Vol. 13, pp. 209–216 (2016)
doi: 10.2142/biophysico.13.0_209

Biophysics and Physicobiology

https://www.jstage.jst.go.jp/browse/biophysico/

Database and Computer Program

myPresto/omegagene: a GPU-accelerated molecular dynamics simulator tailored for enhanced conformational sampling methods with a non-Ewald electrostatic scheme

Kota Kasahara[1], Benson Ma[2], Kota Goto[3], Bhaskar Dasgupta[4,5], Junichi Higo[4], Ikuo Fukuda[4], Tadaaki Mashimo[5], Yutaka Akiyama[3,6] and Haruki Nakamura[4]

[1]*College of Life Sciences, Ritsumeikan University, Kusatsu, Shiga 525-8577, Japan*
[2]*College of Engineering, University of Illinois, Urbana-Champaign, United States*
[3]*School of Computing, Tokyo Institute of Technology, Tokyo 152-8550, Japan*
[4]*Institute for Protein Research, Osaka University, Suita, Osaka 565-0871, Japan*
[5]*Technology Research Association for Next Generation Natural Products Chemistry, Tokyo 135-0064, Japan*
[6]*Molecular Profiling Research Center for Drug Discovery, National Institute of Advanced Industrial Science and Technology, Tokyo 135-0064, Japan*

Received June 28, 2016; accepted August 8, 2016

Molecular dynamics (MD) is a promising computational approach to investigate dynamical behavior of molecular systems at the atomic level. Here, we present a new MD simulation engine named "myPresto/omegagene" that is tailored for enhanced conformational sampling methods with a non-Ewald electrostatic potential scheme. Our enhanced conformational sampling methods, *e.g.*, the virtual-system-coupled multi-canonical MD (V-McMD) method, replace a multi-process parallelized run with multiple independent runs to avoid inter-node communication overhead. In addition, adopting the non-Ewald-based zero-multipole summation method (ZMM) makes it possible to eliminate the Fourier space calculations altogether. The combination of these state-of-the-art techniques realizes efficient and accurate calculations of the conformational ensemble at an equilibrium state. By taking these advantages, myPresto/omegagene is specialized for the single process execution with Graphics Processing Unit (GPU). We performed benchmark simulations for the 20-mer peptide, Trp-cage, with explicit solvent. One of the most thermodynamically stable conformations generated by the V-McMD simulation is very similar to an experimentally solved native conformation. Furthermore, the computation speed is four-times faster than that of our previous simulation engine, myPresto/psygene-G. The new simulator, myPresto/omegagene, is freely available at the following URLs: http://www.protein.osaka-u.ac.jp/rcsfp/pi/omegagene/ and http://presto.protein.osaka-u.ac.jp/myPresto4/.

Corresponding author: Kota Kasahara, College of Life Sciences, Ritsumeikan University, 1-1-1 Noji-higashi, Kusatsu, Shiga 525-8577, Japan.
e-mail: ktkshr@fc.ritsumei.ac.jp

Key words: molecular simulation, software, generalized ensemble, high performance computing, GPGPU

◀ Significance ▶

The molecular dynamics (MD) method is a promising approach to investigate a variety of biophysical phenomena at the atomic level. In particular, elucidation of the conformational ensemble of bio-molecules is of paramount importance. We present here a new MD simulation program, "myPresto/omegagene" which is tailored for efficient computation of enhanced conformational sampling powered by GPU acceleration. "myPresto/omegagene" is unique in that it adopts our original multi-canonical ensemble approach and our non-Ewald electrostatic potential scheme named "zero-multipole summation method" to effectively enhance computations. In addition, "myPresto/omegagene" builds upon a wealth of tools and resources provided by the myPresto suite to enable good user experience for myPresto users.

The molecular dynamics (MD) method is a key technology for dissecting dynamical properties of molecular systems at the atomic level. Along with the rapid growing of the high performance computing technologies, the field of MD simulations has been growing and extensively applied for highly complex and large-scale phenomena. A recent milestone in this field is the rise of the special-purpose hardware, named "ANTON" [1,2]. This hardware achieved millisecond-timescale simulations for a variety of molecular systems [3,4]. Yet another MD-specialized hardware, "MD-GRAPE", was recently developed by Taiji *et al.* in RIKEN [5]. While specialized hardware pushes the limits of the MD method, it is not widely used due to its very high costs and difficulty in modifying the algorithms set in the hardware.

Consequently, developments in the field of MD simulations have mainly targeted general-purpose cluster machines and a variety of MD software have been developed. At the moment, one of the most efficient implementations of the MD is "GROMACS", which comes with implementations of low level single-instruction-multiple-data (SIMD) instructions [6,7]. The software "AMBER" also accomplishes a good performance, especially when powered by NVIDIA GPU-based accelerator [8,9]. In parallel, several MD programs based on unique state-of-the-art physics theories have also been developed. The MD program "GENESIS" takes advantage of a variety of replica exchange MD (REMD) methods, *e.g.*, the temperature REMD and the surface tension REMD, and several original algorithms [10–12]. "MODYLAS" applies a non-Ewald electrostatic potential calculation scheme, the fast multipole method [13]. Another MD program "MARBLE" implements a partial rigid-body method under membrane-specific ensembles [14]. "SCUBA" has a unique algorithm for generating the structural ensemble fitted to low resolution electron microscopy data [15]. These programs are extensively tuned for supercomputer systems, such as the "K computer".

During the past several decades, we developed and maintained "myPresto" package, which consists of a variety of tools for in-silico drug development, including a docking program between a receptor protein and drug candidates ("sievegene") [16–18], a ligand binding pocket finder ("molsite") [19], a small-compound structure library ("ligandbox") [20]. More recently added to myPresto was the MD simulation engine "psygene-G", which implemented two major original features: a multi-GPU accelerated parallelization scheme [21], and the zero-multipole summation method (ZMM) [22–26]. By utilizing many GPUs to parallelize the space-decomposition routine, myPresto/psygene-G accomplishes a scalable computation for a wide range of system sizes and have been applied for a variety of biomolecular systems [27–29]. As the ZMM estimates the electrostatic potential with a cutoff pair potential, the computational cost is drastically reduced compared with Ewald-based methods, and the scalability of the multi-node parallel computation is improved.

In addition, we developed a series of theories for the non-canonical MD simulations, namely enhanced conformational sampling methods. We have developed the multi-canonical ensemble MD simulation method [30], and hereinafter we refer this method as multi-canonical MD (McMD). On the basis of the McMD method, several enhanced conformational sampling methods have been developed and applied to analyze a variety of bio-molecular systems [31–34]. These methods are powerful tools to elucidate physicochemical properties of molecular systems at an equilibrium state. We have previously demonstrated that trajectories generated from many independent McMD runs with different initial conditions can be combined for statistical analysis [35]. This theory allows us to execute many relatively short MD runs independently in parallel instead of a single long simulation.

Our enhanced conformational sampling methods and the ZMM have been implemented in myPresto/psygene-G. However, myPresto/psygene-G was tailored for long-term simulations of large molecular systems using multi-GPU parallelization with the space-decomposition algorithm, and thus it is not suitable for running multiple independent simulations of small systems, which are what enhanced conformational sampling methods target.

Here, we developed a new MD simulation program from scratch, named "myPresto/omegagene" that is tailored for our original McMD methods with the ZMM. myPresto/omegagene is distributed under open-source license from http://www.protein.osaka-u.ac.jp/rcsfp/pi/omegagene/, and it is also available from http://presto.protein.osaka-u.ac.jp/myPresto4/. In this report, Section 1 "Software Details" presents requirements and detailed information of the implementation. Section 2 "Simulations" presents results of two types of test simulations, with a 20-mer peptide, Trp-cage. The first simulation with the micro-canonical ensemble demonstrates a base capability: four-fold acceleration compared with myPresto/psygene-G with an energy conservation property. The second simulation with the virtual-system-coupled McMD (V-McMD) method, which is a variant of the McMD method, shows successful application results: an energy landscape of the Trp-cage at 300 K along with near-native conformations sampled as the most stable cluster.

1. Software Details

1.1. Overview

myPresto/omegagene consists of the two parts: i) the core MD engine written in C++ and CUDA languages, and ii) the toolkit for pre- and post-simulation processing written in Python language (Fig. 1). The installation process is semi-automated, by taking advantage of a standard cross-platform building tool "cmake" (version 3.2 or later). myPresto/omegagene is easily built on a variety of environment, *e.g.*, Linux, Windows, and OS X (also known as MacOS). For

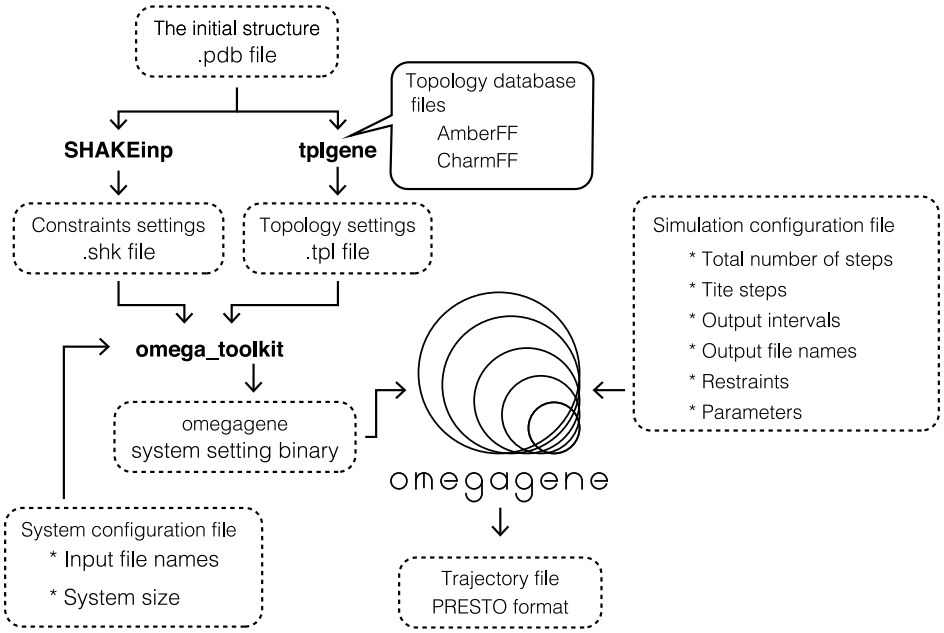

Figure 1 Overview of the myPresto/omegagene ecosystem.

compiling the programs, C++ compilers compatible to C++11 standard is required. The GPU acceleration requires a NVIDIA GPU board with compute capability ≥3.5.

The input and output files formats of myPresto/omegagene are compatible with those of existing tools in the myPresto package. The input molecular topologies of proteins and nucleic acids are generated by a "tplgene" program, a myPresto tool. A list of atom groups with non-holonomic constraints is prepared, by using the "SHAKEinp" program. In addition, some other tools for pre- and post-simulation analyses and data handling are attached as omega_toolkit. For example, the simulation trajectory written in PRESTO format can be converted into other standard binary formats, *e.g.*, .trr format (GROMACS). The initial velocities of atoms in simulation systems are generated by using a Python script included in omega_toolkit. These input files are then integrated into a single binary file, by using another script. Separation of such pre- and post-simulation lightweight tasks from the core engine makes it easier to develop and maintain the tools.

1.2. Implementation

The MD engine of myPresto/omegagene is implemented in C++11 in an object-oriented manner. SIMD parallelization, powered by NVIDIA GPU with CUDA was applied for the cutoff pair-wise potential calculations, which is the bottle-neck of MD computations. As this program utilizes the ZMM for the electrostatic potential calculation, Lennard-Jones and electrostatic potentials are computed in the same kernel function on GPU. A main difference from myPresto/psygene-G is the routine for pairwise potential calculations on GPU. myPresto/omegagene applies an algorithm based on the method proposed by Páll and Pronk [6] which is utilized in GROMACS. To find atom pairs within the cutoff radius (the neighbor search task), myPresto/omegagene groups the atoms in the system into "sub-cells" with the same number of atoms, in contrast to myPresto/psygene-G, which divides the system into sub-cells with a uniform volume [21].

As the SIMD width of NVIDIA GPU is 32 threads (a bunch of 32 threads is defined as a "warp"), the atoms in the system are grouped into 8-atom sub-cells (Fig. 2A). Accordingly, there are 64 atom-atom interactions between atoms from each sub-cell in a pair, and these pair potentials are calculated in bulk by one warp (32 threads) using two loop iterations. Before the pair potential calculations, a search for all pairs of sub-cells whose inter-cell distance are within a distance threshold (the summation over the cutoff radius and an offset value) is performed at every predefined interval number of steps (test calculations for this interval are described in the next section). The neighbor search process enumerates all sub-cell pairs within this cutoff. For example, in Figure 2A, the neighbor sub-cells of the 1st sub-cell are 0th, 2nd, 3rd, and 4th sub-cells. In one iteration of the pair-wise potential calculation algorithm, the m-th warp calculates pairs of the m-th sub-cell and its neighbor sub-cells. In order to balance the number of sub-cell pairs owned by each warp, we developed our original scheduling scheme, in which a sub-cell pair m–n (m and n refer to sub-cell indices) is considered for calculation if either one of following conditions hold: 1) when m is an even number, n is an odd number for $m<n$ and n is an even number for $n>m$; and 2)

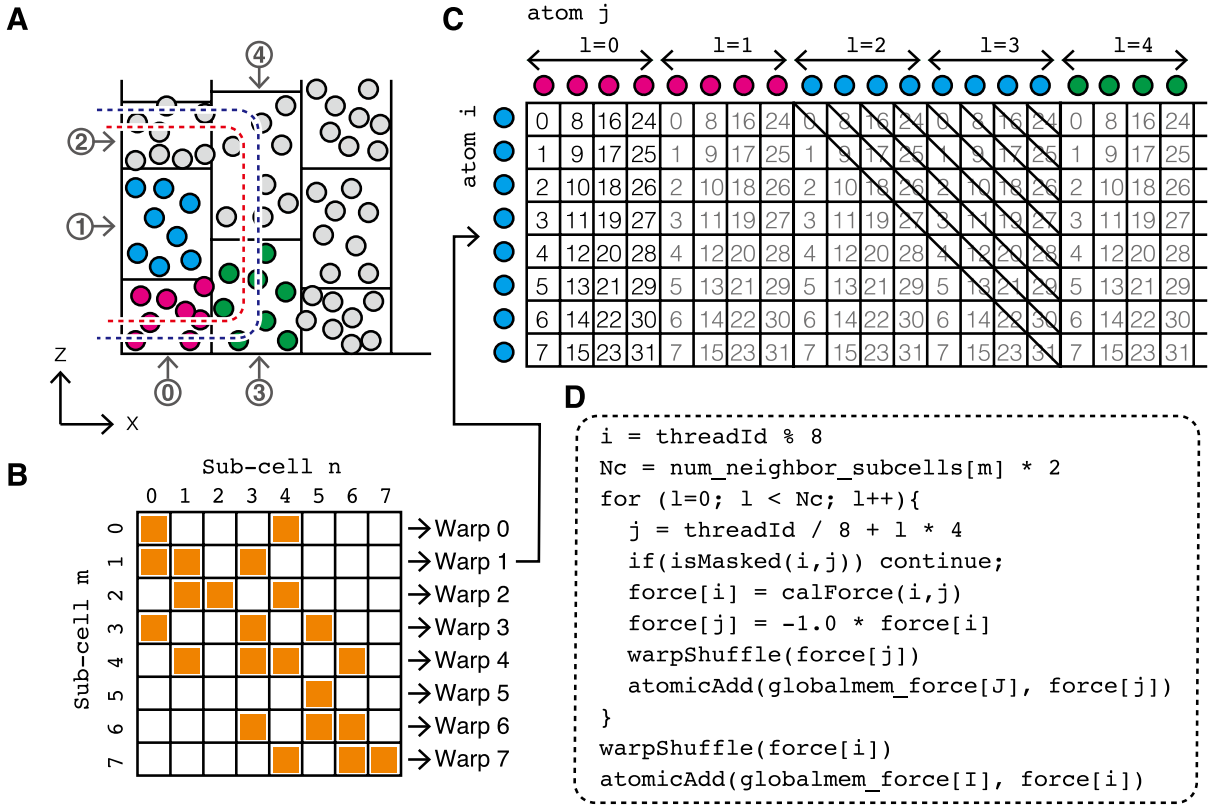

Figure 2 Schematic diagram of the algorithm for the pairwise potential calculations. (A) The system consisting of sub-cells, each of which consists of eight atoms (filled circles). The number in the circle indicates the ID of each sub-cell. The red dashed line indicates the cutoff length from the border of the 1st sub-cell. The blue dashed line represents the cutoff added by the offset value. (B) The matrix describing assignments of each sub-cell pairs to each warp. The pairs marked as the orange squares in each row are processed by a warp. Note that the periodic boundary condition is ignored for simplicity in this figure. (C) Pairs of atoms calculated in each thread in warp 1. The numbers in the matrix denote the thread IDs in the warp, processing the corresponding pair of atoms. Each row indicates one of eight atoms in 1st sub-cell. Each column indicates an atom nearby the 1st sub-cell. The filled circles correspond to the atoms in Figure 2A. (D) A pseudo code of the kernel function. *num_neighbor_cubcells[m]* is the number of neighboring sub-cells of *m-th* sub-cell. *isMasked(i, j)* excludes the *i–j* pairs within four successive covalent bonds and the pairs $i \leq j$ in the same sub-cell. *force[i]* is the 3D array keeping the force of the *i*-th atom along x, y, and z axes. The function *calForce(i, j)* calculates the force of the *i*-th atom acted by the *j*-th atom. *warpShuffle(force[i])* sums up the force values among threads with the same *i* value in the same warp. *atomicAdd* adds the *force[i]* value to the variable *globalmem_force[I]* in the global memory.

when m is an odd number, n is an even number for $n<m$ and n is an odd number for $n>m$; or $m=n$. Using Figure 2B as an example, the 1st warp in the GPU grid calculates pairwise potentials in the sub-cell pairs 1–0, 1–1, and 1–3, while the 2nd and 4th warps calculate pairwise potentials in sub-cell pairs 2–1 and 4–1, respectively. Note that this algorithm does not guarantee a good load balance among warps, because not all sub-cells have the same number of neighbors. While Páll and Pronk's method limits the number of sub-cell pairs calculated in each warp to eight in order to balance the workload of each warp, we did not apply this scheme here, since dividing the loops requires frequent communications between the registers and the global memory with atomic operations. In addition, as it is expected that the distribution of the particle density is not largely biased in usual systems, we assume the impact of this load imbalance is limited.

Then, the pairwise potentials are calculated in each warp (Fig. 2C and D). As each sub-cell has eight atoms, there are 64 atom pairs in each pair of sub-cells, and these pair potentials are calculated in two cycles of the loop, by using 32 threads (one warp). The variable l means the iterator variable of the loop. The first cycle (l=0) of the 1st warp, pairs of eight atoms in the 1st sub-cell and four atoms in the 0th sub-cell are computed. In the next cycle (l=1), pairs of the eight atoms in the 1st sub-cell and the remaining four atoms in the sub-cell 0 are computed. The calculated forces for each atom are gathered with a warp shuffle instruction and are written back to the global memory with an atomic operation. While the calculations of force and energy in each term are executed in the single precision, summation of these terms (*warpShuffle* and *atomicAdd* in Fig. 2D) is done in the double precision. To exclude the *i–j* atom pairs connected within four successive covalent bonds and the atom pairs $i \leq j$ in a sub-cell ($m=n$; shaded pairs in Fig. 2C), a 64-bit bitmask is prepared for each sub-cell pair to switch the pair potential of each atom pair [36].

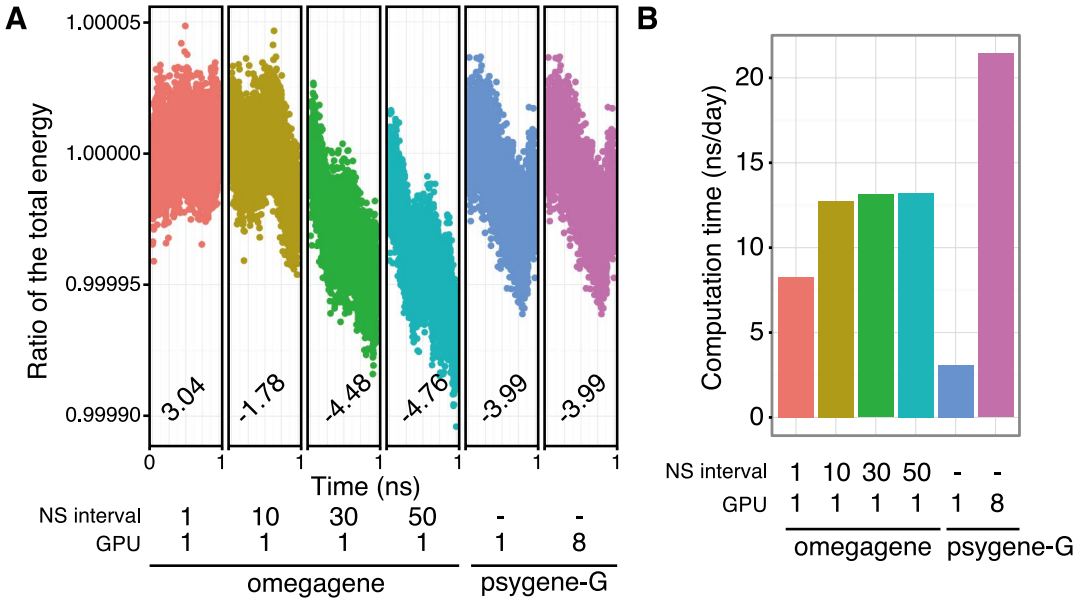

Figure 3 Benchmark results of myPresto/omegagene. (A) Time courses of the total energy of the system in various conditions. "NS interval" means the interval steps of the neighbor search. "GPU" is the number of GPU boards used in the simulation. The vertical axis indicates the ratio of the total energy over that at the first step. The value overlaid on each plot is the slope of fitted linear function. In these calculations, the time step was 2.0 fs, and the SHAKE algorithm was applied. (B) The computation speed in each condition.

2. Simulations

To demonstrate the performance of myPresto/omegagene, we carried out micro-canonical simulations and enhanced conformational sampling with the V-McMD method. The former was performed to evaluate the energy drift and compare the computation speed compared with that of myPresto/psygene-G. The system used in the both simulations includes a 20-mer peptide, Trp-cage with the sequence NLYIQWLKDGGPSSGRPPPS. For the potential parameters, Amber99SB-ILDN force field [37] for the protein, TIP3P water model [38], and the ion model reported by Joung and Cheatham [39] were applied. The electrostatic potential was calculated by ZMM with the zero-dipole condition and dumping parameter $\alpha=0$ [22]. The integration time step was 2.0 fs and covalent bonds with hydrogen atoms were constrained by the SHAKE algorithm [40].

2.1. Micro-canonical Simulation

We ran micro-canonical simulations with myPresto/omegagene for 1.0 ns, using several settings of intervals for the neighbor search, *e.g.*, 1, 10, 30, and 50 steps. The initial structure of Trp-cage was taken from the solution NMR structure (PDB ID: 1L2Y, model 1) [41], immersed into 150 mM NaCl solution (13,277 atoms; Supplementary Fig. S1A). A 12 Å cutoff radius was applied for the Lennard-Jones and electrostatic interactions and a 1 Å offset value was applied for the neighbor search. The same system was also simulated with myPresto/psygene-G for comparison, whose computation speeds were described for many protein systems [21].

We found that while a 10-step interval for the neighbor search improves the computation time and maintains the accuracy of simulation run that runs the neighbor search routine at every step, further elongation of the interval (30 and 50) causes drifts in the total energy (Fig. 3A and B). Thus, we recommend a 10-step interval as the default setting. Although even the conditions with small intervals (1 and 10 steps) in myPresto/omegagene showed a drift of the energy as well as myPresto/psygene-G does, the drift is so small that the results would be acceptable for usual purposes. Under the 10-step interval setting, the computation time with a single GPU is 4.12-times faster than the execution by myPresto/psygene-G with the same resource, and 1.68-times slower than the execution by myPresto/psygene-G with eight GPUs (the $2\times2\times2$ space decomposition). The bottleneck of computations was the pair potential functions on GPU (78.5% of the computation time was occupied; Supplementary Fig. S2) and the other potential calculations do not affect the computation speed because they are performed by CPU behind the GPU task. The integration task including the SHAKE algorithm spent 16.6% of the computation time. The remaining 5.9% was for the neighbor search. Although our main targets of myPresto/omegagene are small molecular systems with enhanced conformational sampling methods, we also tested a large molecular system consisting of an 11-subunit assembly of RNA polymerase (PDB ID: 4qiw; Supplementary Figs. S1C and S2B). The computation speed was approximately 0.27 ns/day for this 670,957-atom system by a single CPU/GPU.

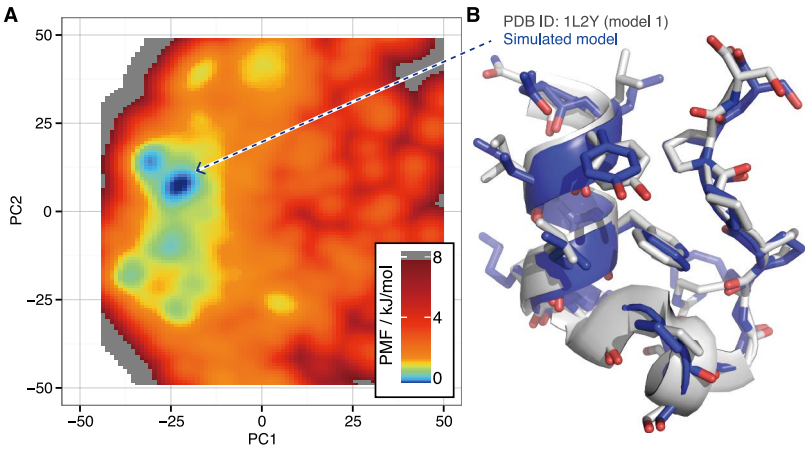

Figure 4 The results of the V-McMD simulation. (A) The simulated PMF landscape at 300 K. The horizontal and vertical axes are the first and second principal components (PC1 and PC2), respectively. The color graduation represents the PMF values (blue means the most stable). The lowest PMF is set to zero kJ/mol. (B) A superimposed picture of the NMR structure of Trp-cage (gray; PDB ID: 1L2Y, model 1) and the most stable conformation in the V-McMD simulation (blue).

2.2. McMD simulation

We applied myPresto/omegagene to conformational sampling with the V-McMD method [31]. The initial structure of Trp-cage was generated using "Modeller" without any template, and randomized by a constant temperature simulation at a high temperature (800 K; Supplementary Fig. S1B). The solution contains only one Cl^- for neutralization of the system (15,596 atoms in total). The same force field used for the earlier micro-canonical simulations was applied.

We briefly explain the protocols of the V-McMD for the present study below. A detailed elaboration of the methodology can be found in the reference [42]. The initial guess of the density of states was first approximated from a series of constant temperature simulations at various temperatures, ranging from 296 K to 629 K. Based on this result, the first iteration of the V-McMD simulation was performed under a biasing force. The energy distribution obtained from this simulation (Supplementary Fig. S3A) was then used to refine the estimated density of states and the biasing force, and this iterative process was performed 34 times (1.8×10^9 steps in total), until the energy distribution became nearly flat. After obtaining the biasing force, which produces the flat energy distribution, 50 independent runs with 1.5×10^7 steps for each were performed as the production runs (Supplementary Fig. S3B). The canonical ensemble at 300 K of this system was generated by reweighting the probability of each snapshot. The conformational ensemble of Trp-cage at 300 K is summarized as the potential of mean force (PMF) landscape, by applying principal component analysis (PCA), where 190 interatomic distances for the all pairs of Cα atoms in the peptide were used for variables to define the variance-covariance matrix [34]. As a result, we obtained the landscape with a major basin and some peripheral shallow basins (Fig. 4A). The most thermodynamically stable basin corresponds to the native structure (Fig. 4B) and the Cα RMSD value between one of the most stable structures and the NMR structure (PDB ID: 1L2Y, model 1) is 0.944 Å. This result demonstrates that myPresto/omegagene reasonably analyzed the conformational ensemble of the 20-mer peptide. In this calculation, the calculation speed was 10.7 ns/day, which is four-times faster than myPresto/psygene-G for the V-McMD with the ZMM.

3. Conclusions

We developed a new MD simulation program, myPresto/omegagene, that is tailored for our original enhanced sampling methods [31,32] and the electrostatic potential scheme [22–26]. This software is freely available to users, and the source code is distributed under an open-source license. In contrast to myPresto/psygene-G, which is powered by multi-GPU parallel computations [21], myPresto/omegagene is tailored for a single process execution with a single GPU, in order to optimize the enhanced conformational sampling methods [35]. Elimination of space-decomposition routines from the code greatly simplifies the codebase and the simple object-oriented structure of the code allows for easy development and maintenance of the software. In addition, the compatibility with the myPresto family provides advantages for applying many existing tools accumulated during the past decades and easy to use for myPresto users.

The evaluation study with the micro-canonical ensemble demonstrated the acceptable properties of the energy drift and efficient computation. Furthermore, we demonstrated that the enhanced conformational sampling simulation successfully reproduced experimentally solved structure of Trp-cage, as the most thermodynamically stable structure in the simulated conformational ensemble. myPresto/omegagene can be applied for the conformational sampling of such a practical molecular system with computation speed about

four-times faster than that of myPresto/psygene-G.

Acknowledgements

We thank Yoshifumi Fukunishi, Shun Sakuraba, Gert-Jan Bekker, Hasitha Muthumala Waidyasooriya, and Masanori Hariyama for fruitful discussions.

This work was supported by Japan Society for the Promotion of Science KAKENHI (Grant Numbers 24118008, 24240044, 16K18526, and 16K05517). This work was performed in part under the Cooperative Research Program of Institute for Protein Research, Osaka University, CR-16-05.

The MD simulations were performed on the TSUBAME2.5 supercomputer at the Tokyo Institute of Technology, provided through the HPCI System Research Project (Project IDs: hp130061, hp140032, hp150015, hp150065, and hp160165). The numerical computations for post-simulation analyses were performed, by using the supercomputer system provided by the National Institute of Genetics, Research Organization of Information and Systems, Japan.

Conflicts of Interest

All the authors declare that they have no conflict of interest.

Author Contribution

KK, YA and HN designed this work. KK, BM, KG, and TM designed the algorithms. KK, BM, and KG developed the software. JH and IF developed the theories. KK, BD, and JH performed calculations. All the authors wrote the paper.

References

[1] Shaw, D. E., Deneroff, M. M., Dror, R. O., Kuskin, J. S., Larson, R. H., Salmon, J. K., *et al.* Anton, a special-purpose machine for molecular dynamics simulation. *ACM SIGARCH Computer Architecture News* **35**, 1–12 (2007).

[2] Shaw, D. E., Grossman, J. P., Bank, J. A., Batson, B., Butts, J. A., Chao, J. C., *et al.* Anton 2: raising the bar for performance and programmability in a special-purpose molecular dynamics supercomputer. *SC '14: Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis* 41–53 (2014).

[3] Jensen, M. Ø., Jogini, V., Borhani, D. W., Leffler, A. E., Dror, R. O. & Shaw, D. E. Mechanism of voltage gating in potassium channels. *Science* **336**, 229–233 (2012).

[4] Shaw, D. E. 166 Millisecond-long molecular dynamics simulations of proteins on a special-purpose machine. *J. Biomol. Struct. Dyn.* **31**, 108 (2013).

[5] Ohmura, I., Morimoto, G., Ohno, Y., Hasegawa, A. & Taiji, M. MDGRAPE-4: a special-purpose computer system for molecular dynamics simulations. *Philos. Trans. A Math. Phys. Eng. Sci.* **372**, 20130387 (2014).

[6] Páll, S. & Hess, B. A flexible algorithm for calculating pair interactions on SIMD architectures. *Comput. Phys. Commun.* **184**, 2641–2650 (2013).

[7] Abraham, M. J., Murtola, T., Schulz, R., Páll, S., Smith, J. C., Hess, B., *et al.* GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **1-2**, 19–25 (2015).

[8] Götz, A. W., Williamson, M. J., Xu, D., Poole, D., Le Grand, S. & Walker, R. C. Routine Microsecond Molecular Dynamics Simulations with AMBER on GPUs. 1. Generalized Born. *J. Chem. Theory Comput.* **8**, 1542–1555 (2012).

[9] Salomon-Ferrer, R., Götz, A. W., Poole, D., Le Grand, S. & Walker, R. C. Routine Microsecond Molecular Dynamics Simulations with AMBER on GPUs. 2. Explicit Solvent Particle Mesh Ewald. *J. Chem. Theory Comput.* **9**, 3878–3888 (2013).

[10] Mori, T., Jung, J. & Sugita, Y. Surface-Tension Replica-Exchange Molecular Dynamics Method for Enhanced Sampling of Biological Membrane Systems. *J. Chem. Theory Comput.* **9**, 5629–5640 (2013).

[11] Jung, J., Mori, T. & Sugita, Y. Efficient lookup table using a linear function of inverse distance squared. *J. Comput. Chem.* **34**, 2412–2420 (2013).

[12] Jung, J., Mori, T. & Sugita, Y. Midpoint cell method for hybrid (MPI+OpenMP) parallelization of molecular dynamics simulations. *J. Comput. Chem.* **35**, 1064–1072 (2014).

[13] Andoh, Y., Yoshii, N., Fujimoto, K., Mizutani, K., Kojima, H., Yamada, A., *et al.* MODYLAS: A Highly Parallelized General-Purpose Molecular Dynamics Simulation Program for Large-Scale Systems with Long-Range Forces Calculated by Fast Multipole Method (FMM) and Highly Scalable Fine-Grained New Parallel Processing Algorithms. *J. Chem. Theory Comput.* **9**, 3201–3209 (2013).

[14] Ikeguchi, M. Partial rigid-body dynamics in NPT, NPAT and NPγT ensembles for proteins and membranes. *J. Comput. Chem.* **25**, 529–541 (2004).

[15] Ishida, H. & Matsumoto, A. Free-Energy Landscape of Reverse tRNA Translocation through the Ribosome Analyzed by Electron Microscopy Density Maps and Molecular Dynamics Simulations. *PLoS ONE* **9**, e101951 (2014).

[16] Fukunishi, Y., Mikami, Y. & Nakamura, H. Similarities among receptor pockets and among compounds: Analysis and application to in silico ligand screening. *J. Mol. Graph. Model.* **24**, 34–45 (2005).

[17] Fukunishi, Y. & Nakamura, H. Integration of Ligand-Based Drug Screening with Structure-Based Drug Screening by Combining Maximum Volume Overlapping Score with Ligand Docking. *Pharmaceuticals* **5**, 1332–1345 (2012).

[18] Fukunishi, Y. & Nakamura, H. Improved Estimation of Protein-Ligand Binding Free Energy by Using the Ligand-Entropy and Mobility of Water Molecules. *Pharmaceuticals* **6**, 604–622 (2013).

[19] Fukunishi, Y. & Nakamura, H. Prediction of ligand-binding sites of proteins by molecular docking calculation for a random ligand library. *Protein Sci.* **20**, 95–106 (2011).

[20] Kawabata, T., Sugihara, Y., Fukunishi, Y. & Nakamura, H. LigandBox: A database for 3D structures of chemical compounds. *BIOPHYSICS* **9**, 113–121 (2013).

[21] Mashimo, T., Fukunishi, Y., Kamiya, N., Takano, Y., Fukuda, I. & Nakamura, H. Molecular Dynamics Simulations Accelerated by GPU for Biological Macromolecules with a Non-Ewald Scheme for Electrostatic Interactions. *J. Chem. Theory Comput.* **9**, 5599–5609 (2013).

[22] Fukuda, I., Yonezawa, Y. & Nakamura, H. Molecular dynamics scheme for precise estimation of electrostatic interaction via zero-dipole summation principle. *J. Chem. Phys.* **134**, 164107 (2011).

[23] Fukuda, I., Kamiya, N., Yonezawa, Y. & Nakamura, H. Simple and accurate scheme to compute electrostatic interaction:

Zero-dipole summation technique for molecular system and application to bulk water. *J. Chem. Phys.* **137**, 054314 (2012).
[24] Fukuda, I. & Nakamura, H. Non-Ewald methods: theory and applications to molecular systems. *Biophys. Rev.* **4**, 161–170 (2012).
[25] Fukuda, I. Zero-multipole summation method for efficiently estimating electrostatic interactions in molecular system. *J. Chem. Phys.* **139**, 174107 (2013).
[26] Fukuda, I., Kamiya, N. & Nakamura, H. The zero-multipole summation method for estimating electrostatic interactions in molecular dynamics: Analysis of the accuracy and application to liquid systems. *J. Chem. Phys.* **140**, 194307 (2014).
[27] Arakawa, T., Kamiya, N., Nakamura, H. & Fukuda, I. Molecular Dynamics Simulations of Double-Stranded DNA in an Explicit Solvent Model with the Zero-Dipole Summation Method. *PLoS ONE* **8**, e76606 (2013).
[28] Kasahara, K., Fukuda, I. & Nakamura, H. A Novel Approach of Dynamic Cross Correlation Analysis on Molecular Dynamics Simulations and Its Application to Ets1 Dimer-DNA Complex. *PLoS ONE* **9**, e112419 (2014).
[29] Nishikawa, Y., Oyama, T., Kamiya, N., Kon, T., Toyoshima, Y. Y., Nakamura, H., *et al.* Structure of the Entire Stalk Region of the Dynein Motor Domain. *J. Mol. Biol.* **426**, 3232–3245 (2014).
[30] Nakajima, N., Nakamura, H. & Kidera, A. Multicanonical Ensemble Generated by Molecular Dynamics Simulation for Enhanced Conformational Sampling of Peptides. *J. Phys. Chem. B* **101**, 817–824 (1997).
[31] Higo, J., Umezawa, K. & Nakamura, H. A virtual-system coupled multicanonical molecular dynamics simulation: Principles and applications to free-energy landscape of protein–protein interaction with an all-atom model in explicit solvent. *J. Chem. Phys.* **138**, 184106 (2013).
[32] Higo, J., Dasgupta, B., Mashimo, T., Kasahara, K., Fukunishi, Y. & Nakamura, H. Virtual-system-coupled adaptive umbrella sampling to compute free-energy landscape for flexible molecular docking. *J. Comput. Chem.* **36**, 1489–1501 (2015).
[33] Iida, S., Nakamura, H. & Higo, J. Enhanced conformational sampling to visualize a free-energy landscape of protein complex formation. *Biochem. J.* **473**, 1651–1662 (2016).
[34] Higo, J., Nishimura, Y. & Nakamura, H. A Free-Energy Landscape for Coupled Folding and Binding of an Intrinsically Disordered Protein in Explicit Solvent from Detailed All-Atom Computations. *J. Am. Chem. Soc.* **133**, 10448–10458 (2011).
[35] Ikebe, J., Umezawa, K., Kamiya, N., Sugihara, T., Yonezawa, Y., Takano, Y., *et al.* Theory for trivial trajectory parallelization of multicanonical molecular dynamics and application to a polypeptide in water. *J. Comput. Chem.* **32**, 1286–1297 (2011).
[36] Anthopoulos, A., Grimstead, I. & Brancale, A. GPU-accelerated molecular mechanics computations. *J. Comput. Chem.* **34**, 2249–2260 (2013).
[37] Lindorff-Larsen, K., Piana, S., Palmo, K., Maragakis, P., Klepeis, J. L., Dror, R. O., *et al.* Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins* **78**, 1950–1958 (2010).
[38] Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**, 926 (1983).
[39] Joung, I. S. & Cheatham, T. E. Determination of Alkali and Halide Monovalent Ion Parameters for Use in Explicitly Solvated Biomolecular Simulations. *J. Phys. Chem. B* **112**, 9020–9041 (2008).
[40] Ryckaert, J. P., Ciccotti, G. & Berendsen, H. J. C. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J. Comput. Phys.* **23**, 327–341 (1977).
[41] Neidigh, J. W., Fesinmeyer, R. M. & Andersen, N. H. Designing a 20-residue protein. *Nat. Struct. Biol.* **9**, 425–430 (2002).
[42] Higo, J., Ikebe, J., Kamiya, N. & Nakamura, H. Enhanced and effective conformational sampling of protein molecular systems for their free energy landscapes. *Biophys. Rev.* **4**, 27–44 (2012).