Usage Report for Fiscal Year 2021

Project Title:

Development of multi-scale simulation methods for large biomolecular systems

Name:

○ Cheng Tan (1), Jaewoon Jung (1,2), Chigusa Kobayashi (1) and Diego Ugarte La Torre (1)

Laboratory at RIKEN:

(1) RIKEN Center for Computational Science, Computational Biophysics Research Team

(2) RIKEN Cluster for Pioneering Research, Theoretical Molecular Science Laboratory

1. Background and purpose of the project, relationship of the project with other projects

Molecular dynamics (MD) simulation is a tool to investigate biological processes with carefully designed computational models and algorithms. In this research project, we developed MD methods in the GENESIS software for simulating large-scale biomolecular systems at the molecular level. We have implemented new models at different resolutions and have optimized the code to gain high performance on various platforms. With our implementation, we have applied MD simulations to study the liquid-liquid phase separation of proteins, transcription factor-DNA interactions, and the calcium ion pump. In this fiscal year, we have released GENESIS 1.7.0/1, containing some of the developed functions and models as free software. In addition to the source code, test sets and tutorials have been prepared. We used Hokusai to test and check the new versions of GENESIS. Specifically, we have conducted four studies as listed below:

1) Development of Hydrogen mass repartitioning (HMR) for a large time step MD (by Jung)
2) Implicit solvent coarse-grained lipid model development in GENESIS for the study of large lipidic systems (by Diego)
3) Residue level Coarse-grained (CG) model development in GENESIS and applications to the phase behavior of protein and nucleic acids (by Tan)
4) Multiple-site model of divalent ions (calcium ion) and applications to calcium ion in solution and calcium ion pump (by Kobayashi)
5) Preparation for Release of new versions (1.7.0/1.7.1) of GENESIS (by Kobayashi, Jung, Tan)

2. Specific usage status of the system and calculation method

1) Hydrogen mass repartitioning (HMR) is a scheme to increase the mass of hydrogen to reduce high-frequency motion while the mass of a heavy atom bonded to a hydrogen atom is reduced to conserve the total mass. We improved this HMR scheme by increasing the stability from a new mass scaling factor and by combining accurate instantaneous temperature and pressure evaluations. Various tests and benchmark performance were done on the Hokusai supercomputer.
2) We have implemented the coarse-grained iSoLF model for POPC and DPPC lipids in GENESIS. Furthermore, we have extended it to DOPC and POPE lipids and fine-tuned the parameters to improve the agreement with various calculated physical properties of

biological membranes. Finally, we tested the model by sampling different bilayer conformations.

3) We have implemented the residue-level CG models for protein (AICG2+), DNA (3SPN.2C), RNA (structure-based), protein-DNA intermolecular interactions (PWMcos and hydrogen-bond) in GENESIS 1.7.1. We also introduced a framework for the intrinsically disordered protein (IDP) and tested several different parameter sets. We then utilized our CG development in GENESIS to study the liquid-liquid phase separation of IDPs and the regulation by a class of heat-resistant obscure proteins (HERO proteins).

4) We have introduced a multiple-site model for divalent ion (calcium ion) (Zhang et al. Nat. Comm. 11, 922 (2020)) into the recent GENESIS developing version. We also apply the models to calcium ion in solution and calcium ion pump.

3. Result

1) To show the improvement of stability, we tested a pyrrole ring structure in AMPA receptors using conventional and our suggested HMR schemes with 5 fs time step. Compared to the original scheme, our new HMR provides much more stable trajectories. Our scheme was tested on various biological systems, including soluble/membrane proteins and lipid bilayers. From the tests, we confirmed that the new scheme gives consistent results, even extending the time step of slow-motion force up to 7 fs. This increases the overall performance more than twice.

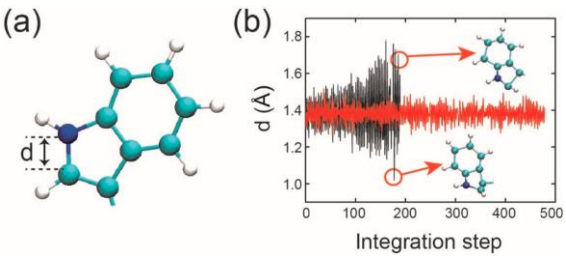

Figure 1. Bond distance in the pyrrole ring (shown in d) during integration step (black line: conventional scheme, red line: our new scheme).

2) (a). We have added electrostatic interactions to the iSoLF model for lipids, enabling the representation of charged lipids. (b). We have extended the iSoLF model to DOPC and POPE lipids. (c). We have tested the parameters by simulating small unilamellar vesicles and multicomponent membranes.

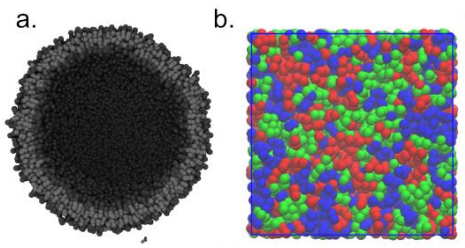

Figure 2. (a) Half-view of a small unilamellar vesicle made of POPC lipids. (b) Top-view of the phase separation in a multicomponent lipid bilayer. POPC, DPPC, and DOPC are represented with the colors red, blue, and green, respectively

3) We have implemented the necessary CG models for the study of liquid-liquid phase separation and released the code in the latest version of GENESIS 1.7.1. We also prepared detailed tutorials and documentation.

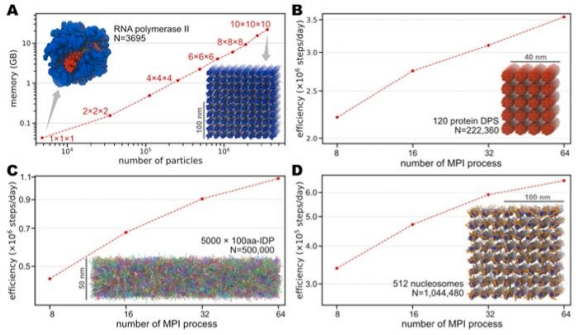

Figure 3. Benchmark of GENESIS CG

simulations on Hokusai. (A) Memory benchmark of the RNA Pol II system. (B-D) CPU benchmark on DPS protein, IDP, and nucleosome systems, respectively. OpenMP thread number was set to 5, while changing MPI process numbers.

We have applied our implementation to study the phase behavior of LLPS-prone IDPs and the Hero proteins, which can regulate the condensation of the LLPS-IDPs. We found that generally speaking, the highly charged IDPs, such as the Hero proteins, can decrease the phase separation probability of other IDPs by introducing electrostatic repulsion interactions between pre-mature condensates and decreasing the contact probability inside the condensates.

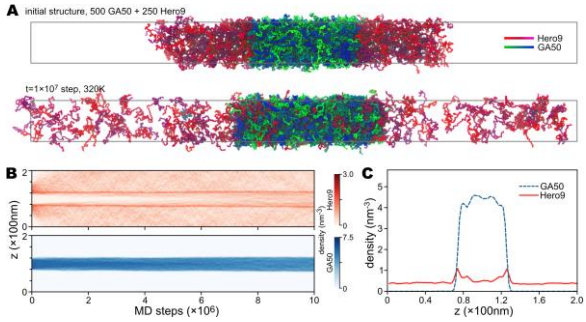

Figure 4. Anti-aggregation effect of highly charged IDPs on LLPS-prone IDPs.

4) The multiple-site model of calcium ion has been introduced into the latest developing version of the code. We have tested the amount of coordination to water molecules and the average bond distance.

4. Conclusion

We have utilized Hokusai in developing multi-scale MD simulation methods in the MD software GENESIS. On the atomic resolution, we have implemented the HMR scheme for the long-time-step MD simulation. We also introduced multi-site models for the divalent cations and applied the models in the study of calcium pumps. On the coarse-grained resolution, we have used Hokusai for testing the implementation of the iSoLF lipid model in GENESIS and improving the existing parameter set. Besides, by developing the residue-level CG models for protein and nucleic acids, we have simulated the phase behaviors of biomolecules. Most of the above developments have been packaged and released in GENESIS 1.7.0/1. We believe our developments and application studies will contribute both simulation tools and scientific insights to the computational biophysics society.

5. Schedule and prospect for the future

In the future, we will further optimize our CG and all-atom implementation to gain better performance. We also have planned to apply our simulations to study more biophysical processes, such as phase-separation, lipid membrane, and cation channels.

6. If no job was executed, specify the reason.

Usage Report for Fiscal Year 2021

Fiscal Year 2021 List of Publications Resulting from the Use of the supercomputer

[Paper accepted by a journal]

1. Chigusa Kobyashi, Yasuhiro Matsunaga, Jaewoon Jung, Yuji Sugita, "Structural and energetic analysis of metastable intermediate states in the E1P–E2P transition of Ca^{2+}-ATPase.", *Proc. Natl. Acad. Sci. USA* 118 e2105507118 (2021)
2. Jaewoon Jung, Kento Kasahara, Chigusa Kobayashi, Hiraku Oshima, Takaharu Mori and Yuji Sugita, "Optimized hydrogen mass repartitioning scheme combined with accurate temperature/pressure evaluations for thermodynamic and kinetic properties of biological systems", *J. Chem. Theory Comput.* 17, 5312-5321 (2021)

[Oral presentation]

1) Chigusa Kobayashi, Yasuhiro Matsunaga, Jaewoon Jung, Yuji Sugita, "Analysis of reaction pathway in E1/E2 transition of SR-Ca2+-ATPase." *The 59th Annual Meeting of the Biophysical Society of Japan*, *Online*, 2021/11/25
2) Chigusa Kobayashi, Yasuhiro Matsunaga, Jaewoon Jung, Yuji Sugita, "分子動力学法シミュレーションによる Ca2+-ATPase の E1P-E2P 転移の中間状態の解析", *The 47th meeting of Japan Bioenergetics Group*（日本生体エネルギー研究会第 47 回討論会）, *Online*, 2021/12/16
3) Jaewoon Jung, 笠原健人, 小林千草, 尾嶋拓, 森貴治, 杉田有治, "Optimized hydrogen mass repartitioning scheme combined with accurate temperature/pressure evaluations for thermodynamic and kinetic properties of biological systems", *第 35 回分子シミュレーション討論会*, *Online*, Nov. 29-Dec. 1, 2021
4) Jaewoon Jung, Chigusa Kobayashi, and Yuji Sugita, "Development of GENESIS MD software on Fugaku supercomputer for understanding large scale biomolecular phenomena in cellular environments", **Pachifichem symposium**, *Online*, Dec.16-Dec. 21, 2021
5) Cheng Tan, Jaewoon Jung, Chigusa Kobayashi, Yuji Sugita. "Implementation of residue level coarse-grained models for biomolecules in GENESIS." *The 13th JLESC Workshop*, *Online*, Thr, Dec 16, 2021

[Poster presentation]

1) Chigusa Kobayashi, Yasuhiro Matsunaga, Jaewoon Jung, Yuji Sugita, "カルシウムイオンポンプの E1/E2 転移における反応経路解析" *The 35th symposium of the molecular simulation society of Japan, Online,* 2021/11/29
2) Cheng Tan, Jaewoon Jung, Chigusa Kobayashi, Yuji Sugita. "Phase behavior of transcription factors studied with molecular dynamics simulations." *The 8th HPCI System Research Project Report Meeting*, *Online*, Fri, Oct 29, 2021
3) Cheng Tan, Jaewoon Jung, Chigusa Kobayashi, Yuji Sugita. "Implementation of residue level coarse-grained models in GENESIS." *The 59th Annual Meeting of the Biophysical Society of Japan*, *Sendai, Japan*, Thr, Nov 25, 2021