

BioSupercomputing Newsletter

2015.3

Vol.12

CONTENTS

Open Up SPECIAL TALK Discussion meeting

**Toward a reproduction of biomolecular functions
in intracellular environments through large-scale
simulation utilizing the high computational
capabilities of the K computer**

Hisashi Ishida Japan Atomic Energy Agency / Isseki Yu RIKEN /
Motoshi Kamiya Kyoto University / Ryo Kanada Kyoto University /
Kazunari Iwamoto RIKEN / Yuichi Kokabu Yokohama City University ————— 2

SCLS Eyes S-cruise From the field of software development

**Reproducing the processes of blood clot formation
and growth by focusing on platelet activation**

Norio Shimamoto The University of Tokyo ————— 6

ZOOM IN Close-up on SCLS Research

Theme 1 Simulations of biomolecules under cellular environments

**Exploring a chemical reaction in an intracellular crowded
environment by the QM/MM method**

Motoshi Kamiya Kyoto University ————— 8

Theme 3 Hierarchical integrated simulation for predictive medicine

**Multi-scale simulation of platelet aggregation
in initial stage of thrombogenesis**

Seiji Shiozaki Tokai University School of Medicine ————— 9

Theme 4 Large-scale analysis of life data

**Simulating cancer evolution to understand principles
generating intratumor heterogeneity**

Atsushi Niida The University of Tokyo ————— 10

SCLS Gotcha!

Remote Interactive Lectures "Foundations of Computational Life Sciences" — 11

Further Sophistication to Make Effective Use
of the High Performance Computing Infrastructure ————— 11



Theme 1: Simulations of biomolecules under cellular environments

Toward a reproduction of biomolecular functions in intracellular environments through large-scale simulation utilizing the high computational capabilities of the K computer

To transition from molecular level simulations to cellular functions, and connect with understanding and predicting biological phenomena

The introduction of the K computer is about to open a path for reproducing the behavior of a biomolecule under conditions close to those of an actual intracellular environment, which has been considered difficult due to the limitations of hitherto computational capabilities. The aim of Theme 1, "Simulations of biomolecules under cellular environments" (leader: Yuji Sugita, Principal Investigator of RIKEN), is to go from molecular level computations to "whole-cell simulation", which is directed to understanding and predicting biological phenomena. Today, we are hearing from six investigators who are undertaking research in the simulation of biomolecules under cellular environments, which is key to achieving this goal.

On the basis of a loose collaboration, the theme of each investigator is tackled

● What type of research are you currently undertaking in Theme 1?

Yu: The aim of the Sugita team is to reproduce, in a computer, the behavior of biomolecules under cellular environments, through large-scale molecular dynamics calculations using the K computer. In this process, we have a bacterium called mycoplasma as the target, and we have created a cell environment model at an atomic level, in which almost all the proteins and metabolites involved in its metabolic network coexist at concentrations that are close to those inside an actual cell. Currently, we are carrying out all-atom molecular dynamics simulations using this model. The interior of a cell is

extremely crowded with biomolecules, and there are a large number of proteins, small molecules, ions, and so on, in a melee, similarly to a packed commuter train. What type of dynamics and interactions the proteins and small molecules display in such a packed environment is an important question in biology as well, but it is not well understood, due to the difficulties in obtaining experimental measurements. The objective is to elucidate this aspect.

Iwamoto: What our team is doing is modeling and simulation of intracellular signal transduction pathways, which take the intracellular environment into consideration. A cell responds suitably to a change in the environment. This

response is called signal transduction, which pathways are biochemical reaction networks among a number of intracellular proteins. A large number of such signal transduction pathways exist inside a cell. Well known among them is the epidermal growth factor (EGF) signal transduction pathway. It has the role of receiving EGF ligand stimulation from the external environment communicating the information to the cell interior, and prompting the cell to grow, divide or differentiate. We are currently undertaking cell simulations pertaining to the networks in this pathway. The cell simulation is a method for simulating a cell virtually in a computer, and one important



Hisashi Ishida

Senior Researcher, Quantum Beam Science Center, Japan Atomic Energy Agency

Participating in the round-table talk again made me realize the importance of multi-scale simulation that links the broadness and hierarchy of the scale at which biological phenomena are handled. At the same time, I strongly felt the various developments in the field of simulation.



Isseki Yu

Research Scientist, RIKEN Theoretical Molecular Science Laboratory

I strongly felt that all the members who participated in the round-table talk were undertaking research toward the elucidation of a theme with passion and a sense of mission. Going forward, I think that it is important to undertake the elucidation of biological phenomena on multiple scales while effectively sharing the data obtained with one another.

Motoshi Kamiya

Program-Specific Researcher, Department of Chemistry, Graduate School of Science, Kyoto University

The calculations I am performing tend to have aspects that inevitably require the size of the system to be modest, so, I felt a sort of admiration in regard to the large-scale computation that the people who participated in the round-table talk are carrying out with the K computer. In addition, I strongly agree with the discussion on the importance of collaboration with experimental people.



Ryo Kanada

Program-Specific Researcher, Department of Biophysics, Division of Biological Sciences, Graduate School of Science, Kyoto University

Listening to everyone's story, I realized once again that my own research was closely related to other people's research. I think that by deepening collaborative relationships and exploiting findings and data obtained on different scales, progress in each other's research can be made more effectively.



Kazunari Iwamoto

Postdoctoral Researcher, RIKEN Quantitative Biology Center

I truly feel that, in order to utilize the K computer effectively, it is important to have a good handle on knowledge regarding parallelization, and on the characteristics of the calculator itself. Meanwhile, today, coordination with experimental systems is also important; so, arrangements may need to be made to have investigators of experimental systems participate.



Yuichi Kokabu

Postdoctoral Researcher, Computational Life Science Laboratory, Graduate School of Medical Life Science, Yokohama City University

Listening to the people who participated in the round-table, I am beginning to see where my research stands in the ongoing large-scale simulations. In addition, thinking about the maximizing the efficiency of the computational capabilities of the K computer, I truly felt that simulation research geared toward gigantic systems will become more and more important.



feature of our method is that the high computational capabilities of the K computer are utilized so that the cell shape and structures are included and diffusion, collision, reaction fluctuation, and so on, are represented at the molecular resolution.

Kamiya: Intracellular signal transduction is mainly carried out through phosphorylation reactions by several kinases. The reaction activity of a kinase is a chemical process that is key to intracellular signal transduction, and elucidating the molecular mechanism

of reaction activity in a cellular environment is important in establishing a molecular platform oriented toward understanding and control of intracellular signal transduction. For this purpose, our team is carrying out research and development related to

Toward a reproduction of biomolecular functions in intracellular environments through large-scale simulation utilizing the high computational capabilities of the K computer

reactivity analysis of kinases in signal transduction pathways. As mentioned by Dr. Yu, the interior of a cell is an extremely crowded environment. In order to elucidate what kind of chemical reaction is occurring in such a crowded environment and how different it is from a conventional reaction occurring in water, we are performing analyses using the QM/MM method, which combines quantum chemical (QM) computation and molecular mechanical (MM) computation.

Kanada: We are using the coarse-grained molecular dynamics simulation program (CafeMol) which has been developed by our research team to carry out research focused on the MAPK system phosphorylation cascade (a series of reactions), which is a representative model for signal transduction pathways. As already mentioned by Dr. Iwamoto and Dr. Kamiya regarding signal transduction, in the mammalian MAPK system cascade, upon stimulation from the exterior, the upstream protein MEK1 (MAPKK) activates the downstream protein ERK2 (MAPK) by phosphorylating it, leading to a transmission of the signal, finally the signal is transmitted to the interior of the nucleus. However, the details of the molecular mechanism for phosphorylation are not understood, and need to be elucidated. Therefore, based on the individual monomeric structures of MEK and ERK, which have been resolved experimentally, we are investigating the MEK-ERK complex structure and the dynamics of its

formation. Another target is related to the ERK protein translocated into the nucleus after phosphorylation. Applying coarse-grained simulation, we would like to elucidate the dynamical behavior of the ERK protein in an intranuclear crowded environment where chromatin is present in high concentration.

Ishida: We are dealing with nucleosomes. The DNA from eukaryotes, including humans, has a fundamental structure called a nucleosome, which is DNA wound around a core of proteins called a histone, and many of them are aligned to form a chromosome, which is housed inside the nucleus in a compact form. This nucleosome is deeply involved in gene regulation, and known to change structure or position in relation to DNA metabolism such as transcription, replication, and repair. In order to elucidate the mechanism of movement, structural stability and so on, of the nucleosome, we are using total atomic models to calculate the extent of DNA winding and unwinding around the histones in each of the systems including conventional canonical histones and variants that constitute the nucleosome. More specifically, we simulate a slow pulling and peeling of the DNA and use the K computer to calculate the structure distribution in so doing, called a free energy profile in expert terms, i.e., whether the free energy of the pulled and peeled DNA is high or low. We are trying to go as far as seeing, at the total atomic level, how different the ease of peeling becomes when a component of the nucleosome is changed.

Kokabu: As explained by Dr. Ishida, the intranuclear DNA has a nucleosome structure in which DNA is wound around histone proteins, and adopts a structure in which several nucleosomes have condensed. In our laboratory, we are undertaking research regarding the functional dynamics of these intranuclear DNA-binding proteins, using molecular simulation, while linking with experimental systems. In this process, we have been undertaking program development for the coarse-grained MD-SAXS method. The SAXS (small-angle X-ray scattering method) is a method in which a protein in solution is irradiated by X-rays, and the overall shape of the molecule is measured from its scattering pattern, and it allows a protein structure that fluctuates to be seen. Meanwhile, MD (molecular dynamics) simulation can predict, on the total atomic level, the behavior of a biological macromolecule such as a protein by solving an equation of motion. However, in the case of a gigantic molecule such as the nucleosome, a total atomic MD simulation would require too much computation time, so that a coarse-grained (CG) MD is used, in which an amino acid is substituted with a particle, to predict the structure, and from this result, the SAXS scattering pattern is computed theoretically. The computation result and the SAXS experimental data are compared and examined; if they correspond, then the simulation result can be deemed appropriate.

Undertaking research with a dream

● What is your impression after using the K computer?

Ishida: We are doing a large computation using 4,800 nodes for an all-atom simulation of peeling a nucleosome consisting of approximately 500,000 atoms. The calculations simulate the behavior of about 50,000,000 atoms in total. Such a simulation is indeed impossible if not for the K computer, and I really have the feeling of doing cutting-edge computation, which could not be carried out so far. While we are currently in the middle of analyses, we are starting to learn about what is needed in order

to do high-precision analyses through ultra-large system simulations.

Yu: Even in such an extremely small bacterium as mycoplasma, when you try to include all the proteins related to a metabolic network at realistic number density ratios, the total number will end up being about 1,500 at least. Since the simulation of one to a few large proteins is common at a conventional laboratory level, I indeed think this research would be impossible without the K computer. In terms of impression after actual use, in addition to the fact that it is crowded with jobs, determining

an optimal number of nodes required a bit of trial-and-error. If the number of nodes secured is too large, the wait time becomes extremely long; conversely, if the number of nodes is too small, although jobs can be placed at a good pace, the trade-off is that the time available for simulation is shorter. I struggled to find the most balanced number of nodes in the beginning.

Kamiya: The K computer is indeed attractive for handling large systems. Regarding MD, we are also using the molecular dynamics simulator "GENESIS", which is used by Dr. Yu's

group as well. It is extremely efficient, and helpful to us. The ability to run MD of a large-scale system at high-speed is indeed important. As for QM/MM, while it can be run, there is the fact that the input has become very complicated, and there are many failures, which is an issue. In addition, since job submission has to be reiterated several times, the wait time to do so is long, which is a real problem.

Kanada: When investigating the dynamics of the ERK protein in an intranuclear crowded environment, we are including in the system a chromatin structure consisting of 20 nucleosomes to perform the calculations. My impression is that, since this is an extremely large system for coarse-grained simulation by CafeMol, utilizing the K computer is extremely helpful to proceed with large-scale calculations efficiently and systematically.

Iwamoto: I too think that the ability to look at large systems in fine detail is what makes the K computer attractive. I am dealing with a whole cell. Ideally, I would like to bring the grid size to the size of one protein molecule, i.e. about 5 nanometers, but there is not enough memory in PC clusters such as those in conventional laboratories. On this point, the K computer allows you to proceed with no problem at all.

Kokabu: In my case, since I use coarse-grained simulation, the to-be-computed system itself is not that large, but in order to examine various conditions, I am doing simulation with 90 conditions for one nucleosome. In addition, even for coarse-grained simulation, enough conformational space cannot be searched in only one round, and the same computation is repeated about 100 times. I feel that using the K computer is extremely important for this purpose as well.

● Going forward, what type of research would you like to carry on?

Ishida: I have been doing simulation

research on biological macromolecular systems that exceed 1,000,000 atoms. Calculators have been developed, which possess high computational capabilities, such as those of the K computer, and simulation accuracy also has been increasing dramatically compared to the past. In this context, I would like to keep my interest in simulation techniques, and carry out research that nicely links the best part of coarse-grained simulation with the best part of all-atom simulation. In addition, I think that ultimately connecting with drug development and such, may be a theme in simulation research, but I personally would like to carry on steadily with fundamental research.

Yu: In contrast, I would like to expand in the direction of medicine and drug development. The research I am currently undertaking has strong basic science aspects, and while this obviously remains important to me, in the future, I would like to contribute to the development of drug design processes that take the cellular environment into account. Currently, computer-assisted drug development focuses mainly on the affinity between a protein and a drug; to this, I would like to add the influence of the intracellular environment. For example, intracellular water has different viscosity and dielectric constant from pure water, and I believe there is also a significant influence on the binding affinity between a target and a drug. If a general cellular environment model could be built by gathering such findings, wouldn't it lead to a more realistic drug development process?

Kamiya: My "root" being chemistry, I would like to continue by properly handling the chemical reactions upon which the field of life science is based. While setting my aim in such a direction, I would like to contribute to society through drug development, or the like.

Kanada: I would like to reproduce the behavior of large systems by utilizing

the simulation program "CafeMol". I am thinking that whole-cell simulation will also become possible with "CafeMol" in the future. By running high performance calculators such as the K computer over an extended time period, using not only protein models but also DNA and lipid models, it should be possible to simulate an entire cell. While my current main targets are proteins related to signal transduction, I would like to contribute also to research on systems including DNA and lipids.

Iwamoto: To be brief, I too would like to achieve whole-cell simulation. However, I do not mean coarse-grained simulation; I would like to achieve whole-cell simulation including all of the numerous functions that a cell possesses in order to respond to various stimuli from the environment, such as those involved in signal transduction and metabolism. About two years ago, an article came out about an American research group having simulated all the mycoplasma reaction systems. Progress is also being made in whole-simulation research with *Escherichia coli* and the like. I would also like to proceed in such a direction.

Kokabu: We happen to collaborate with molecular biologists in the experimental field, and when we learn that a prediction by our simulation turned out to be correct through an actual experiment, we are sometimes overjoyed. For such reasons, I would like to continue placing an importance on collaboration with experimenters. I think that findings through simulation are also extremely useful to molecular biologists conducting experiments. I would also like to do research that can support experimental projects, through simulation.



After the interview, the interviewees answered questions from and talked freely with young researchers and students. Visit the URL below for more information.



After SPECIAL TALK

● See the URL on the right for longer version of this discussion meeting. <http://www.scls.riken.jp/eng/newsletter/Vol.12/openup.html>

Reproducing the processes of blood clot formation and growth by focusing on platelet activation

Multiscale thrombus simulator “EX-THROM”

Researcher, Department of Mechanical Engineering, The University of Tokyo **Norio Shimamoto**



Theme 3 “Hierarchical integrated simulation for predictive medicine” (leader: Prof. Shu Takagi, Graduate School of Engineering, The University of Tokyo) aims to elucidate the complicated onset process of heart diseases and motor dysfunction by integrating simulations of the circulatory, musculoskeletal and central nervous systems to provide optimum medical support to each and every individual. A multiscale thrombus simulator “EX-THROM” is being developed as a part of the project. Developed by expanding “ZZ-THROM”, which was developed in a grand challenge program “Next-Generation Integrated Simulation of Living Matter (ISLiM)” (until March 2013) and is a simulator for reproducing the initial phase of thrombus formation, “EX-THROM” aims to reproduce thrombus growth. We interviewed Dr. Shimamoto, and asked about the simulator and its development.

Focusing on ADP stimulation in thrombus growth

— How does a thrombus actually form?

Shimamoto The mechanism of thrombus formation is basically the same as that of blood coagulation at hemostasis, which is a defensive mechanism against damage to blood vessels. The coagulation system is an important response for maintaining life; but the blood clot may grow, hinder the flow of blood, block up a vessel and cause ischemia or infarction. This is called thrombosis.

Platelets play the key role in thrombus formation. When a vascular wall is damaged due to some reason, it triggers von Willebrand factor (vWF) on the wounded site to bind with protein GPIIb/IIIa on a platelet, resulting in the platelet adhering to the vessel wall. The activated protein GPIIb/IIIa present on the platelet binds with vWF and fibrinogen, and causes the platelet to successively adhere with other platelets and a platelet thrombus, which is an aggregate of platelets, to form. Then, a series of multistage chain reactions by 12 blood coagulation factors takes place. During the process, a network structure called fibrin mesh is formed and entraps blood cells such as erythrocytes, making the thrombus to grow further and a stable “fibrin clot” to be formed.

— Your team is developing EX-THROM. What kind of application software is it?

Shimamoto ZZ-THROM, which is a multisimulator for reproducing the initial phase of thrombosis, was developed in the grand challenge program based on a new fluid-structure coupling method (ZZ-EFSI). It reproduces the processes of platelet adhesion on a damaged vascular wall. We are now developing EX-THROM, which is a multiscale thrombus simulator, to reproduce the activation process of adhered platelets and the subsequent thrombus formation processes by expanding the functions of ZZ-THROM. A number of substances are involved in thrombus formation compositely and at various stages. The processes are still not fully understood. To develop the thrombus simulator, we need to incorporate minute reactions, but we are thinking of narrowing targets first, creating a framework that reproduces the characteristic movements of platelets, and then moving into minute reactions.

— What are the important targets for building the framework?

Shimamoto During the processes of platelet aggregation, platelet-inducing

substances, which binds with receptors on the platelet surface and stimulate and promotes activation of the platelet, play important roles. Adenosine diphosphate (ADP) is one of such substances. It has been experimentally revealed that platelets aggregate and dissociate depending on ADP concentration. Activated platelets release ADP contained in dense granules, and the “ADP positive feedback mechanism”, which is a chain reaction that amplifies activation, occurs. Therefore, we are focusing on and modeling the activation of platelets by ADP, which is a key process in platelet aggregation. In concrete terms, we are reproducing the activation of GPIIb/IIIa, which is induced by ADP binding to ADP receptors P2Y12 and P2Y1 and subsequent activation reactions. Using this model, we are going to investigate how the activations of P2Y12 and GPIIb/IIIa on the platelet surface undergo time sequential changes in a condition where ADP, vWF and fibrinogen diffuse advectively and undergo concentration changes. Obviously, ADP alone does not stimulate the entire process. Various substances are involved in an integrated manner. We will be expanding the functions of the model by adding their involvement.

● We also aim to use the model to evaluate a thrombosis therapeutic agent

— **What are the important points in modeling?**

Shimamoto The important point is to well reproduce the characteristics of the actual phenomena. There are many substances involved; but if we incorporate them all blindly, we would not be able to understand what is happening. On this point, we are receiving advice from Professor Shinya Goto of Tokai University School of Medicine, who is a specialist in thrombosis, as well as other specialists. We are going to create a model consistent with experimental data by also including the reactions of thrombin, serotonin and thromboxane A2 as well as ADP stimulus.

— **We have heard that your study project also aims to use the model to evaluate a thrombosis therapeutic agent.**

Shimamoto Yes. We are now investigating clopidogrel, which blocks P2Y12

of the ADP receptor and inhibits binding, and an inhibitor of GPIIb/IIIa, which binds platelets together.

— **Have you already obtained the computation results from the K computer?**

Shimamoto We are still in the phase of testing calculation. Actual calculation using the K computer will be the next step. We will be calculating the process of platelet activation inter-connectedly with blood flow calculation including erythrocytes, which was developed in ZZ-THROM.

— **What was the most difficult point in modeling?**

Shimamoto I am originally from an engineering school, and had difficulty understanding the medical information. Moreover, much is left unveiled in the mechanism of thrombus formation. It is

also difficult deciding how to refine the model. At present, I'm trying to improve the model so as to fit the knowledge that has experimentally been acquired by Prof. Goto and his team of medical specialists by continuously giving feedback.

— **Please tell us about your prospects.**

Shimamoto We have just advanced from the stage of aggregation to the stage of coagulation. We are now also investigating coagulation. Our goal is to reproduce the actual process of thrombus formation. We also aim to contribute to evaluation of a thrombosis therapeutic agent. For example, if we can reveal how much should be administered and at which timing of thrombus formation by simulation, we would be able to contribute to thrombosis treatment.

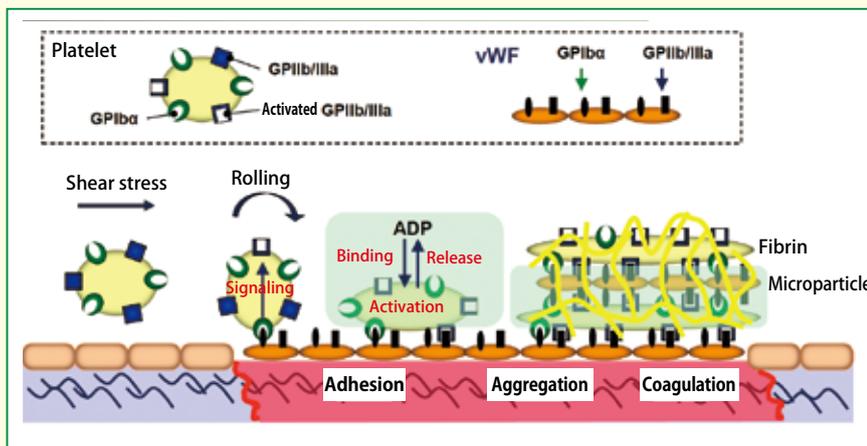


Fig. 1: Mechanism of thrombus formation. Von Willebrand factor (vWF) in the damaged vascular endothelium binds with protein GPIIb/IIIa present on a platelet, resulting in the platelet adhesion. When the platelet is activated by a signal transduction of stimulus, it releases ADP, inducing further activation of itself and other platelets, and the procedure to advance to aggregation and coagulation.

Modification of an illustration cited from *Platelet Biology (in Japanese)* edited by Ikeda and Maruyama (Medical Review Co., Ltd.)

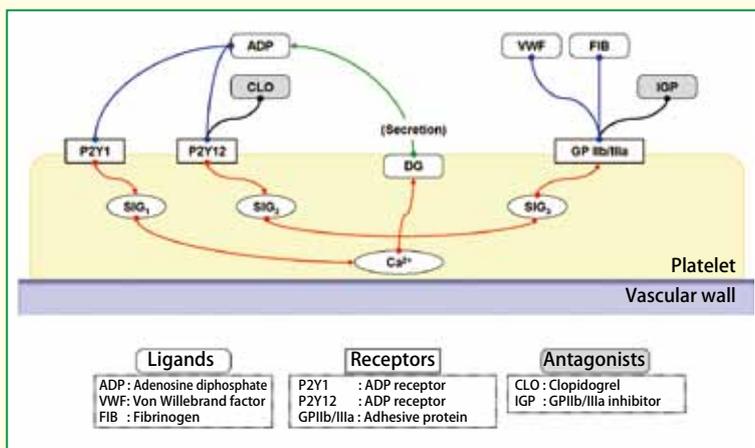


Fig. 2: Reaction model of platelet activation

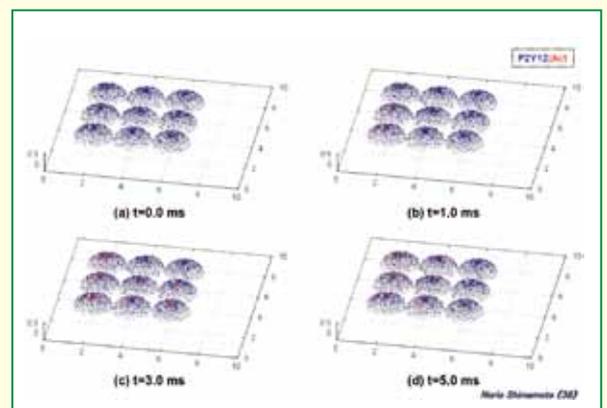


Fig. 3: Calculation example of P2Y12 activation. During the time course in which ADP undergoes advective diffusion, the P2Y12 molecules on the surface of the platelets bind with ADP, and deactivated molecules (blue) become activated (red).

Exploring a chemical reaction in an intracellular crowded environment by the QM/MM method

Theoretical Chemistry Laboratory, Department of Chemistry, Graduate School of Science, Kyoto University

Motoshi Kamiya



Various chemical reactions continuously occur in the living body. Knowledge of those chemical reactions is essential to understand life phenomena and go further. The QM/MM method we use (a combination of quantum chemistry and classical physics to deal with a chemical reaction in a complex system with a high degree of accuracy) is a very important method to gain knowledge of a chemical reaction in a complex environment like a biological system. In 2013, Dr. Karplus, Dr. Levitt and Dr. Warshel shared the Nobel Prize in Chemistry for the development of the QM/MM method. It shows how important and promising it is.

Although the QM/MM method is a very excellent method to deal with chemical reactions, it has a problem of requiring very high computational capacity. Recently developed computers can now deal with an isolated protein floating in water (Fig. 1). However, their capacity is not sufficient to handle *in*

vivo chemical reactions. For example, a real living cell is packed with various proteins (Fig. 2). It is much different from the too ideal environment shown in Fig. 1. However, conventional QM/MM computation does not take account of such a crowded cellular environment at all. On top of that, temperature is another important and sensitive factor for living organisms. Although thermal fluctuation of protein conformation always occurs, its impact has not been handled properly. However now, thanks to the K computer and other large-scale computers, QM/MM computation in the environment shown in Fig. 2 has become possible. We also can deal with temperature efficiently to some extent by the method our group has developed recently (QM/MM RWFE SCF method). Honestly speaking, how things change in the very complex environment shown in Fig. 2 has been poorly understood. Even imagining it is difficult, but now we can study it anyway.

Now we are studying the reaction

by which GTP (guanosine triphosphate commonly found in functional regulation of a living organism) breaks down into GDP (guanosine diphosphate) in a system called the Ras/GAP complex in the environment shown in Fig. 2. Both Ras and GAP proteins are involved in canceration as well as in life. Although we have just started computing and there are many things we still don't understand, we will discover, bit by bit, how a more realistic environment (temperature, crowded environment, etc.) works, and how things are affected by such an environment.

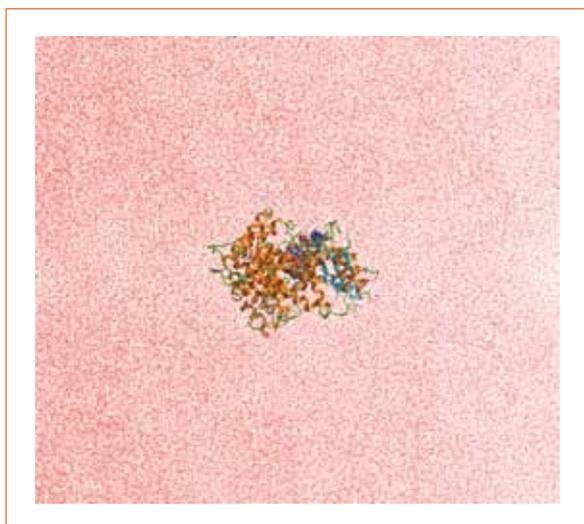


Fig. 1 : Ideal protein (Ras/GAP complex) floating in water (red dot: water molecule)

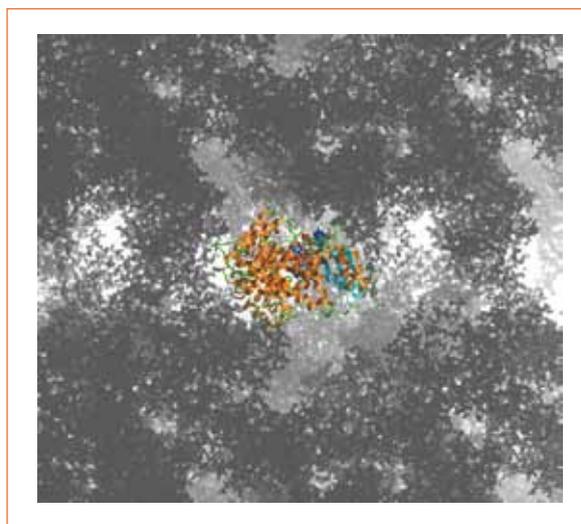


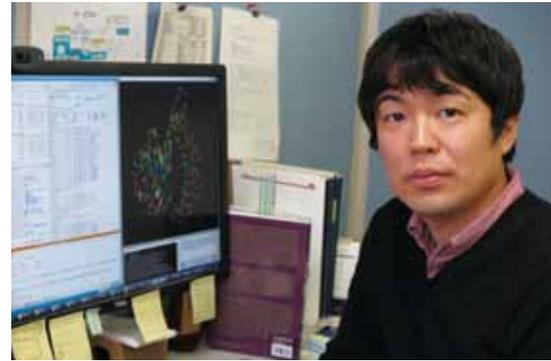
Fig. 2 : Ras/GAP complex in schematic cell environment (water molecules omitted. Gray dots represent other proteins)



Multi-scale simulation of platelet aggregation in initial stage of thrombogenesis

Division of Cardiovascular Medicine, Department of Internal Medicine,
Tokai University School of Medicine

Seiji Shiozaki



When blood vessels are damaged by injury, etc., platelets clump together at the site of injury to stop bleeding. When the inner wall of a blood vessel is damaged due to arterial sclerosis, platelets also agglutinate and form a blood clot inside the blood vessel. Such blood clots cause thrombosis including heart infarction and cerebral infarction, which are major cardiocirculatory diseases and rank fourth as a cause of death in Japanese people. Platelets agglutinate in a blood vessel through interaction between glycoprotein Ib alpha (GPIb alpha) found on platelets and the von Willebrand factor (VWF) adhering to the wall of an injured blood vessel. A few dozen GPIb alpha-VWF couples are formed on the joining surface of the platelet to stick platelets together. On a larger scale, the dynamic action of erythrocytes in blood plays a crucial role in adhesion of platelets.

Fig. 1 shows the binding site structure of the GPIb alpha-VWF couple obtained by the molecular dynamics method. The molecular dynamics method can determine the momentary position of every atom by solving Newton's equation

of motion for all atoms constituting the protein molecules and water. By this method, we can calculate the force working between the proteins and the potential energy of the system we calculate. It enables us to model reaction rates of bond formation and dissociation of a protein couple. Fig. 2 shows a grid model of a platelet surface. In this model, the behavior of a membrane protein on the platelet surface is analyzed by a statistical approach called the Monte Carlo method to find the momentary bond number and adhesion force of the protein. In the calculation area, $1.5 \mu\text{m}$ on a side, which is equivalent to about one fifth part of the platelet surface area, there are 4,000 GPIb alpha and 3,600 VWF binding sites. Each dot in Fig. 2 represents one GPIb alpha molecule. We presumed that GPIb alpha is scattered on the platelet surface whereas VWF is anchored to the blood vessel wall, and that GPIb alpha molecules bound to VWF are neither scattered nor move over the membrane. We also presumed that GPIb alpha near VWF forms a bond with VWF at a certain combination rate, and the

bond is dissociated at a certain bond dissociation rate. As those rates, we can use a model modeled by the molecular dynamics method. By introducing the number of protein bindings and their adhesion force into an Eulerian approach for fluid-hyperelastic body interaction by the finite difference method, it becomes possible to reproduce a phenomenon in which platelets adsorb onto a blood vessel wall in the presence of erythrocytes^[1].

By performing a multi-scale simulation from the protein scale to the bloodstream scale in this way, we are studying further to elucidate the clotting mechanism, and in the future we will construct a useful model for the discovery of an antiplatelet drug, and predictive and personalized medicine.

Reference :

[1] BioSupercomputing Newsletter, Vol. 7, p. 7.

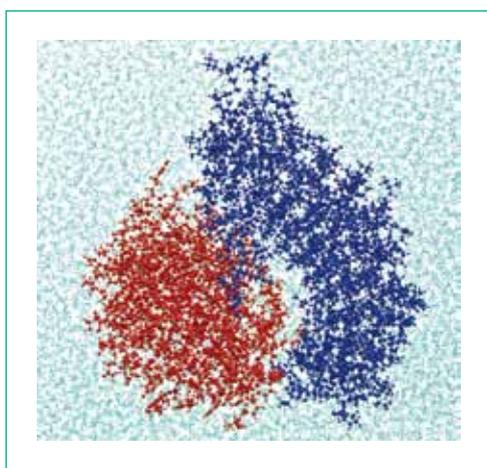


Fig. 1 : Binding site structure of N-terminal of platelet membrane protein GPIb alpha (blue) and A1 domain of VWF (red)

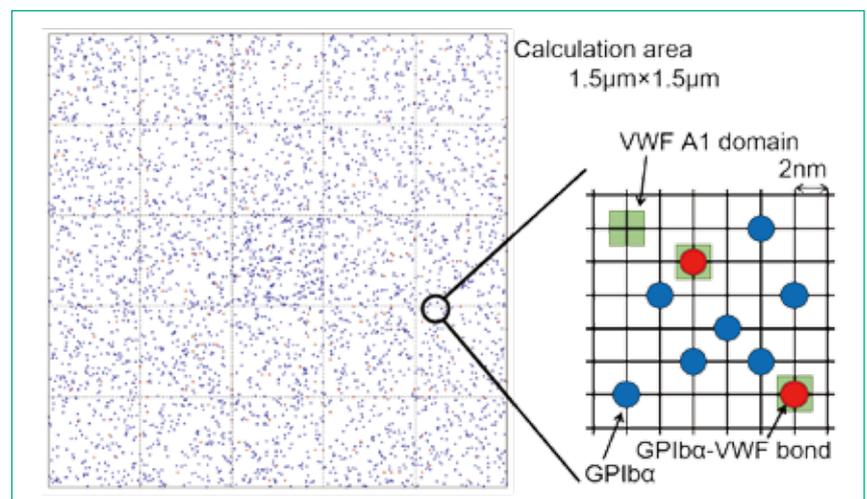


Fig. 2 : Platelet surface grid model. Red dots represent bound GPIb alpha, and blue dots represent free GPIb alpha which moves over the platelet surface.



Simulating cancer evolution to understand principles generating intratumor heterogeneity

Human Genome Center, The Institute of Medical Science,
The University of Tokyo

Atsushi Niida



Cancer results when cells abnormally proliferate by genetic mutation. It is known that every patient has his or her own different cancer-causing genetic mutation. Furthermore, it has been revealed that even a single cancer in a patient has multiple cell populations with various combinations of genetic mutation. This phenomenon is called intratumor heterogeneity, and thought to play a role in resistance to anti-cancer therapy. Provided that a tumor consists of a large number of cell A which responds to treatment and small numbers of treatment-resistant cell B, the tumor temporarily shrinks with treatment due to the decrease of cell A. However, when cell B grows, before long, the tumor recurs. Therefore, it is a clinically important task to clarify the principles generating intratumor heterogeneity.

We first worked with Kyushu University Beppu Hospital to investigate the tumor heterogeneity of colorectal cancer. We took DNA samples from multiple sites of a single colorectal tumor and analyzed their genomes with the Next Generation Sequencer. As a result, we found a genetic mutation shared by all samples.

Meanwhile, it was revealed that each sample has its own genetic mutation, which causes high heterogeneity.

Next, we tried to clarify the principle generating such intratumor heterogeneity, and constructed a simulation model of cancer evolution, the Branching Evolutionary Process (BEP) model. When cells are allowed to grow by mutating multiple genes randomly, cells accumulating driver mutations that increase proliferation rate are selected. As a result, they evolve into a cell population with high proliferating ability. In such an evolutionary process, the cells branch into cell populations with different mutations under a certain condition and acquire heterogeneity.

We simulated cancer evolution under various conditions with the K computer in an exhaustive search for a condition which generates high heterogeneity. We found that if we postulate a high gene mutation rate and the existence of cancer stem cells, we can reproduce the gene mutation pattern with high heterogeneity observed in the above experimental results

using colorectal cancer. In addition, the results of our simulation suggested that driver genes are shared by all tumor cells, while most heterogeneity-producing mutations are neutral mutations which do not affect the cell growth rate.

As above, in our study, some evolutionary principles generating intratumor heterogeneity were elucidated by genome analysis of colorectal cancer and simulation of cancer evolution with the K computer. From now on, we plan to establish a treatment strategy based on a study to overcome treatment-resistant cancers.

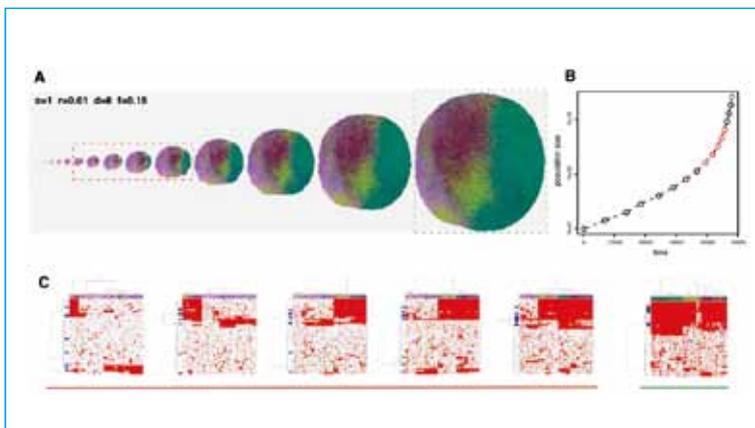


Fig. 1: Visualized simulation of cancer evolution. (A) Growing cancer, (B) Growth curve, (C) Genetic mutation pattern

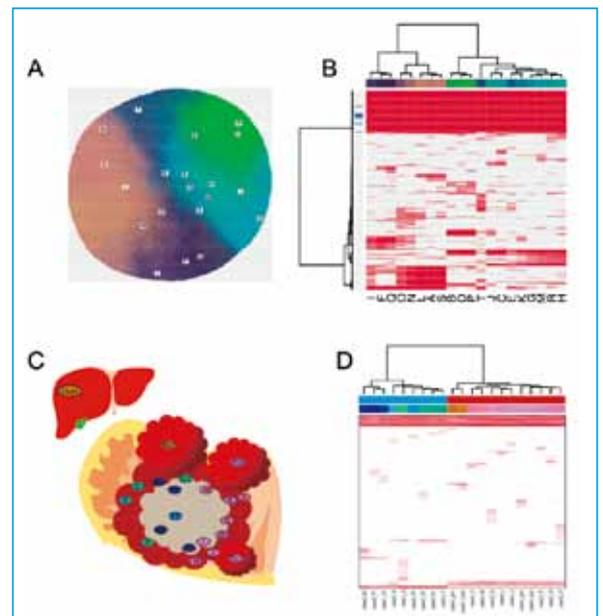


Fig. 2: Genetic mutation patterns with high heterogeneity obtained by simulation of cancer evolution (A, B) and genome analysis of colorectal cancer (C, D)





Remote Interactive Lectures “Foundations of Computational Life Sciences”

From the tangent point of life and engineering sciences to applications in society

Held by : Education Center on Computational Science and Engineering, Kobe University
Co-held by : RIKEN HPCI Program for Computational Life Sciences

Life sciences, such as biology, medicine, pharmacy and agricultural science, developed dramatically in the latter half of the 20th century. In the 21st century, which is called “the era of life sciences,” scientists are required not only to investigate the fundamental theories of physics and chemistry but also to carry out research based on enormous databases related to life sciences, which have accumulated all over the world. It is also increasingly required to apply the fundamental theories of physics, chemistry and engineering in life sciences.

SCLS is giving a series of a course of basic lectures on computational life science, which forms the tangent point of life and engineering sciences, jointly with the Education Center on Computational Science and Engineering of Kobe University for undergraduates, graduate students and working members of society.

The course, which started in October 2014, consists of 15 lectures

in three categories ranging from the tangent point of life and engineering sciences to applications in society; they are “Life Sciences Viewed from the Genome,” “Life Sciences Viewed from the Protein,” and “Life Sciences in Medicine and Drug Discovery.” The lectures are easy to attend as auditing students can attend the lectures via their own personal computer.

All lectures are given by scientists active in various fields at universities, companies, and others. The lecturers overview computational life science, introduce technologies of data mining, mention gene expression databases, and briefly explain methods of calculating and analyzing molecular dynamics simulations of proteins, and the present states and prospects of computational life science in the fields of medicine and drug development.

About 260 undergraduates, graduate students and researchers of companies have registered and are attending the lectures. We believe that they will form a community of

computational life science researchers, transform Kobe, which is now a supercomputer center, into a base for transmitting computational life science information, and propel the development of computational life science into the future.

The FY2014 course will end on February 3 (Tuesday), 2015; but the course is scheduled to be held also in FY 2015. We look forward to your attendance.



Further Sophistication to Make Effective Use of the High Performance Computing Infrastructure

High Performance Computing Development Group, RIKEN HPCI Program for Computational Life Sciences

The High Performance Computing Development Group of RIKEN HPCI Program for Computational Life Sciences, which is a representative organization of SCLS, has its headquarters in RIKEN Advanced Institute for Computational Science in Port Island, Kobe. It is in charge of providing support services to users, and help them use the computing infrastructure and execute their research tasks smoothly. Our work can be broadly classified into two parts: 1) to manage the computing infrastructure efficiently, and 2) to provide research support and assist users to use the K computer. As for the former, we have constructed and are operating an original computer environment called the SCLS supercomputer system, which supplements the High Performance Computing Infrastructure (HPCI), so that researchers who actually work on the task and related medical and drug

development can make full use of the K computer and other constituents of HPCI. The operating policy is determined by a steering committee established to deliberate each and every issue. For the convenience of users, open software has also been ported into the environment, and are being used by many users mainly for preparation before calculation using the K computer and also for post-processing work. As for the latter category of our work, we have actively collected and accumulated know-how, such as advanced parallel computing and programming techniques necessary for using the K computer, at academic conferences and study meetings, and from various kinds of information available on the Internet. The information has been transferred to persons in charge of theme studies and is used for improving the parallel computing performance of the program for analyzing intracellular

signal transduction. To disseminate study results to the public, we also give a short course on the use of the new programs whenever appropriate. To cultivate human resources who have advanced knowledge and skills sufficient for making use of the HPCI environment, we are also holding various kinds of lecture meetings and seminars jointly with RIKEN Advanced Institute for Computational Science, other strategic fields, and the Registered Institution for Facilities Use Promotion of the K computer.



A scene from a lecture meeting on mu2lib

We released a video of the “Multi-scale, Multi-physics Heart Simulator UT-Heart”

We created a video describing the “Multi-scale, Multi-physics Heart Simulator UT-Heart” which runs on the K computer and reproduces the actual movements of the heart. With many varied camera techniques as well as music and sound effects, the video depicts the movements of the heart in an easy-to-understand manner.

The UT-Heart aims to reproduce the heart of a person on a computer, which moves physically and physiologically identically with the real heart, for practical application in the medical setting. The K computer calculates the flows of blood, blood pressure and electrocardiogram, enabling physicians to precisely investigate remedies, correctly predict conditions after treatment, virtually practice heart operations, and thus decide the optimum surgical method.

The video shows heart beats, heart muscle contractions and the movements of myosin molecules in sarcomere in an easy-to-understand manner not only for professionals but also for ordinary people. The video has been released on the Web page, Strategic Programs for Innovative Research Field 1 “Supercomputational Life Science (SCLS).” Please visit the page.



- Dates of release : July 11, 2014 (Japanese version), October 20, 2014 (English version)

SCLS issued a brochure on its research

SCLS is working on four themes: “Simulations of biomolecules under cellular environments,” “Simulation applicable to drug design,” “Hierarchical integrated simulation for predictive medicine,” and “Large-scale analysis of life data.” The brochure introduces researchers and their research, trying to convey the appeal of computational life science.



RIKEN KOBE Campus Open Day (Kobe: RIKEN Advanced Institute for Computational Science)

RIKEN KOBE Campus Open Day was held on Saturday October 25, 2014. At “Kagaku-no-Hiroba” open to the public, five research institutes that are using the K computer exhibited their research in various ways. At the exhibition hall, panels were exhibited describing the research. Visitors of all ages, from children to the elderly, enjoyed making paper crafts shown in the illustration, which helped them understand the hierarchical system of our body.

Besides the exhibitions at “Kagaku-no-Hiroba,” there were tours to see the K computer and brief lectures describing the research. About 2,500 visited the institute on the day.



Bio Japan 2014 World Business Forum

Bio Japan 2014 World Business Forum was held on October 15 (Wednesday) to 17 (Friday), 2014, at Pacifico Yokoyama. SCLS exhibited panels at RIKEN’s booth also this year and showed results of simulations applicable to drug design and other activities.

Information

A short course on GHOST-MP (a parallel sequence similarity search tool) and a drill using the SCLS supercomputer system (compatible with the K computer) will be held.

- Date and time : March 20 (Friday), 2015, 1:00 p.m. - 4:30 p.m.
- Location : 8F, AIST Tokyo Waterfront Annex
- Number of trainees invited : 10

GHOST-MP is a software that parallelizes the GHOSTX’s search algorithm, which is a rapid sequence similarity search tool, for a distributed memory system. A vast amount of short base or amino-acid sequences obtained by a next-generation sequencer, etc., can be searched for sequence similarity to those in an amino-acid sequence database in a short period of time. The software is useful for metagenome analysis, etc. The short course will explain the structure and methods of using GHOST-MP, and participants will practice metagenome data analysis by using GHOST-MP in the drill. For inquiries, please contact: scls-mado@riken.jp.



Strategic Programs for Innovative Research Field 1 Supercomputational Life Science

SPiRE (Strategic Programs for Innovative Research) is a MEXT program aiming to produce ground-breaking results in computer science technology by maximizing the benefits of the HPCI (High Performance Computing Infrastructure) centered on the supercomputer K computer, and encouraging developments in five research fields that need to be strategically addressed.

“Supercomputational Life Science” has conducted research with the mission of understanding and predicting life phenomena based on large-scale simulation and advanced data analysis, and to apply these results to design and implement medicine and medical care with its research.