Next-Generation Integrated Simulation of Living Matter

ISLiM
Integrated Simulation of Living Matter

BioSupercomputing Newsletter

Vol.3 2010.12

Participants in the summer school 2010 for the Integrated Simulation of Living Matter
(See pages 10 and 11 of the text)

CONTENTS

The role of supercomputers is important for integrating various leading-edge research bases for proactive use.

Manager
Clinical Development Planning and Management
Mochida Pharmaceutical Co., Ltd.
Visiting Professor Tohoku University
Kazumi Nishijima

The success rate for development of new drugs is only one part per 25 thousand.

From the standpoint of promoting research and development in the pharmaceutical industry, Mr. Nishijima has been involved in various leading-edge research bases.

● Nishijima (dispense with the Mr.) Firstly, I was involved in Spring-8. In order to make shared use of exclusive beam line (BL) by the Japan Pharmaceutical Manufacturers Association (JPMA), a consortium for the analysis of protein structure by the JPMA was launched by 22 companies (19 companies due to merger, etc., at present). All the Japanese pharmaceutical companies are among the medium-sized businesses in the world. It is quite difficult for one such company to construct an exclusive BL in a world's top level public institution. Therefore, the industry attempted to launch the consortium for use by all the companies in this industry. Thereafter, moreover, many drug discovery companies also came to use the NMR (nuclear magnetic resonance) facility of RIKEN (Yokohama Laboratory), J-PARC (Japan Proton Accelerator Research Complex) in Tokai Village and the radiation light facility PF of the High Energy Accelerator Research Organization (Tsukuba City), and 8 companies joining the consortium have already sent samples to the experimental module "Kibo" of the International Space Station. Finally, the ten quadrillion speed computer "Kei,"and the XFEL (X-ray free energy laser) in Harima Science Park City (Hyogo Prefecture) were promoted in the fourth Basic Program for Science and Technology, and we will also promote an approach to use these facilities for drug discovery in the all-Japan organization.

Proactive utilization of such leading-edge research bases will have an important significance for the pharmaceutical industry, won't it?

● Nishijima Absolutely yes. According to the latest data of the JPMA, there were about 610 thousand types of synthetic (extracted) compounds produced for drug discovery in Japan in 5 years from 2004 to 2008. Among them, there were only 24 compounds approved as new drugs. They are so few in 5 years. The probability is 1 : 25,482. It is just like a miracle. The expenditure of research and development for a new drug is about 70 billion yen per new drug. I think there are less than 10 pharmaceutical companies in Japan that can do the development alone from start to finish. This is because it is difficult to secure the development cost, so there has been an increase in cases in which the development project was changed into a joint development with another company in the middle of the project. Even if the profit is less, the most important purpose is to launch the drug. About 70% of this development cost is spent in clinical studies and the stage thereafter. The cost for nonclinical studies is about 30% , that is, the budget that we aim at exploratory drug discovery utilizing leading-edge research bases accounts for about 30% of the whole. Of course, the assertiveness of divisions involved in clinical studies inside and outside the company may become bigger. However, I think that the part in which corporate efforts are sufficiently reflected is before clinical studies, that is, the upstream part of drug discovery. For example, in the structural analysis and screening test of proteins to discover new compounds, the parts described as searching sources are very important. The upstream flow is very weak despite being important, but since we are doing research and development while imagining that this flow will run into the ocean, we have a high motivation and are very proactive, although money is limited, so there is a considerable possibility that the flow will dry up in mid-course. Even so, it is necessary to address the issue of drug discovery while promoting efficient searching of candidate compounds for development. Therefore, proactive utilization of leading-edge research bases is required.

For overcoming the disease of a low degree of contribution of drugs

Is the method of creating new drugs changing from day to day?

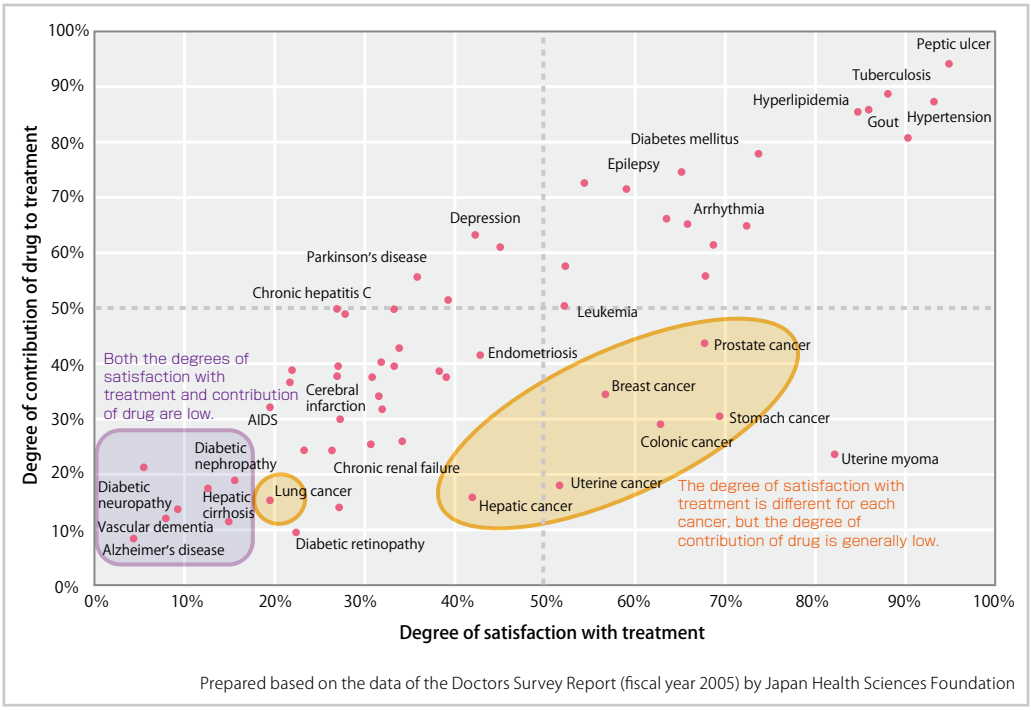
● Nishijima In Japan, it has largely changed from the time when the consortium for structural analysis of proteins was established by the JPMA. Until then, drug discovery consisted of first searching the cause of the symptoms of a disease, but the target was completely based on guesswork. We conducted so-called random screening of a lot of compounds and natural products thinking they might be related to diseases. We searched for new drugs in this manner. The difference of new drug discovery is as follows. First, it became possible to identify receptors and enzymes, the targets of a drug, by structural and functional analysis instead of guesswork, and searching for new drugs can be carried out in a rational manner. It also became possible to adjust the expression of a protein based on genome information, and to verify whether or not it is involved in a disease or is the target of a drug. Therefore, it became possible to design new drugs to treat disease in a rational way. By using molecular imaging, moreover, it is also possible to confirm the action and therapeutic effect of a drug in the body. It can be said that the time has changed from the era of symptomatic drug discovery to that of rational drug discovery aimed at fundamental treatment. Therefore, it is essential to carry out rational drug discovery by integration of leading-edge research bases.

For what diseases will rational drug discovery be promoted in the future?

● Nishijima Here is data summarizing the degree of satisfaction with the treatment of various diseases, and the degree of contribution of drugs to the treatment (Figure 1). This data shows that diseases treated very well with surgery and drugs have already been differentiated from diseases which heal poorly and for which there is no good drug. It can be understood that there have already been good drugs for hypertension, hyperlipidemia, peptic ulcers, tuberculosis, etc. On the other hand, although there are various drugs for dementia, complicated diseases like diabetes mellitus, various cancers, etc., the degree of satisfaction is very low. For example, this is because we barely inhibit the progress of dementia and cannot cure it. For diabetes mellitus, moreover, treatment must be continued indefinitely once it occurs. That is, since we cannot heal them, it is natural that drugs do not contribute essentially. Therefore, rational drug discovery in the future will be performed for overcoming diseases in which the degree of contribution of drugs is low.

In the molecular imaging research strategy promotion program (Phase 2 program) of the Ministry of Education, Culture, Sports, Science and Technology, it was shown that we will address the 2 fields of refractory cancers and dementia.

● Nishijima This is the first reason for specialization in cancer and



Prepared based on the data of the Doctors Survey Report (fiscal year 2005) by Japan Health Sciences Foundation

Figure 1: Degrees of satisfaction with treatment and contribution of drug for various diseases

dementia. In other words, the target site of the drug is not clear in these diseases. In order to overcome this, it is necessary to implement drug discovery by integrating leading-edge research bases instead of conventional methods. For example, this requires a breakthrough in brain science, deepening and integration of the understanding of the brain and nervous system, and progress in nerve and brain cell simulation utilizing supercomputers and the contribution of molecular imaging to diagnosis and drug discovery. We believe that all citizens want to introduce leading-edge research bases into this field for which doctors, patients and pharmaceutical companies need assistance. Moreover, it can be said that this is just where the performance of the ten quadrillion speed computer "Kei" comes in. The mindset of the general public is more important than the pull of future science technology or being top in the world. For example, when you or your family members undergo examination, it takes 2 weeks until the result of diagnosis is obtained by precise image processing, a huge amount of data processing and comparisons, etc. with a conventional computer. If the ten quadrillion speed computer is used, however, it may take 30 minutes to obtain the result. If you are asked which do you choose, anxiously waiting out the 2 weeks or acquiring the result quickly on site, everyone would choose getting the result quickly. I think that the general public will understand that this needs money. In a brain science such as dementia, moreover, the dimension of the drug discovery process is different from lowering blood pressure or killing bacteria invading the body. Because of the problem of recognition, all the problems cannot be tackled by using experimental animals, neither is it easy to open the human brain to try drugs. If so, there is no other way than simulation utilizing supercomputers. As an extreme case, if you were asked, "May I open your head or may I use a supercomputer?", your answer would be obvious (laughing). Actually, there is a field such as dementia, in which it is necessary to perform simulation.

It is important to cultivate human resources to open up the next generation.

Are high-level calculation tools such as supercomputer very important for the practice of rational drug discovery carried forward by the pharmaceutical industry?

● Nishijima If Spring-8 is used, for example, the amount of data increases 10- or 20-times in comparison with the X-ray analysis which was done by our company. The more precise, the clearer the image, so the amount of data does not decrease but increases. Then, we may of course have to consider choices like upgrading the computers in our company, use of supercomputers owned by national research organizations, or cooperation with a university which can use a supercomputer.

For the pharmaceutical industry, the 2 main reasons for using a supercomputer are processing of a huge amount of data and simulation as described previously.

● Nishijima Of course, simulation of the structure of proteins, drug design and disease models will become more important in the future. As concerns diseases with a low degree of satisfaction with the treatment and contribution of drugs, such as dementia, complicated diseases like diabetes mellitus and some cancers, described previously, in particular, research by simulation may become important. Moreover, the processing speed will also be important. A demand for the fastest computer, if possible, is common in government, industry and academia. If possible, moreover, it is important to develop an environment in which other leading-edge research bases and supercomputers are integrated, and everything can be used together. If the user can use a supercomputer without recognizing it as a supercomputer, it may be better. For example, it will be wonderful if an environment could be developed so that data can be collected with SPring-8, a structural and functional analysis is immediately performed with a ten quadrillion speed computer, and virtual screening is also carried out. This is because future drug discovery is not possible with individual leading-edge research bases alone. It is necessary to address this problem by bringing everything together nationwide. In the future, there is no doubt that supercomputers may be involved in the most important part. Since the data obtained in leading-edge research are all precise, the high processing performance of supercomputers is required to utilize the data

Finally, please give us your opinion about operation of the ten quadrillion speed computer after its completion, if any.

● Nishijima After looking at various leading-edge research bases and their operations so far, the matter I feel the most important is cultivation of human resources. I feel that Japanese facilities have no room to expand. We are good at making facilities themselves, but after starting operation, the budget to expand and improve it is not available. Therefore, cultivation of human resources there is not well implemented. I think that the facilities need room for minimal maintainence, as well as for future evolution. To be specific, we need to cultivate young people more for the next generation. For example, in choosing research projects, we should consider cultivating people who will develop the next generation by creating a fund for young researchers in their 20's and 30's, instead of selecting glamorous projects. As concerns the present ten quadrillion speed computer, it will probably be difficult to develop the next generation computer after this with the same way of thinking. A new breakthrough is required. There is no doubt that young researchers in their 20's or 30's may achieve it. We should consider how to cultivate such human resources.

Sonic simulation research in the body which is essential for promotion of ultrasound therapy and development of therapeutic apparatus

Extraordinary researcher, Department of Mechanical Engineering,
School of Engineering, University of Tokyo
Akira Sasaki

In order to promote the HIFU treatment being led by Japan

Dr. Sasaki, you have been in the industrial world up until very recently.

● Sasaki (dispensed with Mr.) I was engaged in the design and development of an ultrasound diagnostic system in the company called as Hitachi Medical Corporation for a long time. A digital ultrasound diagnostic system which appeared in the latter half of the 1980s has evolved at a rapid pace since the mid-1990s, it became possible to provide more information for clinical diagnosis, and when its performance improved, its scope became very large. About 5 years ago, moreover, ultrasound waves came to be utilized not only for diagnosis but also for treatment as a global movement, and because we have to do it in Japan, I was engaged in the development of an ultrasound therapy system. I retired from the company at the mandatory age in March 2010, joined the university in June, and have been addressing the development of HIFU (High Intensity Focused Ultrasound) using ultrasound waves. I think that my work is to play the role of interface between "industry" and "the academic sector" by utilizing my previous experience in the company.

What is the HIFU therapeutic system?

● Sasaki HIFU is a system focusing ultrasound waves on one point in the body and performing treatment by making only the focused part high temperature, which is also called a "high Intensity therapeutic ultrasound" For example, the present cancer treatment consists of drugs and surgery, both of which are associated with severe suffering and stress. However, HIFU does not need to cut the body and causes no suffering as with anticancer drugs. If a patient undergoes treatment in the morning, he/she can return to home in the afternoon. Moreover, radiation therapy can be done only once, but HIFU treatment has the advantage that it can be repeated many times. The research ongoing now is a DDS (Drug Delivery System)-type treatment using ultrasound waves. This is a treatment focusing a micelle containing a drug on a site such as cancer through a blood vessel, and breaking the particle with faint ultrasound waves to perform topical treatment. As described, the HIFU therapeutic system has greater potential. However, at present, Japanese companies are doing the research, but not commercialization or manufacture.

Why are Japanese manufacturers inactive?

● Sasaki The problem of ensuring safety is a major bottleneck. Considering the liability issue in failure of treatment, manufacturers cannot start development so easily. In the case of a drug, it is OK if suitable effects can be obtained considering the risk and benefit ratio but in the case of a treatment, failure is not permitted, and issues of liability and compensation may occur. Another issue concerns the market. An ultrasound diagnostic system is introduced to a lot of medical institutions and has good marketability, but the market for treatment is limited, and the cost will be high. Moreover, a therapeutic apparatus has to undergo clinical evaluation. Considering the cost, period and manpower involved in a clinical trial, it may be particularly unprofitable for the manufacturer. Considering the global movement, however, an Israeli company has already developed this system in about 2002, which was approved by the FDA in 2004 and launched. In the U.S.A., the GE company is dealing with it. Philips will launch it in 2011. Siemens is developing a prototype and will start clinical trials in 2011. In order to establish the field of HIFU therapy, moreover, the Focused Ultrasound Surgery Foundation was already established in Western countries in 2006, and research and development have begun. If Japan does not start development despite the fact that 3 of the world's leading companies are moving ahead with manufacture by shifting from diagnosis to treatment, 100 % of the market will be occupied by overseas companies. Because we have to promote HIFU therapy from Japan in some way, it was decided to establish a consortium in Japan. In order that the Japanese HIFU therapeutic system survives in the future, it is important to build a strategy for Japan to win the world. This is not a time when one domestic manufacturer can win. To improve our chances of success, we will work together in an all-Japan organization.

Making a bridge between "industry" and the "academic sector" is my job.

What is the first issue in developing the HIFU therapeutic system from Japan?

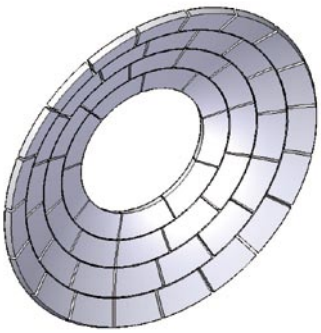
● Sasaki More than anything else, we have to emphasize the safety of treatment. HIFU therapy is a method to directly treat only a target site in the body. In order to perform treatment safely, precise imaging diagnosis, navigation and biological simulation are important keys. First, in order to draw up a solid therapeutic plan, a precise preoperational imaging diagnosis should be made by making effective use of CT, MRI, US (ultrasound diagnostic system) and RI (nuclear medicine diagnosis system). It is required to observe in real time during operation, and to conduct appropriate guided navigation. Unlike surgery which directly opens the body, moreover, this HIFU therapy only looks at images. The efficacy of treatment is also evaluated with images. In order to observe the progress of treatment, a biological simulation is therefore essentially required.

Does a safety issue still remain in the HIFU therapy itself?

● Sasaki Ultrasound waves are characterized in that refraction and reflection occur due to acoustic impedance (density× sound velocity). Like light, ultrasound waves bend when they enter a different medium. Therefore, refraction and a change in sound velocity with the fats, muscles and organs in the body affect the beam characteristics of ultrasound waves and displacement of the focal point. In the bone, moreover, thermal changes occur due to absorption of ultrasound waves, and heat diffusion may also occur at a blood flow-rich site. Moreover, ultrasound waves are influenced in the body by various factors such as movement of organs such as the heart and presence of air bubbles such as intestinal gas which induce phase disturbance or sound disturbance. It is never easy. In order to avoid these influences in the therapeutic plan, it is essential to measure the hurdles exactly, and to perform optimum control. For this, it is necessary to simulate the sound characteristics in the body in advance.

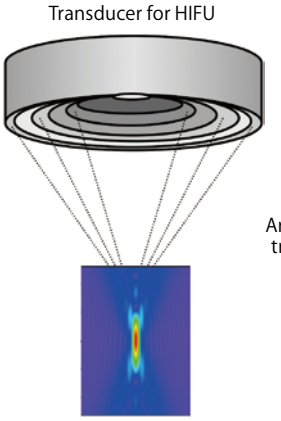
Does the result of such a biological simulation assist the design of HIFU therapy?

● Sasaki One way is optimization design of the ultrasound beam including the transducer generating ultrasound waves. In order to irradiate so that the ultrasound waves are precisely focused on the target site, the processes of penetration through the skin on the body surface, muscles, fats and organs will be important. Moreover, procedures for avoiding intestinal

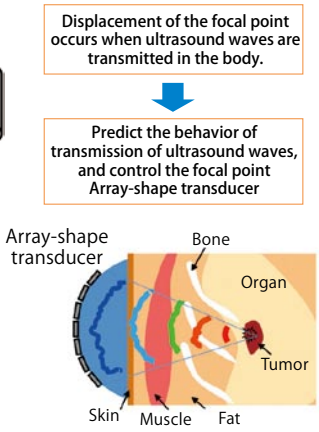


An example of a HIFU transducer of concave 56 channel array type

- Coagulation therapy at a target site at 60 to 90℃ in the focused region by increasing the intensity of ultrasound waves
- Annular Array form: Enlargement of the region of coagulation with Dynamic Focus



Beam pattern
in the focused region



An example of a method to control the focal point in the body by simulation

gas and bones which may become large hurdles are also required. In a transducer consisting of multiple elements, each element is limited so as to avoid irradiation to the place where bones and air bubbles are present in advance, and refractive correction is performed by giving the delay time. By these procedures, a method of irradiating ultrasound waves precisely to the target site has been studied. Moreover, a method to shorten the treatment time has also been studied by predicting the increase in temperature at the coagulation site, and by irradiating while continuously changing the irradiation position. In conventional HIFU therapy, some time to cool off heat after irradiation was required, and it took 2 to 4 hours for treatment, but this is an attempt to shorten the treatment time largely by continuous coagulation while moving. In optimization design of HIFU coagulation, optimization of multi-cauterizing design to shorten the treatment time, and minimization of side effects, simulation will play an important role.

— In order to establish a HIFU therapy system after overcoming various issues, is biological simulation still important?

● Sasaki If navigation and simulation are done correctly, the safety of HIFU therapy will increase. I want to make treatment safe by conducting experiments based on this. In Europe, research has already progressed to an attempt of treating brain tumors through the cranial bone. By calculating the refractive index of cranial bone, they are attempting to focus ultrasound waves on a point in the brain. It is wonderful that treatment can be performed without opening the brain. The idea is wonderful. I think it is very difficult to realize it, but simulation will become truly important in such research.

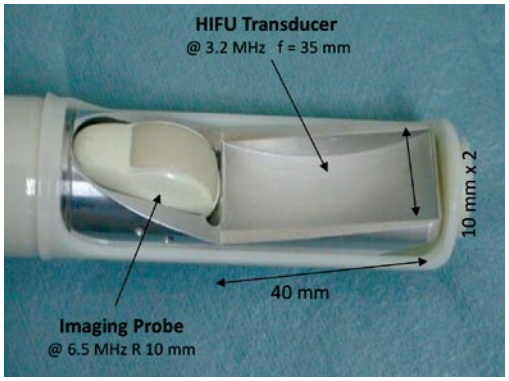
— In promoting the technical development of such HIFU therapy, do you think that computational resources such as a ten quadrillion speed computer will make a significant contribution?

● Sasaki It is absolutely necessary, and I expect it. In the Organ and Body Scale Team, development of a focused ultrasound simulator is actually ongoing, and I would be a fool not to take advantage of this (laughing). As described previously, since this treatment can be observed only on images, a simulation is required for monitoring processes. I think that the research itself cannot be done without simulation.

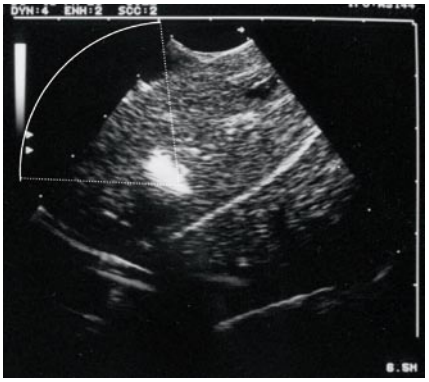
— If the safety of HIFU therapy increases along with the progress of simulation research, will the Japanese manufacturer go on to develop or manufacture a HIFU therapy system?

● Sasaki Yes, it will. Considering the global trend, the time for development and manufacture has come. For this, if the "academic sector" researches what degree of safety has been established, "industry" will take action, and I think that my mission is to make this possible.

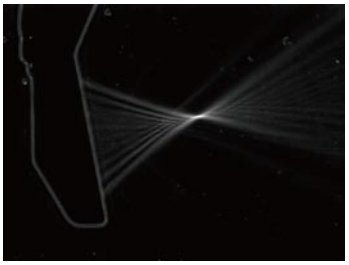
HIFU Transducer for In Vivo Experiment



Example of HIFU Transducer for Prostate



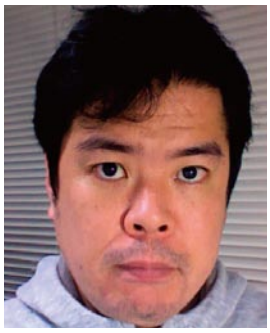
B-mode Image of HIFU Treatment



HIFU Beam Profile

Achievement of a Multiscale Molecule Simulation of QM/MD/CGM

Institute for Protein Research, Osaka University
(From the above) Yasushige Yonezawa, Shusuke Yamanaka, Hiromitsu Shimoyama, Hideki Yamazaki and Haruki Nakamura
RIKEN, Computational Science Research Program
Ikuo Fukuda



Our group is attempting to elucidate the mechanism of functions of biological macromolecules, proteins and nucleic acids, at multiple levels including molecular, atomic and electronic levels using molecular simulation. Actual molecules in living matter show thermally fluctuating actions at room temperature, and the fluctuation has the essential roles on molecular functions. Our purpose is to elucidate the mechanism of such biological molecules including their electronic state while simulating the thermal fluctuation in details.

For this achievement, we have developed a multiscale simulation program, *Platypus* (*PLATform for dYnamic Protein Unified Simulation*), which couples quantum mechanics (QM) with molecular dynamics (MD) which were developed individually. Quantum mechanics to calculate electronic state is a computation method essential for handling enzymatic functions in simulation, and here, computation by the Hartree-Fock (HF) method, the density functional theory (DFT), CASSCF and CIS are available. Molecular dynamics is a simulation method, which reproduces the thermal fluctuation of molecules. *Platypus* is an integral of the program which realizes massive parallelism with the original codes, which has accelerated performance of up to 8192 parallels in the computations of both HF and DFT in the development so far, and particularly, it shows good scalability up to 1000 to 2000 parallels. In addition, *Platypus* loads the MD computation with an independently developed coarse graining model (CGM), with which the multiscale molecule simulation of QM/MD/CGM can be implemented. Moreover, we are developing an algorithm which performs efficient sampling with the QM/MD computation.

As one of the applications of *Platypus*, we introduce research on the cis-trans isomerization mechanism of the proline residue, which is related to the signal molecular control of various life activities. In a study of small peptides including proline in water (reference 1), the transition state between *cis* and *trans* forms the conventionally known pyramid as well as the inverse pyramid, and it was firstly shown that this state fluctuates between these 2 forms (Figure 1). From a study of the Pin1 isomerase, which accelerates the cis-trans isomerization of the proline residue, we succeeded in grasping the mechanism, in which the strain occurring by binding of the peptide including proline to the enzyme active site stabilizes the energy in the transition state so as to accelerate isomerization.

In the molecular dynamics simulation, the time for computation of interatomic long-range forces not based on chemical bonds between molecules accounts for a large amount. By parallelizing this part, the computation time can be considerably shortened. The Ewald method is a general method most often used, but it is known that this Ewald method is not suitable for parallelization of a very large system. We are therefore pursuing research to revise the long-range force potential recently developed by Wolf et al., and develop our original Force Switching-Wolf method (FSw-Wolf method), by which a consistent and steady simulation can be performed instead of the Ewald method. We succeeded in showing by simulation of a fused sodium salt (reference 2 and Figure 2) and a short peptide in water that the FSw-Wolf method can realize the computation at a precision comparable to that of the Ewald method. The FSw-Wolf method is not only a very convenient algorithm, but also it considers only the force between atoms at a relatively short distance. Thus, it is very suitable for large scale parallelized molecular dynamics simulation. Moreover, it is considered that this method will become a key method which can realize high-speed parallel computation by combining with other algorithms.

Reference 1. Yonezawa Y., Nakata K., Sakakura K., Takada T., Nakamura H., *J. Am. Chem. Soc.* 131(12), 4535-4540.
Reference 2. Fukuda I., Yonezawa Y., Nakamura H., *J. Phys. Soc. Jpn.* 77(11), 114301, 2008.

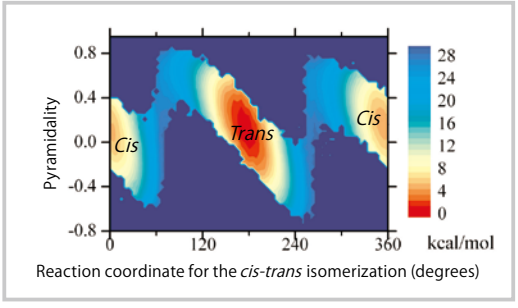


Figure1: Free energy landscape of *cis-trans* isomerization of a peptide including a proline residue, which was elucidated by QM/MD multiscale simulation. A transition state largely fluctuating between cis and states is observed.

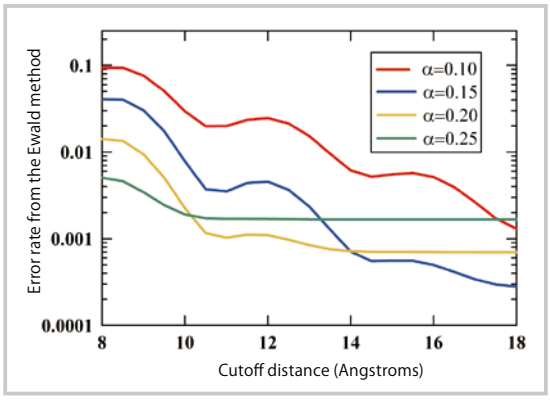


Figure 2: The energy error between the FSw-Wolf method and the Ewald method, plotted as a function of the cutoff distance and the parameter α in the FSw-Wolf method. This shows that the FSw-Wolf method can reproduce an energy very close to that in the Ewald method.

Towards development and experimental demonstration of liver model based on large-scale metabolic simulation at individual cellular level

School of Medicine, Keio University
Ayako Yachie-Kinoshita

Hepatic parenchymal cells (hepatocytes) which occupy a large part of the cellular volume in liver have very extensive and complex metabolic pathways including most metabolic reactions occurring in the body, and play an important role in regulation of vast variety of small molecules. Compartmentation of metabolism by site specificity, in which an enzyme expression pattern is different between the upstream (portal vein) and downstream (central vein) of sinosidal blood flow, is considered to be one of the factors that makes the functional switching of metabolism in liver efficient. However, due to the striking complexity in metabolic dynamics in which many metabolic functions are intricately involved with each other, it is very difficult to comprehensively understand the dynamics of hepatic metabolism and its underlying mechanism, and for many phenomena, analogies are still being drawn from circumstantial evidence. In such a case, model-based analysis and simulation by integration of independent kinetic characteristics is very powerful to understand the system as a whole.

We therefore established an unprecedented large-scale metabolic model including approximately 500 defined substances and about 250 dynamic reactions using the E-Cell Simulation Environment[Note1] based on the reaction kinetics obtained by literature search. With this model, the dynamics of the metabolism in the mitochondria and cytoplasm can be predicted separately and simultaneously(considering the compartmentation of metabolism in the cells), and the effects of exchange of small molecules between hepatocytes in portal vain and central vein regions via a gap junction can be predicted by changing the enzyme expression pattern to set up the regional heterogeneity of the both regions (considering the compartmentation of metabolism between cells). We are also developing the model to reproduce the condition in which hepatocytes with different metabolic properties interlock each other to function by using the Riken Cell Simulator (RICS)[Note 2] (Figure 1).

Using the large-scale metabolic model of hepatocyte, we compared a predicted dynamics under hypoxic and hypoglycemic condition from a couplet model in which two connecting hepatocytes having portal vein and central vein metabolic properties with those from the models in which both hepatocytes have a property of portal vein or central vein only. As a result, by coupling cells with different properties, over-all changes in metabolites and a decrease in ATP energy in response to these stimulations were suppressed, and it was predicted that the robustness of the metabolism would be improved at an organ level.

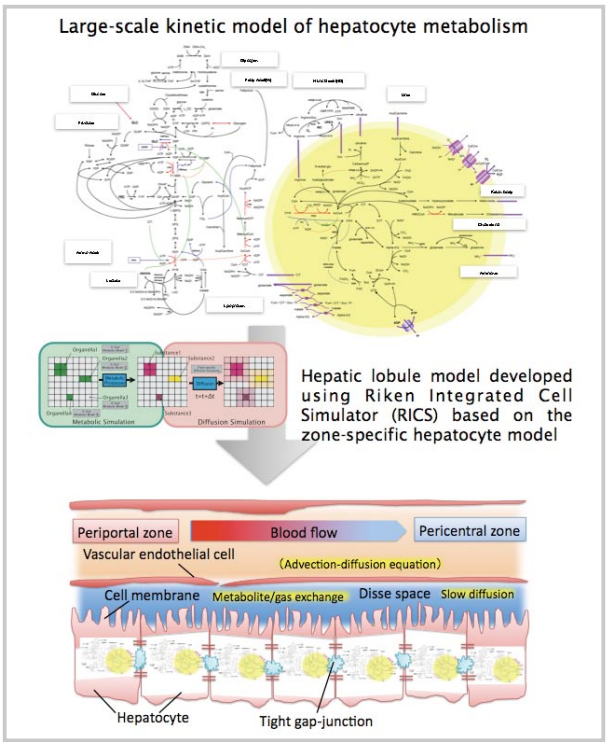


Figure1 : Expansion from exhaustive hepatocellular metabolism model to liver model

This result may contribute not only to understanding how hepatic metabolism is optimized using heterogeneity, but also to understanding of metabolic failure in the intercellular material exchange disorder such as a case of hepatic cirrhosis. Moreover, the *in vivo* experiments in our laboratory have clearly showed that the metabolism changed largely not only in metastatic foci but also the surrounding hepatocytes due to hepatic metastasis of colon cancer. Our hepatic metabolism simulation model can be applied to reproduction and discussion of such a disease model, and prediction of resistance to hepatic ischemia during surgery and spare ability at hepatectomy.

In considering the effects of cell metabolism on organs and the whole body, particularly pathological conditions, on the other hand, it may be essential to establish an experimental system to observe the metabolic conditions without destroying the tissue structure.

Our laboratory therefore focus on an experimental technique of "visualization of metabolism with μm resolution" combining a metabolome analysis with an imaging mass spectrometry[Note3] (Hattori K et al., Antioxid Redox Signal. 13(8): 1157-67 (2010)). At present, improvement of this method is moving ahead with great urgency, and it has become possible to determine more metabolites including aminoacids at higher resolution. By this technique, it may be possible to measure and quantify the difference in metabolism between portal vein and central vein hepatocytes with the *in vivo* spatial arrangement (Figure 2).

In the future, the study to quantify and observe metabolism as the *in vivo* spatial arrangement of tissue is maintained, and to elucidate the results and the meaning of material variation by exhaustive cell simulation (large-scale calculation with a supercomputer) considering spatial information will become important.

Note 1: Simulation environment focuses on phenomena of intracellular molecules, which has been being developed by Institute for Advanced Biosciences of Keio University, and RIKEN.
Note 2: A platform of cell simulation, which is being developed by RIKEN.
The reactions of small molecules considering the spatial information in the cell and the position among cells can be simulated by constructing the cellular model in the space (voxel) of a fixed grid, and establishing diffusion and transfer among voxels.
Note 3: The mass distribution is visualized at a resolution of 10 μm by ionizing metabolites by applying a laser to a section of frozen tissue at fine intervals, by measuring ions with a mass spectrometer connected to it, and then by reconstructing information on masses of the metabolites determined at the respective positions.

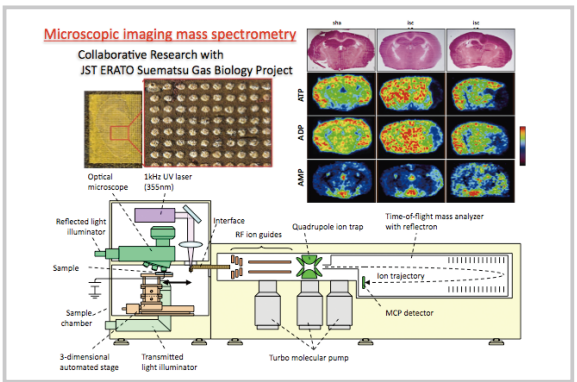


Figure2 : "Visualization of metabolism" using imaging mass spectrometry and metabolome analysis

Exhaustive Protein-Protein Interaction Network Prediction by Using MEGADOCK

Graduate School of Information Science and Engineering,
Tokyo Institute of Technology
(From left, above) Yutaka Akiyama, Yuri Matsuzaki,
Nobuyuki Uchikoga and Masahito Ohue

We are working on prediction of the protein-protein interaction (PPI), one of the important problems of systems biology, by the method of bioinformatics and parallel processing (Figure 1). Usually, the expected role of the computational method using physical chemistry on PPI analysis was to examine the configuration and affinity of interactions concerning the known one-to-one protein-protein interaction in detail. We then developed "MEGADOCK," a novel PPI prediction system based on large-scale parallel calculation, and made it possible to predict a candidate pair of PPI exhaustively from a large amount of protein groups. This is expected to contribute to the discovery of new PPI by collaboration with experiments in the future.

MEGADOCK is a system to predict the presence or absence of interaction using information about the tertiary structure of proteins based on various scores obtained from rigid-body docking. In this calculation, a high-speed evaluation is conducted mainly based on the shape complementarity of the molecular surface without considering the structural change of the protein. We introduced the rPSC (real Pairwise Shape Complementarity) score composed of the terms of shape complementarity and electrostatic interaction assigned to the molecular structures on the voxel space. With a conventional tool, ZDOCK, score is calculated using 3 interactions with 3 complex numbers, whereas in the rPSC, score is calculated using 2 interactions with 1 complex number by expressing the shape complementarity in a real number part and introducing electrostatic interactions into an imaginary number part. The number of three-dimensional fast Fourier transformation (FFT) required for convlolution sum calculation was reduced. When executed by a single CPU, about four times higher calculation speed was achieved with the same precision as ZDOCK.

MEGADOCK is parallelized using the MPI library. When a certain processor was assigned multiple receptor and ligand proteins, one ligand is taken sequentially from the ligand set, transformed by FFT with the certain angle increment, and compared as the innermost loop with all the data in the receptors set. A procedure of making the FFT transformed library concerning known proteins, and read from a hard disk to perform convolution sum calculations was implemented, and a speed increased of up to about 3 times was achieved. With appropriate load balancing, efficient calculation is possible by hundreds of processors or more.

As a benchmarking of MEGADOCK, we firstly performed PPI prediction of 44×44=1,936 combinations on 44 protein complexes. The predicted (red) and native (green) structures were consistent (upper part of Figure 2). In the prediction of PPI pairs, many correct complexes were predicted as shown by the warm color on the diagonal line in the lower part of Figure 2, and a prediction performance (F-measure = 0.415) similar to or higher than in a related study was obtained. As an actual application in systems biology, PPI prediction was performed on the signal transduction pathway of bacterial chemotaxis (89×89=7,921) and the human EGFR signal transduction pathway related to lung cancer (497×497=247,009). Our goal is to perform the calculation of a 1,000×1,000 (mega) class routinely.

References
(1) Matsuzaki Y., Matsuzaki Y., Sato T. and Akiyama Y., *J Bioinform Comput Biol*, 7: 991-1012 (2009).
(2) Ohue, Matsuzaki, Matsuzaki, Sato and Akiyama. IPSJ Transactions on Mathematical Modeling and its Applications (TOM), 3(3): 91-106 (2010).

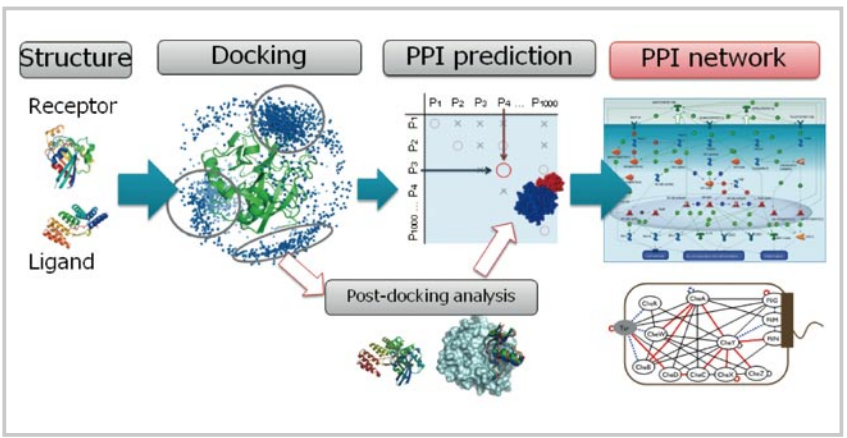


Figure 1: Estimation of the protein interaction network based on PPI prediction

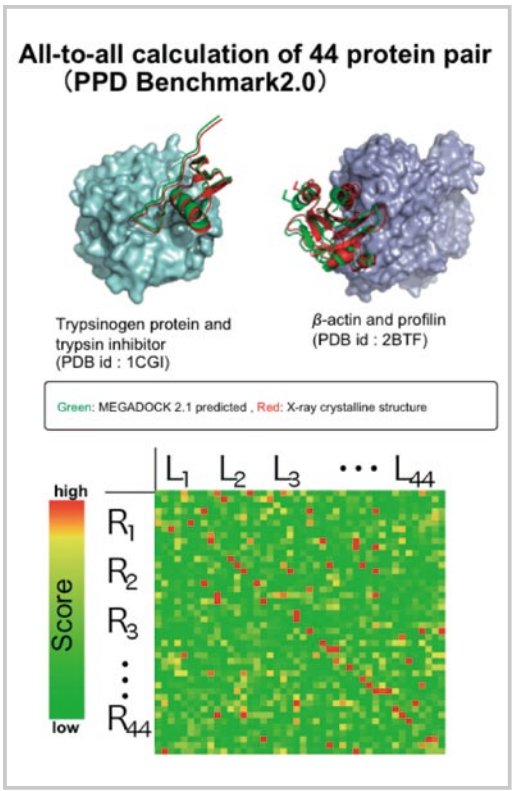


Figure 2: Result of PPI prediction (44×44 benchmark)

Whole Brain Simulation of the Insect Olfactory System

Research Center for Advanced Science and Technology,
The University of Tokyo

Tomoki Kazawa, Stephan Shuichi Haupt

Nerve cells (neurons) are the building blocks necessary for the emergence of intelligence as they process internal and external information, generally encoded in sequences of spikes and form networks in which such information is relayed between them. Input and output parameters of neurons such as gain can change dynamically depending on prior activity history. Neurons are present in all metazoan animals from simple coelenterates like Hydra and have the same general properties across the animal kingdom. Though the basic building blocks are so similar, there is a big difference in the size of nervous systems, human brains harbour $10^{11\text{-}12}$ neurons, insects have $10^{5\text{-}6}$ brain neurons, and the somatic nervous system of the nematode Caenorhabditis elegans contains just 302 neurons. Insects are sufficiently complex to display many sophisticated behaviours, such as learning floral odors and colors. Multimodal information is integrated in the insect brain which can generate a wide variety of complex and plastic behaviours, making insect brains a highly suitable system to investigate the emergence of intelligence. Our approach to understanding such an intelligent system shaped by millions of years of evolution is to attempt to recreate it through simulation. This insect whole brain simulation has to encompass information processing from sensory input to behavioural ouput. We currently focus on the odour-source localisation behaviour of silkmoth *(Bombyx mori)* males as a convenient model system. We are analyzing the brain of the silkmoth accumulating numerous data obtained with a large grid of methods including molecular genetics, electrophysiology, imaging, immunohistochemistry, and behavioural experiments. Our methods cover the full scale of observation levels from the molecular and cellular levels to the network and behaviour levels. These multi-scale data are integrated in our database, which, for example, contains data on >1600 neurons individually identified in terms of 3D-morphology and physiology, currently making it the largest single-species individual neuron database worldwide. An important insight from these data was the understanding of global and regional connectivity in the olfactory system in the brain of the silkmoth.

Since insects can exhibit highly complex behaviours despite having a relatively small number of neurons, it is likely that the relative contribution of each individual neuron is more important than for instance in mammals. In our large-scale simulation of insect neural circuits, we are therefore performing a detailed simulation of each individual neuron using multicompartment models (Figure 1). The 3D-morphology of neurons is acquired by confocal microscopy and can then be approximated by sets of connected cylinders. A cylinder per se has a simple equivalent circuit consisting of voltage and impedance to the outside of the cylinder (membrane resistance and capacitance). Cylinders are connected by resistors that correspond to axial resistance in neurites. Inputs are described as chemical synapses having postsynaptic potentials of 5-20ms duration upon the simulated activation by neurotransmitter. The synaptic activation can induce spikes of 1-2ms width that are modeled by the conventional sodium and potassium currents (Figure 1C). By mapping and combining numerous single neuron simulation models in the brain, a neural circuit simulation can be constructed (Figure 2).

We estimate that a total of about 10000 neurons in the brain are involved the whole information processing from reception of olfactory sensory input to the generation of the control activity patterns for odour-source localisation behaviour. Using representations of neurons consisting of about 100 compartments, we estimated that the computations necessary to simulate information processing in the entire insect brain can be performed in real time on a machine with the performance of a petaflop HPC. If a real-time simulation of the neural circuit from sensory input to motor output can be achieved, it also becomes possible to investigate information processing by modification of the circuit properties in ways that are currently not feasible in actual experiments. Such manipulations are of paramount importance to identify minimal neural circuits to perform specific information processing tasks. We expect this approach will have an impact in medical applications, in particular neurorehabilitation and neural protheses.

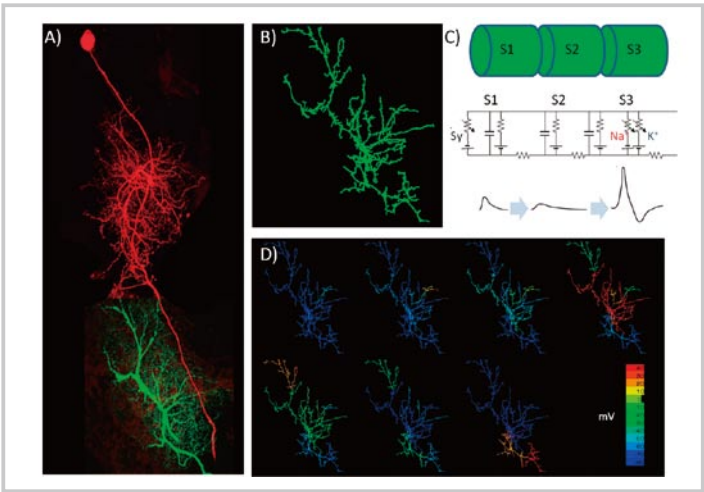


Figure 1 : Simulation with a multicompartment model generated by extraction of neuronal morphology.
A) Double-label image of a GII descending neuron and the Cv1 motor neuron
B) Extracted morphology of the Cv1 motor neuron
C) Multicompartment model and equivalent circuit
D) Single neuron simulation of the Cv1 motor neuron

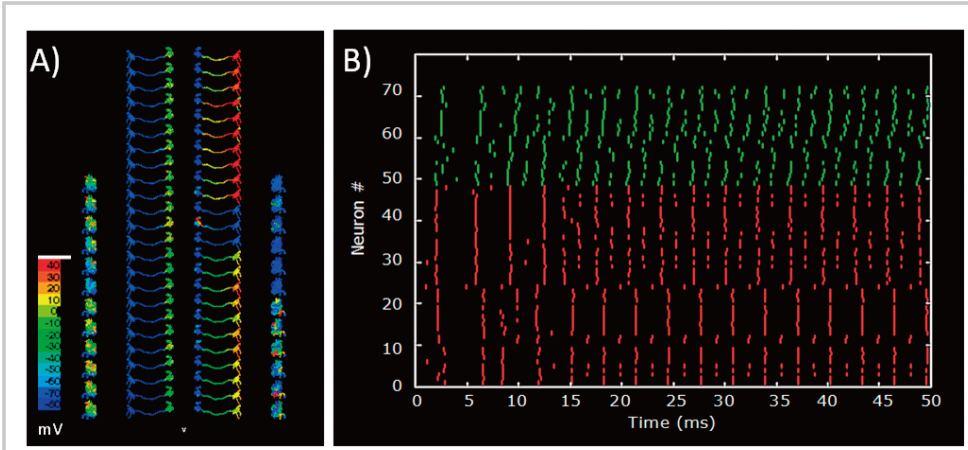


Figure 2 : Neural circuit simulation of the premotor center in the silkmoth brain.
A) An example of the spatial distribution of the membrane potentials of local interneurons (outer side) and bilateral neurons (inner side)
B) Raster plot of spike events of local interneurons (green) and bilateral neurons (red)

The summer school 2010 for the Integrated Simulation of Living Matter was held.

RIKEN, Computational Science Research Program
Yasuhiro Ishimine (Organ and Body Scale WG)
The Institute of Medical Science, The University of Tokyo
Teppei Shimamura (Data Analysis WG)
RIKEN, Computational Science Research Program
Yasuhiro Sunaga (Cell Scale WG)
Kyoto University Graduate School of Informatics
Naoki Honda (Brain and Neural WG)
RIKEN, Computational Science Research Program
Gen Masumoto (High-Performance Computing Team)
RIKEN, Computational Science Research Program
Kosuke Matsunaga (Molecular Scale WG)
(Order of the Japanese syllabary)

At Shonan International Village Center in Hayama Town, Kanagawa Prefecture, the "Summer school 2010 for the Integrated Simulation of Living Matter" was held for 3 days from July 5 to 7, 2010. This summer school was mainly planned and managed by young researchers participating in the project for the next-generation integrated simulation of living matter. A total of 57 members including post-doctoral researchers and graduate school students in related fields, who are participating in this project, presented lectures and exchanged opinions concerning numerical simulation and large-scale data analysis related to life sciences. Here, we will give an overview of the summer school, interweaving responses to questionnaires obtained from participants and details of reports submitted by all participating graduate school students.

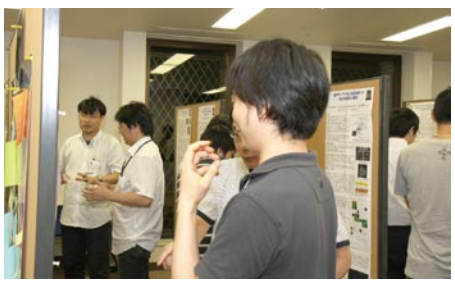
The summer school was basically carried forward in a lecture mode. First, as a "keynote lecture," Koji Kaya, a director of the Computational Science Research Program, gave a summary of the research in which the director himself was engaged, describing his theory concerning the attitude and backbone for research which a first-class researcher should always strive to maintain. Then, the latest research achievements were introduced by a total of 14 researchers representing 6 teams affiliated with the next-generation integrated simulation research promotion group for living matter for 3 days (see Table on p. 11 for details of lecturers and lecture titles). In the afternoon of the second day, a time for general discussion was established, the latest information about the progress and future prospects of the project was provided by Ryutaro Himeno, the sub-director of the Computational Science Research Program, and then a question-and-answer session and exchange of opinions by all the participants was held concerning measures for effective use of supercomputers. Extra time was also provided to exchange information individually, such as for a get-together and poster sessions. The lunch break on the second day was made slightly longer for the same purposes, not only information exchange but also voluntary intercommunion among participants, such as enjoying football outdoors.

According to the results of the questionnaire, the lecturers and content of lectures were basically satisfactory, and impressions were manifested such as "I realized that there were research themes at various levels (from the level one molecule to the level of the tissue) ," "since there were some questions which could not be conceived from an insider's viewpoint - for example, a question about a theory that we ourselves accept implicitly - I could appreciate an outsider's viewpoint," "I found out which project covers the Petacom plan, what people are involved, what I can do using Petacom and what are the problems in using Petacom," and "I understood that, since there may be a large computation effort in performing simulations to understand one factor, next-generation supercomputers have great significance". There were also impressions referring to expansion of future collaborative research, such as "I understood that a lot of researchers have so much in common in that we are dealing with a certain problem of optimization", and "since there are some common areas among teams as far as concerns not only themes but also procedures, I expect that better research would be possible if there were active collaboration".

As for demands for the future, moreover, there were some opinions such as "I want you to hold an invitation lecture of a famous researcher in the field of biology or a person who is doing simulation in the research and development division of a company", and "it may be better to hold it at a meeting like a study session with the people like graduate school students." There were also some unique opinions such as "I want to discuss things in view of what goes on in other fields", and "it might have been interesting if a scientific contest were held in a team session consisting of roommate members in the hotel."

The biggest regret as a secretariat was that there were multiple opinions that we should have paid attention to participation of graduate school students. Some graduate school students said, "I thought it was wonderful that I did not feel any barrier of position and field at all," "it was a wonderful session which was not too rigid or relaxed, and suitable for the title of summer school", and "I heard opinions from people of quite different background, approach, experience and age, who are different from those with whom I usually interact, and this summer school was a good turning point for my future research life," whereas there were some impressions like, "I felt that only a person in a high position or postdoctoral researcher could ask questions", and "I felt atrophied because the participants other than me were postdoctoral researchers." Moreover, there was an impression, "the story about budget and evaluation of project was stimulating for a student who wants to purely pursue science in a negative way." Since there were some opinions like, "individual self-introduction was required," as an answer to the questionnaire, we felt that more significant interchange may be possible by taking time for orientation at the beginning of the summer school to shorten the mental distance between each other and to create a mood to deal with common subjects.

Based on the above experience in this summer school, we expect that a more fulfilling summer school will be held in and after next year.



After participation in the summer school 2010 for the Integrated Simulation of Living Matter

First year of doctor's course,
The University of Tokyo
Graduate School of Science
Ken Saito

I belong to the seminar of Professor Shinya Kuroda of the University of Tokyo, a member of the Brain and Neural System Research and Development team which analyzes intracellular signaling pathways using mathematical models (particularly time-series statistical models). A senior associate in the seminar introduced this summer school to me, and I participated in the summer school because I wanted to meet people in different fields who are related to my research theme, but with whom it is difficult to exchange opinions in my daily research life.

I have been interested in the work of the Data Analysis Fusion Research and Development Team estimating the structure of gene networks using time line data about gene expression, such as SiGN and LiSDAS. Because I have difficulty in appreciating the significance of explanations of signaling pathways using mathematical models, I wanted to hear other opinions about it. When I actually heard the lecture at the summer school, it was impressive that a lot of researchers wanted to apply it in medical practice through modeling. I also heard the story of researchers who have the same directionality as mine, and it was a very valuable experience for me.

Since other lectures handled the theme across various levels from brain and organs to molecules, I was anxious about keeping up with the contents before participation, but actually, the lectures were explained so that the unprofessional people can understand, so I could interestingly hear the lecture of the people far from my specialty to the end. It was particularly good that the presentation topic was explained carefully, especially the motivation for the research and why a supercomputer is required. I had been struggling with the question, "why supercomputers are required for analyzing the functions of organisms," but now I could sweep away that question.

Particularly, the presentation by the High-performance Computing Team was even more impressive than I had expected. Since the mathematical model that I use now does not require the performance of a computer so much, I was not so interested in the extra speed of parallel computation, but I heard the tuning methodology behind this at the summer school, and this prompted me to have more motivation to study it seriously. In the future, I want to describe thousands of molecular behaviors in cells with one model in an integrated manner, and to utilize the knowledge learned at this summer school.

The matter that I expect from the summer school to be held next year and so on is aggressive participation of students. In all, only 8 students including me participated in this summer school. I think that there is no student who does not know of the "supercomputer project" which came up in conversation at the screening process for government revitalization, but I think that there are few students who really understand what is going on in terms of research themes actually using supercomputers. In order to understand the positioning of scientific research as a national project, I thought it would be very meaningful if more students participate in such an event.

Finally, since not only research, but there were also many recreation events like football and night networking, etc., at this summer school, I think that participants could have a real opportunity to discuss and exchange opinions. I would like express my thanks to the people in the Secretariat of the summer school.

Program of the summer school 2010 for the Integrated Simulation of Living Matter

[Day 1] July 5 (Monday)

Opening speech

Koji Kaya (Director of the Integrated Simulation Software of Living Matter Research and Development Program)
"Explanation of purport and keynote lecture"

Yoshinori Tamada (Extraordinary Assistant Professor, Human Genome Center, Institute of Medical Science, The University of Tokyo)
"Estimation of gene network from gene expression data using a Bayesian networks by nonlinear regression and its parallel algorithm"

Keiji Misawa (Researcher, Data Analysis Fusion Team, Computational Science Research Program, RIKEN (The Institute of Physical and Chemical Research))
"ParaHaplo - Program package to quickly implement analysis related to genome-wide haplotype by parallelization"

Yuri Matsuzaki (Industry-Academia-Government collaborative researcher, Tokyo Institute of Technology Graduate School of Information Science and Engineering, Department of Computer Science)
"Development of system for estimation of protein-protein interaction network based on steric structure"

Ryo Yoshida (Assistant Professor, Institute of Statistical Mathematics)
"LiSDAS: Life Science Data Assimilation Systems"

Get-together

Poster session

[Day 2] July 6 (Tuesday)

Jun Igarashi (Special researcher, Brain and Neural Systems Team, Computational Science Research Program, RIKEN (The Institute of Physical and Chemical Research))
"Basis of realistic neural circuit and large-scale computation"

Naoki Honda (Researcher, Kyoto University Graduate School of Informatics)
"Simulation of cytoskeleton-dependent change in cell morphology"

Tomotaka Kasaguchi (Doctor Researcher, Yokohama City University Graduate School of Nanobioscience)
"Approach for experimental procedures from an MD"

Tsutomu Yamane (Extraordinary Assistant Professor, Yokohama City University Graduate School of Nanobioscience)
"Multiple drug discharging transporter AcrB"

Naoyuki Miyashita (Researcher, Molecular Scale Team, Computational Science Research Program, RIKEN (The Institute of Physical and Chemical Research))
"Sampling method applying molecular dynamics – Development of Replica Exchange Interface (REIN) –"

Yosuke Ohno (Senior Researcher, High-performance Computing Team, Computational Research Program, RIKEN (The Institute of Physical and Chemical Research))
"Hardware and tuning"

Hiroshi Koyama (Senior Researcher, High-performance Computing Team, Computational Research Program, RIKEN (The Institute of Physical and Chemical Research))
"The reason why I hate HPC (High Performance Computing)"

General discussion

Poster session

[Day 3] July 7 (Wednesday)

Yasuhiro Sunaga (Researcher, Cell Scale Team, Computational Research Program, RIKEN (The Institute of Physical and Chemical Research))
"Intracellular biochemical reaction simulator considering organelles"

Youhei Nanazawa (Researcher, Department of Cardiovascular Medicine, Tokai University Hospital)
"Development of hemostasis and thrombosis simulator connecting cells with an organ model"

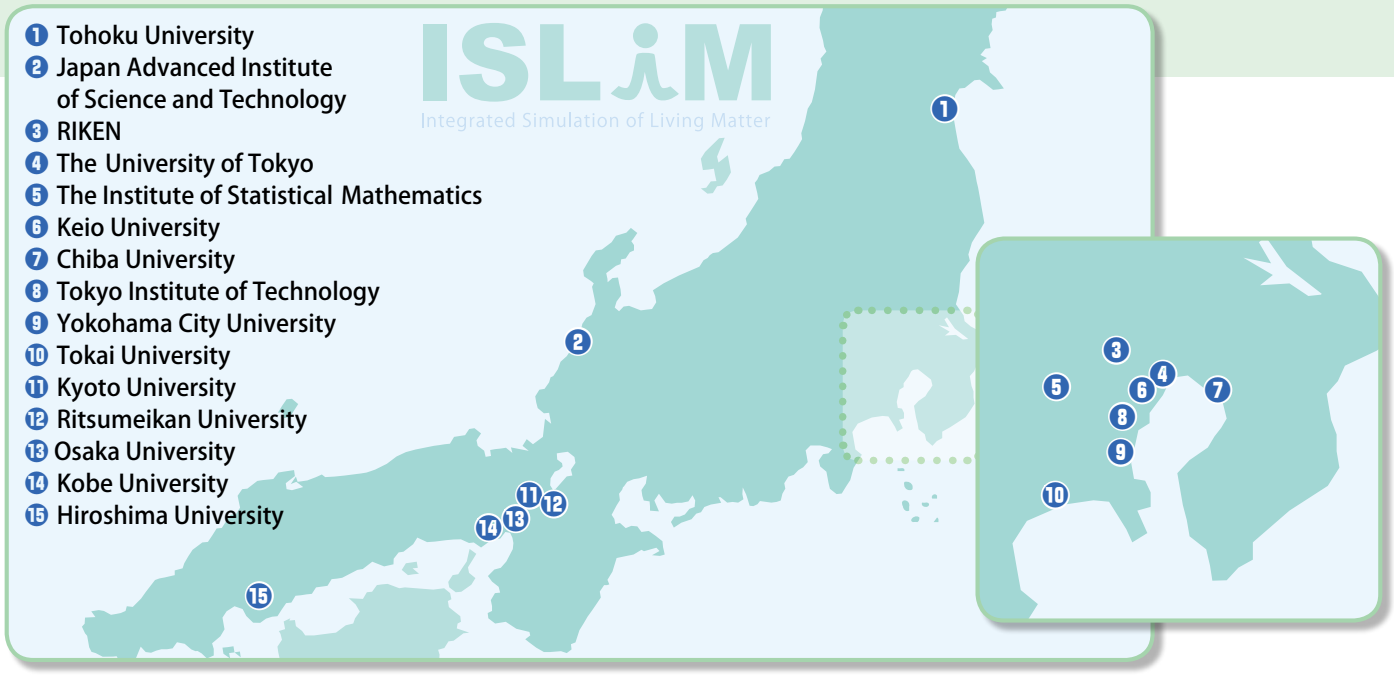
Kazuyasu Sugiyama (Associate Professor, Faculty of Engineering, The University of Tokyo)
"Development of procedures for coupled analysis of fluid structure based on the fixed mesh method"

Takanori Sugihara (Researcher, High-performance Computing Team, Computational Research Program, RIKEN (The Institute of Physical and Chemical Research))
"Determination of performance for application of next-generation computers"

Closing speech

Break up

ISLiM Participating Institutions



Event information

The 3rd BioSupercomputing Symposium

- Date : Feb 21 (Mon) and 22 (Tue), 2011 (Reception: night of February 21 (Monday))
- Location : RIKEN Advanced Institute for Computational Science (Kobe)

※ The program details and registration of participation will be notified on the Web page as soon as they are confirmed.

Topics

The ten quadrillion speed computer “京(Kei)” was started to built.

The nickname “Kei” in Japanese was chosen from public applicants for the Next-Generation Supercomputer in July, and RIKEN will refer to the Next-Generation Supercomputer as the K computer in English in the future. The first 8 racks of K computer were set up on Wednesday, September 29 (photograph on the right). The computation speed of these 8 racks is about 98.3 TFlops (teraflops) as the theoretical peak performance. More than 800 racks are to be arranged in the computer room at the completion.

The K computer with about 0.5% of its total capacity was ranked 169 among the Top500 list presented in November, 2010.

BioSupercomputing Newsletter Vol.3 2010.12

Issued : December 2010

Computational Science Research Program,
RIKEN Research Cluster for Innovation

2-1, Hirosawa, Wako, Saitama, 351-0198 Japan
TEL : +81-48-462-1488 FAX : +81-48-462-1220
http ://www.csrp.riken.jp