Supplemental Material

I. Fragment libraries

As we described in Sec. II.F, two library types were used in our LBMC study: for proteins, peptide-plane configurations were employed, whereas residues were use for peptides. Peptide-plane configurations span between alpha-carbons, and residue configurations span between peptide bonds. The libraries differ in the degrees of freedom included with the fragments and the treatment of dummy atoms. Bellow we describe each library type in detail. All libraries and corresponding energies are available on our website: www.epdb.pitt.edu.

I.A. Peptide plane based libraries

The peptide plane fragments include all of the relative degrees of freedom (with respect to the subsequent fragment) except for the ψ dihedral. The ψ dihedral is not included to allow a sufficiently high density of configurations for generation of neighbor lists. We found that when the ψ dihedral was included with peptide planes, it was very hard to find "good" neighbor configurations because the density of configurations significantly decreased in the expanded configuration space (i.e., φ-ψ plane instead of φ only). To allow the incorporation of a separate Ramachandran potential in the future, the peptide plane libraries were modified to be conditional on the φ dihedral angle – i.e., to be uniformly distributed in φ with a suitable energy correction.

The dummy atoms include alpha-carbon at the beginning of the peptide-plane configuration and carbonyl carbon at the end. The alpha-carbon had the same forcefield parameters as in the corresponding fragment atom and the valence set to one. The van der Waals and partial charge parameters of the carbonyl carbon were set to zero and the valence was set to one.

Three peptide-plane types are employed, corresponding to Ala, Gly and Pro residues. In our simulations the library size used was $2.9*10^5$ for Ala, $2*10^5$ for trans-Pro, $2*10^5$ for cis-Pro and $3.6*10^5$ for Gly. The cis-Pro library was implemented to allow the sampling of proteins containing a Pro residue in cis conformation. For this initial study, peptide planes corresponding to cis-Pro were swapped only with the cis-Pro library and trans-Pro planes were swapped only with the trans-Pro library. For all libraries, the neighbor lists were generated to contain 10 configurations.

Libraries of peptide plane configurations were generated using Langevin dynamics as implemented in the Tinker v. 4.2 software package [1] with OPLS-UA forcefield [2] and implicit GB/SA solvent [3] at 298 K.

I.B. Residue-based libraries

The residue-based fragments include all the degrees of freedom comprising the entire amino acid residue from N-H to the C=O group. Three types of fragments were used in this study, corresponding to Ace, Ala, and Nme residues.

The dummy atoms used at the N-terminus of a fragment are carbonyl carbon, carbonyl oxygen, and terminal alpha-carbon with valence set to one. The dummy atoms used at the

C-terminus include amide nitrogen, amide hydrogen, and terminal alpha-carbon, also with valence set to one. The dummy atoms were assigned the same forcefield parameters as used in the corresponding fragment atoms. Thus dummy-atom interactions affect the library distribution and energy. These interactions are accounted for in the modified acceptance criterion as described in Sec. II.F.

The residue fragments are distributed according to the OPLS-AA forcefield [4] in a simple solvent model with uniform dielectric of 60 at 298 K. The fragment libraries were generated by sampling the internal coordinates independently as described in Ref. [5]. Each residue library employed 10^4 configurations.

References

(1) Ponder, J. W.; Richard, F. M. J. Comput. Chem. 1987, 8, 1016-1024.

(2) Jorgensen, W. L.; Tirado-Rives, J. J. Am. Chem. Soc. 1988, 110, 1657-1666.

(3) Qiu, D.; Shenkin, P. S.; Hollinger, F. P.; Still, W. C. J. Phys. Chem. A 1997, 101, 3005-3014.

(4) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. J. Am. Chem. Soc. 1996, 118, 11225-11236.

(5) Zhang, X.; Mamonov, A. B.; Zuckerman, D. M. J. Comput. Chem. 2009, in press.