Electronic Supplementary Material (ESI) for Physical Chemistry Chemical Physics.

Toward Structure Prediction of Cyclic Peptides

Hongtao Yu and Yu-Shan Lin*

Department of Chemistry, Tufts University, Medford, Massachusetts 02155, United States

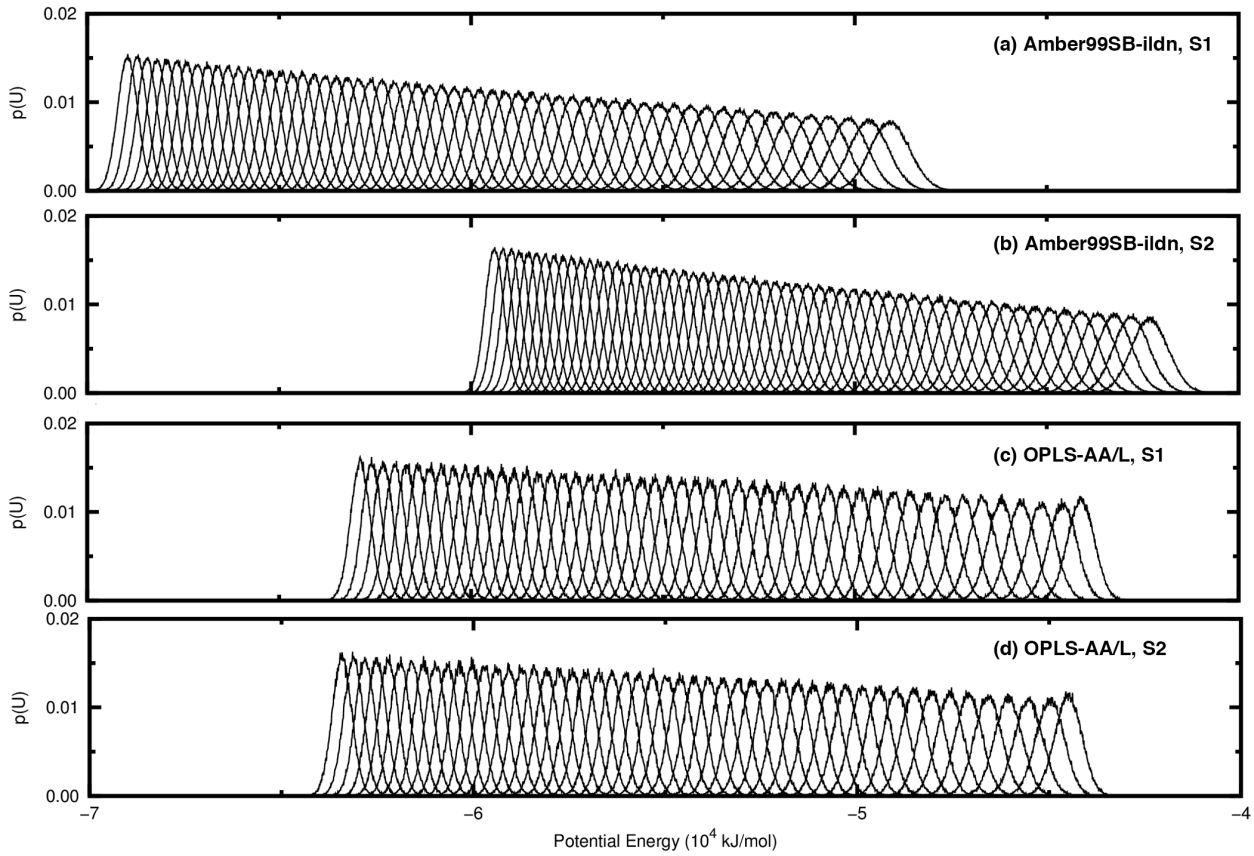

Fig. S1. Potential energy distributions for each replica in the REMD simulations. (a) Amber99SB-ildn simulation using input structure S1; (b) Amber99SB-ildn simulation using input structure S2; (c) OPLS-AA/L simulation using input structure S1; (d) OPLS-AA/L simulation using input structure S2.

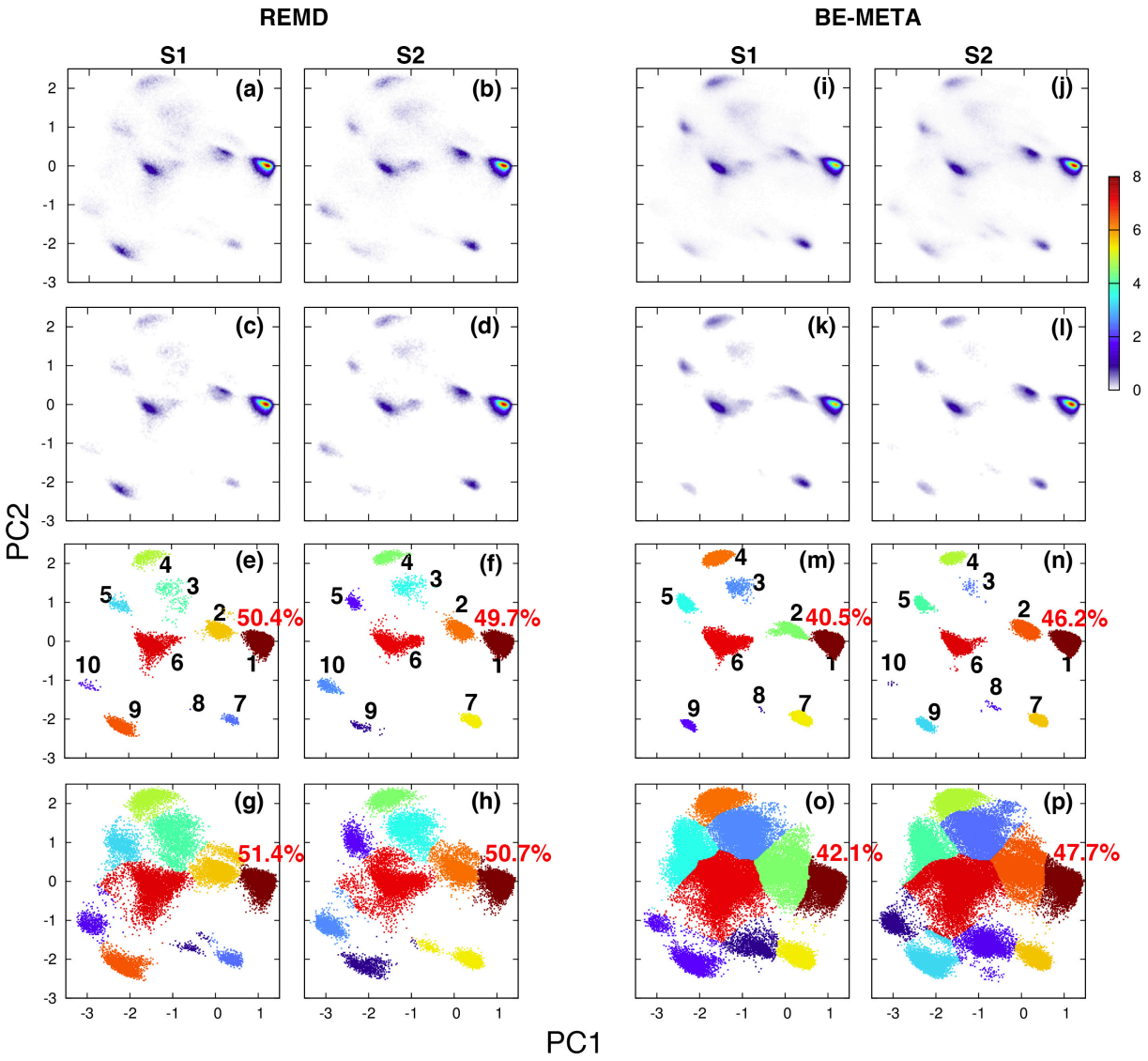

Fig. S2. Dihedral PCA and cluster analysis results for cNPF1 using the OPLS-AA/L force field. Columns from left to right: REMD results with initial structure S1, REMD results with initial structure S2, BE-META results with initial structure S1, and BE-META results with initial structure S2. First the original 2D principal subspace (top row) was divided into 200×200 grids. Before clustering, grids with data populations lower than 0.1 were removed (second row). In the resulting cluster analysis results (third row), the clusters are colored based on their populations with the largest cluster colored in dark red and the smallest cluster colored in dark blue. To estimate the upper bound of the error resulted from discarding the grids with data populations lower than 0.1, we assigned each discarded grid a cluster ID based on the assignation of the clustered grid it is closest to (fourth row).

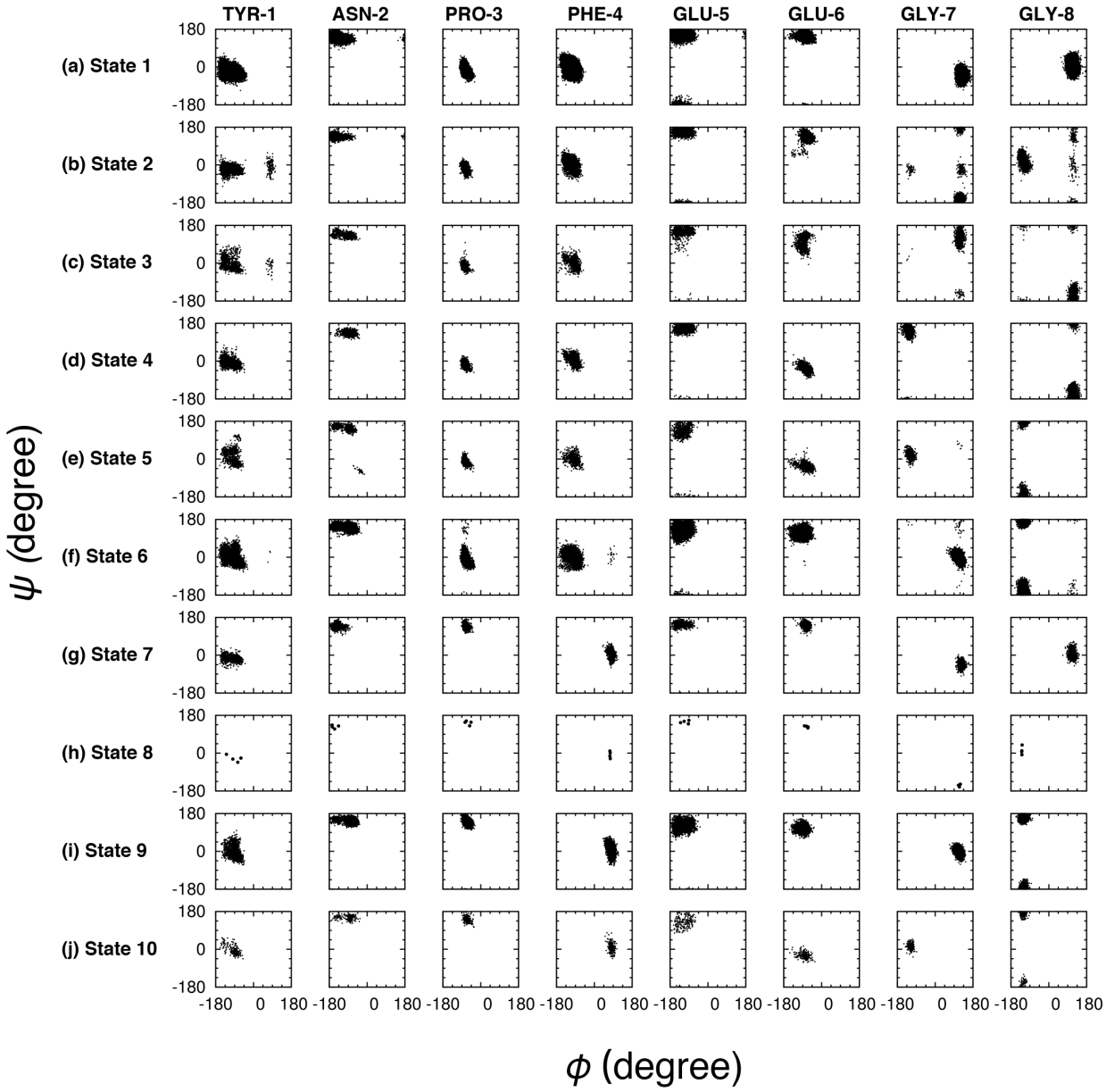

Fig. S3. φ/ψ distributions of the 8 residues in cNPF1 for the 10 states identified from clustering the OPLS-AA/L REMD simulations (**Fig. S2**).

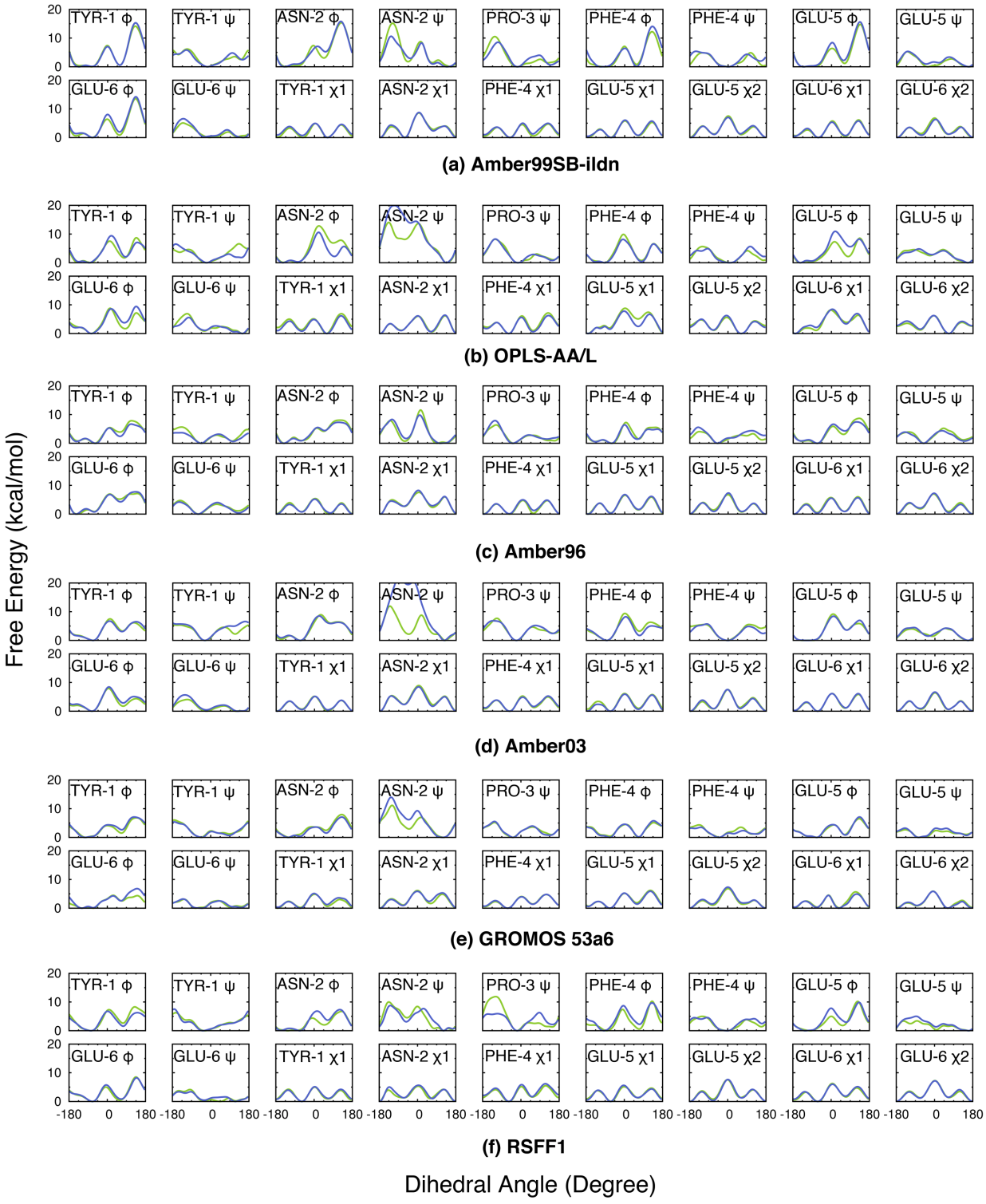

Fig. S4. Free energy profiles along the 18 collective coordinates calculated from S1 (green) and S2 (blue) BE-META simulations using different force fields. All simulations were 300 ns in length.

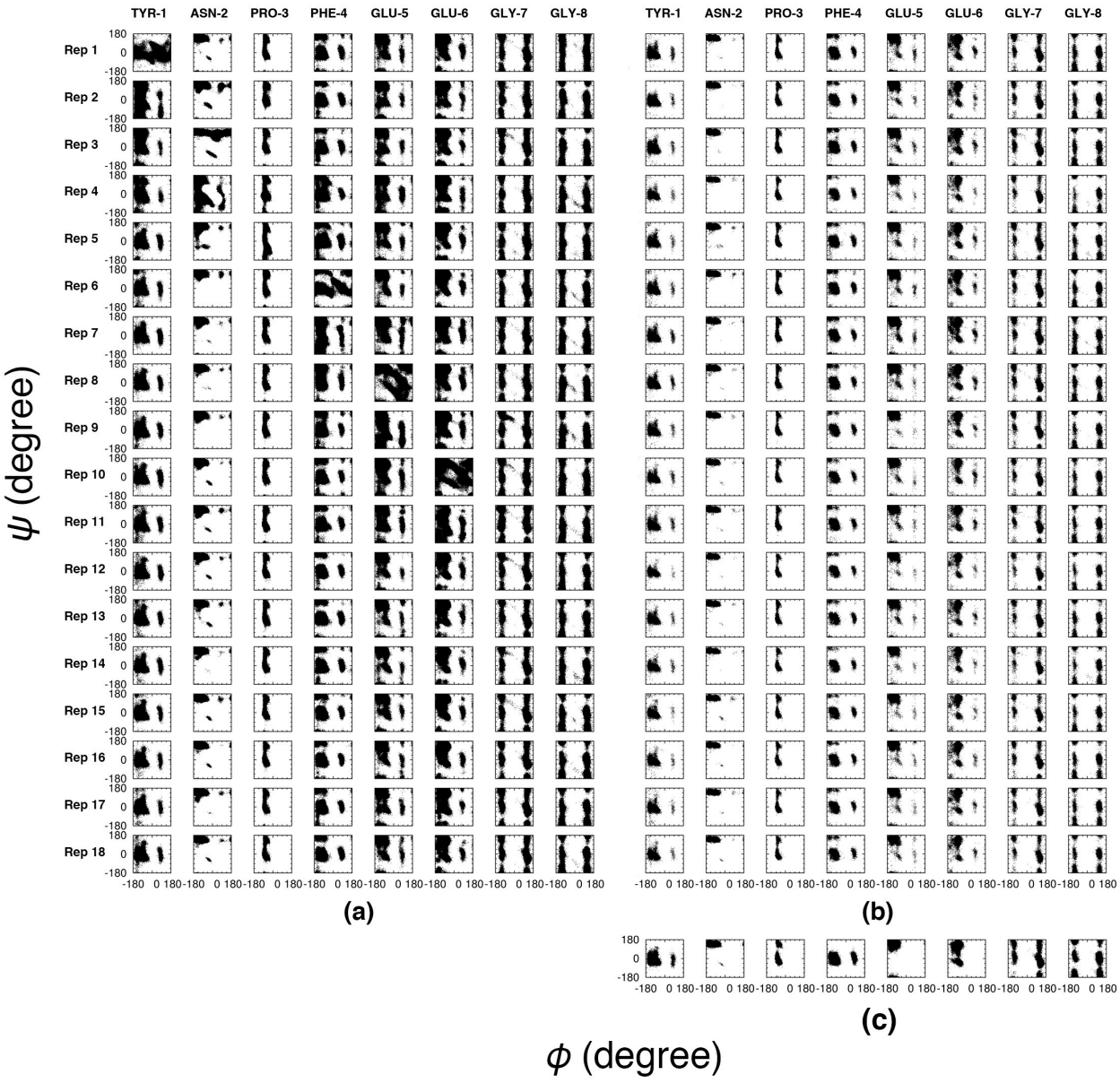

Fig. S5. φ/ψ distributions of the 8 residues of cNPF1 in the OPLS-AA/L S1 simulation calculated using (a) the last 100 ns of the original BE-META trajectories，(b) the last 100 ns of the Boltzmann reweighted BE-META trajectories, and (c) the last 50 ns of REMD trajectory.

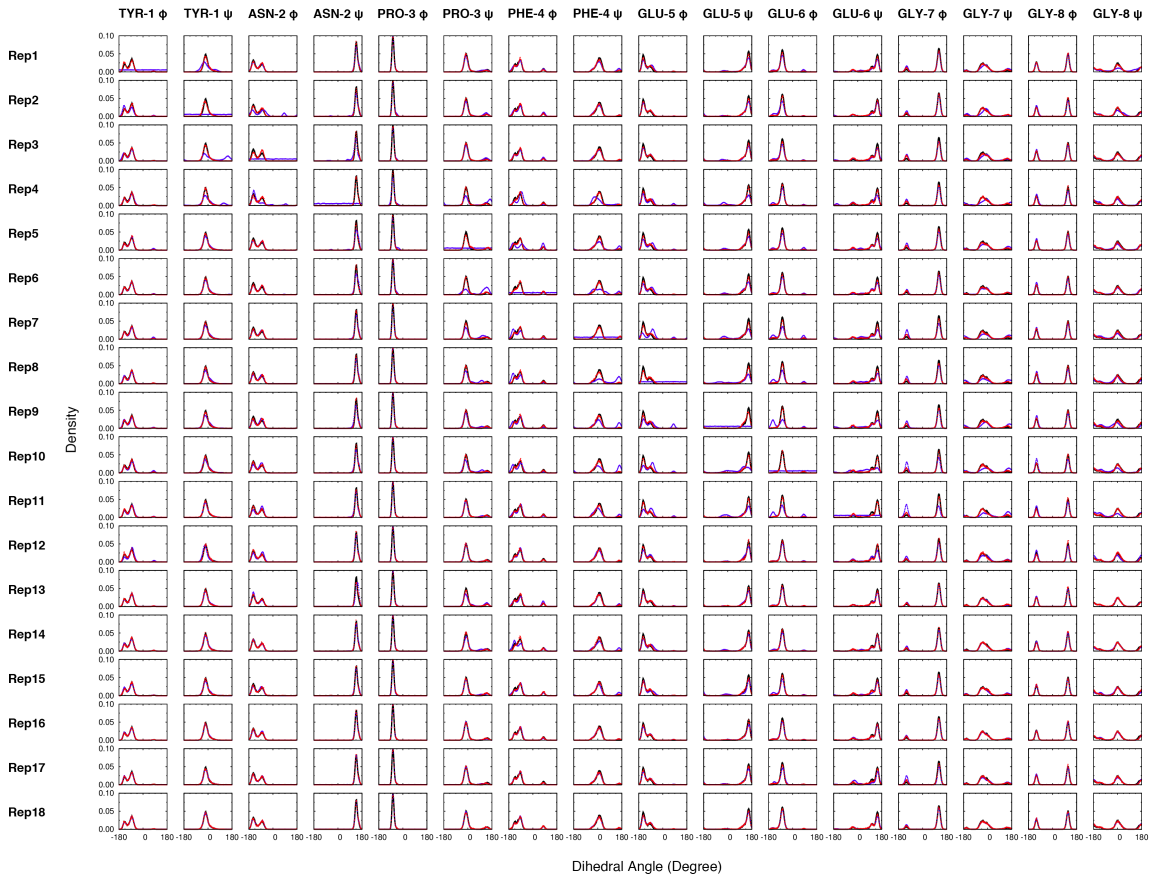

Fig. S6. Dihedral distributions of the 8 residues of cNPF1 in the OPLS-AA/L S1 simulation calculating using the last 50ns of REMD trajectory (black), the last 100ns of the original BE-META trajectories (blue), and the last 100ns of the Boltzmann reweighted BE-META trajectories (red).

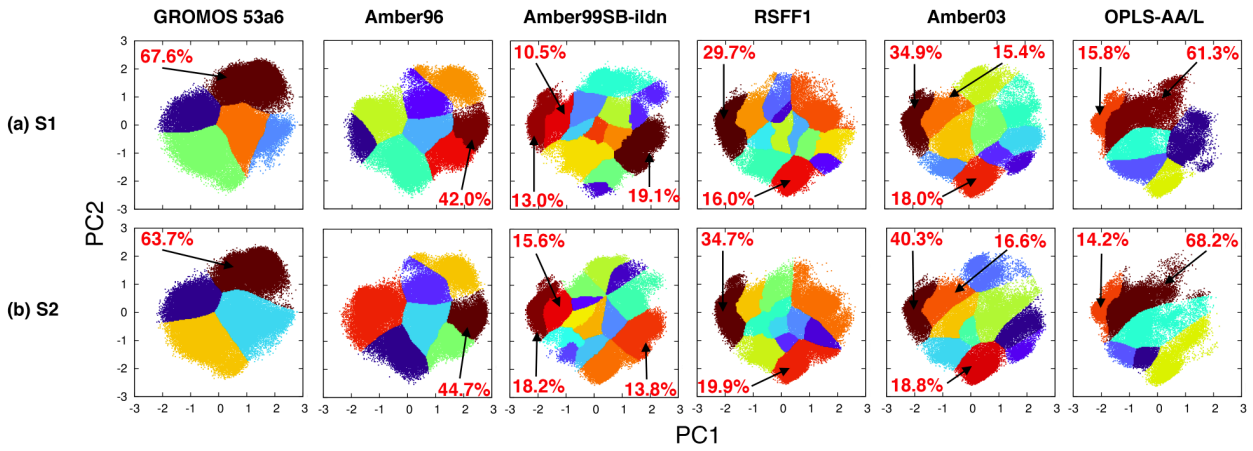

Fig. S7. Clustering of the grids in 2D principal subspace for the twelve weighted BE-META trajectories (6 force fields × two input structures S1 and S2), all grids assigned. Before clustering, the original 2D principal subspace was divided into 200×200 grids and grids with data populations lower than 0.1 were removed. To estimate the upper bound of the error resulted from discarding the grids with data populations lower than 0.1, we assigned each discarded grid a cluster ID based on the assignation of the clustered grid it is closest to.