Supporting Information

Performance Comparison of Systematic Methods for Rigorous Definition of Coarse-Grained Sites of Large Biomolecules

Yuwei Zhang,[†,‡] Zexing Cao,[†] John Zenghui Zhang,[‡,§] Fei Xia*[,‡,§]

[†]*State Key Laboratory of Physical Chemistry of Solid Surfaces and Fujian Provincial Key Laboratory of Theoretical and Computational Chemistry, College of Chemistry and Chemistry Engineering, Xiamen University, Xiamen 361005, China*

[‡]*School of Chemistry and Molecular Engineering, East China Normal University, Shanghai 200062, China.*

[§]*NYU-ECNU Center for Computational Chemistry at NYU Shanghai, Shanghai 200062, China*

**Corresponding Author E-mail: fxia@chem.ecnu.edu.cn.*

Section S1. ANM Model

ANM is constructed by connecting the Cα atoms in protein with harmonic springs. The potential of ANM is shown in Eq.(1) below:

$$V_{ANM} = \sum_{i}^{M} \sum_{j>i}^{M} \frac{k}{2} \cdot \left(r_{ij} - r_{ij}^{0}\right)^2 \cdot H\left(r_c - r_{ij}^{0}\right) \quad (1)$$

where M is the number of Cα atoms, k and r_c are the uniform spring constant and cut-off distance, r_{ij} and r_{ij}^{0} are the instantaneous distance and equilibrium distance between the Cα atoms i and j, respectively. $H(x)$ is the Heaviside step function. When $x > 0$, $H(x) = 1$; Otherwise $H(x) = 0$. In our work, k and r_c adopt the empirical values of 100 $kcal/(nm)^2$ and 1.5 nm, respectively. When $r_{ij} > r_c$, there is no interaction between the Cα atoms i and j. For ANM containing M Cα atoms, $3M$-6 eigenvectors are generated by normal mode analysis. These eigenvectors are used for the coarse-graining in ED-CG or the evaluation of distance fluctuations used in FM-CG.

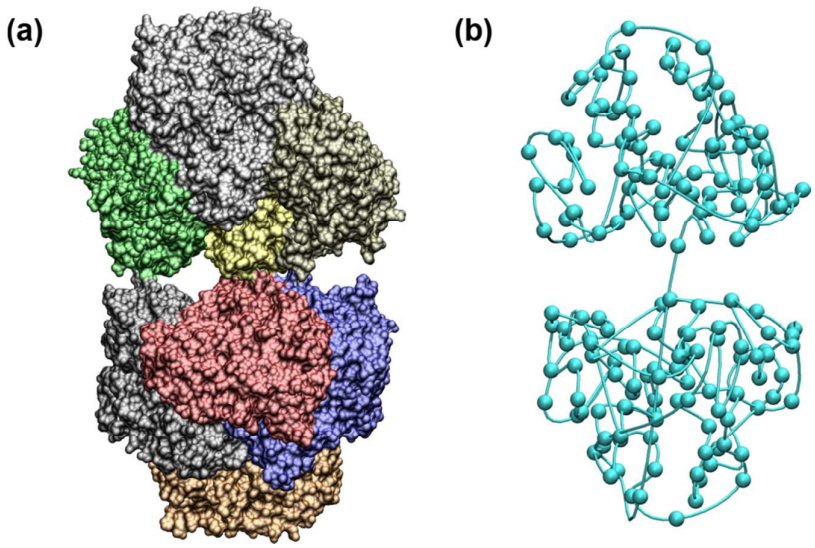

Figure S1 (a) The contour map of the all-atom model of Formolase protein (PDB 4QPZ) with 4497 residues. (b) The optimal 181 CG sites derived by SLIO based on ED-CG at the time cost of 2.4 seconds.

Table S1. Optimized values of χ^2 (nm^2) by SASD and SOBC based on ED-CG for F-actin with the CG number $N = 4$. The samplings of SASD and SOBC are 100, while SLIO performs only once with the value 82.958308 nm^2 obtained. The optimal values of SASD and SOBC are 84.722629 and 82.958683 nm^2, respectively.

Results of SASD:

χ^2_{SASD} (nm^2)				
89.808226	94.211313	95.562618	92.270700	94.895964
91.045174	93.947219	93.439861	95.875256	95.546767
89.861590	93.730051	88.092014	87.026986	86.869723
93.425717	88.809727	87.840673	92.469413	96.114167
90.565028	93.878252	90.956111	90.385555	92.330416
87.476881	94.166396	93.481424	94.881853	91.979100
96.258870	90.230059	97.227651	89.872963	92.102583
92.224629	92.327777	93.649151	95.900138	90.187839
87.705686	95.042634	91.017615	100.763582	91.776656
93.353101	84.722629	97.312729	89.152472	97.450517
95.863463	91.841703	98.266525	94.297916	92.536582
92.333974	86.968455	95.563620	95.034926	89.467820
92.641051	94.593792	95.247934	94.068368	88.611423
92.772178	95.364367	91.658780	97.182721	84.722629
85.824775	96.398809	95.911638	85.463245	94.344740
91.750514	87.623170	93.538861	94.050640	91.096140
89.420799	93.177133	85.339832	90.776770	94.057826
93.285402	88.023774	90.897943	96.012150	88.961745
94.546844	96.532384	88.040830	96.467800	93.231397
91.779362	91.178068	89.496630	85.583636	94.548622

Results of SOBC:

χ^2_{SOBC} (nm^2)				
85.685299	85.728465	83.146230	88.521408	89.246474
89.002393	83.270981	90.614003	84.110614	89.320758
88.723320	83.992084	92.759306	86.510129	83.615513
91.646134	91.629505	88.146170	88.018341	82.958683
89.275418	87.863836	87.221481	92.111703	83.601150
89.007621	83.668183	84.086510	85.184588	87.215276
87.557487	88.877657	89.445404	95.058621	94.937346
86.848383	90.012111	83.857571	84.930223	93.852160
83.577801	83.718436	83.314077	90.041934	92.159442
88.494590	85.690390	92.946934	92.631321	85.870463
87.259617	87.412423	95.608812	90.356861	88.341180
84.155757	85.532280	82.976035	84.347394	85.901304
92.611789	86.149543	87.912743	83.823700	92.415069
84.918394	84.803603	92.757755	87.534998	89.093351
87.851806	87.815916	87.705643	83.339770	89.012362
90.167817	84.166282	87.449007	91.026705	84.939740
86.414106	91.812152	90.577712	86.530481	92.281292
89.485595	85.976164	86.027175	85.496629	84.889453
87.444966	88.788842	90.104843	86.149957	91.052835
89.235177	88.103428	85.264507	86.310972	86.371548