Bio-simulation

Consortium Representative

Go Nobuhiro — Forum on the Bio-Simulation (Japan Atomic Energy Research Institute)

Representatives for each subgroup

Fumitoshi Sato — Institute of Industrial Science, University of Tokyo

Minoru Saito — Faculty of Science and Technology, Hirosaki University

Yuko Okamoto — Institute for Molecular Science

Yutaka Akiyama — National Institute of Advanced Industrial Science and Technology

Hisashi Ishida — Japan Atomic Energy Research Institute

In the field of biology, a consortium of bio-simulation researchers, composed mainly of biological researchers in Japan, was formed in 2003.

1. Aims of the consortium

This consortium voluntarily coordinates and selects projects to be executed on the Earth Simulator and then makes recommendation to the Earth Simulator Center.

2. The activities of the consortium

The consortium is composed of an organizing committee (consisting of a chairman, several organizers and officers) and general members. Any researcher who wants to use the Earth Simulator can apply for a recommendation from the consortium by submitting an application form on which the purpose and content of the calculation, and the reasons why the use of the Earth Simulator is required should be stated. The application form is sent to all the members of the consortium who then discuss the content and significance of the proposal. According to the comments made by the members, the committee then decides whether or not to recommend the proposal. If the proposal is considered to be suitable, a letter of recommendation is sent to the chief of the Earth Simulator Center.

3. Subgroups in the bio-simulation project

The whole project is composed of five subthemes:

Subtheme1: All-electron calculation on very large-sized protein by density functional method (Fumitoshi Sato, University of Tokyo)

Subtheme2: Realistic simulations of the structural changes of proteins (Minoru Saito, Hirosaki University)

Subtheme3: Protein folding simulations from the first principles (Yuko Okamoto, IMS)

Subtheme4: The Molecular dynamics simulations of the conformational transition of prion protein from its cellular form to the anomalous form using the Earth Simulator (Yutaka Akiyama, AIST)

Subtheme5: Analysis of the function of a large-scale supra-biomolecule system by molecular dynamics simulation (Hisashi Ishida, JAERI)

4. Activities among the subgroups

As the calculations for these five different subthemes are based on different scales of time and space, this project covers a wide range of functions within the field of biology. To optimize the use of the Earth Simulator, every month each subgroup reports on their progress and the amount of computational time they have used. Based on each subgroup's requirements, the amount of computational time for the following month is allocated. This enables all the subgroups to proceed with their research efficiently.

5. Present status and future plans

Present status: the project of bio-simulation is progressing successfully.

Subgroup1 is developing a gaussian-based density functional method program for proteins, called ProteinDF which can treat a whole protein as a molecule and calculate more than 10,000 canonical orbitals (100 million elements). This SPMD program was successfully reconstructed to achieve MPMD without generating any dynamic processes. At present the vectorization ratio of 92% and the parallelization ratio of 90% have been reached with 8 nodes.

Subgroup2 have tuned up COSMOS90 on the earth simulator and successfully obtained the performance speed (0.005 sec/step for 16034 atoms using 64 processors). In this measurement, long-range Coulomb interactions are explicit-

ly calculated by the PPPC method. This performance speed is faster than other typical software, CHARMM, AMBER7, and NAMD2.4.

Subgroup3 is applying a powerful molecular dynamics simulation algorithm called Replica-Exchange Molecular Dynamics (REMD) to the folding simulation of a small protein. At present the vectorization ratio of 96.5% and the parallelization efficiency ratio of 81.0% have been achieved for the system of protein G (about 50,000 atoms) even when 896 CPUs are used on the Earth Simulator.

Subgroup4 is parallelizing and vectorizing the molecular dynamics simulation programs, AMBER and MolTreC. These programs are distinguished from calculation of long range interaction (AMBER:EWALD, MolTreC:Tree Code). Currently, the vectorization ratios are 98% and the parallelization ratios are 97.4%

Subgroup5 is developing an integrated molecular simulation system for biological macromolecules, called PABIOS which can treat more than a million particles. At present the vectorization ratio of 95.5% and the parallelization efficiency ratio of 50.5% have been reached even when 144CPUs are used on the Earth Simulator.

Future plans: from now on, we are going to develop and tune our programs further and execute them massively to elucidate the function of proteins.

Subgroup1 is going to develop and tune ProteinDF further, and perform an all-electron calculation on a large-sized protein of 30,000 canonical orbitals (1 billion elements) in the next year.

Subgroup2 begins MD simulations of hemoglobin on the earth simulator to demonstrate its tertiary structural change computationally. The reliability of simulation largely depends on force field parameters used for the simulation. We carefully test the parameters by performing short simulations of hemoglobin repeatedly.

Subgroup3 is going to apply REMD to the folding simulation of protein G in explicit water (the number of amino acids is 56 and the total number of atoms including water is about 50,000). We plan to use 896 CPUs on the Earth Simulator.

Subgroup4 is modeling hexameric prion protein from this year's results. Then, we will perform a large scale molecular dynamics simulation on hexameric prion protein to deeper understanding of conformational transition of its cellular form to the anomalous form.

Subgroup5 is going to develop and tune PABIOS further, and perform a large-scale molecular dynamics simulation of a Holliday junction (biomolecular complex consisting of four strands of DNA and four protein molecules to execute recombination of homologous DNA strands) in order to understand how branch migration occurs.