Chapter 9

Computational Biophysics Research Team

9.1 Members

Yuji Sugita (Team Leader (Concurrent))*, **

Osamu Miyashita (Senior Research Scientist)

Jaewoon Jung (Research Scientist)

Chigusa Kobayashi (Research Scientist)

Yasuhiro Matsunaga (Research Scientist)

Hiromi Kano (Assistant)

Takaharu Mori (Research Scientist (Concurrent))**

Isseki Yu (Research Scientist (Concurrent))**

Raimondas Galvelis (Postdoctoral Researcher (Concurrent))**

Takao Yoda (Visiting Scientist) ***

Mitsunori Ikeguchi (Visiting Scientist)****

Naoyuki Miyashita (Visiting Scientist)

Michael Feig (Visiting Scientist)

* The main affiliation of these people is Laboratory for Biomolecular Function Simulation, Computational Biology Research Core, RIKEN Quantitative Biology Center.

** The main affiliation is RIKEN Theoretical Molecular Science Laboratory.

*** The main affiliation is Nagahama Bio Institute.

**** The main affiliation is Yokohama City University.

9.2 Research Activities

In this team, we have developed GENESIS(Generalized Ensemble Simulation System) for molecular dynamics simulations. The key features of GENESIS are that it is highly parallelized for K and other massively parallel supercomputers and that GENESIS contains a lot of enhanced conformational sampling methods and various molecular models for multi-scale and multi-resolution simulations. We have already open the code of GENESIS as free software under the license of GPLv2 and will update it every two year by adding new functions and optimizing the code into K or other computational platforms. These activities are necessary, in particular, for biological applications, since many interesting biological phenomena happen on the milliseconds or slower but current all-atom MD simulations cover only 1-10 microseconds on the general-purpose supercomputers or GPU clusters. We intend to spread GENESIS into academia as well as industries as a basic MD program that is useful for research and development.

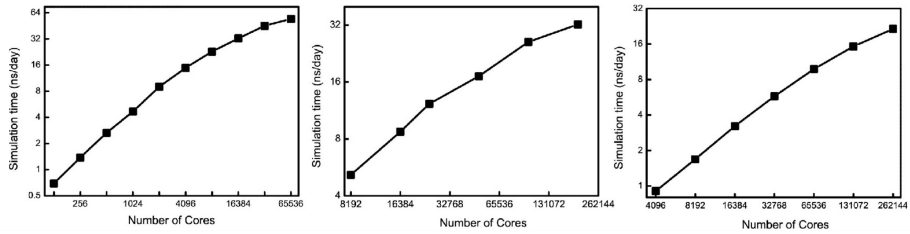

Figure 9.1: GENESIS performance of 1M (left), 8.5M (middle), and 28M (right) systems

9.3 Research Results and Achievements

9.3.1 Developement of GENESIS

We have already optimized GENESIS for large scale MD simulations on K computer. In the given fiscal year, we further optimized it by increasing parallel efficiency and enlarging the available number of processors. First, we make use of a multiple-program, multiple-data approach by separating computational resources responsible for real space and reciprocal space interactions. Second, we assign multiple time step integrator where time-consuming parts are skipped regularly based on the multiple-program and multiple-data approach. Our new implementation was tested on the K computer, and we could obtain very good performance results for big systems consisting of 1 million, 8.5 million, and 28 million atoms systems just increasing the parallel efficiencies. One MD cycle with the PME calculations for systems containing 1 million, 8.5 million, and 28 million atoms could be finished within 2.8 ms, 5.4 ms, and 8 ms (Figure 9.1).

9.3.2 Multi-resolution simulation methods for reactions couple with large conformational changes

Recently, experimental studies proposed that large conformational changes of proteins play important roles on biological functions. The conformational changes can originate as domain motions, where rigid structural units (domains) change their positions and/or orientations with respect to each other through flexible hinges or loops. It is difficult to investigate atomistic details of multi-domain proteins by experimental studies. In addition, it is still difficult to simulate using all-atom MD due to the slow time-scale. To overcome the difficulties, we have developed multi-resolution simulation method including the following three steps; 1. Analysis for "dynamic domains" and the magnitude of local domain motions in a protein through "Motion Tree", a tree diagram that describes conformational changes in a hierarchical manner from two structures. (Koike et al., *J. Mol. Biol.*, 2014) 2. Development of a structure-based coarse-grained (CG) model enables a stable and efficient MD simulation from the information of domain motion obtained by "Motion Tree" [5]. The CG model provides a stable trajectory that is comparable to experimental studies and long-time all-atom MD simulations. 3. Performing sampling simulations with the CG model and investigate conformational changes in response to reactions in biological systems. We examine how many CVs are required to capture the correct transition-state structure during the open-to-close motion of adenylate kinase using a coarse-grained model in the mean forces string method to search the minimum free-energy pathway [4].

9.3.3 Systematic evaluation of collective variable choice for describing conformational changes of a protein

Collective variables (CVs) are often used in molecular dynamics simulations based on enhanced sampling algorithms to investigate large conformational changes of a protein. The choice of CVs in these simulations is essential because it affects simulation results, and impacts on the free-energy profile, the minimum free-energy pathway (MFEP), and the transition-state structure. Here, we examine how many CVs are required to capture the correct transition-state structure during the open-to-close motion of Adenylate Kinase using a coarse-grained model in the mean forces string method to search the MFEP. Various numbers of large amplitude principal components (PCs) are tested as CVs in the simulations. The incorporation of local coordinates into CVs, which is possible

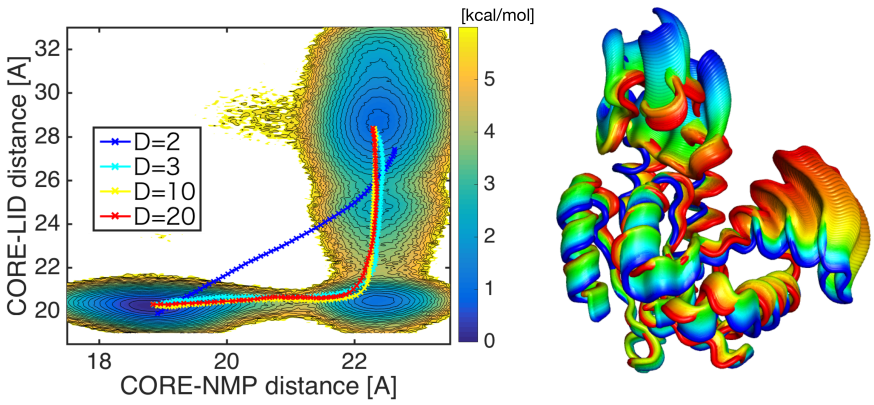

Figure 9.2: Free-energy landscape in the distances between domains' centers of mass. Lines indicate minimum free energy paths calculated in 2D (dark blue), 3D (light blue), 10D (yellow), and 20D (red) principal component spaces.

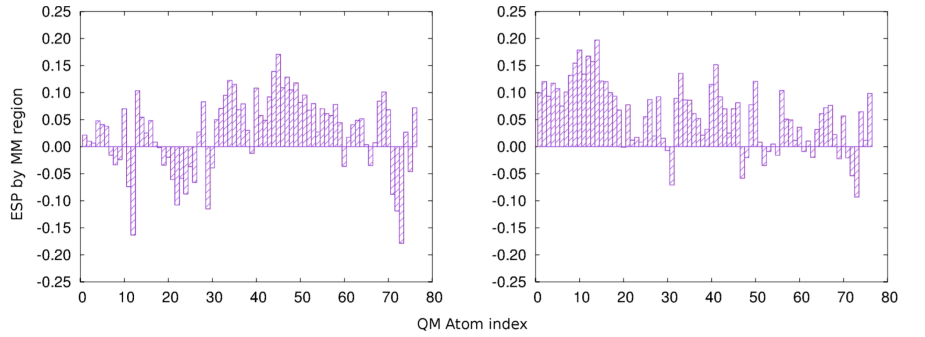

Figure 9.3: Electrostatic potential on QM atoms of the reactant state in the model crowding environment (left) and solution (right). QM atom index 43-55 corresponds to triphosphate moiety of GTP.

in higher dimensional CV spaces, is important for capturing a reliable MFEP. The Bayesian measure proposed by Best and Hummer is sensitive to the choice of CVs, showing sharp peaks when the transition-state structure is captured. We thus evaluate the required number of CVs needed in enhanced sampling simulations for describing protein conformational changes (Figure 9.2 [7]).

9.3.4 Molecular crowding effect on GTP hydrolysis reaction in Ras-GAP complex

Macromolecular crowding effects have essential role in biomolecular system. Such effects have been extensively investigated experimentally, and also in classical Molecular Dynamics (MD) calculations. However, in the quantum chemistry level, those effects are not investigated due to the computational costs and methodological difficulties. In this study, we studied the molecular crowding effect on the GTP hydrolysis reaction in Ras-GAP complex by QM/MM RWFE method, which can take crowding effects into account with a reasonable computational cost. We modeled a crowding environment by adding 7 BSAs to the system as a crowder, and refined the reactant and transition states of the hydrolysis reaction by QM/MM RWFE method, where MD calculations were performed by GENESIS at K-computer. The structural difference around GTP were not significant between solution and crowding environment. However, there was a large difference in the electrostatic potential (ESP) imposed by the surroundings as shown in Figure 9.3. This large ESP change suggests that there must be significant differences in the free energy barrier between crowding and solution environments.

9.4 Schedule and Future Plan

So far, GENESIS has been optimized mainly on K computer. In this year or later, we consider other platforms than K, such as intel CPU cluster, nvidia GPU processor, and post K. Since these CPU (or GPU) architectures are quite different with each other, a single MD kernel does not work well for all the different computational platforms. So, GENESIS will have multiple kernels that are optimized to one of the computational platforms. The disadvantage of this approach is that we have more effort on programming, reducing potential bugs for each kernel, and so on. It should be hard task for our team, but there is no other good ways to improve the performance of GENESIS in multiple platforms.

We would like to simulate more and more large biological systems for investigating slow biological dynamics. For this purpose, we need to develop multi-scale and multi-resolution programs that are scalable on K or post-K computers. Currently, GENESIS/SPDYN is useful for all-atom MD simulations on these supercomputers, but does not show good performance on CG-modeling and simulations of biological systems due to the small number of particles and load-balance problems. We need a new program that is suitable for such CG-modeling and simulations by introducing a different parallelization scheme. Such new program, which we call CGDYN, will be developed soon.

Another important aspect is the introduction of quantum effect to investigate the chemical reactions in enzymes. Bond-formation or breaking can not be simulated by using classical force fields, but should be investigated by using ab initio Quantum theory. Considering the large system size in biological systems, only possible approach is to use QM/MM hybrid calculations. We have a basic QM/MM code for computing potential energies of QM/MM systems and optimizing the systems based on the hybrid QM/MM potential energy functions. We plan to extend the calculations for larger periodic boundary systems and to allow the reaction calculations in proteins.

9.5 Publications

Journal Articles

[1] R. Galvelis and Y. Sugita. “Replica State Exchange Metadynamics for Improving the Convergence of Free Energy Estimates”. In: *J. Comp. Chem.* 36 (2015), pp. 1446–1455.

[2] J. Jung et al. “GENESIS: A hybrid-parallel and multi-scale molecular dynamics simulator with enhanced sampling algorithms for biomolecular and cellular simulations”. In: *WIREs Comput. Mol. Sci.* 5 (2015), pp. 310–323.

[3] J. Jung et al. “Parallel implementation of 3D FFT with volumetric decomposition schemes for efficient molecular dynamics simulations”. In: *Comput. Phys. Commun.* 200 (2016), pp. 57–65.

[4] C. Kobayashi et al. “Domain Motion Enhanced (DoME) Model for Efficient Conformational Sampling of Multidomain Proteins”. In: *J. Phys. Chem. B* 119 (2015), pp. 14584–14593.

[5] C. Kobayashi et al. “Hierarchical domain-motion analysis of conformational changes in sarcoplasmic reticulum Ca2+-ATPase”. In: *Proteins* 83 (2015), pp. 746–756.

[6] Y. Matsunaga, A. Kidera, and Y. Sugita. “Sequential data assimilation for single-molecule FRET photon-counting data”. In: *J. Chem. Phys.* 142 (2015), p. 214115.

[7] Y. Matsunaga et al. “Dimensionality of Collective Variables for Describing Conformational Changes of a Multi-Domain Protein”. In: *J. Phys. Chem. Lett.* 7 (2016), pp. 1446–1451.

Invited Talks

[8] J. Jung. *Development of GENESIS for Large scale MD simulations.* The 4th JLESC workshop. 2015.

[9] J. Jung. *Development of GENESIS for Multiscale Molecular Dynamics simulations.* 1st Multiscale Computational Biology Workshop. 2015.

[10] J. Jung. *Parallelization of Molecular Dynamics.* The 13th CMSI Lecture. 2015.

[11] J. Jung and Y. Sugita. *Multiple time step integrator in molecular dynamics (MD) program GENESIS.* Kyoto Workshop. 2015.

[12] C. Kobayashi. *Introduction to GENESIS*. Chem-Bio Informatics Society Annual Meeting. 2015.

[13] C. Kobayashi et al. *Development of multi-resolution simulation methods for reactions with large conformational changes in biological system*. The 53th Annual Meeting of Biophysical Society of Japan. 2015.

[14] C. Kobayashi et al. *Development of multi-resolution simulation methods for reactions with large conformational changes in biological system*. The 41th Workshop of Japan Bio-energy. 2015.

[15] Y. Matsunaga. *Drug extrusion mechanism of multidrug exporter AcrB studied by the string method*. The 53th Annual Meeting of Biophysical Society of Japan. 2015.

[16] Y. Sugita. *Development and parallelization of MD software GENESIS*. Lecture at Hong Kong University of Science and Technology. 2015.

[17] Y. Sugita. *Free-energy calculations of conformational changes in membrane proteins by multi-resolution methods*. The 25th Hot Spring Harbor International Symposium. 2015.

[18] Y. Sugita. *Free-energy calculations using a new MD simulator*. Snowmass Free-Energy Conference. 2015.

[19] Y. Sugita. *Large-scale all-atom MD simulations of biomolecules under cellular environments*. Seminar at Hong Kong University of Science and Technology. 2015.

[20] Y. Sugita. *Molecular Simulations of Membrane Transporting Proteins*. RIKEN Symposium on Metal in Biology. 2015.

[21] Y. Sugita. *Parallel MD simulations for enhanced conformational sampling of biological systems*. Mini symposium on Rare Events in Complex Physical Systems at the 8th International Congress on Industrial and Applied Mathematics, Beijing. 2015.

[22] Y. Sugita. *Recent development of replica-exchange molecular dynamics simulation methods*. The 251st ACS national meeting and exposition on From Dynamics to Function and Back Again: Adventures in Simulating Biomolecules. 2015.

[23] I. Yu et al. *Dynamics, Stability, and Interactions of Proteins and Metabolites in Bacterial Cytoplasm: All-atom Molecular Dynamics Study*. The 53th Annual Meeting of Biophysical Society of Japan. 2015.

Posters and Presentations

[24] J. Jung and Y. Sugita. *Parallel implementation of molecular dynamics by combining multiple program/multiple data with multiple time step integrator*. The 6th AICS International Symposium. 2016.

[25] J. Jung et al. *Development of multiple timestep integrator in isothermal and isobaric conditions for efficient MD simulations of biological systems*. The 53th Annual Meeting of Biophysical Society of Japan. 2015.

[26] J. Jung et al. *Efficient strong scale parallelization of molecular dynamics on hybrid CPU/GPUs for large scale simulations*. The 29th Annual Meeting of Molecular Simulation Society of Japan. 2015.

[27] M. Kamiya et al. *Molecular crowding effect on GTP hydrolysis reaction in Ras-GAP complex*. SCLS Symposium 2015. 2015.

[28] M. Kamiya et al. *Molecular crowding effect on GTP hydrolysis reaction in Ras-GAP complex*. The 29th Annual Meeting of Molecular Simulation Society of Japan. 2015.

[29] M. Kamiya et al. *Molecular crowding effect on GTP hydrolysis reaction in Ras-GAP complex*. The 6th AICS International Symposium. 2016.

[30] C. Kobayashi et al. *Development of multi-resolution simulation methods or model of reaction with large conformational changes in biological system*. The 6th AICS International Symposium. 2016.

[31] Y. Matsunaga and Y. Sugita. *Sequential data assimilation for single-molecule FRET photon-counting data*. The 6th AICS International Symposium. 2016.

Patents and Deliverables

[32] *GENESIS http://www.riken.jp/TMS2012/cbp/en/research/software/genesis/index.html.*