December 8, 2003

RT0559
Life Sciences 7 pages

Research Report

Integration of Heterogeneous Bio-simulators

Makoto Kano

IBM Research, Tokyo Research Laboratory
IBM Japan, Ltd.
1623-14 Shimotsuruma, Yamato
Kanagawa 242-8502, Japan

Research Division
Almaden - Austin - Beijing - Haifa - India - T. J. Watson - Tokyo - Zurich

Integration of Heterogeneous Bio-simulators

Makoto Kano
IBM Tokyo Research Laboratory
1623-14, Shimotsuruma, Yamato-shi, Kanagawa-ken 242-8502, JAPAN
mkano@jp.ibm.com

Abstract

Comprehensive bio-simulation from genes to a body is one of the most important research topics in 21st century. However, no single model or simulator can cover whole-body bio-system. Therefore, it is essential to develop environment for integrating heterogeneous bio-simulators. In this article, three types of key issues for integration of bio-simulators are discussed: (1) a wide variety of scales (time and spatial) should be handled, (2)various physical phenomena should be handled, (3) there are needs for distributed management of simulators. In addition, survey results of current status of bio-simulation environments are also reported.

1. Introduction

Recent advances of biological technologies have enabled us to have full range of biological information from genes to body. The Human Genome Project [1] ended in 2003 with the completion of the human genetic sequence. The achievement of the project gave us footholds for understanding biological phenomena at genomic level. On the other hand, various medical instrumental technologies such as NMR, MRI and CT have been developed. Structures of proteins and dynamics & electrophysiological behaviors of tissues & organs have been steadily elucidated. Therefore, the next key issue is how to integrate the wide variety of information. Especially, comprehensive bio-simulation from genes to body is one of the most important research topics [2] (Fig.1).

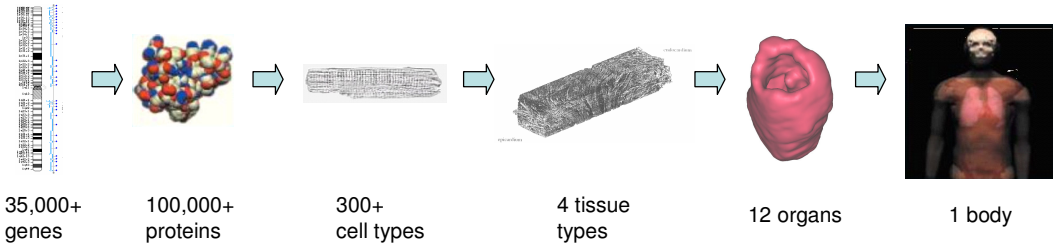

Figure1. Bio-system from genes to body

Since 1990s, many researchers all over the world have made challenges for simulating cells. Simulation of cells has become gradually sophisticated and many biological phenomena at cell level have been reproduced in silico. However, current bio-simulators are limited to application in education. In other words, bio-simulators have not been applied to medical care. There are great expectations that bio-simulators can be applied to medical care, especially for a prediction of drug therapy efficacy and adverse effects. For this objective, it is necessary to construct whole body simulation from cell-simulations. However, it is difficult for a single simulator to cover from genes to body. Therefore, for achieving comprehensive bio-simulation, integration of heterogeneous bio-simulators is essential.

This article reported survey results of current status of bio-simulators and stated several important points for developing an environment for integration of heterogeneous bio-simulators.

2. Current Status of Bio-Simulation Environments

2.1 Various Architectures for Bio-Simulations

Currently, many bio-simulators have been developed and are publicly available. E-CELL [3] (Fig.2), Gepasi [4] and other many simulators provide function for simulating biochemical phenomena described in differential equations. StochSim [5] is a stochastic simulator for phenomena where density of molecules cannot be considered uniform. Genomic Object Net [6] (Fig.3) is based on Hybrid Petri Net and aims at facilitation of describing biological facts and biological phenomena. Virtual Cell [7] focuses on associating biochemical and electrophysiological data describing individual reactions with experimental microscopic image data describing their subcellular locations.

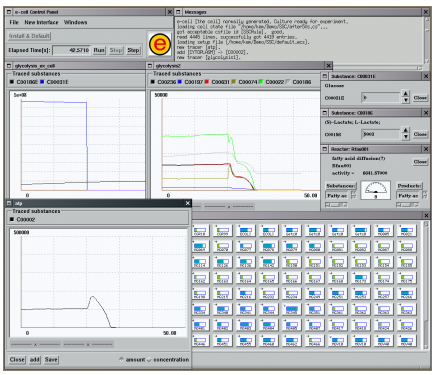

Figure2. E-CELL

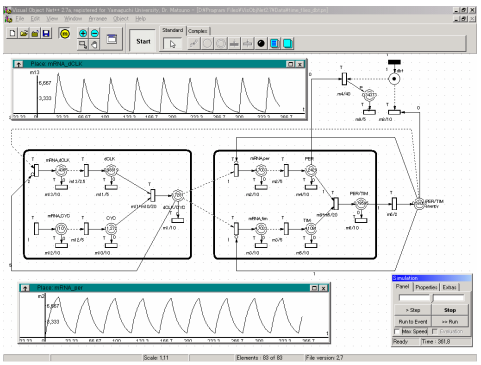

Figure3. Genomic Object Net

Since most of these simulators support standard data format such as Systems Biological Markup Language (SBML) [8] or CellML, models generated on one simulator are supposed to be able to work on other simulators. However, since each simulator has its own specific function, this catch

phrase is not always achieved. Although many simulators aim at handling all biological phenomena in a single simulator, it would be unrealistic. Actually, in above certain level of institutions, they are individually developing their own simulators for their purpose.

Systems Biological Workbench (SBW) [9] is a modular, broker-based, message-passing framework for communication between applications. On SBW, distributed applications can communicate via SSH. However, it is difficult to integrate heterogeneous simulators on SBW. What SBW provides is the function for remote procedure calls. SBW does not provide the function of absorbing the differences of scales of simulators. For achieving comprehensive bio-simulation, it is necessary to integrate various simulators which cover different spatial and time scales.

The group of Auckland University is proposing hierarchy of models to handle multi-scale simulation from genes to body [2] (Fig.4). Since it is difficult to cover all range by single model, each model is developed for a more limited range of spatial and temporal scales such as cell level, tissue level, organ level and so on. At any one level, several parameters are set as black boxes. Conventionally, these parameters are defined by experiments. However, they pointed out that these parameters should be interpretable in terms of more detailed models and simulators. That is, each model loads "black box" parameters from DB, and saves parameters for a more macroscopic model after its simulation. At each level, standard markup language to share mathematical models are now under development. For example, tissue simulators will load several parameters from DB, which are the results of cell simulators.

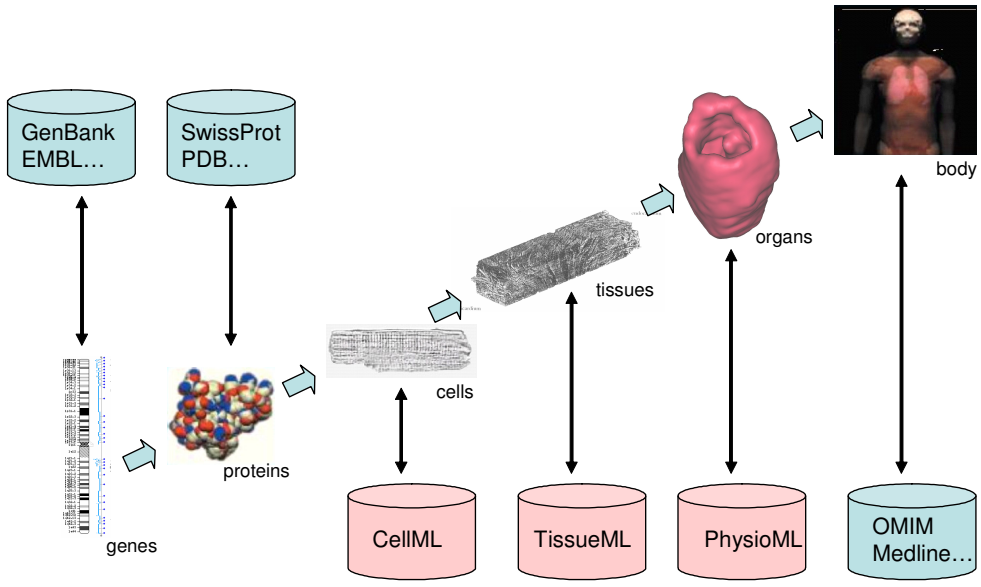

Figure4. Hierarchy of models for handling multi-scale simulation

2.2 Applying Bio-simulations to Medical Cares

The first target of bio-simulators for medical care would be cardiac abnormalities, because it is relatively easy to measure dynamics and electrophysiological behaviors of a heart compared to those of other organs. The group of Auckland University has developed a heart simulator for reproducing contractions and electrocardiographic waves. In this year, the group of Kyoto University commenced a project of the development of a heart simulator for predicting drug therapy efficacy and adverse effects [10]. We are participating in this project. In this project, the results of simulation of phenomena in cells are integrated as an organ. They have developed a heart cell simulator called KYOTO-model (Fig. 5) and it can simulate changes in ion densities and voltages in cells very well. Phenomena in an organ, such as contractions and electrocardiographic waves, are reconstructed from the simulated phenomena in cells. We are planning to contribute to this project by developing an environment for integrating heterogeneous simulators.

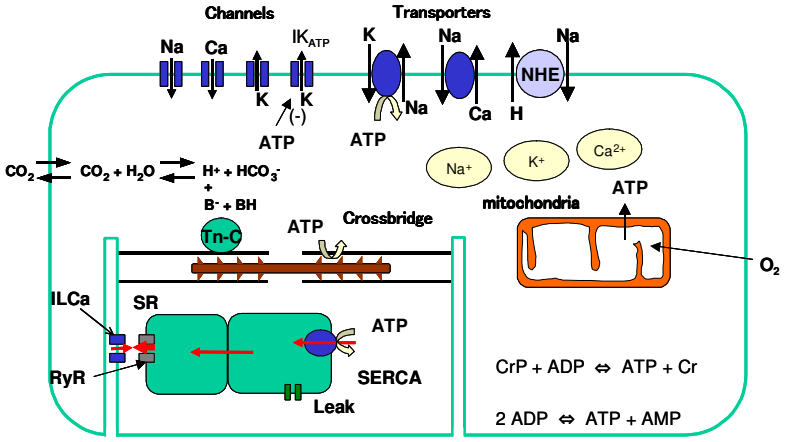

Figure5. KYOTO Model

3. Key-Issues on Integration of Heterogeneous Simulators

3.1 Integration of Different Scale Simulators

For handling comprehensive bio-simulation, simulators have to cover a factor of 9th of ten in a spatial scale from 1nm (a protein) to 1m (an intact body). In addition, a factor of 15th in a time scale from 1ms (Brownian motion) to 10^9s (= 70 years: a human life) also has to be covered. It would be difficult to develop a single simulator which can handle all scales. Current approach to handle wide range scales is to adopt hierarchy of models [2]. However, in this architecture, parameters passed are static values. There are needs that several parameters with dynamic changes should be passed to higher models. For example, to restructure electrocardiogram from results of cell level simulations, parameters with periodic changes should be passed from cell level to organ

level. In addition, the hierarchical model is for one-way communication from microscopic level to macroscopic level. This architecture cannot manage feedback from a macroscopic level to a microscopic level. There are cases where two-way communication is required. For example, heart contraction will change concentration of calcium ion in each cell. In other words, results of organ simulation effect on cell level simulation. Therefore, simple hierarchy of models would be insufficient.

3.2 Integration of Simulators Handling Different Types of Physics

In comprehensive whole body bio-simulation, we have to consider many kinds of physics such as mechanics, electrophysiology and fluid. The suitable analytic methodology would be different with respect to each phenomenon. Therefore, most of simulator focus on limited phenomena and consider other parameters produced by other phenomena as boundary conditions (="black box"). However, it is often the case that users get the urge to simulate black box parameters in detail later. Therefore, it is required to develop a frame-work for easily switching between static values and dynamically simulated values.

3.3 Needs for distributed management

For a prediction of drug therapy efficacy and adverse effects, it is necessary to simulate all organs. That is, it is important to simulate how drugs of hearts effect on pancreas or other organs. To develop and manage organ simulators, expertise in each organ is required. However, it would be difficult to manage every organ models in a single institute. Therefore, distributed management is required for whole body simulation.

4. Discussion:

In this chapter, we discuss several ideas for developing an environment for integration of heterogeneous simulators.

(1) State Manager

The most important component which a platform for integration of bio-simulators should have would be a state manager. Each simulator registers their simulating parameters on a state manager and copies the simulated parameters to state manager at every time points. Then, these parameters are loaded by other simulator as black box parameters from state manager at every time points. Each simulator does not need to know whether black box parameters are static values or dynamically simulated values.

(2) Scheduling Manager

The time-step of each simulator is defined as the maximum length which allows us to consider changes of parameters uniform. Schedule manager handles different time-steps of simulators.

(3) Deployment of Simulators

For achieving a whole body simulation, distributed management of simulators is required. We have the luck to be able to distribute organ simulators. Since there are few interactions between them, organ simulators could be loosely connected. On the other hand, cell simulators should be tightly connected. That is, it would be difficult to distribute cell simulators.

To prove this concept, we plan to integrate simulators as described in figure 7. KYOTO model mainly simulates ion channels in cells and E-cell mainly simulates metabolome. Since the home grounds of KYOTO model and E-cell are different, integration of them would improve simulation results. We integrate them in a local site, because there will be a lot of interactions between them. The group of Kyoto University plans to develop not only a heart simulator but also a pancreas simulator. We integrate a heart simulator and a pancreas simulator via Web Services.

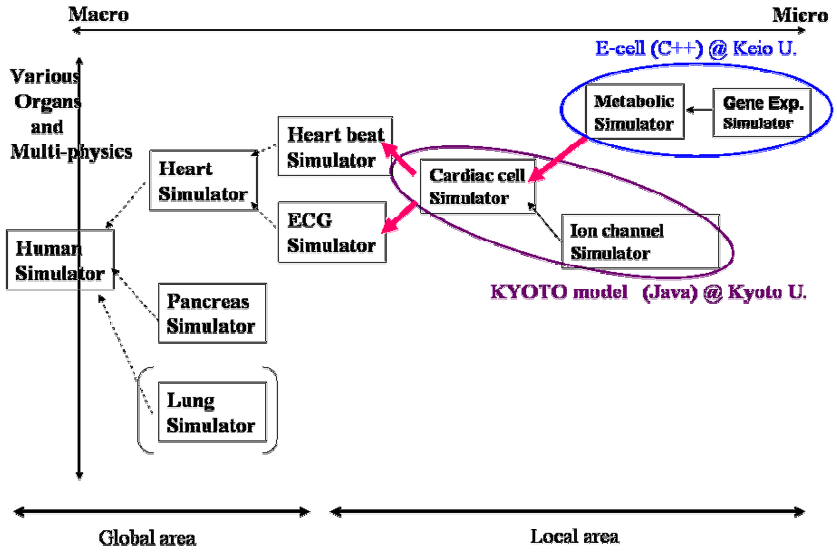

Figure7. Integration of heterogeneous bio-simulators

5. Conclusion

In this article, we have clarified key issues in developing environment for the integration of heterogeneous bio-simulators. First is that a wide variety of scales (time and spatial) should be handled. Second is that various physical phenomena should be handled. Third is that there are needs for distributed management of simulators.

Comprehensive simulation from genes to a body is one of the most important research topics in 21st century. However, no single model or simulator can cover all whole-body bio-system. Therefore, integration of heterogeneous bio-simulators is essential in the field of systems biology.

References

[1] http://www.ornl.gov/sci/techresources/Human_Genome/home.shtml

[2] Hunter PJ, Borg TK. Integration from proteins to organs: the Physiome Project. Nat Rev Mol Cell Biol. 2003 Mar;4(3):237-43.

[3] Tomita M, Hashimoto K, Takahashi K, Shimizu TS, Matsuzaki Y, Miyoshi F, Saito K, Tanida S, Yugi K, Venter JC, Hutchison CA 3rd. E-CELL: software environment for whole-cell simulation. Bioinformatics. 1999 Jan;15(1):72-84

[4] Mendes, P. (1997) Biochemistry by numbers: simulation of biochemical pathways with Gepasi 3. Trends Biochem. Sci. 22, 361-363.

[5] Le Novere N, Shimizu TS. STOCHSIM: modelling of stochastic biomolecular processes. Bioinformatics. 2001 Jun;17(6):575-6.

[6] Matsuno, H., Doi, A., Nagasaki, M., Miyano, S. 2000. Hybrid Petri Net representation of gene regulatory network. Proc. Pacific Symposium on Biocomputing 2000, pp.338-349.

[7] http://www.nrcam.uchc.edu/

[8] http://sbml.org/

[9] http://sbw.sourceforge.net/

[10] http://www.bme.sys.i.kyoto-u.ac.jp/biosim/index.html