Perspectives on Systems Biology*

Hiroaki Kitano
Sony Computer Science Laboratories, Inc.
3-14-13 Higashi-gotanda, Shinagawa, Tokyo 141-0022 Japan
and
ERATO Kitano Symbiotic Systems Project,
Japan Science and Technology Corporation
Suite 6A, M31 6-31-15 Jinguu-mae, Shibuya, Tokyo 150-0001 Japan

1 Introduction

The ultimate goal of biology is to explain every detail and principle of biological systems. Biological systems refer to various forms of natural life, such as bacteria, cells, individual creatures. Since the discovery of the structure of DNA in 1953 [Watson and Click, 53], the field of molecular biology has emerged and has made enormous progress. Molecular biology enables us to understand biological systems grounded on physical systems; that is, on molecular machines composed of proteins. Many biological processes — such as those of heredity, development, disease — can now discussed on a molecular basis, and the basic mechanisms of the such processes can be made clear. Such mechanisms include replication, transcription, translation, etc. Genes and the functions of their transcription products have been identified. The symbolic accomplishment along this line of research is the complete sequencing of DNA. DNA sequences were completely decoded for a numbers of organisms — such as *mycoplasma, E. coli, C. elegans, Drosophila,* and the sequencing of human DNA is expected to complete within a couple of years. The identification of genes from these sequences has also underway with astonishing speed, and studies deepening our understanding of protein and their interactions are also in progress. Parallel to such efforts, numbers of methods for disturbing biological systems selectively, such as loss-of-function knock-out of specific genes, have been and are being invented. For a particular species, *C. elegans*, an easy and efficient disruption method called RNA interference (RNAi) was invented, and a project to systematically and exhaustively knock-out various genes is underway.

*This paper originally appeared in *New Generation Computing* Vol. 18, 199-216 (2000) Springer-Verlag.

There is no doubt that our understanding of the molecular-level mechanisms of biological systems will progress at an even faster pace, but this will not provide us with an understanding of biological systems as systems. Genes and proteins are components of the systems. While understanding what constitutes the system is necessary for understanding the system, it is not sufficient. A series of methods and technique has to be developed that are specifically geared to provide system-level understanding.

Systems biology is a new field of biology that aims at a system-level understanding of biological systems. If we are to understand biological systems as systems, we must understand (1) the structures of the systems (both components and their structural relationships), (2) behaviors and their characteristics at different points in the parameter space, (3) methods controlling the states and behaviors of the system, and (4) methods by which systems with desired functions are designed and built.

The scope of systems biology is potentially very broad and different sets of techniques may be deployed for different research targets. Nevertheless, two of the main targets will be genetic and metabolic network systems, because they are the systems controlling the fundamental mechanisms that govern biological systems. Gene regulatory networks and metabolic networks are highly complex networks with extensive feedback loops. We must develop methods for understanding and controlling such complex and large-scale networks.

Fortunately, extensive data is available for some of well studied model animals, such as *C. elegans*. A complete cell lineage has already identified [Sulston et al., 83, Sulston and Horvitz, 77], the topology of the neural system has been fully described [White et al., 86], and the DNA sequence has been fully identified [C. elegans, 98]. A project using *in situ* hybridization [Tabara et al., 96] to provide a full description of gene expression patterns during development is underway, and the construction of a systematic and exhaustive library of mutants has begun. In addition, a series of new projects for measurement of neural activity *in vivo* have started as has a project for the automatic construction of a cell lineage in real time by using advanced image processing combined with special microscopy [Yasuda, et al., 99].

While the effort focused on *C. elegans* is a symbolic example of efforts directed at a comprehensive and exhaustive understanding of biological systems, similar efforts can be expected to be focused on a range of biological systems in the near future. Although at this moment these studies are limited to the understanding of components of the system and of their local relationships with other components, the combination of such exhaustive experimental work and computational and theoretical research would be a viable foundation for systems biology.

The next stage of understanding need to be accomplished at the system-level, where the focuse is on how the components work together and behave as a system. There are numbers of specific issues that need to be addressed at each of the following four levels:

System Structure Identification Identification of structure of the system — such as regulatory relationship of genes and interaction of protein that forms signal transduction and metabolism pathway — need to be accomplished. Both the topological relation of the network of components and the parameters for each relation need to be identified. The use of high-throughput expression data, RT-PCR, and other methods monitoring biological processes are critical resources. Nevertheless, methods to identify the structure of the gene and metabolism networks from these data are still to be established.

Identification of gene regulatory networks in multicellular organisms is even more complex than for a single cell, because such a network involves extensive cell-cell communication and physical configuration in three dimensional space. Structure identification for multicellular organisms inevitably involves not only the structure of gene regulatory networks and metabolism networks, but also requires that the physical structures of the whole animals be known with the same precision that their cellular structures are known. Obviously, new instrumentation systems to collect the necessary data need to be developed.

System Behavior Analysis Once a system structure is identified to a certain degree, the system's behavior needs to be understood. Various analysis methods can be used to serve various purposes. For example, one may wish to know the sensitivity of certain behaviors to external perturbation and to know how quickly the state of the system returns to the normal state after it is perturbed. The analysis required for answering these questions not only reveals system-level characteristics, but also provides important insights for medical treatments because it reveals how cells responds to a range of concentrations of chemicals. Therapeutic effects can thus be maximized with minimizing side effects.

System Control If we are to apply the insight obtained from an understanding system structure and behavior, we must establish method for controlling the state of biological systems. How can we transform cells that are malfunctioning into healthy cells? How can we control cancer cells to turn them into normal cells or cause apoptosis? Can we control the differentiation status of a specific cell such that it transforms into a stem cell and then control it to differentiate into a desired cell type? Technologies providing such control would benefit human health enormously.

System Design Ultimately, we would like to establish technologies that allow us to design biological systems curing diseases. One futuristic example would be to design and grow organ from the tissue of the patient. Such an approach, which may be called "partial organ cloning," would be enormously useful for treating diseases that require the transplantation of organs. There may also be some engineering applications using biological materials for robotics or computation.

By using materials that are able to maintain and repair themselves, industrial systems will undergo a revolutionary transition.

2 Characteristics of Biological Systems

Before specific approaches to understanding biological systems can be discussed, there must be some discussions of how the characteristics of biological systems compare with those of other complex systems. In research on most complex systems, it is assumed that large number of simple components emerge to exhibit complex behaviors. Such a phenomenon is termed emergence. In many case, the components of the system are assumed to be homogeneous. Biological systems are indeed composed of very large numbers of cells, proteins, and genes, but these components are not at all homogeneous or simple. Biological systems are best characterized by the following three structural characteristics:

Heterogeneity of components: The components of a biological system are heterogeneous. At the genetic level, it consists of thousands of genes, each of which has different regulatory relation and each product of which has a different function. The system cannot be simply considered a large number of homogeneous components, nor can its behavior be approximated using average behavior.

Complexity of components: Each component is itself complex. Each gene has a complex regulatory structure and its product has its own complex structure and dynamics. Each protein has a different structure and a different function. And protein do not exist in isolation. They form complex and even larger structures such as microtubles, cell membrane, and other substructures of cells. Diverse functions of proteins are essential for biological systems.

Selectivity of interactions: Interaction among components are highly selective. Which gene regulates which other genes is highly specific, and the interactions of proteins are also highly specific. This specificity ensures that complex and diverse components can be created, and it ensures that various substructures of cells can be maintained.

These structural characteristics are essential features of biological systems. Although one might wish to model the system as networks of simple and homogeneous elements, such an abstraction fails to capture the essence of the system's properties. While the conventional approaches analyzing the average behaviors of the system may provide some insights, we need to establish methodologies that can cope with large-scale networks of complex and heterogeneous elements.

3 Design Patterns and Control Principles

As it can be assumed from the fact that the structures of biological systems are formed through evolution — that is, through the accumulation of the effects of random events with selection pressure — there is no guarantee that existing biological systems are optimally designed for the various functions they exhibit. Thus, it is not possible to infer the structure of a system from the function of the system. Instead of pursuing design principles that dictate how a system shall be optimally designed for the desired function, we should try to identify patterns of design so that we can create a library of design patterns that are used is biological systems and develop methods that can quickly identify which of these patterns is used for a specific biological system. Design patterns in living things can be identified in various levels. The most important design pattern that we focus are patterns of genetic and metabolism network because they are the basis of various biological processes and responsible for fundamental characteristics of the living creatures.

The generation of the design patterns underlying biological systems was, of course, not completely random. Although evolution is a stochastic process, selection pressure chose certain classes of circuits that are likely to be functional in some aspects. Various forms of feedback loops, redundancy path, and modular design are incorporated in many of the circuits. While the structures of the circuits and their components may vary, the number of underlying control mechanisms can be reduced through evolution and these mechanisms can be conserved. Although evolutionarily conserved genes are the focus of interest at this moment, evolutionarily conserved control circuits will be the major interest in systems biology research.

There are interesting analogies between biological systems and engineering systems. Both kinds of systems are designed incrementally through some sort of evolutionary processes and are generally sub-optimal for the given task. And both attain higher levels of robustness and stability as their complexity increases. Mycoplasma, which has only about 400 genes is a minimal self-sustained organism and can survive only in a consistent environment. *E. coli* has evolved to have nearly 4,000 genes and can survive in a varying environment. Most the additional genes contribute to robustness against internal and external purtabations. Similarly, the first airplane built by the Wright brothers had only a handful of parts, but a modern jet like the Boeing 777 has millions of parts. Such a great increase in the number of components is for the most part, a result of efforts to ensure the stability and robustness of the airplane's operation.

In engineering systems, robustness and stability are achieved by the use of feedback, redundancy, and modular design. Feedback is a sophisticated control system that closes the loop of the signal circuits and attains the desired control of the systems.A negative feedback system detects a difference between the desired output and the actual output and compensates the difference by adjusting the input. This is one of the most widely used methods to increase the stability

and robustness of the system. Redundancy is a method widely used to improve a system's robustness against damages to its components by using multiple pathways to accomplish the system's function. Modular design prevents damage from spreading without limit and also eases the evolutionary up-grading some of the system components.

4 System Structure Identification

To understand a biological system, we must first identify its structure. To identify a gene regulation network, for example, one must identify all the components of the network, the function of each component, the interactions of all components, and all associated parameters. All the experimental data available should be used becausethis task is clearly not a trivial one. The difficulty is that the structure of such a network cannot necessarily be inferred from experimental data based on some principles or universal rules because biological systems evolved through a stochastic process. In most cases, there are multiple solutions to the problems posed by a given set of data, and one must find computational and experimental methods to identify which of these solution is the one applying to the specific biological system under consideration.

4.1 A Choice From Multiple Solutions

The trap of multiple solutions can be illustrated in a simple example of stripe-pattern formation. Various forms of stripe patterns are formed in the process of development, and how such strips are formed is an interesting research topic. For example, Kondo and Asai demonstrated that stripe patterns in a marine angelfish *Pomacanthus* are generated by the turing wave [Kondo and Asai 95]. A stripe pattern is also formed in the early embryogenesis of *Drosophila melanogaster*. Seven vertical stripes are formed by transcription products of even-skipped (eve) gene. Staut Kaufmann once claimed that this is also formed by the turing wave, but it later became clear that each of the seven stripe is controlled independently by a set of regulator genes. Even if two phenotypically similar patterns are formed, there is no guarantee that they are formed by the same mechanisms. We should consider that biological systems exploit all possible mechanisms that can support desired functions. This means that there may be several different mechanisms that can create similar phenotypes.

Figure 1(A) shows an expression level of gene A along the anteior-posterior axis. Three evenly spaced stripes can be formed by a Turing wave (Fig. 1(B)), or by independently controlled gene regulation, similar to *eve* in *Drosophila* (Fig. 1(C)). While which one of two possible gene regulatory networks is actually used cannot be determined by looking only at wild-type, it can be distinguished by creating loss-of-function knockout of gene B. A Turing wave pattern will disap-

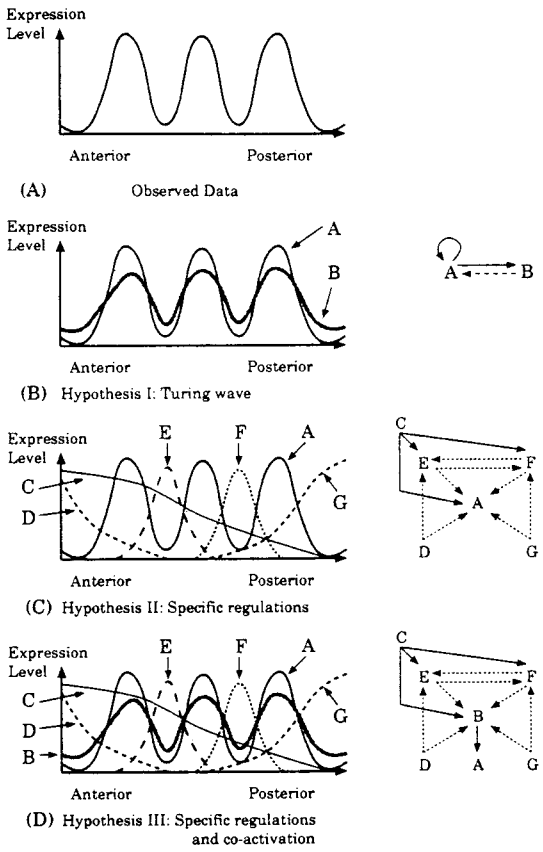

Figure 1: Possible Genetic Regulatory Network for Stripe Patterns

pear if gene B is knocked out, whereas the stripe patterns will be unaffected (or only one or two of them will be affected) if the stripes are controlled independently. Gene knockout thus provides data for identification of gene regulation networks. Situation is rather complicated, however, if the network is structured as indicated by the expression-level plot in Fig. 1(D). In this network, gene A is activated by gene B. Thus, if we knock out gene B, stripe pattern will disappear just as a Tunring wave pattern disappears when gene B is knocked out. Nevertheless, if we can knock out genes that control gene B independently, only one or two stripe will be eliminated so the pattern can be distinguished from a Turing wave pattern.

4.2 Appraoches for Structure Identification

Several attempts has already been made to identify gene regulatory network from experimental data. These attempts have taken the following approaches:

Bottom-Up Approach: The bottom-up approach tries to construct a gene regulatory network from independent experimental data, mostly obtained through a literature search and the rest obtained from experiments designed to provide information about specific aspects of the network of the interest. Some early examples of this approach were modeling of a lambda phage decision circuit [McAdams and Shapiro, 95], the early embryogenesis of *Drosophila* [Reinitz, et al., 95] [Hamahashi and Kitano, 98] [Kitano et al., 97], leg formation [Kyoda and Kitano 99a], wing formation [Kyoda and Kitano 99b], eye formation on ommatidia clusters formation and R-cell differentiation [Morohashi and Kitano, 98], and a model of eye formation based on reaction-diffusion model [Ueda and Kitano, 98]. This approach is suitable when most of genes and their regulatory relationships are relatively well understood. In some cases, biochemical constants can be measured so that very precise simulations can be performed. When most parameters are available, the main purpose of the research is to build a precise simulation model so that dynamical properties of the system can be analyzed by changing parameters that cannot be changed in the actual system, and so that available knowledge can be tested to see if it leads to simulation results consistent with experimental data.

Top-Down Approach: The top-down approach tries to make use of data gathered using a high-throughput DNA micro-array and other new measurement technologies. There have already been attempts to infer genetic network structures from DNA micro-array data gathered using a clustering technique to explore the yeast cell cycle [Brown and Botstein, 99, DeRisi, et al., 97, Spellman, et al., 98] and the development of the mouse central neural system [D'haeseleer, et al., 99]. Clustering methods are suitable for handling large-scale profile data but do not directly provide network structures. It only provides

clusters of genes that are co-expressed with similar temporal patterns. Some heuristics must be imposed if we are to infer network structures from data gathered in experiments using such methods. Alternative methods to directly infer network structures directly from expression profiles are now being developed [Morohashi and Kitano, 99, Liang, et al., 99] and from an extensive gene disruption data[Akutsu et al., 99]. Easy-to-understand visualization is often required [Michaels, et al., 98], and this poses serious computational challenges. Methods that enable the structure of genetic networks to be inferred from the smallest possible amount of expression profile data need to be developed.

Hybrid Approach: A hybrid of the bottom-up and top-down approachs is a promising and practical method. It is unlikely that no knowledge should be assumed in the process of gene network inference from the experimental data. In practical cases, it can be assumed that genes and their interactions are already understood rather well, and all that needs to be identifed is the rest of the network. The use of trustworth knowledge can significantly reduces the kinds of network structures feasible.

4.3 Computational Challenges

There are many challenges common to these approaches. One of the computational challenges can be defined as follows:

> Given a set of expression profile data and a gene network, find a set of simulation parameters that generates the expression profile.

This situation, however, is too much simplified and the challanege can serve only for proof-of-concept level studies. The actual situation is much more complex, and the true computational challenge can be defined as follows:

> Given a set of noisy expression profiles, experimental data, and partially correct networks, find a set of plausible gene regulatory network topologies and their associated parameters.

Not only finding the structure of the network, but also finding a set of parameter is important because all computational results have to be tested against actual experimental results. In many cases, parameter set has to be estimated from experimental data. Various parameter optimization methods, such as genetic algorithms and simulated annealing, are used to find a set of parameters that can generate simulation results consistent with experimental data [Hamahashi and Kitano, 99]. It must be noted that there may be multiple parameter sets that generate simulation results equally well fitted to experimental data. An important feature of parameter optimization algorithms to be used in this approach is that they find as many local minima (including a global minimum) as possible, rather than finding single global

minimum. Combined with the parameter search, there must be a mechanism to generate hypotheses about genetic and metabolic interactions. Even in the most well-investigated biological systems, not all the network structure is identified. One of the most important roles of the bottom-up approach is to predict unknown genetic interactions consistent with available knowledge and data. There have already been preliminary attempts to predict unknown genes and their interactions [Morohashi and Kitano, 98, Kyoda and Kitano 99a, Kyoda and Kitano 99b]. These attempts have involved manual searches for possible unknown interactions from which simulation results consistent with experimental data can be obtained. An exhaustive search of all possible space of network structures have not been performed. Research on an automatic hypotheses generator is now underway [Akutsu et al., 99].

4.4 Measurement Issues

Computational efforts alone will never identify the structure of gene and metabolic networks. There are numbers of aspects of the measurement technique that need to be improved.

First, the accuracy of the DNA micro-array technique and other measurement techniques needs to be improved drastically. While RT-PCR provides a more accurate measurement when it is calibrated properly, it cannot measure as efficiently as a DNA micro-array can. Second, measurements are performed for cultures of cells. While such measurements are suitable for studying homogeneous cell cultures, single-cell measurement techniques are necessary for most research in developmental biology and in the investigation of homogeneous cell cultures. Third, not only mRNA levels, but also protein concentrations need to be measured, preferably simultaneously. In addition, intracellular localization patterns should be measured.

5 System Behavior Analysis

Once we understand the structure of a system, research will focus on the dynamic behaviors of the system. How does it adapt to the changes in the environment, such as change in the levels and kinds of nutrients available? How does it maintain its integrity when subjected to damages such as DNA damage and mutations. If we want a system-level understanding, we need to understand the robustness and stability of the system.

This is a very interesting issue from both biological and engineering viewpoints. There is relationship between the robustness of a system and complexity of that system. Consider again the example of an airplane. If atmospheric air flow is stable and the airplane does not need to change its courses, altitude, or weight balance, and does not need to take off or land, it can be build using only a handful of components. Modern jet airliners, however, have millions

of components, mainly to improve their stability and robustness. One of the major reason for increasing the complexity of engineering systems is to increase their stability and robustness. Is this also the case for biological systems?

Mycoplasma is one of the smallest self-sustaining organisms and has only about 400 genes. It can live only under narrowly specific conditions, and thus very vulnerable to environmental fluctuations. *E. coli*, on the other hand, has over 4,000 genes and can live in varying environments. *E. coli* has evolved genetic and biochemical circuits for various stress responses and basic behavioral strategies, such as chemotaxis [Alon, et al., 99, Barkai and Leibler, 97]. These response circuits forms a class of negative feedback loop. Similar mechanisms also exists in eukaryotic cells.

The major methods used to improved the robustness and stability of engineering system are feedback control, redundancy, and modular design. Is this also the case for biological systems? and if so how do these metods work in biological systems?

5.1 Feedback

One of the simplest examples of how a biological system exploits feedbacks can be seen in the lambda phage fate decision circuit[McAdams and Shapiro, 95]. Lambda phage exploits a feedback mechanism to stabilize the committed state and to enable the switching of its pathways. When lambda phage infects *E. coli*, it chooses one of two pathways: lysogeny or lysogen. While a stochastic process is involved in the early stage of commitment, two positive and negative feedback loops involving CI and Cro plays critical roles in the stable maintenance of the committed decision (Fig. 2 and Fig. 3). Whether to maintain feedback or not is determined by amount of activator bound to the O_R region, and the activator itself cuts off feedback signal if this amount exceeds a certain level. This is an interesting molecular switch that is not found elsewhere. Overall, the concentration of Cro is maintained at a certain level by using positive feedback and negative feedback (Fig. 4). It is important that we can identify such mechanisms and create a library of them if we are to understand patterns of genetic circuit designs.

Another example demonstrating a critical role of the feedback system is seen in control of the growth of human cells. Growth control is one of the most critical cellular functions and the feedback circuit involved in p53 presents a clear example how feedback is used (Fig. 5). When DNA is damaged, a DNA-dependent kinase (DNA-PK) is activated and promotes phosphorylation of a specific locus of the p53 protein. When this locus is phosphorylated, p53 no longer forms a complex with MDM2 and it from being dissolved The phosphorylation locus depends on what kind of stress is imposed on the DNA. In one case, phosphorylated p53 promotes transcription of p21, and causes G1 arrest. In another case, it promotes pig-3 activation and results in apoptosis. For those cells that entered G1 arrest, DNA-PK activity is lost as soon as DNA is repaired. The

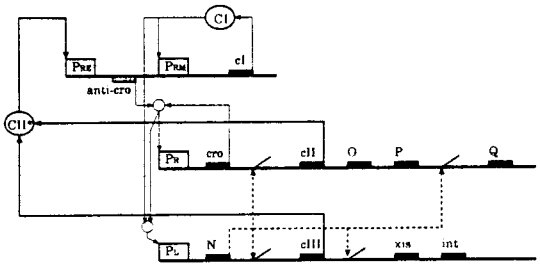

Figure 2: Lambda Decision Circuit

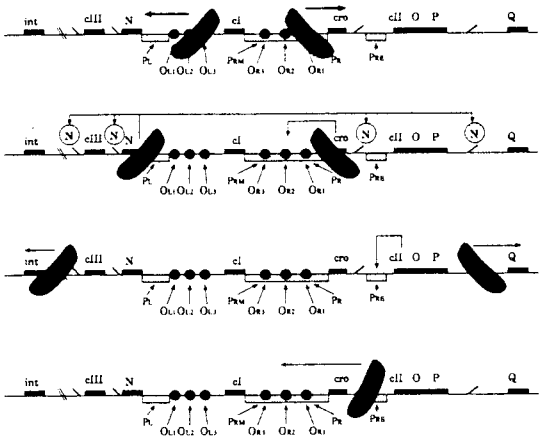

Figure 3: Binding Site for Lambda Switch

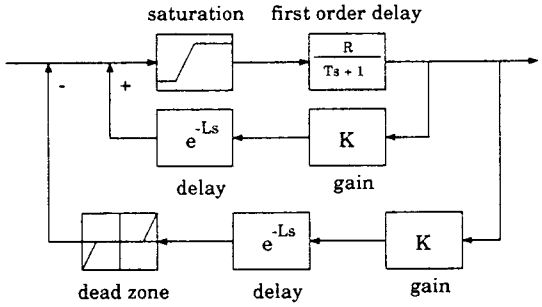

Figure 4: A Block Diagram for Double Feedback Loops

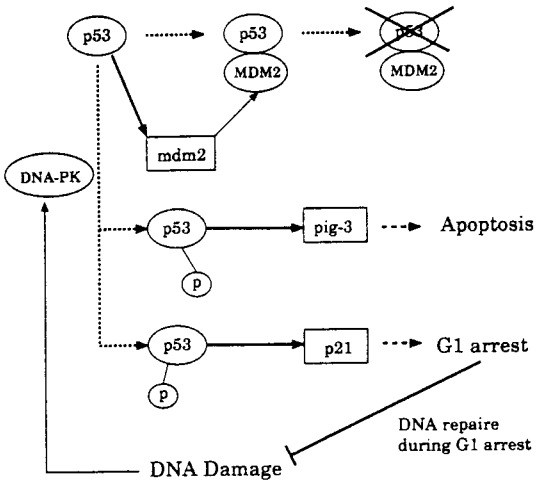

Figure 5: p53 related feedback loop

loss of DNA-PK activity decreases phosphorylation of p53, so that p53 will bind with MDM2 and dissolve.

Without phosphorylation, p53 protein promotes mdm-2 transcription. It is interesting that mdm-2 protein forms a complex to deactivate p53 protein. This is another negative feedback loop embedded in this system.

5.2 Redundancy

Redundancy also plays important role in assuring robustness of a system. It is critical for coping with accidental damage to components of the system. If there are four independent signal transmission sub-systems, the system functions normally if one or two of them are damaged. In fact, there are four independent hydraulic control systems in the Boeing 777, so it is highly robust. The MAP kinase cascade involves extensive crosstalk among collateral pathways. Even if one of these pathways is disabled as a result of mutation or some other cause, the function of the MAP kinase pathway as a whole can be maintained because the other pathways still carry signals (Fig. 6).

Once we understand stability and robustness of the system, we should be able to understand how to control and transform cells. We should be able to address such questions as how can we transform cells malfunctioning into normal cells? and how can we predict the risk of diseases and treat those diseases preemptively.

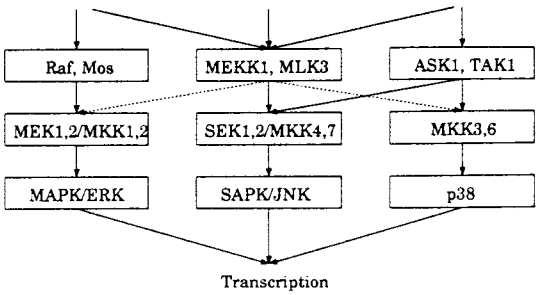

Figure 6: Redundancy in MAP Kinase Cascade

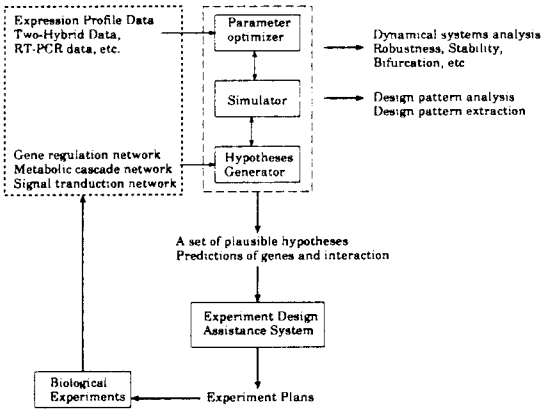

Figure 7: Software Tools for Systems Biology

6 Software Platform

A set of software systems to assist systems biology research need to be developed and integrated. Such systems will include software for data collection, for simulators, for parameter optimization, for data visualization system, and for various analytical tools. While there are many independent efforts to develop some of this software, there has been no effort to create a common platform that integrates these software modules. Recently, a group of researchers initiated an effort to define software platform for systems biology. Although there are a number of issues relevant to a software platform for systems biology that need to be addressed, the rest of the section describes only some illustrative issues.

Simulation of the behavior of gene and metabolism networks plays an important role in systems biology research, and there are several devoted to simu-

lator development [Mendes and Kell, 98, Tomita et al., 96, Kyoda, et al., 2000, Nagasaki, et al., 99]. Because of the complexity of the network behavior and the large number of components involved, it is almost impossible to understand behaviors of such networks intuitively. In addition, accurate simulation models are essential for analyzing system dynamics by changing the parameters and structure of the gene and metabolism networks. Although such analysis is necessary for understanding dynamics, the parameters and structures of actual biological systems cannoted freely changed. Simulation is a essential tool not only for understanding the behavior of the existing systems but also for designing new ones. Various forms of simulation are used when complex engineering systems are designed, and it is unthinkable today that any significantly complex engineering systems can be designed and built without simulation. VLSI design undergoes serious design simulation that creates one of the major market for supercomputers. Commercial aviation provides another example. The Boeing 777 design was fully based on simulation and digital pre-fabrication. Once we entered that stage of designing and actively controlling the biological systems, simulation would be the central method of design process.

If simulation is to be a viable methodology for studying biological systems, we need to develop highly functional, accurate, and user-friendly simulator systems. Simulators and associated software systems often require so much computing power that they need to be run on highly parallel cluster machines, such as Beowulf-cluster [Okuno et al., 99]. Although there are some simulators, no system meets the needs of a broad range of biology research, where simulators must be able to simulate gene expression, metabolism, and signal transduction for single cells and for multiple cells. The kind of simulator we need must be able to simulate both high concentrations of proteins that can be described by differential equations, and low concentrations of proteins that need to be handled stochastically. Some efforts on simulating a stochastic process [McAdams and Arkin, 98] and integrating it with high-concentration level simulation are underway. In addition, no existing simulator incorporates localization within a cell.

In some cases, it is necessary to model not only gene regulatory networks and metabolic networks, but also the high-level structure of chromosome, such as heterochromatin structures. In the study of aging, there have been attempts to model heterochromatin dynamics [Kitano and Imai, 98, Imai and Kitano, 98]. Nevertheless, how to model such dynamics and how to estimate the structure from sparse data and our current level of understanding are major challenges.

The simulator need to be coupled with parameter optimization tools, a hypothesis generator, and a group of analysis tools. And the algorithms underlying these software systems need to be designed for biological research. One example, mentioned earlier here, is that the parameter optimizer needs to find as many local minima (including a global minimum) as possible because multiple solutions are possible and only one of them is actually used. The assumption that the most optimal solution is the one used in the actual system does not

hold for biological systems. Most parameter optimization methods are designed to find the global optima for engineering design and for problem solving. While existing algorithms can be starting points, they must be modified for biological research. Similar arguments also holds for other software tools.

Ultimately, the software tools used for modeling diseases and for simulating organ growth and control needto provide a comprehensive and highly integrated simulation and analysis environment.

7 Conclusion

Systems Biology is an emerging field in biology. It aims at system-level understanding of biological systems. System-level understanding requires a range of new analysis techniques, measurement technologies, experimental methods, software tools, and concepts for looking at biological systems. The field is a new one and has a long ways to go before it will provide a deep understanding of the biological systems. Nevertheless, the author believes that systems biology will be the dominant paradigm in biology and that it can be expected to provide a number of medical applications as well as scientific discoveries.

References

[Akutsu et al., 99] Akutsu, T., Miyano, S. and Kuhara, S., "Identification of Genetic Networks from a Small Number of Gene Expression Patterns under the Boolian Network Model," *Proc. of Pacific Symposium on Biocomputing 99*, pp17-28, World Scientific, 1999.

[Alon, et al., 99] Alon, U., Surette, M.G., Barkai, N., and Leibler, S., (1999) "Robustness in bacterial chemotaxis," *Nature*, Vol. 397, pp 168- 171.

[Barkai and Leibler, 97] Barkai, N., and Leibler, S., (1997) "Robustness in simple biochemical networks," *Nature*, Vol. 387, pp 913- 917.

[Brown and Botstein, 99] Brown, P.O. and Botstein, D., "Exploring the New World of the Genome with DNA Microarrays," *Nature Genetics*, Vol. 21, 33-37, 1999.

[C. elegans, 98] The C. elegans Sequencing Consortium, (1998). "Genome Sequence of the Nematode C. elegans: A Platform for Investigating biology," *Science*, Vol. 282, 2012-2018.

[DeRisi, et al., 97] DeRisi, J.L. and Lyer, V.R. and Brown, P.O., "Exploring the Metabolic and Genetic Control of Gene Expression on a Genomic Scale,", *Science*, Vol. 278, 680-686, 1997.

[D'haeseleer, et al., 99] D'haeseleer, P. and Wen, X. and Fuhrman, S. and Somogyi, R., "Linear Modeling of mRNA Expression Levels During CNS Development and Injury," *Proc. of Pacific Symposium on Biocomputing*, 41-52, 1999.

[Hamahashi and Kitano, 98] Hamahashi, S. and Kitano, H., "Simulation of Fly Embryogenesis," *Proc. of International Conference on Artificial Life*, 1998.

[Hamahashi and Kitano, 99] Hamahashi, S. and Kitano, H., "Paramter Optimization in Hierarchical Structure," *Proc. of European Conference on Artificial Life*, 1999.

[Imai and Kitano, 98] Imai, S., and Kitano, H. "Heterochromatin Island and Their Dynamic Reorganization: A Hypothesis for Three Distinctive Features of Cellular Aging," *Experimental Gerontology*, Vol. 33, No. 6, 555-570, 1998.

[Kitano and Imai, 98] Kitano, H. and Imai, S., (1998) "The Two-Process Model of Cellular Aging," *Experimental Gerontology*, Vol. 33, No. 5, 393-419, 1998.

[Kitano et al., 97] Kitano, H., Hamahashi, S., Takao, K., and Imai, S., (1997) "Virtual Biology Laboratory: A New Approach of Computational Biology," *Proc. of European Conference on Artificial Life*, 1997.

[Kitano, et al. 98] Kitano, H., Hamahashi, S., and Luke, S., "The Perfect C. elegans Project: An Initial Report," *Artificial Life*, Vol. 4: 141-156, 1998.

[Kondo and Asai 95] Kondo, S. and Asaii, R., "A reaction-diffusion wave on the skin of the marine angelfish *Pomacanthus*," *Nature*, 376, 765-768, 1995.

[Kyoda and Kitano 99a] Kyoda, K. and Kitano, H., "Simulation of Genetic Interaction for Drosophila Leg Formation," *Pacific Symposium on Biocomputing 99*, World Scientific, 1999.

[Kyoda and Kitano 99b] Kyoda, K. and Kitano, H., "Simulation of Axis Determination of Drosophila Wing," *Proc. of European Conference on Artificial Life 99*, The MIT Press,, 1999.

[Kyoda, et al., 2000] Kyoda, K., Muraki, M., and Kitano, H., "Construction of a Generalized Simulator for Multi-Cellular Organisms and Its Application to SMAD Signal Transduction," *Pacific Symposium on Biocomputing 2000*, World Scientific, 2000.

[Liang, et al., 99] Liang, S. and Fuhrman, S. and Somogyi, R., "REVEAL, A General Reverse Engineering Algorithm for Inference of Genetic Network Architectures," *Proc. of Pacific Symposium on Biocomputing 99*, 18-29, 1999.

[McAdams and Arkin, 98] McAdams, H.H. and Arkin, A., "Simulation of Prokaryotic Genetic Circuits," *Annu. Rev. Biophys. Biomol. Struct.*, Vol. 27, 199-224, 1998.

[McAdams and Shapiro, 95] McAdams, H. and Shapiro, L. (1995) "Circuit Simulation of Genetic Networks," *Science*, Vol. 269, pp 650-656.

[Mendes and Kell, 98] Mendes, P. and Kell, D.B., "Non-linear Optimization of Biochemical Pathways: Applications to Metabolic Engineering and Parameter Estimation," *Bioinformatics*, Vol. 14, No. 10, 869-883, 1998.

[Michaels, et al., 98] Michaels, G.S., Carr, D.B., Askenazi, M., Fuhrman, S., Wen, X., and Somogyi, R., "Cluster Analysis and Data Visualization of Large-scale Gene Expression Data," *Proc. of Pacific Symposium on Biocomputing'98*, 42-53, 1998.

[Morohashi and Kitano, 99] Morohashi, M. and Kitano, H., (1999) "Identifying Gene Regulatory Networks from Time Series Expression Data by *in silico* Sampling and Screening," *Proc. of European Conference on Artificial Life*, 1999.

[Morohashi and Kitano, 98] Morohashi, M. and Kitano, H., "A Method for Reconstructing Genetic Regulatory Network for Drosophila Eye Formation," *Proc. of International Conference on Artificial Life*, 1998.

[Nagasaki, et al., 99] Nagasaki, M., Onmani, S., Miyano, S., and Kitano, H., "Bio-Calculus: Its Concept and Molecular Interaction," *Proc. of Genome Informatics Workshop 1999*, Tokyo, 1999.

[Okuno et al., 99] Okuno, G. H., Kyoda, K., Morohashi, M. and Kitano, H., "An Initial Assesment of ERATO-1 Beowulf-Class Cluster," *Proc. of International Workshop on Parallel and Distributed Computing for Symbolic and Irregular Applications*, Sendai, 1999.

[Reinitz, et al., 95] Reinitz, J. and Mjolsness, E. and Sharp, D.H., "Model for Cooperative Control of Positional Information in *Drosophila* by Bicoid and Maternal Hunchback," *The Journal of Experimental Zoology*, Vol. 271, 47-56, 1995.

[Spellman, et al., 98] Spellman, P.T., Sherlock, G., Zhang, M.Q., Iyer, V.R., Anders, K., Eisen, M.. Brown, P.O., Botstein, D., and Futcher, B., "Comprehensive Identification of Cell Cycle-regulated Genes of the Yeast *Saccharomyces cerevisiae* by Microarray Hybridization," *Molecular Biology of the Cell*, Vol. 9, 3273-3297, 1998.

[Sulston and Horvitz, 77] Sulston, J. E. & Horvitz, H. R. (1977). "Post-embryonic Cell Lineage of the Nematode, *Caenorhabditis elegans*" *Devl Biol.* **56**, 110-156.

[Sulston et al., 83] Sulston, J. E., Schierenberg, E., White, J. G. & Thomson, J. N. (1983). "The Embryonic Cell Lineage of the Nematode *Caenohabditis elegans*," *Dev. Biol.* **100**, 64-119.

[Tabara et al., 96] Tabara, H., Motohashi, T. & Kohara, Y. (1996). "A Multi-Well Version of *in situ* Hybridization on Whole Mount Embryos of *Caenorhabditis elegans*," *Nucleic Acids Res.*, **24** 2119-2124.

[Tomita et al., 96] Tomita, M., Shimizu, K., Matsuzaki, Y., Miyoshi, F., Saito, K., Tanida, S., Yugi, K., Venter, C., Hutchison, C., "E-Cell: Software environment for whole cell simulation," *Bioinformtics*, Vol. 15, No. 1, 72-84, 1999.

[Ueda and Kitano, 98] Ueda, H. and Kitano, H., "A Generalized Reaction-Diffusion Simulator for Pattern Formation in Biological Systems," *Proc. of International Conference on Artificial Life*, 1998.

[Watson and Click, 53] Watson, J. D., and Click, F. H. (1953) "Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid," *Nature*, 171: 737-738.

[White et al., 86] White, J. G., Southgate, E., Thomson, J. N. & Brenner, S. (1986). "The Structure of the Nervous System of the Nematode *Caenorhabditis elegans*," *Phil. Trans. R. Soc.* **314**, 1-340.

[Yasuda, et al., 99] Yasuda, T., Bannai, H., Onami, S., Miyano, S., and Kitano, H., "Towards Automatic Construction of Cell-Lineage of *C. elegans* from Normarski DIC Microscope Images," *Proc. of Genome Informatics Workshop 1999*, Tokyo, 1999.