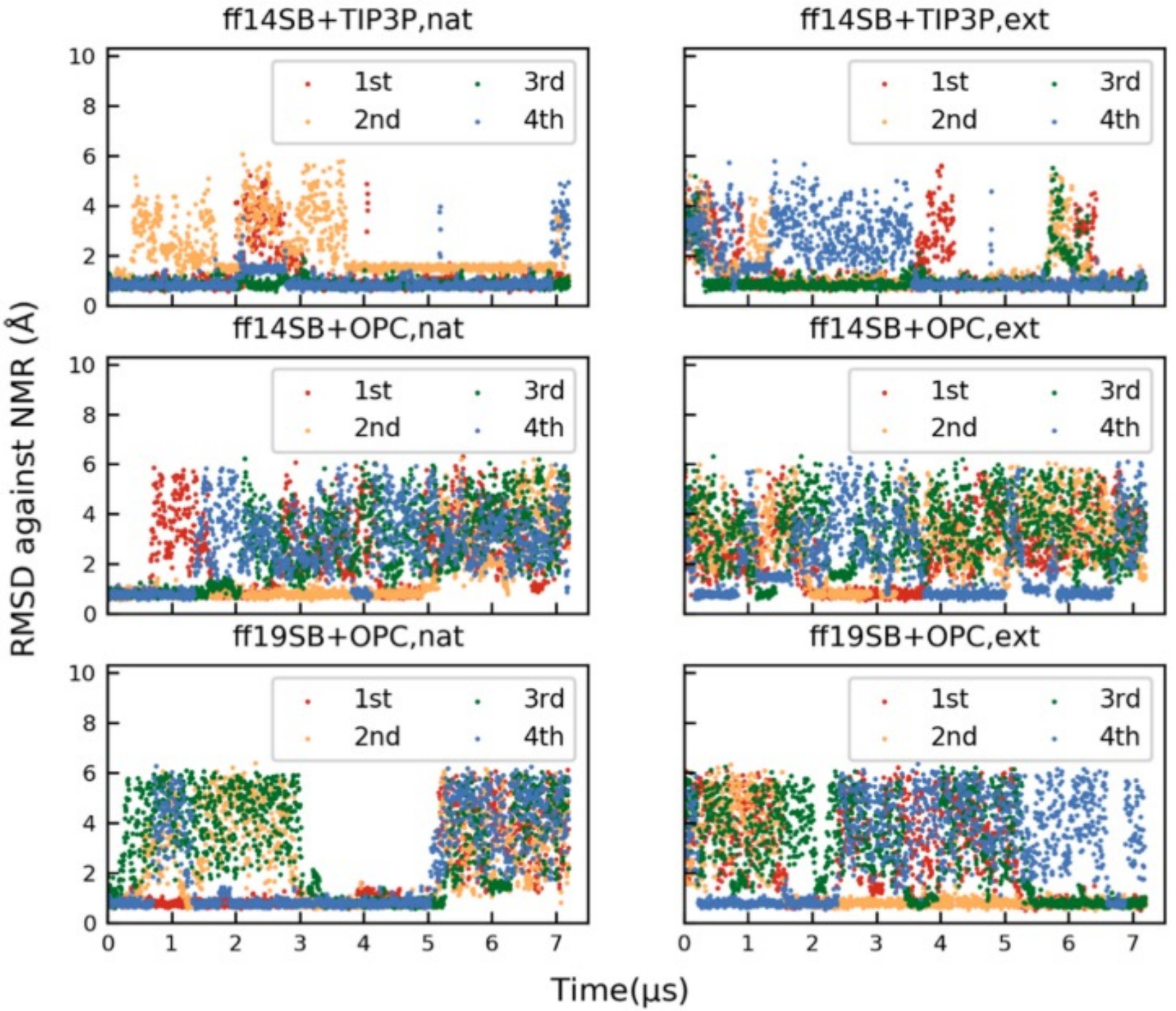


DOI: 10.1021/acs.jctc.9b00591
J. Chem. Theory Comput. 2020, 16, 528−552

Beta hairpins

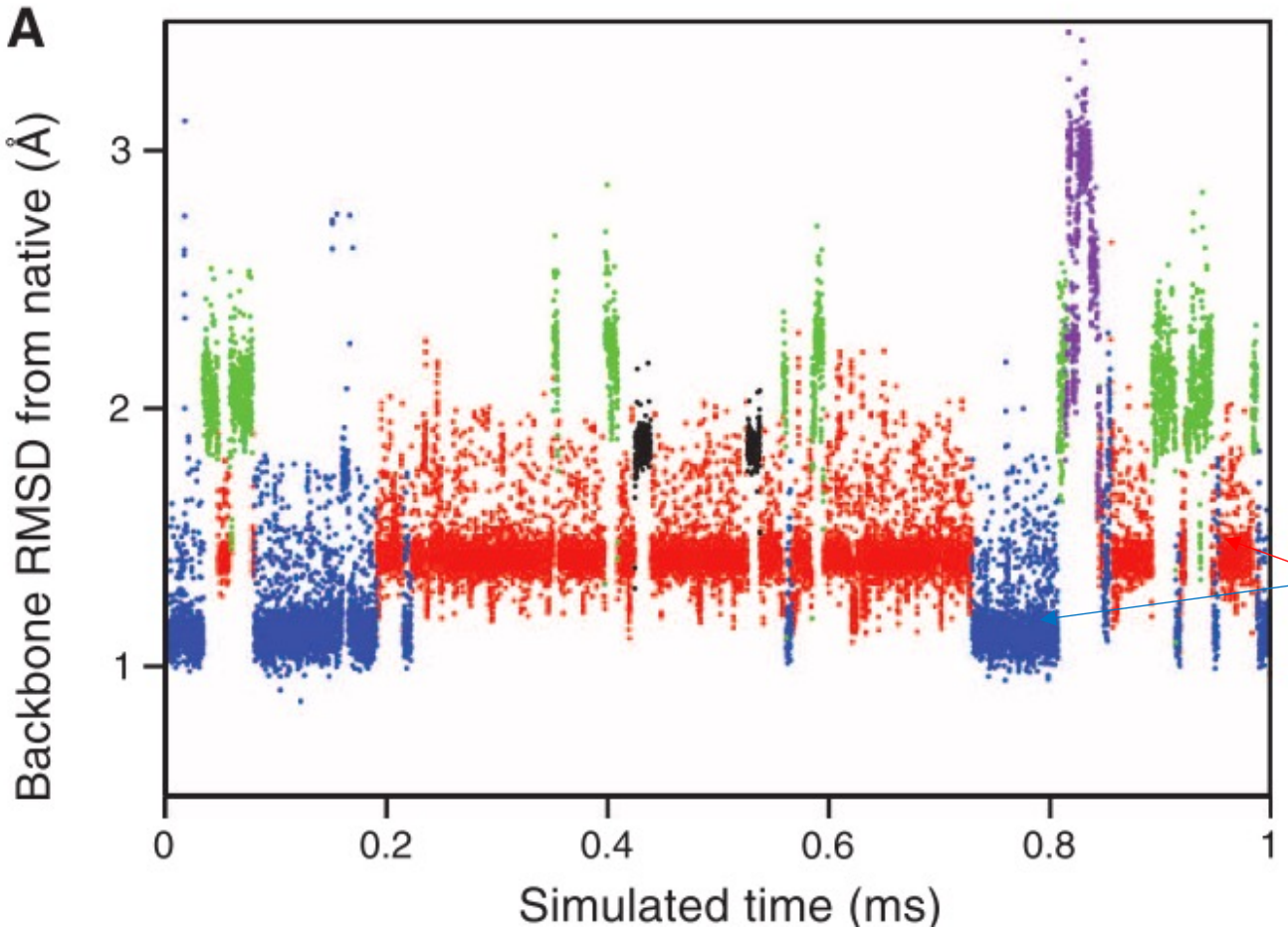
PDB 2RVD



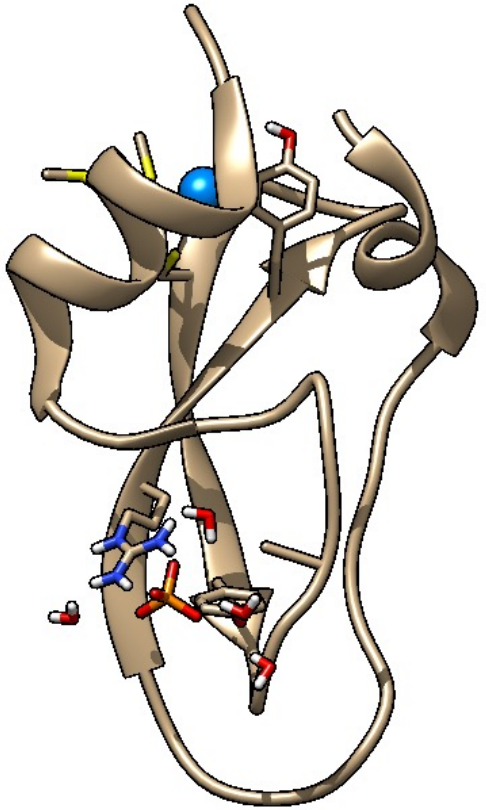


DOI: 10.1021/acs.jctc.9b00591
J. Chem. Theory Comput. 2020, 16, 528−552

Long time-scale dynamics (1ms simulation)



Expected states (based on experiments), however blue should be more populated. Suggests an error of ~2RT (~1.2 kcal/mol).



http://science.sciencemag.org/content/330/6002/341.full

Tutorial 5.1

- Build is simple, done in tleap
- Minimize energy (go quickly to a local minimum) and equilibrate (bring the simulation to life by adding thermal energy, raising the temperature)
  - There are many ways to do this but generally with explicit water you restrain the heavy atoms in the biomolecules and let the ions and waters move around during minimization
  - Then run MD with restraints
  - Then minimize again

Tutorial 5.1

- Run the simulation:
  - Sander is the free engine, pmemd requires a license
  - pmemd.cuda runs fast on GPUs
  - there are .mpi versions that can run parallell on multiple gpus at once, however multiple GPUs doesn’t gain you much.
- Analyzing simulations is hard:
  - CPPTRAJ is one tool that you can use to analyze (https://amberhub.chpc.utah.edu/cpptraj-cookbook/cpptraj-one-liners/)