論文内容の要旨

論文題目　Free energy landscape analysis of biomolecules
by massive parallel multi-scale simulation
(超並列マルチスケールシミュレーションによる生体分子の自由エネルギー地形解析)

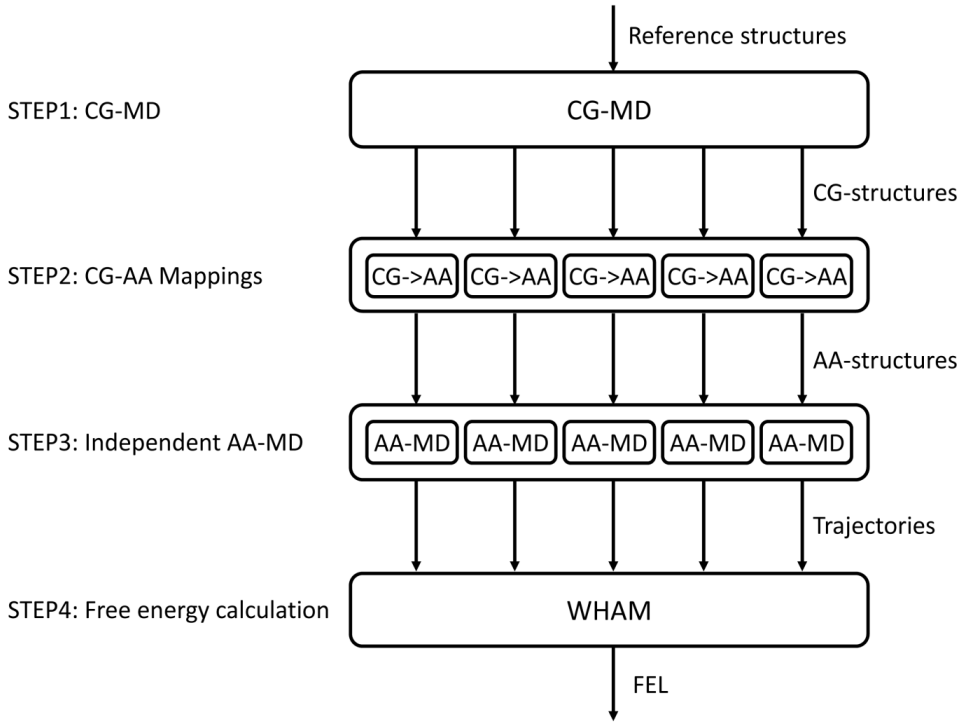
氏名　原 田　隆 平

Free energy landscape (FEL), is an important outcomes of molecular dynamics (MD) simulations, providing essential information for understanding the relations between conformational changes and functions. However, the FEL of biomolecular systems has multiple local energy minimum states, which makes it difficult to explore broad conformational space using conventional MD (CMD) simulations. Biomolecular systems cannot escape from one of multiple local energy minimum states separated by high energy barriers in a limited computational time. This is called multiple energy minima problem. Trapping into local energy minimum states typically leads to an inaccurate picture of the FEL. Hence, the sampling efficiency of the broader conformational space is essential for calculating the FEL.

In this thesis, I propose a novel multi-scale molecular dynamics simulation method for calculating free-energy landscapes of biomolecular systems, namely Multi-Scale Free Energy Landscape calculation method (MSFEL) to solve the multiple energy minima problem. This method is designed to simultaneously achieve high accuracy and sampling efficiency using the combination of coarse-grained (CG-) and all-atom (AA-) models. Firstly, the CG-model is used for calculating a rough energy landscape over broad conformational space. Subsequently, the AA-model is used to increase the sampling accuracy of local energy landscapes. This multi-scale approach attempts to achieve high sampling efficiency in the CG-model and high-resolution sampling in the AA-model. MSFEL has also an advantage of compatibility with massively parallel computing as time consuming AA-MD simulations are performed completely in parallel. Therefore, MSFEL will be one of efficient free energy calculation methods using huge computer resources in the future.

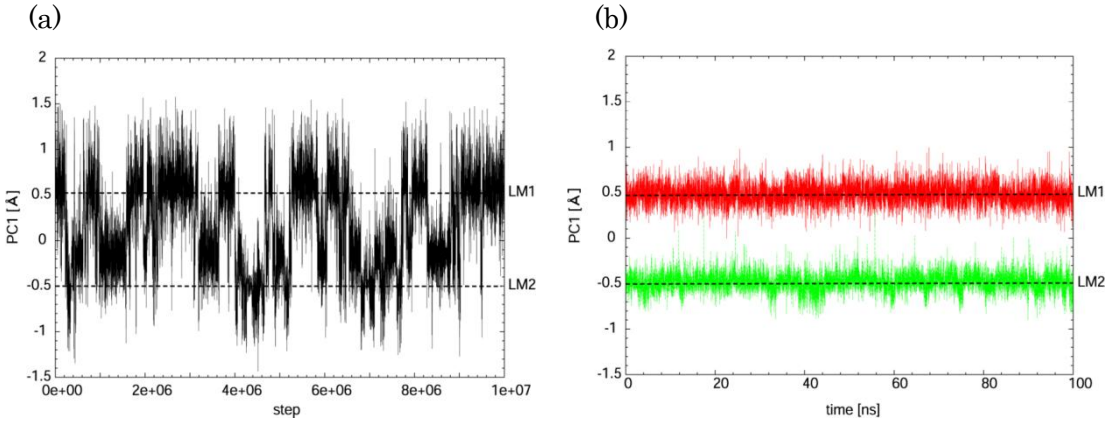
MSFEL consists of the following four steps. STEP1: Wide conformational space is sampled by the CG-MD simulation. STEP2: Multiple CG-structures are selected and AA-model structures are generated by using databases. STEP3: Non-interactive massively parallel AA-MD simulations are performed to investigate the local energy landscapes. STEP4: The multiple trajectories obtained from the independent AA-MD simulations are combined with the weighted histogram analysis method (WHAM) and the FEL is calculated. Figure 1 shows the flowchart of MSFEL.

Figure 1. The flowchart of MSFEL



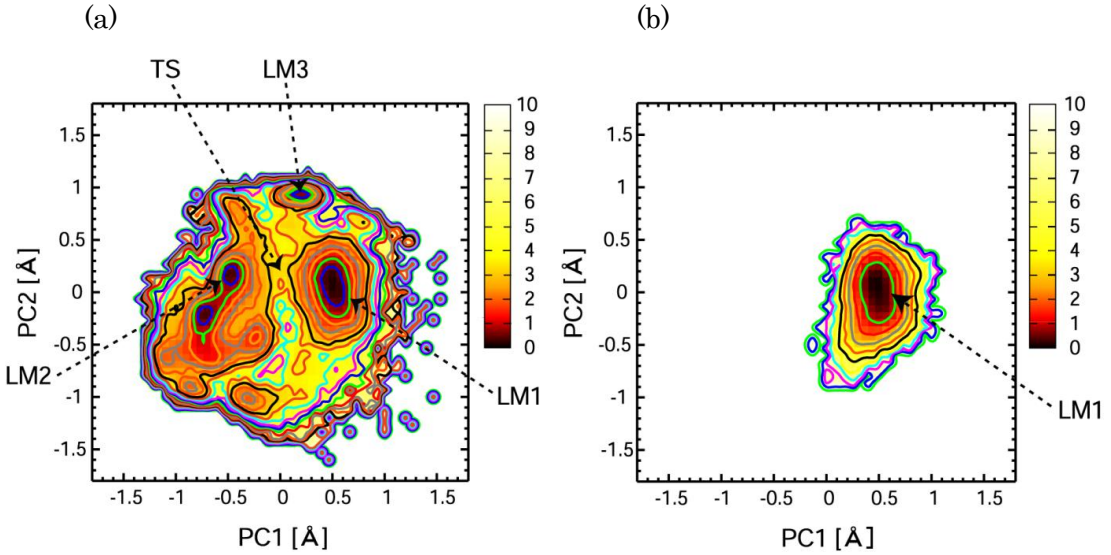
For a benchmark test, we firstly applied MSFEL to short peptide, Met-enkephalin (5-residue) in vacuum. Figure 2(a) shows the 100-ns trajectories of Met-enkephalin in vacuum simulated with all-atom CMD simulation started from two local energy minimum states, LM1 and LM2 projected on to the first principal component axis. As expected from the preceding works, Met-enkephalin is shown to be trapped in the initial states at least for 100-ns. Figure 2(b) shows the CG-MD trajectory projected on to the same first principal axis as in Figure 2(a), indicating frequent conformational transitions between LM1 and LM2 whereas conventional AA-MD failed to attain the transition within 100-ns. Thus more efficient conformational sampling is achieved in the CG-MD simulation.

Figure 2. The time series of Met-enkephalin dynamics in vacuum projected on to the first principal component axis



After CG-MD simulation, we modeled 100 AA-structures from the CG-structures and then performed 1-ns×100 independent AA-MD simulations in parallel. In the free energy calculation, the total probability distribution is calculated from 100 trajectories obtained in the independent AA-MD simulations with the WHAM. The FELs obtained from MSFEL and the CMD starting from LM1 projected onto the first and second principal axes are shown in Figure 3(a) and (b), respectively.

Figure 3. The FEL of Met-enkephalin in vacuum projected onto the first and the second principal axes.

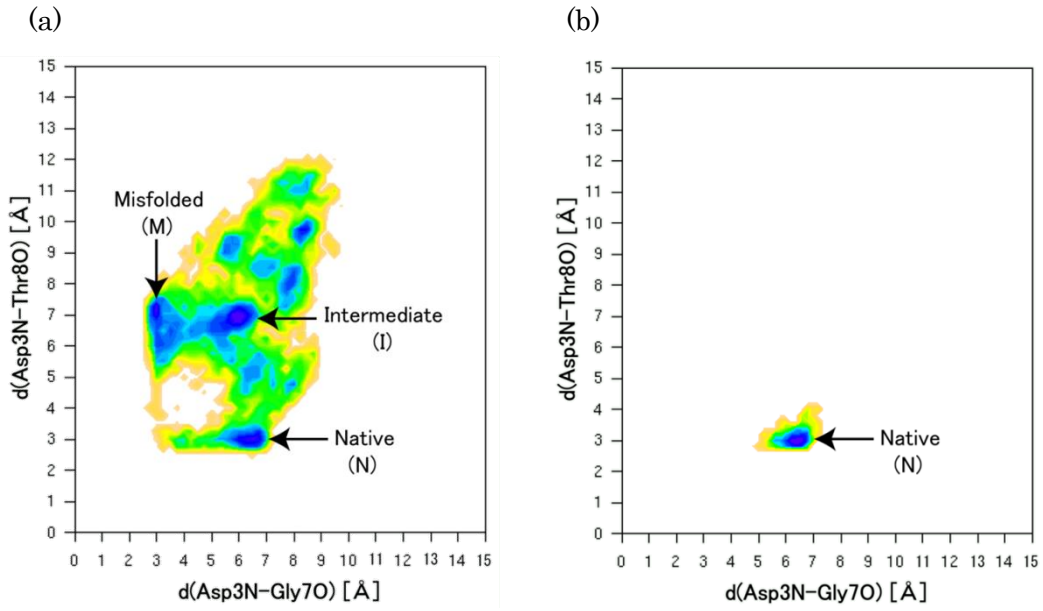


Due to trapping in LM1, the CMD failed to calculate accurate FEL. In contrast, MSFEL is successful in identifying LM1 and LM2 as well as the transition state TS between

LM1 and LM2. In addition to these known states, a novel local energy minimum state LM3 that has not been reported was found by MSFEL.

For other applications, we applied MSFEL to the following typical mini-proteins, tryptophan zipper (12-residue), chignolin (10-residue) and villin headpiece subdomain (35-residue) to calculate folding FELs. The results of each application also showed high computational efficiencies than the CMD simulations. Furthermore, MSFEL could also detect folding pathways and intermediate states during folding processes on each application. Figure4 (a) and (b) show one of the applications, the folding FEL of chignolin obtained from MSFEL and the CMD (100-ns) simulation, respectively. Compared to the CMD simulation, the misfolded and the intermediate conformations were also sampled by applying MSFEL.

Figure 4. The folding FEL of Chignolin projected onto the donor and the acceptor distance of main-chain atoms.



In conclusion, the benchmark test and the applications of MSFEL demonstrated that this method is an efficient and useful method in the free energy landscape calculations of biomolecules the FEL obtained from MSFEL is useful information for understanding the relevant biological functions.