論 文 の 内 容 の 要 旨

論文題目　Association and dissociation simulations of bio-molecular complex using parallel cascade selection molecular dynamics

（並列カスケード選択分子動力学法による生体分子の会合・解離シミュレーション）

氏　名　チャン　フ　ズイ

Sampling conformations of protein complexes during association and dissociation processes is a crucial step to estimate the binding free energy and other kinetic properties from association/dissociation pathways. This is a challenging problem for the classical Molecular Dynamics (MD) simulation because the time scale of these processes exceeds the limit of current computation. Therefore, enhanced sampling techniques play an important role to generate sufficient data for the free energy analysis. For example, Steered Molecular Dynamics (SMD) with Umbrella Sampling (US) [Ramirez et al., Methods Enzymol. (2016)], Replica Exchange Umbrella Sampling (REUS) [Sugita et al., J. Chem. Phys (2000)], Targeted MD (TMD) [Schlitter et al., J. Mol. Graph. (1994)], Parallel Cascade Selection Molecular Dynamics (PaCS-MD) [Harada and Kitao J. Chem Phys. (2011)] and other methods not listed here are used for this purpose. Recently, Yamashita and Fujitani showed that protein structures were distorted when dissociation of lysozyme (enzyme) and HyHEL-10 (inhibitor) was simulated by SMD using a steering force applied to the center of mass (COM) of the protein, which led overestimation of the potential of mean force (PMF) with the following US. This can be considered as the artifact caused by SMD. In contrast to SMD, PaCS-MD performs conformational sampling by cycles of distinct multiple Molecular Dynamics (MD) simulations without applying any bias force to the system. It enhances the sampling by selecting the MD snapshots closest to the destination state and by restarting the MD simulations from the selected snapshots with the velocity re-randomization. PaCS-MD was shown to be very successful in efficient sampling of protein domain motions. Here in this thesis, we describe unbiased association and dissociation simulations by PaCS-MD.

We first show that PaCS-MD dissociated a small ligand, tri-N-acetyl-D-glucosamine (triNAG), from hen egg white lysozyme (LYZ) very efficiently. We performed PaCS-MD trials with 3 different simulation settings: PaCS-MD$^{10,0.1}$ (ten 0.1 ns MDs per cycle), PaCS-MD$^{100,0.1}$ (hundred 0.1 ns MDs) and PaCS-MD^{10,1} (ten 1.0 ns MDs). We found that PaCS-MD is 5 times faster than SMD. In combination with Markov State Model (MSM), we calculated the binding free energy directly from the PaCS-MD trajectories. In comparison, binding free energy was also calculated by the analysis of SMD trajectories using the Jarzynski equality [Jarzynski., Phys. Rev. Lett. (1997)]. Although SMD/Jarzynski overestimated the binding free energy, PaCS-MD/MSM yielded the results in good agreement with experimental results. We also examined the effects of the number of replicas, the length of each MD, the velocity re-randomization, and the selection of snapshots on PaCS-MD sampling. We found that the increase of the number of replicas reduced the number of cycles required for dissociation because the probability of observing rare events is proportional to the number of replicas. The velocity re-randomization enhances the sampling in the

bound state as it acts as a perturbation to raise the occurrence of rare events (dissociation).

We next applied PaCS-MD to the dissociation of MDM2 protein and trans-activation domain of p53 (TAD-p53). Binding free energy of MDM2/TAD-p53 calculated by PaCS-MD/MSM was $40.5 \pm 1.7\ \mathrm{kJ/mol}$, which almost agrees with experimental value 37.7 ± 1.7. kcal/mol. Our result is better than the value calculated by the MMGBSA method, 68.2 kJ/mol [Dastidar et al., JACS (2008)]. We found the binding free energy is strongly dependent on the dissociation pathway of TAD-p53, which is related to the dissociation of the key residues PHE19 and TRP23 of TAD-p53 involved in π-π stacking interactions between TAD-p53 and MDM2.π

We also employed PaCS-MD for simulating association and dissociation process of MDM2/TAD-p53, which can be considered as a flexible-body docking simulation. We used switching conditions between the dissociation and association simulations as follows: if the association simulation does not make any progress for continuous 20 ps, it will switch to the dissociation simulation. When the inter COM distance between MDM2 and TAD-p53 reaches 2.0 nm longer than the last switching point, the association simulation will start. We performed 274 cycles of PaCS-MD and examined whether generated structures of TAD-p53 and MDM2 complex are similar to the crystal complex structure and found that the minimum RMSD was 0.429 nm. In addition, TAD-p53 could bind to the correct binding interface without the guiding force. We further examined 4 representative structures selected from all the bound conformations. Although the two key π-π stacking interactions were not formed in these structures, residual contacts are in agreement with those in the crystal structure with the binding interface RMSD of 0.243 nm. To predict the bound conformation without prior-knowledge of the crystal structure, we examined if the conformation similar to the correct bound conformation can be identified as the lowest free energy structure. We built MSM based on the trajectories of distance RMSD (dRMSD) from the initial conformation of MDM2/TAD-p53 in the unbound state and calculated the Potential of Mean Force (PMF). We found that dRMSD of the lowest PMF position was 4.21 nm, which the corresponding structure was identical to the structure with the lowest interface RMSD from the crystal structure. Therefore, we can select the best structure based on the calculated RMSD.

In conclusion, PaCS-MD algorithm was shown to be an efficient unbiased enhanced sampling tool which can be applied to bio-molecular complexes and is highly suitable for distributed computing. Overall, PaCS-MD is faster in computational time than the other biased sampling techniques. We are currently making an effort to apply PaCS-MD for reducing total simulation time of flexible-body docking simulation.