Study on Atomistic Behavior of Macromolecules by Molecular Dynamics Simulation and its Enlargement in Scale

Tomohiro Seko, Yuki Sasaki, Kazuhiko Yanagida, Kazuki Inami, and Nobuyuki Nakayama; Fuji Xerox Co., Ltd.; Kanagawa, Japan

Abstract

Properties of materials are constituted by molecular structures. The molecular dynamics simulation technique has an advantage in clarifying the relationship between molecular structure and properties through analysis of atomistic structure and dynamics, which are difficult to measure by experiment. On the other hand, major problems with applying this technique to larger and more complicated molecules consist of controlling the calculation time, and the larger scale of molecules.

In this study, first, a relationship between melting properties and molecular structure is investigated by using the atomistic molecular dynamics simulation to verify effectiveness of this technique. The relationship between the reciprocal of self diffusion coefficients obtained from the simulation and 1/2 FT, which is an experimental indicator for melting properties, shows good agreement. This result indicates that atomistic molecular behavior affects melting properties, and implies that polymers with Bisphenol A structure have higher molecular dynamics.

Next, a reduction of calculation cost and an enlargement of molecular simulation are examined by following 2 approaches¬the coarse-grained united-atom model and the supercomputer K computer. The coarse-grained united-atom model shows good parallel scalability. For polycarbonates, it has the equivalent accuracy of the full atomistic model, and is more than 15 times faster than the full atomistic model with the same number of molecules and cores. These techniques expanded the spatial scale of molecular dynamics simulation by one digit larger than ever before.

Introduction

In the electrophotographic system, functional materials like toner and photoreceptor are commonly used. The properties of these materials are constituted by molecular structure. However, it is difficult to experimentally observe the molecular structure and properties at the atomic level.

Molecular simulation technology has been used in recent years as one approach to overcome this challenge. At pharmaceutical companies, screening is conducted by simulating chemical interaction between the target protein and a candidate drug for the purpose of reducing development costs and period [1]. The simulation technique has an advantage in analyzing the atomistic structure and behavior, which is difficult to measure by experiment. However, application of this technique to larger and more complicated molecules is difficult because full atomistic molecular dynamics require a long period of time to calculate movement and interactions of all atoms.

Thus, this technique is limited to simulations of a small ensemble of molecules, such as polymers with low relative molecular mass. An extension of a spatial scale for simulation is required for analyzing a higher-order structure characteristic of macromolecules such as aggregation, entanglement and phase separation.

Based on such a background, in this study, first, usability of molecular dynamics simulation for melting properties of polymers is investigated. By using an equilibrium structure and radial distribution functions, characteristics of molecular structure, depending on the type of monomer or polymer, are estimated. Then, from a view of dynamics contributing to melting properties, a relationship between self diffusion coefficients and measured 1/2 FT is investigated. Second, for the purpose of applying molecular dynamics simulation to larger molecules, reduction of calculation costs and the possibility of enlargement of the spatial scale is examined. By using the two approaches of a coarse-grained united-atom model and the K computer, computational accuracy, and scalability and calculation performance for up to 1024 parallel processor cores in relation to the full atomistic model were verified.

Molecular Dynamics Simulation

Molecular dynamics (MD) is a simulation technique to calculate molecular structure and dynamics by solving the classical motion equation for each atom based on interactions between atoms. The procedure for a full atomistic molecular dynamics simulation is as follows. The first step is to calculate the acting force $\mathbf{F}_i(t)$ in the $\mathbf{r}_i$ coordinate for each atom using equation (1). U in equation (1) is called the molecular potential and is defined by several forces at work between atoms. The molecular potential contains two types of forces derived from chemically bonded force and non-bonded force. Equation (2) represents the constituents of the molecular potential, where the bond potential U_{bond}, the angle potential U_{angle} and the torsion potential $U_{torsion}$ belong to the category of chemically bonded force, and the van der Waals potential U_{vdw} and Coulomb potential $U_{coulomb}$ belong to non-bonded force.

$$\mathbf{F}_i(t) = -\frac{dU}{d\mathbf{r}_i} \tag{1}$$

$$U = U_{bond} + U_{angle} + U_{torsion} + U_{vdW} + U_{coulomb} \tag{2}$$

The bond potential in equation (3) represents a stretching vibration between two atoms that are connected by a covalent bond as shown in Figure 1 (a). The angle potential in equation (4) represents displacement between an equilibrated angle (θ_{eq}) and 3 atoms as in Figure 1 (b). The torsion potential in equation (5) shows the torsion of the angle created by 4 atoms as in Figure. 1 (c). The van der Waals potential in equation (6) and Coulomb potential in equation (7) shown in Figure 1 (d) are non-bonded forces among atoms to be considered. Parameters such as Kr in these potentials are defined for each atom and bond type, i.e., sp^3 carbon. Furthermore, different parameter sets are provided for different molecules such as organic compounds, proteins, and inorganic compounds.

$$U_{bond} = \sum_{bonds} K_r \left(r - r_{eq}\right)^2 \tag{3}$$

$$U_{angle} = \sum_{angles} K_\theta \left(\theta - \theta_{eq}\right)^2 \quad (4)$$

$$U_{torsion} = \sum_{torsions} \frac{V_n}{2}\left(1 + \cos\left(n\phi - \gamma\right)\right) \quad (5)$$

$$U_{vdW} = \sum_{i<j}^{atoms}\left(\frac{A_{ij}}{R_{ij}^{12}} - \frac{B_{ij}}{R_{ij}^{6}}\right) \quad (6)$$

$$U_{coulomb} = \sum_{i<j}^{atoms}\frac{q_i q_j}{\varepsilon R_{ij}} \quad (7)$$

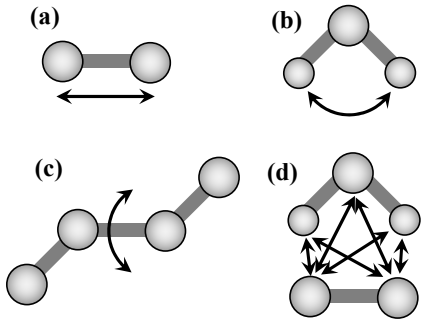

Figure 1*. Defines of each potentials, (a) bond, (b) angle, (c) torsion (d) van der Waals and coulomb.*

In the next step, the velocity $\mathbf{v}_i(t)$ of each atom is calculated using equation (8). Here, t- Δt signifies the time only one step (time increase Δt) before the present time t (current step). The coordinate of an atom after Δt is updated using equation (9). Repeating the series of calculations in equation (1), (8) and (9) allows us to update the force and the coordinate of the atoms, infer molecular behavior through the elapse of time. In actual MD calculations, equation (1) is expanded based on statistical mechanics so that temperature and pressure can be handled.

$$\mathbf{v}_i\left(t\right) = \mathbf{v}_i\left(t - \Delta t\right) + \frac{\Delta t}{m_i}\frac{\mathbf{F}_i\left(t\right) + \mathbf{F}_i\left(t - \Delta t\right)}{2} \quad (8)$$

$$\mathbf{r}_i\left(t + \Delta t\right) = \mathbf{r}_i\left(t\right) + \Delta t\mathbf{v}_i\left(t\right) + \frac{\Delta t^2}{2m_i}\mathbf{F}_i\left(t\right) \quad (9)$$

Polymer Structure and Melting Property

Sample Polymer and Simulation Model

Sample polymers in this study consist of the alcohol and acid monomers shown in Figure 2, and their compositions are listed in Table 1. The VSOP ver. 1.8 SP1 and J-OCTA® ver. 1.8 were used as programs for simulation. The full atomistic model, which depicts all atoms and their interactions clearly, is deployed for modeling monomer structures shown in Figure 2. The simulation model has 30 polymers, and molecular weight Mn is set to about 4400 based on Table 1. The full atomistic model uses the molecular potential GAFF (General AMBER Force Field) [2].

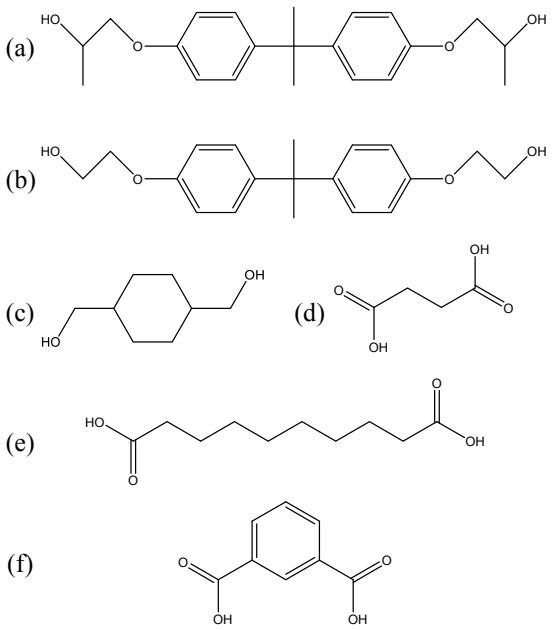

Figure 2*. Structures of monomer. (a) BPA-PO, (b) BPA-EO, (c) CHDM, (d) SA, (e) SBA, (f) IPA.*

Table 1: Composition of polymers.

	Alcohol monomer	Acid monomer
Sample 1	BPA-PO (30) BPA-EO (20)	SA (35), IPA (15)
Sample 2	CHDM (50)	SA (18), IPA (32)
Sample 3	CHDM (50)	SBA (7), IPA (43)

Figures inside brackets mean monomer ratio.

Result and Discussion for Molecular Structure and Dynamics

Melting is the phenomenon of polymers being liberated from kinetic restraints and dynamics being activated. It is thought that molecular structure greatly affects the microscopic dynamics of polymers. The effect that differences between the polymers of BPA and CHDM, and of short chain monomers (SA, C2 chains) and long chains (SBA, C8 chains), as shown in Fig. 2 and Table 1, have on polymer structure and its dynamics were evaluated with a focus on equilibrium structure, radial distribution function, and self diffusion coefficient.

Figure 3 shows the calculated equilibrium structure for three sample polymers. Each sample polymer has an evenly distributed molecular structure, and remarkable differences such as aggregation derived from some alcohol and acid monomer were not found.

Next, radial distribution functions for each sample polymer are compared. The radial distribution function is the ratio of the number density of an atom, which exists at distance *r* from the atom, to the overall averaged number density. Figure 4 shows the radial distribution functions of carbon atoms for three sample polymers. The radial distribution functions of Samples 2 and 3, which have the same alcohol monomer CHDM, are consistent with each other. On the other hand, the radial distribution functions of Sample 1 differ

from those of Samples 2 and 3 at a peak of about 0.3 nm, which indicates that differences in polymer structure at a short-range distance are affected by alcohol monomers.

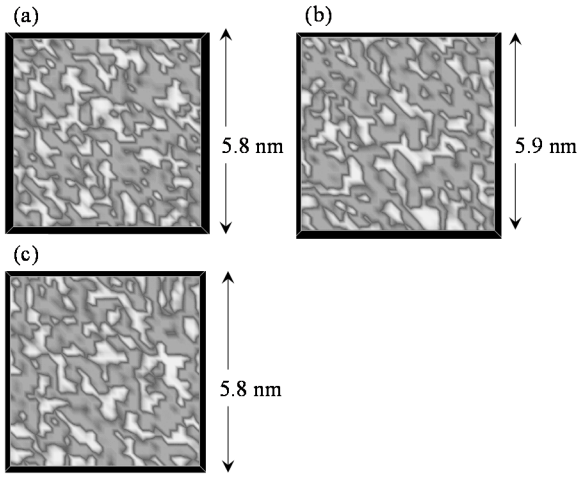

***Figure 3**. Equilibrium structure for sample polymers. (a) Sample 1 (b) Sample 2 (c) Sample 3. Dark gray and light gray indicate polymer and vacancy, respectively.*

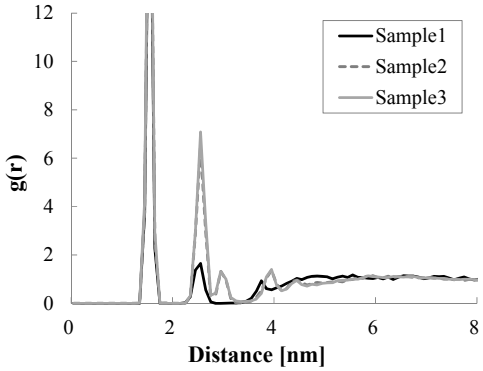

***Figure 4**. Radial distribution function for sample monomers.*

The self diffusion coefficient is used to evaluate atomistic behavior for melting properties. The self diffusion coefficient D is a mean square displacement of each atom for unit time expressed in the Einstein equation (10), and is an indicator of molecular dynamics.

$$D = \lim_{t\to\infty} \frac{1}{6t} \left\langle \left| \mathbf{r}_i(t) - \mathbf{r}_i(0) \right|^2 \right\rangle \qquad (10)$$

The self diffusion coefficients were obtained based on the following procedure. The temperature of the sample polymer in the molecular dynamics simulation was increased from 300 K to more than a glass transition temperature T_g, and then quenched to 300 K to calculate a volume at each point of temperature. The T_g in this simulation was estimated as a point of inflection of a temperature-dependent plot for volume, and then the self diffusion coefficients were calculated near T_g.

Figure 5 shows the relationship between the reciprocal of self diffusion coefficients obtained from the simulation and 1/2 FT. Here, 1/2 FT is an experimental indicator for melting property, and is the average temperature of the beginning and end of outflow of melted polymer measured by the flow tester method. Both have good agreement, with a correlation coefficient of 0.99. This result indicates that atomistic molecular behavior affects melting property, and the self diffusion coefficients can be an indicator.

In Figure 5, Sample 1, which has BPA-PO and BPA-EO with Bisphenol A structure, shows a smaller 1/2 FT and reciprocal of self diffusion coefficients than Samples 2 and 3 with a cyclic compound of the CHDM monomer. This implies that polymers with a Bisphenol A structure have higher molecular dynamics. Compared to Sample 2 with a short chain monomer (SA, C2 chains), Sample 3 with long chain monomer (SBA, C8 chains) has a higher 1/2 FT and reciprocal of self diffusion coefficients. It is considered that a long chain structure restrains polymer dynamics.

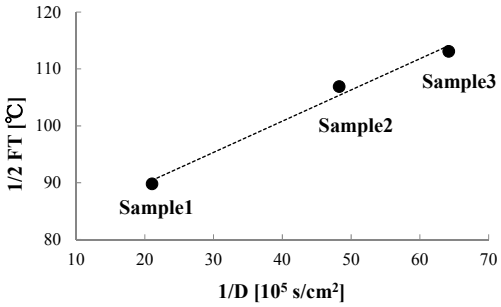

***Figure 5**. Relation between reciprocal of self diffusion coefficient estimated by simulation and measured temperature.*

Enlargement of Molecular Simulation

Methodology of Simulation

Simulation model: coarse-grained united atom model

In a huge molecule with a larger molecular weight than the sample polymer, higher-order structure characteristics such as entanglement and domains are significantly apparent, and affect melting properties. However, the size of such a higher-order structure is more than several 10 nm, and it is difficult to perform a full atomistic molecular dynamics simulation due to calculation costs. Thus, the two approaches of the coarse-grained atom model [3] and using a supercomputer were examined in order to reduce calculation costs and enlarge the spatial scale so that high-order structures can be handled using atomic simulation.

As the first approach, the coarse-grained united-atom model was used. This is a simplified model of the full atomistic model created based on the following chemical concerns [3]. Carbon atoms form the skeleton of macromolecules, directly affect the motion and entanglement of polymers. On the other hand, hydrogen atoms are generally attached to one carbon atom and do not form a main skeleton. Therefore, the coarse-grained united-atom model avoids treating hydrogen atoms explicitly by grouping carbon atoms and hydrogen atoms shown in Figure 6 with the aim of simulating large atomistic conformations with fewer calculations. In this coarse-grained approximation, the radius of hydrogen atoms is considered as the radius of the carbon-hydrogen atom group. The coarse-grained united-atom model makes it possible to reduce the number of atoms to be handled in the motion equations of equation (8) and

(9). In Figure 6, the number of atoms to calculate is reduced from 10 to 3. Furthermore, this model allows us to not consider the calculations for electrostatic interactions in equation (7), which occupy the majority of calculation time in full atomistic model. These simplifications realize faster calculations than the full atomistic model.

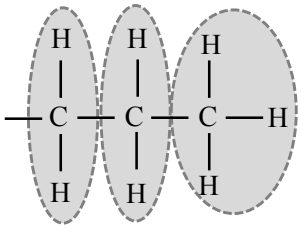

***Figure 6**. Concept of united atom coarse-grained model. A gray circle means group of carbon and hydrogen atoms for coarse-grained approximation.*

The polycarbonate shown in Figure 7, which has the same diphenylmethane structure as BPA-PO and BPA-EO monomers, is used for examining an enlargement of the molecular dynamics simulation. Since polycarbonates are one of the more popular engineering plastics, they are applied toward general-purpose molecular potential. The full atomistic model uses the molecular potential GAFF [2], while the coarse-grained united-atom model uses DREIDING [4].

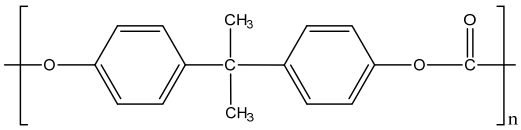

***Figure 7**. Molecular structure of the polycarbonate.*

Supercomputer: K computer

In molecular dynamics simulations, parallel computation is necessary to expand the scale of calculations, and the supercomputer is an effective second approach for that purpose. In this study, the K computer [5] was used to perform the simulation. The K computer is one of the most powerful supercomputers for highly-parallel simulations. It was developed by RIKEN and Fujitsu Co., Ltd., and was recognized as the fastest computer in the world in 2011. It has over 80,000 CPUs and 700,000 cores, and can perform 10 Peta flops calculations. Each CPU is connected by a unique interconnect networking interface called TOFU [5] (TOrus FUsion). A conventional interconnect is linked linearly; on the other hand, TOFU consists of a hybridization of circularly-connected torus and linearly-connected mesh as shown in Figure 8.

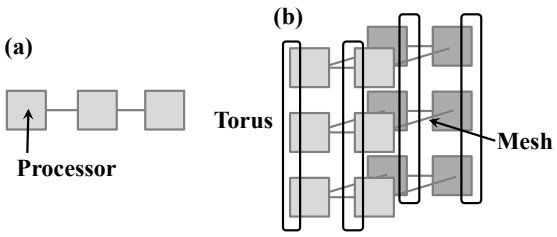

***Figure 8**. Concept of TOFU inter connect network.*

Result and Discussion for Enlarged Simulation

United atom coarse-grained model for polycarbonate

The full atomistic and coarse-grained united-atom models of polycarbonates are created. The models consist of 1,322 atoms and 762 atoms, respectively, for a chain with 40 monomers.

A simulation with these models for 10 chains was performed for an equilibrium calculation consisting of 1 million steps (1 ns), and the calculated densities are shown in Table 2. The calculated values are within 6% of the value obtained by experiment, showing good agreement, and accuracy within 3%.

Table 2: Calculated densities

Model	Density [g/cm^3]
Full atomistic	1.13
United atom coarse-grained	1.16
Experiment	1.20

Scalability of parallel computing for full atomistic and coarse-grained model

Scalabilities of parallel computing for the full atomistic model and the coarse-grained united-atom model were investigated. First, the parallel computing time based on full atomistic model was measured for up to about 2.9 million atoms, observing how long it takes for computers with processor core of 64 to 1,024 processor cores to calculate 1,000 steps (1 ps). Parallel computation was performed based on flat MPI. Figure 9 shows the dependency of the calculation time on the number of atoms and processor cores.

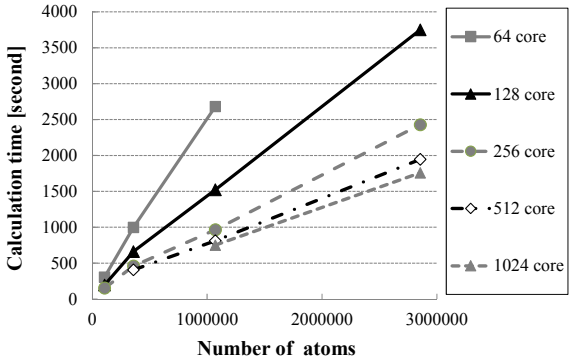

***Figure 9**. Parallel calculation time depending on the number of atoms and processor cores for full atomistic model.*

Generally, the multi-body problem of calculating electrostatic force requires calculation time of a square order of the number of atoms, and this characteristic represents a bottleneck in the full atomistic molecular dynamics simulation. Several algorithms for fast and highly accurate calculation of electrostatic force have been developed to overcome this difficulty. The PPPM (Particle-Particle Particle-Mesh) [6] method used in this study is one of these algorithms. The PPPM method calculates the electrostatic interactions between atoms by dividing space between real space near the focused atom and reciprocal space as shown in Figure 10, which realizes scalability of linear order for the number of atoms. The dependence of the number of atoms in Figure 9 shows a linear scalability from 64 to 1024 parallel cores.

However, the calculation times for when 256 or more cores are used are nearly the same, and a tendency toward saturation with a

large number of parallel cores is observed. The relative efficiency as a result of parallelism of a system of about 2.9 million atoms with 1,024 cores is only 27% in relation to 128 cores, which is not sufficient for accelerating simulations of large-scale molecules, which require a long period of time for calculations.

This is because as shown in Figure 11, reciprocal space calculation in the PPPM method does not decrease with an increase in the number of parallel cores. The dependency of calculation time on processor cores for the reciprocal space in PPPM is inferred to be due to the following. In accordance with an increase in the number of processor cores, the majority of FFT processing is accounted for by the communication time; the processing time does not decrease with an increase in the number of parallel cores, and parallel scalability worsens. In addition, as shown in Figure 10, although the reciprocal space is calculated by partitioning it using mesh, in the PPPM method, when the number of parallel cores becomes larger than the number of such meshes, the calculation grain size becomes smaller than the number of parallel cores, and parallel effects cannot be obtained (300 meshes in a system with approximately 2.9 million atoms). As a result, due to synergistic effects, when there are more than 512 cores, the processing time for reciprocal space calculations in the PPPM method becomes longer than when there are less than 512 cores.

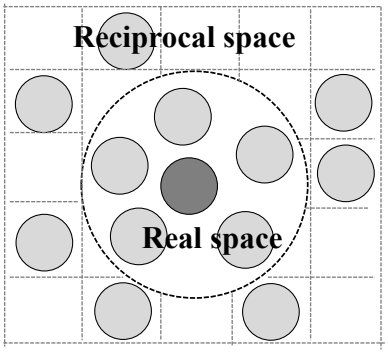

***Figure 10**. Concept of the PPPM method. A dashed circle means a boundary between real space and reciprocal space for this algorithm.*

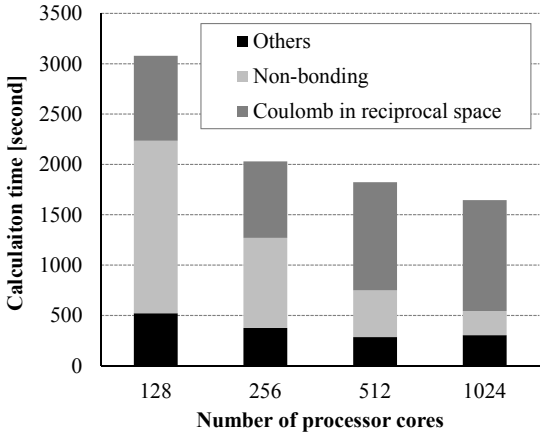

***Figure 11**. Profile for parallel calculation time depending on processor cores for full atomistic model.*

Next, the time required for a 10,000 step (10 ps) calculation with 128 to 1,024 CPUs for a system of up to about 3.9 million particles was measured using the coarse-grained united-atom model as shown in Figure 12. The coarse-grained united-atom model is confirmed to have linear dependence of particles and parallel scalability. This model shows a relative parallel efficiency of 67 % with 1,024 cores for a system of 3.9 million particles, which is twice more efficient as compared to the full atomistic model. The coarse-grained united-atom model was more than 15 times faster than the full atomistic model with the same number of molecules and cores.

Good parallel scalability of the coarse-grained united-atom model is due to a tendency of decreased calculation time for each routine with an increase in processor cores, as shown in Figure 13. A routine for the reciprocal space in PPPM is not included in the coarse-grained united-atom model.

Here, handling of electrostatic interactions in the coarse-grained united-atom model is considered. For example, DNA molecules have a double helix structure based on hydrogen bonds, which connect bases on two chains. Thus, a full atomistic model with electrostatic interactions is essential to simulate such a highly polarized structure. On the other hand, in a molecule that is considered to have relatively smaller polarization, such as polycarbonate, application of electrostatic interactions is negligible to the coarse-grained united-atom model.

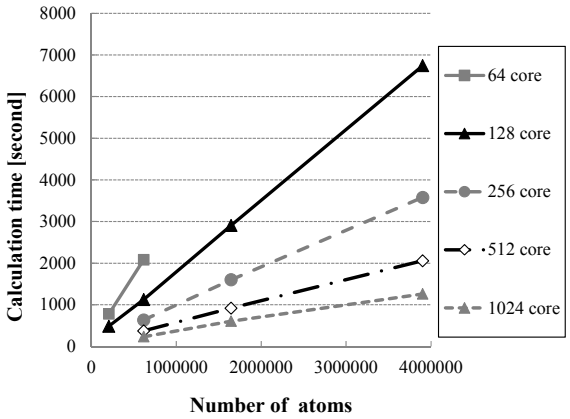

***Figure 12**. Parallel calculation time depending on the number of atoms and processor cores for coarse-grained united-atom model.*

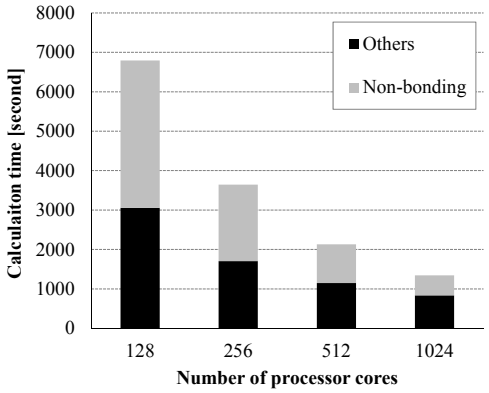

***Figure 13**. Profile for parallel calculation time depending on processor cores for coarse-grained united-atom model.*

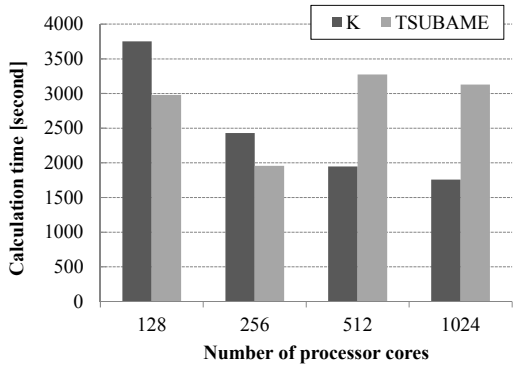

***Figure 14**. Parallel calculation time depending on processor cores by K and TSUBAME supercomputer for the full atomistic model with 2.9 million atoms.*

Although the parallel performance of the full atomistic model shows a tendency to worsen, an increase in calculation time with an increase in processor cores was not observed. On the other hands, results obtained by using the supercomputer TSUBAME, which is equipped with general network interconnect (Infini Band), showed an increase in calculation time on more than 512 processor cores. This result indicates that TOFU has an advantage over general interconnect on a condition of more than a several hundred processor cores, which require high performance of communication ability.

MD simulation for large size of molecules

A large-scale molecular simulation based on the coarse-grained united-atom method was implemented using the K computer. 10.3M step (7.5 ns) calculation was performed for 827,968 atoms constitutive of 544 chains of 80 monomers, based on 512 parallel cores. The density obtained was 1.19 g/cm^3, which agrees well with the value obtained by experiment. The structure obtained is shown in Figure 15. Compared to previous work, the spatial scale of calculations was expanded by one digit larger than ever before in volume. This result is expected to lead to a clarification of the relationship between a higher-order structure, such as entanglement of molecular chains, and melting properties derived from dynamics for huge molecules with larger molecular weight.

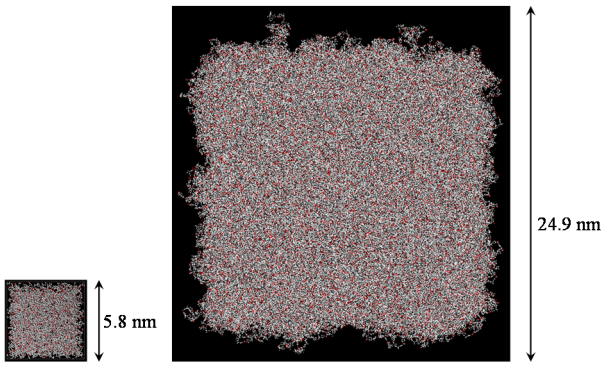

***Figure 15**. Molecular structures. (left) former work. (right) This work. It is consisted of 827,968 atoms by the united atom coarse-grained model.*

Conclusion

First, usability of molecular dynamics simulation was investigated to clarify molecular structure and properties. The equilibrium structure, radial distribution function and self diffusion coefficients for sample polymers were evaluated to clarify the relation between melting properties and microscopic structure and behavior. The equilibrium structures for the three sample polymers did not show a remarkable difference. The radial distribution functions showed a difference, depending on the alcohol monomer structure. A relation between the reciprocal of self diffusion coefficients obtained from the simulation and 1/2 FT have good agreement. This result indicates that the atomistic molecular behavior affects melting property, and implies that polymers with Bisphenol A structure have higher molecular dynamics.

Next, a reduction of calculation cost and an enlargement of molecular simulation are necessary to handle huge molecules with higher-order structure by larger molecular weight. Then, an enlargement of molecular simulation by several 10 nm was examined for a polycarbonate with the same diphenylmethane structure as BPA by following two approaches−the coarse-grained united-atom model and use of a supercomputer. The full atomistic model tended to worsen in terms of parallel efficiency at a scale of 1,024 parallel cores. On the other hand, the coarse-grained united-atom model showed good parallel scalability. The coarse-grained united-atom model had an accuracy equivalent to the full atomistic model, and was more than 15 times faster than the full atomistic model with the same number of molecules and cores. These techniques expanded the spatial scale of molecular dynamics simulation by one digit larger than ever before in volume.

Acknowledgement

Part of this research used computational resources of the K computer provided by the RIKEN Advanced Institute for Computational Science through the HPCI System Research project (Project ID:hp140242), and of the TSUBAME supercomputer provided by Tokyo Institute of Technology (Project ID:14ISG).

References

[1] T. Karaki, “Jisedai supakon EXA ga nihon wo kaeru”, Shogakukan, pp. 120-127, 2014 [in Japanese].

[2] Wang J., Wolf R. M., Caldwell J. W., Kollman, P. A. and Case D. A., “Development and testing of a general AMBER force field,” Journal of Computer Chemistry, **25**, pp. 1157-1174, 2004.

[3] V. A. Harmandaris and K. Kremer, “Dynamics of Polystyrene Melts through Hierarchical Multiscale Simulation”, Macromolecules, **42**, pp.791-802, 2009.

[4] S. L. Mayo, B. D. Olafson, and W. A. Goddard, “DREIDING: a generic force field for molecular simulations”, Journal of Physical Chemistry, **94,** pp. 8997-8909, 1990.

[5] Ryutaro Himeno, “An Illustrated Guide to High Performance Computing”, Kodansya, 2012 [in Japanese].

[6] Shin kagaku gijyutu suishin kyoukai, “Koubunshi simulation”,Kagaku kougyo nippousay, pp. 44, 2014 [in Japanese].

Author Biography

Tomohiro Seko *holds a MS degree* in Chemistry *from Okayama* Univ. *(*1998*). In 1998, he joined Fuji Xerox Co., Ltd., where he has been engaged in a research on electrophotography process simulation and is currently studying on computational chemistry.*