18

Efficiency, Robustness, and Stochasticity of Gene Regulatory Networks in Systems Biology: λ Switch as a Working Example

Xiaomei Zhu, Lan Yin, Leroy Hood, David Galas, and Ping Ao

Summary

Phage λ is one of the most studied biological models in modern molecular biology. Over the past 50 years, quantitative experimental knowledge on this biological model has been accumulated at all levels: physics, chemistry, genomics, proteomics, functions, and more. All of its components are known in great detail. The theoretical task has been to integrate its components to make the organism work quantitatively and in a harmonic manner. This tests our biological understanding, and would lay a solid foundation for further explorations and applications, which is an obvious goal of systems biology. One of the outstanding challenges in doing this has been the so-called stability puzzle of the λ switch; the biologically observed robustness and the difficulty in mathematical reconstruction based on known experimental values. In this chapter, we review the recent theoretical and experimental efforts on tackling this problem. An emphasis is put on the minimum quantitative modeling, where a successful numerical agreement between experiments and modeling has been achieved. A novel method, tentatively named stochastic dynamical structure analysis, emerged from such study, and it is also discussed within a broad modeling perspective.

Key Words: Phage λ; genetic switch; robustness; efficiency; cooperation; stochastic processes; dynamical landscape; systems biology.

1. Introduction

The completion of the Human Genome Project prompts biological and medical research into a new phase, one that has never been experienced

in biology. It is evident that a vast uncharted territory lies ahead, with tremendous promise in store (1). Great questions with important conceptual and practical implications have been asked and discussed (2–4). Speculations on the general principles underlying those great questions, and general methodologies to solve them, have been extensively debated since the beginning of this century (5,6). One of the authors of this chapter has been steadily promoting such exposition and contributing to this trend (7). Such efforts are needed not only as "a call to arms," they also help to define the various emerging fields. Nevertheless, in practical research, a full range of endeavors has to be explored. New tools will be invented to solve new problems and to take on old problems. In this chapter, we therefore turn our attention to the other side of consideration, not as an indication to underestimate the value of grand themes, but as an example to balance the grandeur. Instead of asking general questions and receiving limited answers, we wish to ask limited questions on a limited system and to find as complete answers as possible, along with a few general answers. We have been attempting this for the past few years. Such a methodology has been very effective since the dawn of modern science, and was first exemplified by Galileo. Specifically, we will focus our attention to the robustness and stability of a genetic switch (8,9) in phage λ, arguably the biological model that jump-started modern molecular biology (10).

In the modern information age, switch-like structures are building blocks in all architectures. It is the realization of the binary digit, the unit of information, and the "atom" today. As biology has been increasingly viewed as an information science (11–13), it would be desirable to have a thorough understanding of this building block. Indeed, detailed analyses have demonstrated that the response of a complexity network is often dealt with by various switches (14) and that genetic networks were shown to have the computational ability (15). By drawing a close analogy to the integration circuitry in an electronic wiring board, this methodology has been successfully employed in the modeling of genetic regulation during the earlier developmental stages in sea urchins (16). Currently, the study of switching in biology has been ranging from responses to environmental changes (17,18), developmental biology (19–21), neural networks (22,23), physiological response (24,25), genetic regulation (26–29), signal transductions (30), memory effect (31,32), olfactory perception (33), synthetic biology (34), biotechnological applications (35–38), to photosynthesis (39) and many other areas (40–43). Even in cell cycle processes, if viewing such a process not as driving by a cycling engine, but as what is controlled by a traffic light, the switch-like structure is likely to play a dominant role (44–46). Switch has indeed established itself as one of the fundamental elements in biological processes and as a paradigm for both experimental and theoretical studies in biology.

Why then has so much effort been expended on studying a particular virus genetic switch, the λ switch? To paraphrase Ptashne (8), this is a fair question desiring a clarification at the beginning. After all, every case in biology is at least partly accidental and specific, the workings of every organism having been determined by its evolutionary history, and the

precise description we give of a process in one organism will probably not apply in detail to another. Thus, both robustness and stochasticity in biostructure must be included and carefully studied. This has been well illustrated in the context of the fundamental biological processes, such as those mentioned in the previous paragraph. As already indicated above, at various stages of development, depending in part on environmental signals, cells choose to use one or another set of genes, and thereby to proceed along one or another developmental pathway. It would be of great value to know what molecular mechanisms determine these choices. Hence, the λ life cycle is indeed a prototype for this problem, with the structure of feedback loops and the effect of stochasticity. In addition, we have a nearly complete understanding of all its parts; its genome was in fact known (47) long before the completion of the Human Genome Project, and the corresponding quantitative knowledge has been accumulated at all levels: physics, chemistry, DNA, protein, and functions (8,48,49). Despite such a long history of quantitative studies, the stability and robustness of the λ switch remained as one of the outstanding puzzles for computational biology at least until 2004 (9,50,51). The theoretical challenge has been to put all its components together as a harmonic working organism, one of the major tasks of systems biology.

In addition, one might wonder at the value of using quantitative and detailed modeling. Biological theories are generally known for their descriptive nature. For example, when Darwin presented his evolutionary theory, no single equation had been used. It was rather remarkable that though one of Darwin's main predictions, the age of Earth, was in direct conflict with known physics at his time, it was physics, not Darwin's theory that later went through a fundamental transformation to resolve this glaring contradiction, to the good of both physics and biology. Nevertheless, it would be wrong to conclude that a quantitative method would be of no use in biology. In fact, some subfields in biology, such as physiology and population genetics, are among the most mathematical in the natural sciences (3). As biology is becoming an information science (11–13), more subfields would be likely to do so in the future. The important question is: What would be the right framework of mathematical description (9,52)? It is true that an excessive use of mathematical language, which might be attractive to a modeler, generally does not enhance the understanding of a specific biological phenomenon (53). For example, with excessive parameters, any phenomenon can be described by a set of equations. Such a situation is not acceptable under Ockham's razor. The other extreme is to look for an effective description, with the hope of capturing the biological essence. The latter description is necessarily gross and qualitative, although extremely popular and particularly successful in biology. However, many features are obviously left behind by such an approach. It would be desirable to have a detailed quantitative study that can bridge those two approaches. The phage λ genetic switch provides precisely one of the excellent opportunities in biology to do so (54); one side is an on/off Boolean-type description for the genetic switch, and the other side is the detailed physical and chemical equations.

The rest of the review is organized as follows. Salient biological experimental studies on phage λ switch are summarized in section 2. Its key biochemical modeling elements are summarized in section 3. The stochastic dynamical structure analysis method is discussed in section 4, within the minimal quantitative model of phage λ. Calculated results and the comparison to biological data are discussed in section 5. In section 6, we summarize what has been done and place the minimum quantitative modeling methodology in a broader context. In section 7, the research effort on λ switch is put into an optimistic outlook.

2. Phage λ Genetic Switch

2.1. Phage λ Life Cycle

Bacteriophage λ is a virus that grows on a bacterium (8,55,56). It is one of the simplest living organisms. Almost all its parts have been known for the past 50 years. The genome of phage λ consists of a single DNA molecule wrapped in a protein coat. Upon infection of the host *Escherichia coli* cell, the phage λ injects its genome inside the bacterium and leaves the protein coat outside. Inside the bacterium, it chooses one of two modes of growth (Figure 1). Phage λ uses molecular-genetic apparatus of the cell for running and executing its own ontogenetic

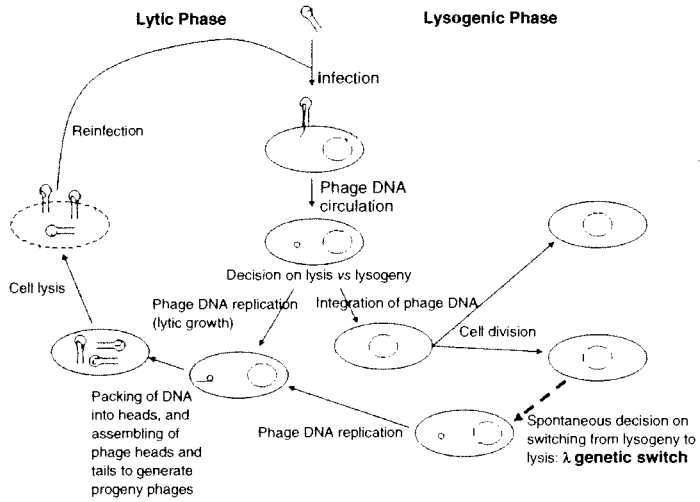

Figure 1. The schematic diagram of phage λ life cycle. The stochastic dynamics of a rare spontaneous decision on switching from the lysogenic phase to the lytic phase is the focus of this chapter. The induction by UV light and other SOS processes are not considered here.

subprograms to produce new λ phage particles, resulting in the lysis of the cell. Or, it can establish dormant residency in the lysogenic state, integrating its genome into the DNA of its host and replicate as a part of the host genome. In these two different life cycles, different sets of phage genes are expressed as a result of molecular interactions. Realistic modeling of the robustness and stability of such a process has remained one of the most important challenges in biocomputation and bioinformatics.

Because the analogous molecular interactions to phage λ are likely to underlie many developmental (8) and epigenetic (56) processes, one wishes to acquire deep understanding of the regulation of major biologic functions on the molecular level through the study of the genetic switch of phage λ. One of these functions is the programming of the epigenetic states; the ways the phage decides if it is going to follow a lysogenic or lytic growing state. Over the past five decades, extensive biological investigations have provided a fairly good qualitative picture in this respect. There exists a plausible scenario to guide the understanding of the experimental observations (8).

The maintenance and operation of the genetic switch is another function performed by the gene regulatory network (8,55,56). The phage growing in the lysogenic state remains latent, unless it is provoked. For example, switching to the lytic state happens when a signal is sent to activate RecA proteins, which cleave CI monomer, sending the phage into lytic growth. It is observed that the phage λ genetic switch is both highly stable and highly efficient. When the phage grows in a lysogenic state, it remains latent for many generations. Spontaneous induction happens less than once in 1 million cell divisions (Figure 1). Once the phage is exposed to an appropriate signal, it changes to the lytic state at a rate of almost 100%. Such a coexistence of stability and efficiency of the genetic switch in phage λ has been considered a mystery from the theoretical and mathematical modeling viewpoint.

2.2. Modeling Effort

There have been continuous mathematical and numerical activities on modeling phage λ. The rationale is rather straightforward; the biological functions should emerge as the systems properties from the model based on the molecular mechanism of phage regulatory elements and their independently measured parameters. The elegant physical–chemical model formulated by Shea and Ackers (48) for gene regulation of phage λ has become the foundation for later studies. However, soon afterward, Reinitz and Vaisnys (57) pointed out that the inconsistency between the theoretical results and experimental data may suggest additional cooperativity. Arkin et al. (58) performed stochastic simulation on phage λ development for the decision of lysogeny in the very early stage, demonstrating that this process is stochastic. Recently, Aurell and Sneppen (59) analyzed the robustness of phage λ genetic switch, using a method based on the Onsager–Machlup functional (58), and concluded that their theoretical analysis could not reproduce the robustness of the

phage λ genetic switch. Their further study confirmed earlier results (51). Similar modeling was found mathematically from a different perspective (50).

The coexistence of the switch stability and switching efficiency is an apparent inconsistency for the following reasons. The lysogenic state is exceptionally stable. The fluctuations in the growth environment, the so-called extrinsic fluctuations, and the intrinsic fluctuation in the genetic switch, which are caused by the discrete nature of chemical reactions, do not easily and accidentally flip the switch. Then when the phage is "threatened," how can the switching process become so complete with so little outside intervention? The question about internal inconsistencies in these models naturally arises; whether the easily operable induction, or highly efficient switching, in Shea and Ackers' work (48) is a result of sacrificing the robustness of the genetic switch. Phrased differently, if a model were so constructed that it faithfully reproduces the observed robustness of the genetic switch, would it lose the efficiency of the switch? Undoubtedly, a credible model of phage λ should reproduce the properties of robustness, stability, and efficiency of the genetic switch simultaneously. From such a model we should also be able to calculate the observed quantities of phage development, such as the protein numbers and lysogenization frequencies. We hope to show that a foundation for such a mathematical framework against the experimental data is there, thanks to recent theoretical efforts on phage λ (9,48,57–59).

2.3. Modeling Strategy

Our procedure is to first summarize a minimal quantitative model for the phage λ genetic switch, a model that is motivated by both first principles and biological observations. We then ask the question of whether or not this minimal modeling can be successfully used to quantitatively reproduce various experimental results, and whether it is qualitatively correct in biology. If successful, the necessary modifications of molecular parameters in the modeling may be viewed as the *in vivo* and *in vitro* differences. Additional or different molecular processes inside a cell should be responsible for such differences. Some of them may be identifiable by current experimental techniques. If the answer to the above question would be negative, we would conclude that the minimum quantitative modeling would not be enough. More biological causes should be looked for instead. We will show that the answer so far is positive. By combining a newly developed powerful nonlinear dynamics analysis method, which takes the stochastic force into account (61–63) and classifies the stochastic dynamical structure into four different elements, with the previously established physical–chemical model (48), a novel mathematical framework was formulated to calculate the following quantitative characteristics of epigenetic states and developmental paths (9,64): the protein numbers in one bacterium, the protein number distributions, the lifetime of each state, and the lysogenization frequencies of mutants using the wild type as reference. We should emphasize that our review is focused on a specific biological system, though we have made an effort to put such work in perspective.

3. Towards Quantitative Modeling

3.1. Binding Configurations

The genetic switch controlling and maintaining the function of phage λ consists of two regulatory genes, *cI* and *cro*, and the regulatory regions O_R and O_L on the λ DNA. Established lysogeny is maintained by the protein CI, which blocks operators O_R and O_L, preventing transcription of all lytic genes, including *cro* (8,55,56). In lysogeny, the CI number functions as an indicator of the state of the bacterium; if DNA is damaged, e.g., by UV light, the protease activity of RecA is activated, leading to degradation of CI. A small CI number allows for transcription of the lytic genes, starting with *cro*, the product of which is the protein Cro.

The decision making, or the switching, is centered on operator O_R, and consists of three binding sites, O_{R1}, O_{R2}, and O_{R3}, each of which can be occupied by either a Cro dimer or a CI dimer (55,56). As illustrated in Figure 2, these three binding sites control the activity of two promoters P_{RM} and P_R for *cI* and *cro* transcriptions, respectively. The transcription of *cro* starts at P_R, which partially overlaps O_{R1} and O_{R2}. The transcription of *cI* starts at P_{RM}, which overlaps O_{R3}. The affinity of RNA polymerase for the two promoters, and subsequent production of the two proteins, depends on how Cro and CI bound to the three operator sites, and thereby establishes lysogeny with approximately 500 CI molecules per bacterium. If, however, the CI number becomes sufficiently small, the increased production of Cro flips the switch to lysis.

There have been numerous quantitative experimental studies on the stability in the switching of bacteriophage λ. Recently, the frequency of spontaneous induction in strains deleted for the *recA* gene has been

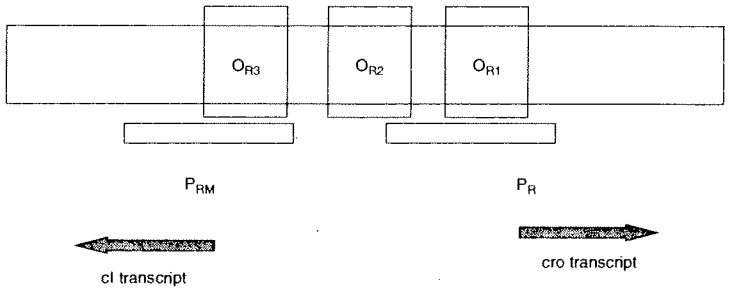

Figure 2. The O_R controlling region of the phage λ genetic switch. The mathematical studies (9,64) indicate that the cooperative binding of CI dimers at the O_{R1} and O_{R2} sites is a key to the robustness of the genetic switch. Such a cooperative binding enhances the positive CI feedback. When the CI positive feedback is turned on by the existing CI dimers, CI proteins are synthesized. The phage evolves to the lysogenic state. Otherwise, Cro proteins are synthesized and the phage evolves to the lytic state.

reported independently by three groups (65–67), which was reviewed by Aurell et al. (67). They all confirmed two earlier important observations: that there is a switching behavior and that the switch is stable. In addition, they all obtained consistent numerical values for the switching frequency, in spite of the use of different strain backgrounds done on different continents and at different times. However, computational and mathematical attempts to quantitatively understand this behavior have not been successful, even permitting the possibility that the wild type may be more stable (59,67).

More recent data (9, 64) suggest that the wild type may be two orders of magnitude more stable than previously observed (66); the switching rate to the lytic state may be less than 4×10^{-9} per minute. In addition to the call for more experimental studies, this puts the theoretical modeling in a more challenging position. This wild-type data was used as the main input to further fix the model in the works of Zhu et al. (9). The previous data were also discussed to illustrate a pronounced exponential sensitivity in such a modeling, which is summarized in the following sections.

The CI and Cro protein molecules in the cell are assumed to be in homeostatic equilibrium. There are not always the same numbers of CI and Cro dimers bound to the operators at any particular time. These numbers are fluctuating, and the equilibrium assumption should give the size of these fluctuations. The key inputs are CI and Cro dimerization constants, and the Gibbs free energies for their bindings to the three operator sites O_R1, O_R2, and O_R3 (68–74) (*see* the legends of Tables 1 and 2 for a more detailed description).

Following Ackers et al. (75) and Aurell et al. (67), we encode a state s of CI and/or Cro bound to O_R by three numbers (i,j,k) referring to O_{R3}, O_{R2}, and O_{R1}, respectively. The coding for s is 0 if the corresponding site is free, 1 if the site is occupied by a CI dimer, and 2 if the site is occupied by a Cro dimer. The probability of a state s with i(s) CI dimers and j(s) Cro dimers bound to O_R is in the grand canonical approach of Shea and Ackers (48)

$$p_R(s) = Z^{-1} [CI]^{i(s)} [Cro]^{j(s)} [RNAp]^{k(s)} \exp(-\Delta G(s)/RT). \qquad (1)$$

For example, if CI occupies O_{R1} and Cro occupies O_{R2} and O_{R3}, we have i(s) = 1, j(s) = 2, k(s) = 0, and $p_R(s) = p_R(221)$. RNA polymerase (RNAp) can occupy either O_{R1} and O_{R2}, or O_R; and O_{R3}, not other configurations. There are a total 40 states represented by *s* (Table 1). The normalization constant *Z* is determined by summing over *s*: $Z = \Sigma_s [CI]^{i(s)} [Cro]^{j(s)} [RNAp]^{k(s)} \exp(-\Delta G(s)/RT)$. Here, [] denotes the corresponding protein dimer concentration in the bacterium, ΔG(s) the binding energy for binding configuration s, R the gas constant, and T the temperature.

3.2. Deterministic Model

We further simplify the expression of $p_R(s)$ by noticing that CI and Cro control the operator (8,55,56). If O_{R1} and O_{R2} are unoccupied by either CI or Cro, RNAp binds to them with a probability determined by RNAp binding energy. The idea that RNAp first binds to O_{R1} and O_{R2}, followed

Table 1. The 40 configurations corresponding to right operator.

State	P_{RM}	O_R3	O_R2	O_R1	P_R	i(s)	j(s)	k(s)
1						0	0	0
2				R_2		1	0	0
3			R_2			1	0	0
4		R_2				1	0	0
5			R_2	R_2		2	0	0
6		R_2	R_2			2	0	0
7		R_2		R_2		2	0	0
8		R_2	R_2	R_2		3	0	0
9				C_2		0	1	0
10			C_2			0	1	0
11		C_2				0	1	0
12			C_2	C_2		0	2	0
13		C_2	C_2			0	2	0
14		C_2		C_2		0	2	0
15		C_2	C_2	C_2		0	3	0
16			C_2	R_2		1	1	0
17		C_2		R_2		1	1	0
18		C_2	C_2	R_2		1	2	0
19		C_2	R_2			1	1	0
20			R	C_2		1	1	0
21		C_2	R_2	C_2		1	2	0
22		R_2	C_2			1	1	0
23		R_2		C_2		1	1	0
24		R_2	C_2	C_2		1	2	0
25		C_2	R_2	R_2		2	1	0
26		R_2	C_2	R_2		2	1	0
27		R_2	R_2	C_2		2	1	0
28					RNA_P	0	0	1
29		R_2			RNA_P	1	0	1
30		C_2			RNA_P	0	1	1
31	RNA_P		R_2			1	0	1
32	RNA_P		R_2	R_2		2	0	1
33	RNA_P		R_2	C_2		1	1	1
34	RNA_P					0	0	1
35	RNA_P			R_2		1	0	1
36	RNA_P			C_2		0	1	1
37	RNA_P		C_2			0	1	1
38	RNA_P		C_2	R_2		1	1	1
39	RNA_P		C_2	C_2		0	2	1
40	RNA_P				RNA_P	0	0	2

R_2 stands for CI (λ repressor) dimer and C_2 for Cro dimer.
(*Sources*: Darling et al. (74), Ackers et al. (75), Capp et al. (76))

by blocking CI and Cro binding, is excluded based on the assumption that only CI and Cro controls the regulatory behavior. In addition to experimental observation, this assumption is justifiable if the time scale associated with CI and Cro binding is shorter than the RNAp binding. Except for an overall constant, which we include into the rate of transcription, the RNAp binding is no longer relevant. We therefore take it out of the expression $p_R(s)$. The total number of states is reduced to 27. This simplification was first used by Aurell and Sneppen (59). We will drop the subscript R for binding probability p_R. We should point out

Table 2. Parameters used in the modeling.

RT	**0.617 kcal/mol**
Effective bacterial volume	0.7×10^{-15} **l**
E_{cro}	20
E_{cI}	1
T_{RM}	0.115/s
$T_{RM}{}^a$	0.0105/s
T_R	0.30/s
τ_{CI}	2.9×10^3 s
τ_{Cro}	5.2×10^3 s
Converting factor between protein number and concentration	1.5×10^{-11}
***in vitro* free energy differences for wild-type λ**	
ΔG (001)	**−12.5 kcal/mol**
ΔG (010)	**−10.5 kcal/mol**
ΔG (100)	**−9.5 kcal/mol**
ΔG (011)	**−25.7 kcal/mol**
ΔG (110)	**−22.0 kcal/mol**
ΔG (111)	**−35.4 kcal/mol**
ΔG (002)	**−14.4 kcal/mol**
ΔG (020)	**−13.1 kcal/mol**
ΔG (200)	**−15.5 kcal/mol**
ΔG(cooperative)	**−2.7 kcal/mol**
dimerization energy	
ΔG_{CI2}	**−11.1 kcal/mol**
ΔG_{Cro2}	**−7.0 kcal/mol**
***in vitro* free energy differences for O_R3' binding**	
ΔG (100)	**−10.5 kcal/mol**
ΔG (200)	**−13.7 kcal/mol**
***in vivo* free energy differences—*in vitro* free energy differences**	
ΔG_{CI}	**−2.5 kcal/mol**
ΔG_{Cro}	**−4.0 kcal/mol**
ΔG (cooperative)	**−3.7 kcal/mol**

CI dimer affinities to O_{R3}, O_{R2}, and O_{R3} are from Darling et al. (73,74). Cro dimer affinities to O_{R1} and O_{R2} are from Takeda et al. (69,79), Jana et al. (71), Kim et al. (70), and Aurell et al. (67). The CI dimerization energy is taken from Koblan and Ackers (80), Cro dimerization energy from Jana et al. (71,72). The bacterial volume is taken from Bremmer and Dennis (99). E_{CI} and E_{Cro} are taken from Shean and Gottesman (100). Ringquist et al. (101), and Kennell and Riezman (102). The *in vitro* parameters have been summarized by Aurell et al. (67), which we largely follow. However, we here point out two differences: (a) our effective bacterial volume, estimating from the typical size of the bacterium, assuming a tube of about 0.7 μm in diameter and 2 μm in length, is approximately a factor 3 smaller; and (b) the normalization factor for the concentrations, calculated against the number of water molecules, is approximately a factor of 60 smaller. The difference between *in vivo* and *in vitro* values is consistent with qualitative experimental observation (8). The wild-type data (9) is used to determine the *in vivo* and *in vitro* difference.

that previous experimental and theoretical results had been concisely reviewed by Aurell et al. (67), whose convention we shall follow.

The dimer and monomer concentrations are determined by the formation and de-association of dimers, which gives the relation of dimer concentration to the total concentration of proteins as follows:

$$[CI] = [N_{CI}]/2 + \exp(\Delta G_{CI}/RT)/8 - ([N_{CI}]\exp(\Delta G_{CI}/RT)/8 + \exp(2\Delta G_{CI}/RT)/64)^{1/2} \quad (2)$$

Here, $\Delta G_{CI} = -11.1$ kcal/mol is the dimer association free energy for CI.

A similar expression for [Cro] is as follows:

$$[Cro] = [N_{Cro}]/2 + \exp(\Delta G_{cro}/RT)/8 - ([N_{Cro}]\exp(\Delta G_{Cro}T)/8 + \exp(2\Delta G_{Cro}T)/64)^{1/2}. \quad (3)$$

Here, $\Delta G_{Cro} = -7$ kcal/mol is the dimer association free energy for Cro. $[N_{CI}]$ and $[N_{Cro}]$ are the monomer concentrations of CI and Cro, respectively.

CI and Cro are produced from mRNA transcripts of CI and Cro, which are initiated from promotor sites P_{RM} and P_R. The rate of transcription initiation from P_{RM} when stimulated by CI bound to O_{R2} is denoted T_{RM}, and when not stimulated it is denoted T^u_{RM}. The number of CI molecules produced per transcript is E_{cI}. The overall expected rate of CI production is as follows:

$$f_{CI}(N_{CI}, N_{Cro}) = T_{RM}E_{cI}[p(010) + p(011) + p(012)] + T^u_{RM}E_{cI}[p(000) + p(001) + p(002) + p(020) + p(021) + p(022)]. \quad (4)$$

Here, N_{CI} and N_{Cro} are the protein numbers for CI and Cro inside the bacterium respectively. The converting factor between the protein concentration and the corresponding protein inside the bacterium is listed in Table 2. Similarly, the overall expected rate of Cro production is

$$f_{Cro}(N_{CI}, N_{Cro}) = T_R E_{cro}[p(000) + p(100) + p(200)]. \quad (5)$$

We use T_{RM}, E_{cI}, E_{cro}, and T^u_{RM} from Aurell and Sneppen (59), which were deduced from the resulting protein numbers in lysogenic and lytic states.

The free energies ΔG(s) are determined from *in vitro* studies, that is, they are obtained outside of the living bacterium. The *in vivo* conditions, inside a living bacterium, could be different. The measured protein–DNA affinities could sensitively depend on the ions present in the buffer solutions, as well as other factors. This observation will be important in our comparison between theoretical results and experimental data. On the other hand, the *in vivo* effects of such changes should be compensated for, e.g., changed KCl concentrations are attributable to putrescine (76), other ions, and crowding effects (77). We note that Record et al. (77) already observed that there may exist a significant difference between *in vivo* and *in vitro* molecular parameters. The data quoted in Darling et al., (73,74) was obtained at KCl concentration of 200 mM, which resembles *in vivo* conditions. Therefore, though we expect a difference between the *in vivo* and *in vitro* data, the difference may not be large, typically within 20%–30% of the *in vitro* values.

The mathematical model that describes the genetic regulation in Figure 2 is a set of coupled equations for the time rate of change of numbers of CI and Cro in a cell (57):

$$dN_{CI}(t)/dt = F_{CI}(N_{CI}(t), N_{Cro}(t))$$
$$dN_{Cro}(t)/dt = F_{Cro}(N_{CI}(t), N_{Cro}(t)), \quad (6)$$

where the net production rates are

$$F_{CI} = f_{CI}(N_{CI}, N_{Cro}) - N_{CI}/\tau_{CI}$$
$$F_{Cro} = f_{Cro}(N_{CI}, N_{Cro}) - N_{Cro}/\tau_{Cro}. \quad (7)$$

Equations (6) and (7) represent the minimum deterministic model. Here, dN/dt is the rate N changes. The production terms f_{CI} and f_{Cro} are functions of CI and Cro numbers in the bacterium. With no Cro in the system, the curve of f_{CI} versus CI number has been experimentally measured (78). As reviewed in Aurell et al. (67) these measurements are consistent with the best available data on protein–DNA affinities (69,70,79), dimerization constants (80), initiation rates of transcriptions of the genes, and the efficiency of translation of the mRNA transcripts into protein molecules. The decay constant τ_{CI} is an effective lifetime, proportional to the bacterial lifetime, as CI molecules are not actively degraded in lysogeny, whereas τ_{Cro} is approximately 30% smaller (81). We comment that there is considerably more experimental uncertainty in the binding of Cro, both to other Cro and to DNA, than the binding of CI; e.g., the work of Darling et al. (73,74). As a minimal mathematical model of the switch, we take τ_{CI} and τ_{Cro} from data and deduce f_{CI} and f_{Cro} at a nonzero number of both CI and Cro with a standard set of assumed values of all binding constants, which are summarized by Aurell et al. (67) and are adopted here (Table 2, with differences in cell volume and converting factor, as well as the *in vivo* and *in vitro* differences).

3.3. Positive and Negative Feedbacks: *In Vivo* versus *In Vitro*

Both positive and negative feedbacks are employed in this genetic switch. For CI (Cro), it has a positive feedback effect on itself and a negative feedback effect on the production of Cro (CI) (8). Evidently, these feedbacks are systems effects; breaking them down into disintegrated parts would cause the feedback effects to disappear. They emerge only when a proper integration is done. Such a systems effect is well known in engineering (82).

In establishing the minimum deterministic model, another major implicit assumption is on the time scales. We have assumed that the dimerization process is a fast process on the scale of Cro and CI production, and hence can be treated as algebraic constraints. The dimerization has been subjected to continuous experimental (83) and theoretical (84) studies. It was concluded by *in vitro* experiments (83) that the Cro dimerization is slower than that of CI. Nevertheless, the Cro dimerization time is on the order of fractions of a minute (83), which is much smaller than the typical time, on the order of 20 min, used in our modeling (Table 2). We may be able to apply the useful quasi-steady-state approximation (85). Thus, the algebraic constraints appear to be a reasonable assumption for such a minimum modeling as that used in previous works (9,59,64,67). Other cellular processes have also been implicitly assumed

to be fast. All their residual effects will be treated as extrinsic stochastic effect contribution, and will be incorporated into the minimum modeling in the name of intrinsic versus extrinsic noises, which will be discussed in the following sections.

We further specify the meaning of minimum deterministic modeling. First, such a model should be viewed as what the system might be, not as what it must be. Many features are not explicitly contained in it, such as the nonspecific binding (86) and the looping (87,88). The nonspecific binding was already demonstrated not to be crucial, but the looping may well be, which we will come back to later. Nevertheless, we point out that by assuming the minimum deterministic modeling, we tentatively and tactically assume it has captured all the essential features of the λ switch by aggregating molecular processes around it. In doing so, it suggests another understanding of the difference between *in vivo* and *in vitro*; all the parameters we adopted from *in vitro* measurement would indeed take a different value *in vivo*, because there exist numerous other biological processes inside the cell that contribute to this difference. If the minimum model is essentially correct, it should be able to account for experimental data in a quantitative manner, along with predictions that need to be further tested. We will show that we have indeed achieved this goal after several decades of theoretical efforts.

4. Stochastic Dynamical Modeling

4.1. Minimum Quantitative Model

Stochasticity is ubiquitous in biology. For this modeling, it is particularly easy to motivate it. If the numbers of CI and Cro were macroscopically large, then equation (1) would be an entirely accurate description of the dynamics, because the fluctuation in numbers is an order of $N^{1/2}$ and the correction is an order of $1/N^{1/2}$, which would be negligibly small when N was very large. However, the numbers are only in the range of hundreds. Hence, the fluctuation is not negligible. The actual protein production process is influenced by many chance events, such as the time it takes for a CI or a Cro in solution to find a free operator site, or the time it takes an RNA polymerase molecule to find and attach itself to an available promoter, suggesting more stochastic sources. As a minimal model of the network with finite N noise, we therefore consider the following system of two coupled stochastic differential equations, with two independent standard Gaussian and white noise sources:

$$\begin{aligned} dN_{CI}/dt &= F_{CI} + \zeta_{CI}(t) \\ dN_{Cro}/dt &= F_{Cro} + \zeta_{Cro}(t) \end{aligned} \tag{8}$$

We further assume that the means of the noise terms are zero, i.e., $\langle\zeta_{CI}(t)\rangle = \langle\zeta_{Cro}(t)\rangle = 0$, with the variance

$$\begin{aligned} \langle\zeta_{CI}(t)\zeta_{CI}(t')\rangle &= 2D_{CI}\delta(t - t') \\ \langle\zeta_{Cro}(t)\zeta_{Cro}(t')\rangle &= 2D_{Cro}\delta(t - t') \\ \langle\zeta_{CI}(t)\zeta_{Cro}(t')\rangle &= 0 \end{aligned} \tag{9}$$

Equations (8) and (9) consist of the present minimum quantitative model. Here, the symbol $< \ldots >$ denotes the average over noise. Equation (9) defines a 2×2 diffusion matrix D. The noise strength may contain contributions from the production and decay rates, assuming each is dominated by one single independent reaction, as used by Aurell and Sneppen (59). Such a noise may be called the "intrinsic" noise. Other noise sources, "extrinsic" noises, also exist (89–92). We treat the noise to incorporate both intrinsic and extrinsic sources: All are assumed to be Gaussian and white. The consistency of this assumption should be tested experimentally, as we will do in the following paragraphs. Certain probability events, however, may not behave as Gaussian and white in the present context of modeling, which can be determined by separate biological experiments, such as the $p_{RM}240$ mutation (9) to be discussed later in the text.

It has been demonstrated (61–63) that there exists a unique decomposition, such that the stochastic differential equation, equation (8), can be transformed into the following form, the four dynamical element structure:

$$[\Lambda(\mathbf{N}) + \Omega(\mathbf{N})]d\mathbf{N}/dt = -\nabla U(\mathbf{N}) + \xi(t), \tag{10}$$

with the semipositive definite symmetric 2×2 matrix Λ defining the dissipation (degradation), the antisymmetric 2×2 matrix Ω defining the transverse force, the single-valued function U defining the potential landscape, and the noise vector $\xi(t)$, and the two-dimensional vectors:

$$\begin{aligned} \mathbf{N}^{\tau} &= (N_{CI}, N_{Cro}); \\ \nabla &= (\partial/\partial N_{CI}, \partial/\partial N_{Cro}); \\ \xi^{\tau} &= (\xi_{CI}, \xi_{Cro}), \end{aligned} \tag{11}$$

Here, τ means the transpose of the vector. The connection between the noise ξ and the matrix Λ is similar to that of ζ and D of equation (9):

$$\begin{aligned} \langle\xi(t)\rangle &= 0 \\ \langle\xi_{CI}(t)\xi_{CI}(t')\rangle &= 2\Lambda_{CI}\delta(t - t') \\ \langle\xi_{Cro}(t)\xi_{Cro}(t')\rangle &= 2\Lambda_{Cro}\delta(t - t') \\ \langle\xi_{CI}(t)\xi_{Cro}(t')\rangle &= 0 \end{aligned} \tag{12}$$

The decomposition from equations (8) and (9) to equation (10) and (12) is determined by the following set of equations:

$$\nabla \times [(\Lambda + \Omega)\mathbf{F}] = 0 \tag{13}$$

$$(\Lambda + \Omega)D(\Lambda - \Omega) = \Lambda. \tag{14}$$

One may solve for Λ, Ω in terms of $\mathbf{F} = (F_{CI}, F_{Cro})^{\tau}$ and D from equations (13) and (14). Indeed, this can be formally done. Once Λ and Ω are known, the requirement that equation (10) can be reduced to equation (8) gives $(\Lambda + \Omega)\,\mathbf{F} = -\nabla U(\mathbf{N})$, which is used to obtain U. In general, this decomposition is an involved mathematical and numerical endeavor. Further simplification follows from the simplification of friction matrix. Typically, the diffusion matrix D is unknown biologically. There are not

enough measurements to fix the noise explicitly. Therefore, we may treat the semipositive, definite symmetric matrix A as parameters to be determined experimentally. In our calculation, we assume that D is a diagonal matrix. Following from equation (14), A is a diagonal matrix for two-dimensional case. The experimentally measured fraction of $recA^{-1}$ lysogens that have switched to lytic state is used to determine the elements of A.

Here, we would like to give an intuitive interpretation of the mathematical procedure. Equation (8) corresponds to the dynamics of a fictitious massless particle moving in two-dimensional space formed by the two protein numbers N_{CI} and N_{Cro}, with both deterministic and random forces. It is easy to check that, in general, $\nabla \times \mathbf{F}(\mathbf{r}) \neq 0$ and $\nabla \cdot \mathbf{F}(\mathbf{r}) \neq 0$. Therefore, $\mathbf{F}(\mathbf{r})$ cannot be simply represented by the gradient of a scalar potential, due to both the force transverse to the direction of motion and force of friction. The simplest case in two-dimensional motion when both transverse force and friction exist is an electrically charged particle moving in the presence of both magnetic and electric fields, which is precisely in the form of equation (10).

Proceeding from equation (10), we note that we may interpret the semi-positive definite symmetric A matrix as the friction matrix, and the antisymmetric matrix Ω as the result of a "magnetic" field. The friction matrix represents the dissipation in physics. It is analogous to the degradation in biology. The scalar function U takes the role of a potential function, which would determine the final steady distribution of the phage. The global equilibrium will be reached when the final distribution function is given by

$$\rho(N_{CI}, N_{Cro}) = \exp(-U(\mathbf{N}))/\int dN_{CI} \int dN_{Cro} \exp(-U). \qquad (15)$$

The potential U, the landscape of the system, is depicted in Figure 3 (*cf* Figure 5).

The phage sees two minima and one saddle point in the potential landscape. Those two minima correspond to the lytic and lysogenic states. Once the phage is at one of the minimum, the probability rate for it to move into another minimum is given by the Kramers rate formula in the form (93,94):

$$P = \omega_0 \exp(-\Delta U_b) \qquad (16)$$

with the potential barrier height $\Delta U_b = (U_{saddle} - U_{initial\ minimum})$, the difference in potential between the saddle point and the initial minimum, and the time scale, the attempt frequency ν_0, determined by the friction, the curvatures of potential and values of transverse force around the saddle and the local minimum. We remark here that the attempt frequency is, in general, a complicated function of dynamical quantities in equation (10). Its form will be determined empirically in this chapter. We refer readers to Hanggi et al. (94) for the general mathematical discussions.

4.2. Stochastic Dynamical Structure Analysis

Equation (10) gives the dynamical structure of the gene regulatory network in terms of its four dynamical components: the friction, the potential gradient of the driving force, the transverse force, and the stochastic force. Such a dynamical structural classification serves two main purposes. It provides a concise description for the main features of the genetic switch by itself and it provides a quantitative measure to compare different gene regulatory networks, for instance, between the wild phage and its mutant. Such an analysis of experimental data based on equation (10) will be tentatively named the dynamical structure analysis.

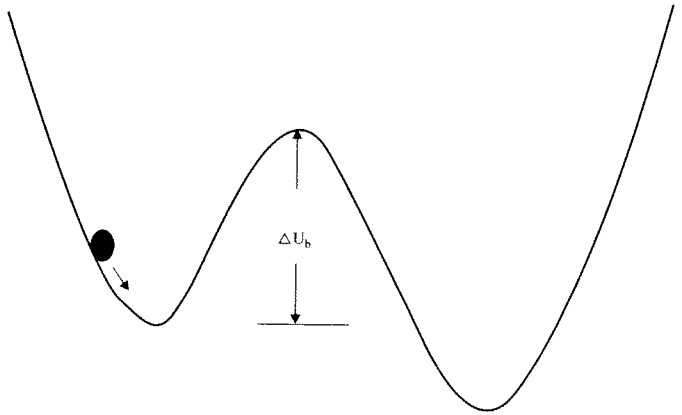

Figure 3. Illustration of the dynamical structure of a genetic switch. The dynamic state of the network is represented by a particle whose position is given by instantaneous protein numbers. The potential function maps a landscape in the protein number space. For a genetic switch, there are two potential minima corresponding to two epigenetic states. The area around each of the minima forms the attractive basin. The state of the network always tends to relax to one of the minima. The fluctuation may bring the network from one minimum to another with a rate given by Kramers rate formula (93,94).

The potential may be interpreted as the landscape map of the phage development. Each of the epigenetic states is represented by a potential minimum and its surrounding area forms an attractive basin (Figure 3). The dissipation represented by the friction gives rise to the adaptivity of the phage in the landscape defined by the potential. The phage always has the tendency to approach the bottom of the nearby attractive basin. The potential change near the minimum, together with the friction, gives the time scale of relaxation: The time it takes to reach equilibrium after the epigenetic state is perturbed. Once we know the friction and the potential around the minimum, we have a good grasp of the relaxation time, $\tau = \eta/U''$; here η is the strength of friction and U'' is the second derivative of potential, both in one-dimensional approximation along a relevant axis. The relaxation time is independent of the amplitude of the perturbation near the potential minimum when U'' is a constant.

Two remarks are in order here. First, the meaning of friction matrix is the same as in mechanics; if there is no external driving force, the system tends to stop at its nearby minimal position. The closest corresponding

concept in biology is "degradation." There is always a natural protein state under given conditions. Second, it turns out that the transverse force is not a dominant factor in the present switch-like behavior. However, its existence is the necessary condition for oscillatory biological behaviors, which will not be discussed further in this chapter.

Another time scale provided by the potential is the lifetime of the epigenetic state, which is given by the Kramers rate formula, equation (16), through the potential barrier height. Such a scale measures the stability of the epigenetic state in the presence of a fluctuating environment. In the case of phage λ, the lifetime for the lysogenic state is very long, unless the phage is mutated at its operator sites. When the phage is provoked, the height of the potential barrier separating lysogenic and lytic states is reduced. The lifetime of the lysogenic state is drastically reduced because of its exponential dependence on the barrier height and switching takes place. Looking at it from a different angle, the stochastic force gives the phage ability to search around the potential landscape by passing through saddle points, and it drives the switching event. The Kramers rate formula is a quantitative measure of this optimization ability.

5. Quantitative Comparison Between Theory and Experiment

5.1. Determining *In Vivo* Parameters

First, we need to decide the free energies to be used in the theoretical model. Without exception, all the binding energies measured so far for phage λ are determined from *in vitro* studies. The difference between the *in vivo* condition and the *in vitro* condition could include the ion concentration in the buffer solutions and the spatial configuration of the genomic DNA, for instance, looping (95–97). The relative large change of the cooperative energy from *in vitro* to *in vivo* in Table 1 may be partly due to the looping effect, though there is no direct consideration of looping in the present model. We note that in the *in vivo* conditions, all the operators are in the same kind of environment, including the ion condition and the DNA configuration. The reason for the latter is that the operators are located close to each other in the genome. If there is a bending of the genomic DNA that increases or decreases DNA–protein bindings, these closely located and short operator sites will most likely experience the same amount of change. Therefore, we assume that in addition to the *in vitro* DNA–protein binding energy, overall binding energy differences are added to all the CI and Cro protein respectively:

$$\textit{in vivo}\ \text{binding energy for CI (Cro)} = \textit{in vivo}\ \text{binding energy for CI} + \Delta G_{CI}(\Delta G_{Cro}).$$

To determine $\Delta G_{CI}(\Delta G_{Cro})$, we need more experimental input than the *in vitro* measurement. To avoid unnecessary uncertainty in the model, we try to include a minimal number of parameters. The cooperative binding between two CI dimers is included. The cooperative bindings between two Cro dimers, between CI and Cro dimers, the unspecific CI and Cro bindings are not included. Our later calculation verifies that CI cooperative binding is essential to the genetic switch properties, while the bindings we ignore do not have significant influence on the calculated results. There are three parameters we need to adjust: the difference between *in vivo* and *in vivo* binding energy for CI (ΔG_{CI}), for Cro (ΔG_{Cro}) and for the cooperativity of CI dimers (ΔG(cooperative)). We first use the CI numbers of both wild-type and mutant $\lambda O_R 121$ to determine ΔG_{CI}, then we determine ΔG(cooperative) and ΔG_{Cro} by requiring that both the lytic and lysogenic states of wild type are equally stable, calculated from Kramers rate formula. The adjusted *in vivo* binding energies and other parameters we use for the modeling are given in Table 1. Using these adjusted parameters, the robustness of the phage's genetic switch is reproduced (shown in Figure 4).

The mutant $\lambda O_R 3'23'$ studied by Little et al. (66) was characterized by Hochschild et al. (98) for binding to $O_R 3'$. To produce the desired protein level, we found that the binding energy between $O_R 3'$ and Cro protein is 1.8 kcal/mol smaller than that of the $O_R 3$ and Cro protein, which is consistent with the result of Hochschild et al. The CI binding energy from $O_R 3$ to $O_R 3'$ is slightly increased, 1 kcal/mol, which is also consistent with the measurement.

We assume that friction matrix λ is a diagonal constant matrix. Similar to Aurell and Sneppen (59), we assume the stochastic fluctuations in equation (2) scale with the square root of protein number divided by relaxation time: $D_{CI} = \text{Const} \times \tau_{CI}/N_{CI,lysogen}$, and $D_{Cro} = \text{Const} \times \tau_{Cro}/N_{Cro,lysis}$, where $N_{CI,lysogen}$ is the CI number at the lysogenic state and $N_{Cro,lysis}$ is the Cro number at the lytic state. The constant is to be determined by experiments. In equation (7), we note that if the antisymmetric matrix Ω is small, that is, $|\det(\Omega)| << \det(\Lambda)$, then Λ is the inverse of D. We calculate Ω assuming $\Lambda = D^{-1}$ and find that, indeed, in the regions of concern, i.e., the potential valley connecting two potential minima through the saddle points, Ω is negligible. The final parameters we have used are

$$\Lambda_{11} = 0.056 \times \tau_{CI}/N_{CI,lysogen}$$
$$\Lambda_{22} = 0.040 \times \tau_{Cro}/N_{Cro,lysis}. \quad (17)$$

5.2. Stochastic Dynamical Structure Analysis of λ Switch

The original problem, described by equation (8), may be interpreted as a set of two-dimensional differential equations describing a particle motion, if we view the protein number N_{CI} and N_{Cro} as the coordinates and the particle position to be $(N_{CI}(t)\ N_{Cro}(t))$ at time t. There is a deterministic force $\mathbf{F}^T = (F_{CI}, F_{Cro})$ and a stochastic force acting on such a particle. The deterministic force has the characteristics of a friction, a potential force, and a transverse force at the same time. The decomposition we have discussed earlier, equation (10) allows us to separate these components. We discuss them here.

The wild-type phage λ and some of its mutants sees two minima and one saddle point in the potential energy landscape (Figure 5). Those two minima correspond to the lytic and lysogenic states (*cf* Figure 3). The

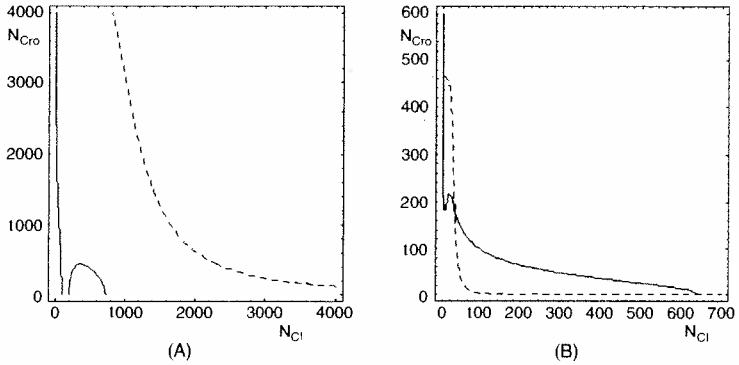

Figure 4. Lines of $d\langle N_{CI}\rangle/dt = 0$ (solid) and $d\langle N_{Cro}\rangle/dt = 0$ (dashed). Here, $\langle\rangle$ is the average to stochastic force, for (a) the wild-type phage, $\lambda O_R 321$, with parameters taken directly from *in vitro* measurement, and (b) the wild-type phage with parameters adjusted allowing *in vivo* and *in vitro* differences. For mutants $\lambda O_R 121$ $\lambda O_R 323$, *see* Zhu et al. (9). For b, these two lines have three intersections. These three fixed point in equation (2) coincide with the potential extrema, minima, and saddle point in equation (10).

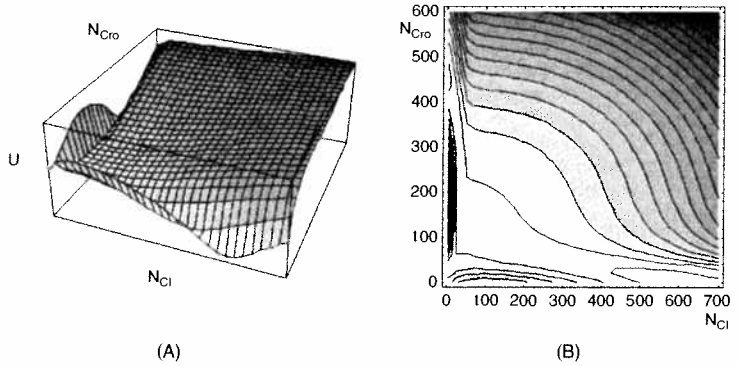

Figure 5. The potential U of wild-type phage plotted on a logarithmic scale (a) and as a contour map (b). There are two potential minima corresponding to the lysogenic and lytic states. Connecting these two states is a narrow potential valley. The highest point along this valley is the saddle point. The most probable state of the phage is at either of the potential minima. The fluctuation may bring the phage from the original potential minimum, moving along the valley and across the saddle point to reach another potential minimum. The rate for such a switching event is given by the Kramers rate formula.

positions of the potential minima give the average protein number for the lytic and lysogenic states. There is a relatively narrow valley connecting these two minima. The highest point along this valley is the saddle point. Because the areas with large potential are not easily accessible and the low-lying potential region forms a valley, we may visualize the potential along the valley and illustrate it in a one-dimensional graph, as shown in Figures 3 and 6.

The antisymmetric matrix Ω may be represented by a single scalar, B, along the z direction $\Omega\mathbf{F} = Bz \times \mathbf{F}$, assuming $x = N_{CI}$, $y = N_{Cro}$. The transverse field B for the wild type is obtained by numerically solving equation (13). This field is small except at the region along two axes. Along the two axes, the transverse B field has no effect because the motion is guided by the steep potential to a valley. Once the phage evolves away from origin, when both CI and Cro number are small, in the later development the transverse force may be taken out of the equation (10) without changing the dynamics of the phage. In both the calculation of the relaxation time and the lifetime of lysogenic state, we may ignore the transverse force for the above reason.

The protein number distributions of Cro and CI have also been calculated by Zhu et al. (9). We refer readers there for details, as well as for the analysis of other quantities, such as the robustness and stability.

5.3. Switch Efficiency

Efficiency is an important feature that, so far, has received relatively less attention in literature. We present the discussion in some detail here.

The analysis of robustness of phage λ genetic switch demonstrates that its epigenetic states are stable against the variations in parameters and robust against major changes in terms of mutations. Then how does the switching take place? From the theoretical point of view, there are two channels that the phage can be induced from lysogenic growth to lytic growth. In reality, phage seems to use both of these strategies. For clarity, we begin by discussing these two channels separately.

The first channel of induction is to increase the noise level of CI protein number, while keeping all the other conditions intact. Mathematically, it means to increase ζ_{CI} in equation (8) and D_{CI} in equation (9), while keeping all the other terms in equation (8) and equation (9) unchanged. The friction matrix Λ is changed through the decomposition procedure. As a result, the potential energy U is also changed. Therefore, for a different noise level, the phage moves in a different potential landscape. Such a change of noise level has a drastic effect. It changes the minima of the potential well of lysogens by making it shallower. As a good approximation, the barrier height of the lysogen potential well scales inversely with the noise strength. Doubling the noise level reduces the potential barrier by half. As a result, the increased noise level drastically decreases the lifetime of the lysogenic state, as shown in Figure 6. The lifetime of the lytic state, on the other hand, remains unchanged. The combination of these two changes in the potential landscape brings the phage to lytic growth efficiently.

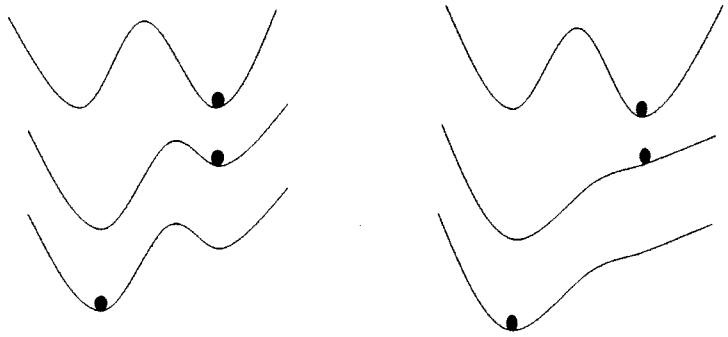

Figure 6. (A) Illustration of the switching mechanism from the current work. Before switching, the phage grows in lysogenic state. The potential barrier separating the lysogenic state and the lytic state is high. When *recA* is activated, this barrier is lowered. The lifetime of lysogenic state reduces drastically, and the phage switches to lytic state. (B) Switching mechanism of Shea and Ackers (48). In their work, fluctuation was not included. Switching was possible only when the lysogenic state is no longer a potential minimum. When stochastic effect is included, the switching happens when the lysogenic potential minimum becomes too shallow to confine fluctuation.

The second channel is through the deterministic terms in equation (8). For the deterministic terms, e.g., introducing CI monomer cleavage is equivalent to substitute N_{CI} in equation (8) with αN_{CI}, α is a factor that represents cleavage strength. If α is smaller than 0.02, we find that lysogenic state is no longer stable, i.e., no longer a potential minima. The interpretation of such a small α is that almost every CI monomer is cleaved. If α is small, let's say 0.1, meaning 90% CI monomers are cleaved, the lysogenic state is still stable, with a lifetime almost unchanged. Apparently, a uniform CI cleavage alone without introducing extra noise to CI levels is not an efficient way for induction.

Phage may have used both of these two channels. The second channel is obviously used because RecA cleaves CI monomers. The strong indication that the first channel is also used comes from the observations that without external stimulus, the $recA^+$ phage shows a much shorter lifetime for the lysogenic state compared to $recA^-$ phage. Such a significant reduction of lysogen lifetime without activating RecA proteins on an observable scale can be explained by doubling of the CI noise level. Figure 6 gives schematic explanations of the switching process.

In the early work by Shea and Ackers (48), stochastic effect was not included. In their model, even a shallow lysogenic potential minimum would confine the phage to continue growing in lysogenic state. Switching happens only when the lysogenic potential minimum disappears completely. For the parameters they used, they found that when 20% of CI monomers were cleaved, such a switch would happen. As pointed out by Aurell et al. (67), in these early works, the genetic switch modeling results do not show the observed robustness. After we require that the genetic switch should demonstrate the observed robustness, the disappearance of the lysogenic potential minimum is pushed down to 2%. However, the actual switching happens before the disappearance of lysogen potential minimum, when the lysogen potential minimum is too shallow to confine the fluctuations. If 10% of CI monomer is cleaved, at least we expect a 10-fold increase in D_{CI} due to the reduced CI monomer numbers. The potential barrier for the lysogenic state reduces to less than 1, and therefore becomes too shallow to allow continual lysogenic growth.

5.4. Quantitative Comparison with the Experiment of Little et al.

We summarize the calculation results related to the measurements of Little et al. (66) in Table 3 because their data are most up-to-date and systematic. In their experiment, they measured the free phage per lysogenic cell for both $recA^+$ and $recA^-$ phage, but did not convert the $recA^+$ into fraction of lysogens that switched to the lytic state. If we assume the burst size for both the $recA^+$ and the $recA^-$ phage are similar, our calculation for the $RecA^*$ protein agrees with their measurements quantitatively.

In Table 3, the bistability of the gene switch in phage λ is assumed, and the protein levels in the lytic state are calculated. This is, of course, not the case for the wild type, hence, it posits a question to test the calculated Cro level experimentally. One way to realize the bistability may be by suppressing the lyses, achieving the so-called antiimmune phenotype (103,104).

As discussed in section 4 of the formulation of the present stochastic model, we have made the simplified assumption of treating all chance or probability events as Gaussian white noise. This assumption affects two testable biological quantities: the lifetime of lysogenic state (equation

Table 3. Comparison between the calculation and the experimental data (in parentheses) of Little et al. (66).

Phage genotype	Relative CI level in lysogen Theoretical (experimental)	Relative Cro level in lysis Theoretical	Switching frequency to lytic state ($recA^-$) per minute Theoretical (experimental*)	Switching frequency to lytic state ($recA^+$) per minute Theoretical
λ^+	100% (100%)	100%	1×10^{-9} (2×10^{-9})	1×10^{-5}
$\lambda O_R 121$	20% (25–30%)	100%	3×10^{-6} (3×10^{-6})	3×10^{-5}
$\lambda O_R 323$	70% (60–75%)	70%	7×10^{-5} (2×10^{-5})	1×10^{-4}
$\lambda O_R 3'23'$	50% (50–60%)	130%	1×10^{-7} (5×10^{-7})	2×10^{-5}

Here, * indicates that the estimated wild-type data (9) is used. The wild-type biological data were used to find out the difference between *in vivo* and *in vitro* molecular parameters, as listed in Table 1. The relative CI level and switch rate of $\lambda O_R 121$ were used to fine-tune parameters. Rest of theoretical entries are then calculated directly from our model.

[16]), and the shape of CI number distribution in lysogenic state (equation [15] and Figure 7). Simultaneous measuring of both can be used as a consistent check to the Gaussian white noise assumption. For example, we have treated the effect of *recA*$^+$ to switching dynamics as that of a Gaussian white noise to simplify our calculation, in the same reasoning of minimal modeling approach in this chapter. In fact, we have assumed that with *recA*$^+$, the total noise strength doubles. Using this assumption, we calculated the lytic switching rates, represented by the last column of Table 3. The CI distribution with *recA*$^+$ should be twice as broad as in the case with *recA*$^-$. Both results are subject to further experimental testing.

There may be some chance events that cannot be treated as Gaussian white noise in the present formulation. One example has been already suggested in biological experiments (9), the $p_{RM}240$ mutation, which greatly weakens the promoter, and therefore the ability to produce CI as well. This mutation makes the lysogens barely stable, and is estimated to be responsible for at least 99% of observed lytic switching in the wild type. We have used this input for both Tables 2 and 3. We have recalculated the switching rates to lytic state of all strands, assuming the same minimal model, with the same forms of functions for the switching rate, but with the previous experimental data (66). The switching rates obtained in this way are: wild type(λ^+), 2×10^{-7}; $\lambda O_R 121$, $\lambda O_R 2 \times 10^{-6}$; $\lambda O_R 323$, 7×10^{-5}; and $\lambda O_R 3'23'$, 5×10^{-7}. Indeed, the stability of the wild type decreases by more than 2 orders of magnitude. The overall noise strength is increased by 60% for the wild type, resulting in a broader CI distribution in the lysogenic state. There is no appreciable change in other quantities, such as the protein level. The only noticeable overall change in molecular parameters is the *in vivo* cooperative energy, from –6.4 kcal/mol to –6.7 kcal/mol. A good overall quantitative agreement exists between modeling and experiment.

It is a fact that any mathematical modeling in natural science should have empirical input to completely fix its mathematical structure. For the modeling of phage λ, there is an already large body of molecular data, which enables us to nearly pin down our model. The additional freedom in our parameters is fixed by data from wild type, such as the switching frequency. Above the less-than-expected sensitivity of our mathematical structure to this frequency that a few percentage of change in molecular parameters can result in 2 orders of magnitude of change in frequency is a remarkable demonstration of the internal consistency of our modeling. It demonstrates that the switching is exponentially sensitive to some molecular parameters. In addition to more theoretical effort to go beyond our present minimal modeling, it is clear that more experiments are needed in this direction to test the present model: The precise *in vivo* molecular parameters and the distributions and time-correlation of protein numbers in our model should be viewed as predictions.

5.5. Experimental Determination of Dynamical Elements

We have introduced four dynamical quantities for a gene regulatory network: friction, potential, the transverse force, and the stochastic force.

The friction and the strength of the stochastic force are related. For the genetic switch, the transverse force is irrelevant to the dynamic properties. Therefore, the two crucial quantities for a genetic switch are the friction and the potential. Those quantities can be calculated from the more microscopic modeling with molecular parameters. The present quantitative success lies in the allowance of the *in vivo* and *in vitro* differences, and of various noise contributions. However, those four quantities may be directly determined biologically.

There are three different types of experimental data to determine the dynamical elements of the local potential function for a genetic switch, the degradation (the friction in physical sciences), and the barrier height (Figure 7). The first type is the protein distribution around each of the epigenetic states. It is given mathematically by

$$\rho(\mathbf{N}) = \rho_0 \exp(-U(\mathbf{N})), \tag{18}$$

where ρ_0 is a normalization constant. The protein distribution is explicitly measurable. Once $\rho(\mathbf{N})$ is measured experimentally, the local potential $U(\mathbf{N})$ near the potential minima can be determined as: $U(\mathbf{N}) = -\ln(\rho(\mathbf{N})) + \ln(\rho_0)$.

The second type of experimental data is the relaxation time, a measure of how long it would take the system to return to its local equilibrium after a small perturbation. It is determined by both the potential near the minimum and the degradation,

$$\tau = \eta/U''. \tag{19}$$

Here, η is the strength of friction, which gives the friction matrix Λ along the path of relaxation. U'' is the second derivative of potential. Because potential can be obtained from $\rho(\mathbf{N})$, relaxation time can be used to obtain friction: $\eta = \tau U''$.

The third type of experimental data is the lifetime of epigenetic states, the measure of the switching rate from one state to another. The probability of phage evolving from one epigenetic state of growth to another is given by the Kramers rate formula, our equation (16), $P = \omega_0 \exp(-\Delta U_b)$, where ΔU_b is barrier height and ω_0 is the attempt frequency. ω_0 is given by the friction and the curvature of the potential barrier. The curvature of the potential is related to the height of the potential barrier and the shape of the potential near its minimum. Therefore, ΔU_b can be determined from the lifetime of its epigenetic state: $\Delta U_b = \ln(\omega_0) - \ln(P)$.

The genetic switch for phage λ is a complex dynamical system. It took decades of ingenious experimental research and laborious work to collect the parameters needed for this mathematical modeling. For a more complicated system, resources and time may limit the ability to study each molecular element in detail. A method that is less demanding on the details, yet still can capture the main features, is of great interest. Dynamical structure analysis provides guidance to build such a phenomenological model, as illustrated in Figure 7.

We emphasize that the quantities introduced in dynamical structure theory, the friction, the potential gradient, the transverse force, and the stochastic force associating with the friction, are all measurable quantities at the given description level. These are quantities similar to temperature,

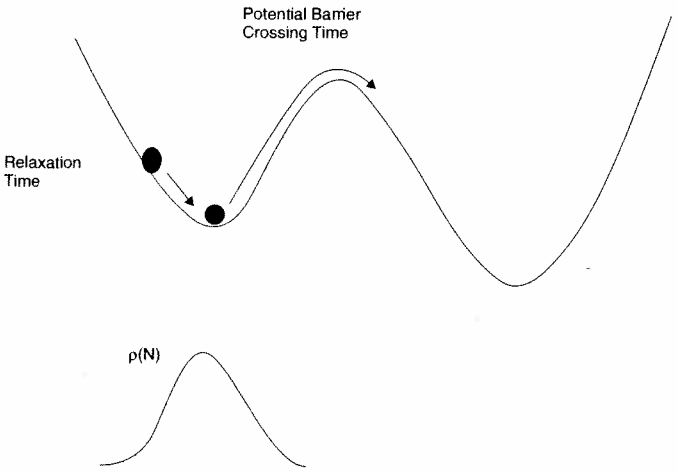

Figure 7. Three types of experiments directly probe the dynamical structure of a genetic switch and determine the dynamic components. The measurement of protein number distribution determines the potential function at each of the epigenetic states. The additional information on the relaxation time determines the strength of friction. The lifetime of each epigenetic state determines the height and shape of the potential barrier.

pressure, and free energy in thermodynamics, which can be determined by microscopic details, but can also be measured independent of those details. Once the relationships between these quantities are established, as shown in equation (10), we are ready to write down an effective equation of motion for the network without resorting to details.

5.6. Stochasticity, Robustness, Cooperation, and Efficiency

Starting from realization that noise is important in the initiation of transcription in phage λ (58), stochasticity has been increasingly viewed as one of the most important elements in the dynamical modeling of biological processes (105–107). Both the intrinsic and extrinsic noises are shown to exist in biological processes and are analyzed theoretically (89–92). The parameterization in the form of equation (17) is a way to account for both noise contributions.

Robustness has been viewed as one of the central features in biological processes (108–110). Numerous recent studies have established its importance (111–116). Combined with stochasticity, the work reviewed here (9,64) established a quantitative criterion, equation (16), for the robustness. The potential landscape function, U(**N**) in equation (10), which emerged from the stochastic dynamics, provides a graphic representation of the robustness. Those results, again, confirm the importance of noise.

Finally, we wish to point out that a complete rigid system, that is, an absolute robust system with no flexibility, is not viable from the evolutionary point of view (117,118). Such a structure would not survive the stringent evolutionary process. This may be illustrated in the switch efficiency discussed above. It is also implied by the ability of the phage λ to switch from the lysogenic to lytic state when provoked (8), a feature that can be captured based on the present minimum model, though no explicit and detailed mathematical analysis has been published yet along this direction. Thus, the stochasticity seems to provide the critical link to understanding both robustness and flexibility. We believe such a feature can, indeed, be understood from the evolutionary point of view (118).

6. Perspective on Mathematical Modeling

6.1. Major Prediction of the Minimum Quantitative Modeling

We have shown that, thanks to continuous theoretical and experimental efforts, the minimum quantitative modeling has achieved the status of quantitative agreement with experimental biological data. New predictions, such as protein distributions, are made and discussed in Zhu et al. (9). There is one prediction on cooperative energy that stands out as an excellent indicator for success of both recent theoretical and experimental efforts.

Since the 1980s, it has been found that it is rather difficult to model the stability of the λ switch with known parameter constraints (48,57), even allowing the possibility of up to 30% difference between *in vivo* and *in vitro* parameter values (51). It has been found that the cooperative energy would play an important role (119,120). Thus, it has been hypothesized that additional effects beyond the minimum model would be needed. One of the most promising ones is a stronger cooperative effect. Indeed, independent of theoretical need, an additional effect, the looping, has been found experimentally (88,96).

A brief account of this effort may be relevant. Four years ago, four of the present authors began a mathematical study on phage λ. Although we tried all methods known to us at that time, we could not solve the stability puzzle. Effectively, we were in the same situation as that reported by Reinitz and Vaisnys and Aurell et al. (57,67). One of the major problems for us was that even allowing the possibility to vary the parameter values drastically in the name of *in vivo* and *in vitro* difference, the parameter space appears too big for an effective research. This difficulty was partially verified in retrospect in Bakk et al. (51), where the change of parameter values appeared too small to explain the "experimental observed robustness" within the minimum quantitative model. Out of this frustration, it was realized 3 years ago that one must have an effective quantification criterion for stability. It turned out that the landscape idea, rooted deeply in both physics and biology, appears to be such a candidate. Driven by this need for quantification, a mathematically consistent construction method for such a landscape function was quickly discovered. With this new method, it was relatively easy to explore bigger parameter space. One critical parameter, the cooperative energy, was then found to double its

value to have the desired stability, the value in bold face for ΔG(cooperative) of −3.7 kcal/mol in Table 2. We note this value is about twice of that tried in Bakk et al. (51).

Interestingly, such a value was indeed observed in an independent biological experiment (88). In writing this review, we further noticed that such a big value was suggested to be possible in an independent theoretical investigation based on thermodynamic consideration (87). Given all of the independent efforts, theoretical (9,87) and experimental (88,96), successful (64) and failed (51), the authors believe that such an agreement between the theoretical prediction (64) and experimental value (88) on the cooperative energy may not be accidental. It indicates that the minimum quantitative model may have, indeed, captured the essential biological features of this genetic switch, with its first nontrivial and verified prediction.

6.2. Relation to Other Modeling Methodologies

There has been a tremendous amount of literature on biomodeling and biocomputation. It is impossible to give an adequate survey of various methodologies in this chapter. Nevertheless, we would like to present the following two classification schemes, according to mathematical and scientific structures, to place our method in a broader context (Figure 8).

From a mathematical point of view, a modeling may be classified according to whether it is discrete or continuous and whether it is deterministic or stochastic. The classic modeling of deterministic and discrete is the Boolean logic circuit (121). The works of Shea and Ackers (48), as well as others (57,122,123), are the finest examples of deterministic and continuous modeling. Examples of stochastic and continuous modeling are Arkin et al. (58) and Aurell and Sneppen (59). Naturally, there are methods of combining various features. One such example is the hybrid of continuous and discrete modeling of Tchuraev and Galimzyanov (124). Of those modeling methods, according to this classification, the simplest one is that based on the Boolean logic circuit. It is clearly an approximation, but in many cases it serves specific biological purposes well. It is, in fact, currently the dominant modeling and presentation methodology in biology. The most difficult, but most detailed, modeling is the continuous and stochastic formulation. Many of its predictions are necessarily probabilistic in nature, corresponding nicely to biological phenomena. Our present method belongs to this last category. Nevertheless, we should point out that no method would be definitely better than the rest. The choice of modeling method must be appropriate to the biological questions being addressed.

From the scientific structure point of view, the modeling may be classified into first principle modeling and completely empirical modeling. It is believed that chemical reactions and other physical processes lie behind various biological processes. Hence, it should be possible to predict biological functions based on the physical-chemical principles. This first principle modeling methodology has been followed by Shea and Ackers (48), by Reinitz and Vaisnys (57), and by many others (58,59,122). Ours is also of this type. The advantages of first principle modeling are that it shows the unity of the sciences and that additional

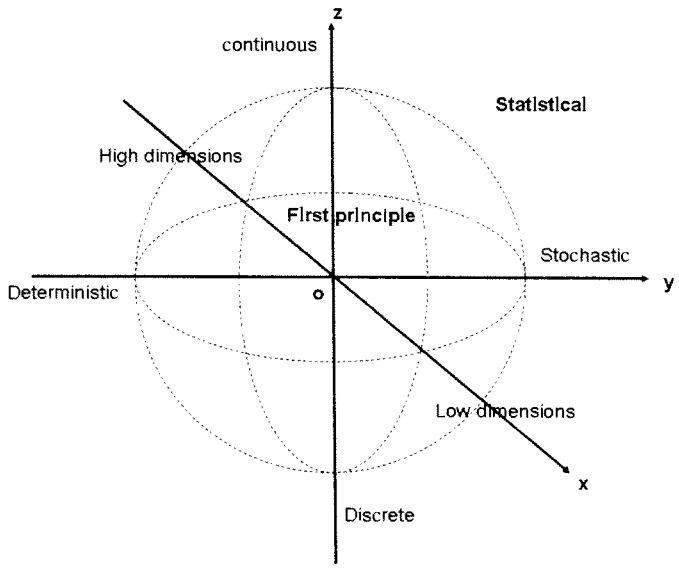

Figure 8. The schematic diagram of modeling methodologies. The coordinates are classified according to mathematical descriptions, which define a three-dimensional space: many degrees of freedom (high dimensions) to few degrees of freedom (low dimensions) (x direction); from deterministic to stochastic (y direction); and from discrete to continuous (z direction). The broken lines, which define a sphere, are classified according to scientific understandings: from first principles (inside) to statistical methods (outside). Both scientific methodologies encompass all mathematical descriptions.

insight and information can be obtained from the lower level scientific descriptions. The evident disadvantage of this modeling, in addition to the difficulty of specifying all needed microparameters, is that higher level processes often show emerging phenomena. It is difficult to make predictions based on the properties of the system's components. The outstanding stability puzzle of phage λ genetic switch (51,59), one of the simplest possible living genetic switches, and the enormous effort (9,48,51,57–59,67) to quantitatively model its behavior, clearly illustrate this situation.

The other extreme, compared with first principle modeling, is to treat the system in question autonomously, inferring its properties completely from empirical studies. Statistical analysis methods play a dominant role in this approach. The advantage of this modeling is that it establishes the independent role of the investigating scientific layer. It is consistent with the view that at each scientific level autonomous laws can be uncovered. Equipped with biological insights, it has been employed successfully in numerous biological studies. The proposals of Waddington (125) on developmental landscape and of Monod and Jacob (126) on gene

regulation mechanism are fine examples. In reality, particularly in molecular biology, what is typically encountered is in between those two extremes, as demonstrated by the Boolean logic circuit modeling (121) and by others (122–124,127). Interestingly, even in the empirical modeling setting, we have briefly discussed the direct and transparent connection of our method to empirical data in subsection 5.5, though it is rooted in first principle modeling. This suggests that our method can be used in this extreme example, too. Further investigation in this direction should be carried out.

The statistical analysis also suggests an important issue in mathematical modeling: the number of variables and the associated problem of the "curse of dimension." Regarding the matching of data to appropriate modeling methodologies, there is another issue of "parsimony of experimental data," which is particularly acute in real-time modeling at present. We will not discuss those issues here.

To summarize the unique characteristics of the present novel method, the stochastic dynamical structure analysis, mathematically, is a continuous and stochastic formulation with four dynamical elements. In terms of scientific structure, it is equivalent to first principle modeling. Nevertheless, it can be directly related to empirical data to establish its autonomy.

6.3. Literature Sampling

First, the book by Ptashne (8) on phage λ is a must read; an excellent review on experimental work up to 2004 can be found. Because both the λ repressor and the *lac* operon are instrumental in the current molecular and synthetic biological studies, the book by Mullller-Hill (128) is another must read. A good summary of earlier study on phage λ can be found in Hendrix et al. (129). A broader and recent general review can be found in Birge (130).

For a recent phage λ study reviewed from the switch point of view, Wegrzyn and Wegrzyn (131) is a good start. Five switches were identified there for the developmental process. The stability puzzle was put into sharp focus in Little et al. (66). Looping study has been theoretically studied in Vilar and Leibler (87), and experimentally in Doff et al. (88) and Revet et al. (96). The phage was studied from an evolutionary point of view in Svenningsen et al. (133). More studies on the role of Cro can be found in Jia et al. (83) and Bundschuh et al. (84). More interesting dynamical behaviors were reported experimentally in Svenningsen et al. (133) and Kobiler et al. (134). The effect of degradation time on stability was considered theoretically in Buchler et al. (135). Various other features on genetic switch have been recently explored (136–141).

7. Third Age of Phage

Because of its enormous biomass in the biosphere on Earth, the importance of phage has already been recognized in the current ecological study (47,142,143), and the title of this section is borrowed from Mann's study (143). The presentation of theoretical effort in this chapter has shown that the phage has been playing an important role in the study of

fundamental biology in the post-Human Genomic era, too, along with the experimental effort (145).

The quantitative and detailed modeling of gene regulatory networks is evidently at its beginning. There are numerical possibilities to go beyond the minimal quantitative modeling reviewed here. For example, even for the phage λ genetic switch, it is a simplification. It would be desirable to have quantitative demonstration on how some of the *in vivo* and *in vitro* differences arise by incorporating more degrees of freedom, such as the left operon and what would be the quantitative differences. A further extension would be to model all five switches of the λ developmental processes (131), to obtain a comprehensive quantitative understanding of the whole process. Deep biological questions, such as why the phage chooses such a structure or what evolution principles guided this choice (66,127), have not yet been discussed adequately by any measure. Nevertheless, we do wish to point out that the study of phage λ genetic switch has revealed a novel mathematical structure (9) and has already put one of the deep-rooted concepts in biology, the landscape (125,145–147), back on a firm mathematical and biological ground. Numerous recent quantitative phage studies (9,49–51,58,59,66,67, 73–75,131–141) have pushed our understanding of systems biology onto another level, complementary to high-throughput and large-scale analyses. It is optimistic that its study will generate more new biological understandings and will have an influence beyond biology, as biology has already inspired the theories of general systems (148) and cybernetics (149). In view of its past successes (e.g., one may count how many Nobel Prize winners in physiology, medicine and chemistry have appeared at http://www.asm.org/division/m/blurbs/secrets.html#top.), we are confident that more discoveries are waiting ahead.

Acknowledgments: We thank J.J. Collins, J.W. Little, B. Muller-Hill, S.M. Stoylar, and G. Wegrzyn for critical comments and M. Mossing for an update on literature. This work was supported in part by the Institute for Systems Biology (L. Hood and D. Galas) and by a National Institutes of Health grant HG002894 (P. Ao).

References

1. Collins FS, Green ED, Guttmacher AE, Guyer MS. A vision for the future of genomics research: a blueprint for the genomic era. *Nature* 2003; 422:835–847.
2. Kitano H. Systems biology: a brief review. *Science* 2002;295:1662–1664.
3. Cohen JE. Mathematics is biology's next microscope, only better; biology is mathematics' next physics, only better. *PloS Biology* 2004;2:2017–2023.
4. Mesarovic MD, Sreenath SN, Keene JD. Search for organizing principles: understanding in systems biology. *Syst Biol* 2004;1:19–27.
5. Kirschner MW. The meaning of systems biology. *Cell* 2005;121:503–504.
6. Williamson MP. Systems biology: will it work? *Biochem Soc Transact* 2005; 33:503–506.
7. Hood L. Systems biology: integrating technology, biology, and computation. *Mech Aging Dev* 2003;124:9–16.
8. Ptashne M. A Genetic Switch: Phage λ revisited. 3rd edition. 2004. Cold Spring Harbor; Cold Spring Harbor Laboratory Press.

9. Zhu XM, Yin L, Hood L, Ao P. Robustness, stability and efficiency of phage lambda genetic switch: dynamical structure analysis. *J Bioinform Comput Biol* 2004;2:785–817.
10. Cairns J, Stent GS, Watson JD. Phage and the Origins of Molecular Biology. expanded edition. Cold Spring Harbor: Cold Spring Harbor Laboratory Press; 1992.
11. Olson MV. The human genome project: a player's perspective. *J Mol Biol* 2002;319:931–942.
12. Lander E. Biology as information. In: Research in Computational Molecular Biology, Lecture Notes in Bioinformatics 3,500. Miyano S, Mesirov J, Kasif S, Istrail S, Pevzner P, Waterman M, eds. Berlin: Springer; 2005.
13. Hood L, Galas D, Dewey G, et al. Biological Information and the Emergence of Systems Biology. Roberts and Co; 2006.
14. Bar-Yam Y, Epstein IR. Response of complex networks to stimuli. *Proc Natl Acad Sci USA* 2004;101:4341–4345.
15. Ben-Hur A, Sigelmann HT. Computation in gene networks. *Chaos* 2004; 14:145–151.
16. Levine M, Davidson EH. Gene regulatory networks for development. *Proc Natl Acad Sci USA* 2005;102:4936–4942.
17. Schroder A, Persson L, de Roos AM. Direct experimental evidence for alternative stable states: a review. *OIKOS* 2005;110:3–19.
18. Balaban NQ, Merrin J, Chait R, et al. Bacterial persistence as a phenotypic switch. *Science* 2004;305:1622–1625.
19. Laurent M, Kellershohn N. Multistability: a major means of differentiation and evolution in biological systems. *TIBS* 1999;24:418–422.
20. Freeman M, Gurdon JB. Regulatory principles of developmental signaling. *Annu Rev Cell Dev Biol* 2002;18:515–539.
21. Kurakin A. Self-organization vs watchmaker: stochastic gene expression and cell differentiation. *Dev Genes Evol* 2005;215:46–52.
22. Hayer A, Bhalla US. Molecular switches at the synapse emerge from receptor and kinase traffic. *PloS Comp Biol* 2005;1:137–154.
23. O'Conner DH, Wittenberg GM, Wang SS-H. Graded bidirectional synaptic plasticity is composed of switch-like unitary events. *Proc Natl Acad Sci USA* 2005;102:9679–9684.
24. Markevich NI, Hoek JB, Kholodenko BN. Signaling switches and bistability arising from multisite phosphorylation in protein kinase cascades. *J Cell Biol* 2004;164:353–359.
25. Qian H, Reluga TC. Nonequilibrium thermodynamics and nonlinear kinetics in a cellular signaling switch. *Phys Rev Lett* 2005;94:028101.
26. Miller P, Zhabotinsky AM, Lisman JE, Wang X-J. The stability of a stochastic CaMKII switch: dependence on the number of enzyme molecules and protein turnover. *PloS Biol* 2005;3:705–717.
27. Herndav AD, Braaten BA, Low DA. The mechanism by which DNA Adenine methylase and PapI activate the Pap epigenetic switch. *Mol Cell* 2003;12:947–957.
28. Travers A. Transcriptional switches: the role of mass action. *Phys Life Rev* 2004;1:57–69.
29. Loayza D, de Lange T. Telomerase regulation at the telomere: a binary switch. *Cell* 2004;117:279–280.
30. Biggar SR, Crabtree GR. Cell signaling can direct either binary or graded transcriptional response. *EMBO J* 2001;20:3167–3176.
31. Casadesus J, D'Ari R. Memory in bacteria and phage. *Bioessays* 2002; 24:512–518.

32. Acar M, Becskei A, van Oudenaarden A. Enhancement of cellular memory by reducing stochastic transitions. *Nature* 2005;435:228–232.
33. Shykind BM, Rohani SC, O'Donnell S, et al. Gene switching and the stability of odorant receptor gene choice. *Cell* 2004;117:801–815.
34. Kaern M, Elston TC, Blake WJ, Collins JJ. Stochasticity in gene expression: from theories to phenotypes. *Nature Rev Genet* 2005;6:451–464.
35. Lidstrom ME, Meldrum DR. Life on a chip. *Nature Rev Microbiol* 2003;1: 158–164.
36. Kramer BP, Viretta AU, Baba MD-E, et al. An engineered epigenetic transgene switch in mammalian cells. *Nature Biotech* 2004;22:867–870.
37. Ozbudak EM, Thattai M, Lim HN, et al. Multistability in the lactose utilization network of *Escherichia coli*. *Nature* 2004;427:737–740.
38. Sauer M. Reversible molecular photoswitches: a key technology for nanoscience and fluorescence imaging. *Proc Natl Acad Sci USA* 2005;102: 9433–9434.
39. Habuchi S, Ando R, Dedecker P, et al. Reversible single-molecule photoswitching in the GFP-like fluorescent protein Dronpa. *Proc Natl Acad Sci USA* 2005;102:9511–9516.
40. Cherry JL, Adler FR. How to make a biological switch. *J Theor Biol* 2000; 203:117–133.
41. Slepchenko BM, Terasaki M. Bio-switches: what makes them robust? *Curr Opin Genet Dev* 2004;14:428–434.
42. Goutsian J, Kim S. A nonlinear discrete dynamical model for transcriptional regulation: construction and properties. *Biophys J* 2004;86:1922–1945.
43. Warren PB, ten Wolde PR. Chemical models of genetic toggle switches. *J Phys Chem B* 2005;109:6812–6823.
44. Ferrell JE, Machleder EM. The biochemical basis of an all-or-none cell fate switch in *Xenopus* oocytes. *Science* 1998;280:895–898.
45. Angeli D, Ferrell JE, Sontag ED. Detection of multistability, bifurcations, and hysteresis in a large class of biological positive-feed back systems. *Proc Natl Acad Sci USA* 2004;101:1822–1827.
46. Chen KC, Calzone L, Csikasz-Nagy A, et al. Integrative analysis of cell cycle control in budding yeast. *Mol Biol Cell* 2004;15:3841–3862.
47. Hendrix RW. Bacteriophage genomics. *Curr Opin Microbiol* 2003;6: 506–511.
48. Shea MA, Ackers GK. The OR control system of bacteriophage lambda—A physical-chemical model for gene regulation. *J Mol Biol* 2005;181:211–230.
49. Dodd IB, Shearwin KE, Egan JB. Revisted gene regulation in bacteriophage lambda. *Curr Opin Genet Dev* 2005;15:145–152.
50. Santillan M, Mackey MC. Why the lysogenic state of phage lambda is so stable: a mathematical modeling approach. *Biophys J* 2004;86:75–86.
51. Bakk A, Metzler R, Sneppen K. Sensitivity of O_R in phage lambda. *Biophys J* 2004;86:58–66.
52. Vilar JMG, Guet CC, Leibler S. Modeling network dynamics: the *lac* operon, a case study. *J Cell Biol* 2003;161:471–476.
53. May RM. Uses and abuses of mathematics in biology. *Science* 2004;303: 790–793.
54. Hwa T. A genetic switch. *Science* 2004;305:345.
55. Ptashne M, Jeffrey A, Johnson AD, Mauer R, Meyer BJ, Pabo CO, Roberts TM, Sauer RT. How the λ repressor and Cro work. *Cell* 1980;19:1–11.
56. Riggs AD, Porter TN. Overview of epigenetic mechanisms. in Russo VEA, Martienssen RA, and Riggs AD (ed), *Epigenetic Mechanisms of Gene Regulation*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press; 1996:29–45.

57. Reinitz J, Vaisnys JR. Theoretical and experimental analysis of the phage lambda genetic switch implies missing levels of cooperativity. *J Theor Biol* 1990;145:295–318.
58. Arkin A, Ross J, McAdams HH. Stochastic kinetic analysis of developmental pathway bifurcation in phage lambda-infected *Escherichia coli* cells. *Genetics* 1998;149:1633–1648.
59. Aurell E, Sneppen K. Epigenetics as a first exit problem. *Phys Rev Lett* 2002; 88:048101.
60. Freidlin MI, Wentzell AD. Random perturbations of Dynamical Systems, 2nd edition. Berlin: Springer; 1998.
61. Ao P. Stochastic force defined evolution in dynamical systems and complex networks. In: Zhao XG, Jiang S, Yu XJ, eds. Computational Physics, Proceedings of the Joint Conference of ICCP6 and CCP2003. Paramus, NJ: Rinton Press; 2005:12–18). (eprint–physics/0302081: http://it.arxiv.org/PS_cache/physics/pdf/0302/0302081.pdf).
62. Ao P. Potential in stochastic differential equations: novel construction. *J Phys A* 2004;37:L25–L30.
63. Kwon C, Ao P, Thouless DJ. Structure of stochastic dynamics near fixed points. *Proc Natl Acad Sci USA* 2005;102:13029–13033.
64. Zhu XM, Yin L, Hood L, Ao P. Calculating biological behaviors of epigenetic states in phage λ life cycle. *Funct Integr Genomics* 2004;4:188–195.
65. Rozanov DV, D'Ari R, Sineoky SP. RecA-independent pathways of lambdoid prophage induction in *Escherichia coli*. *J Bacteriol* 1998;180: 6306–6315.
66. Little JW, Shepley DP, Wert DW. Robustness of a gene regulatory circuit. *EMBO J* 1999;18:4299–4307.
67. Aurell E, Brown S, Johanson J, Sneppen K. Stability puzzle in phage λ. *Phys Rev E* 2002;65:051914–1–9.
68. Koblan KS, Ackers GK. Site-Specific Enthalpic Regulation of DNA Transcription at Bacteriophage λ O_R. *Biochemistry* 1992;31:57–65.
69. Takeda Y, Sarai A, Rivera VM. Analysis of the sequence-specific interactions between Cro repressor and operator DNA by systematic base substitution experiments. *Proc Natl Acad Sci USA* 1989;86:439–443.
70. Kim JG, Takeda Y, Matthews BW, Anderson WF. Kinetic Studies on Cro Repressor-Operator DNA Interaction. *J Mol Biol* 1987;196:149–158.
71. Jana R, Hazbun TR, Mollah AKMM, Mossing MC. A folded monomeric intermediate in the formation of Lambda Cro dimer-DNA complexes. *J Mol Biol* 1997;273:402–416.
72. Jana R, Hazbun TR, Fields JD, Mossing MC. Single-chain lambda Cro repressors confirm high intrinsic dimer-DNA affinity. *Biochemistry* 1998;37: 6446–6455.
73. Darling PJ, Holt JM, Ackers GK. Coupled energetics of λ cro repressor self-assembly and site-specific DNA operator binding I: Analysis of cro dimerization from nanomolar to micromolar concentrations. *Biochemistry* 2000;39:11500–11507.
74. Darling PJ, Holt JM, Ackers GK. Coupled energetics of λ cro repressor self-assembly and site-specific DNA operator binding II: cooperative interactions of Cro dimers. *J Mol Biol* 2000;302:625–638.
75. Ackers GK, Johnson AD, Shea MA. Quantitative model for gene regulation by λ phage repressor. *Proc Natl Acad Sci USA* 1982;79:1129–1133.
76. Capp MW, Cayley DS, Zhang W, et al. Compensating Effects of Opposing Changes in Putrescine (2+) and K^+ Concentrations on *lac* Repressor- *lac* Operator Binding: *in vitro* Thermodynamic Analysis and *in vivo* Relevance. *J Mol Biol* 1996;258:25–36.

77. Record Jr. MT, Courtenay ES, Cayley S, Guttman HJ. Biophysical compensation mechanisms buffering *E. coli* protein-nucleic acid interactions against changing environments. *Trends Biochem Sci* 1998;23:190–194.
78. Hawley DK, McClure WR. Mechanism of activation of transcription initiation from the lambda P_{RM} promoter. *J Mol Biol* 1982;157:493–525.
79. Takeda Y, Ross PD, Mudd CP. Thermodynamics of Cro protein-DNA interactions. *Proc Natl Acad Sci USA* 1992;89:8180–8184.
80. Koblan KS, Ackers GK. Energetics of Subunit Dimerization in bacteriophage lambda cI Repressor: Linkage to Protons, Temperature and KCL. *Biochemistry* 1991;30:7817–7821.
81. Pakula AA, Young VB, Sauer RT. Bacteriophage λ cro mutations: Effects on activity and intracellular degradation. *Proc Natl Acad Sci USA* 1986;83: 8829–8833.
82. Bechhoefer J. Feedback for physicists: A tutorial essay on control. *Rev Mod Phys* 2005;77:783–836.
83. Jia H, Satumba WJ, Bidwell III GL, Mossing MC. Slow assembly and disassembly of lambda Cro repressor dimers. *J Mol Biol* 2005;350:919–929.
84. Bundschuh R, Hayot F, Jayaprakash C. The role of dimerization in noise reduction of simple genetic networks. *J Theor Biol* 2003;220:261–269.
85. Briggs GE, Haldane JBS. A note on the kinetic of enzyme action. *Biochem J* 1925;19:338–339.
86. Bakk A, Metzler R. Nonspecific binding of the O_R Repressors CI and Cro of bacteriophage lambda. *J Theor Biol* 2004;231:525–533.
87. Vilar JMG, Leibler S. DNA looping and physical constraints on transcription regulation. *J Mol Biol* 2003;331:981–989.
88. Dodd IB, Shearwin KE, Perkins AJ, Burr T, Hochschild A, Egan JB. Cooperativity in long-range gene regulation by the lambda CI repressor. *Gene Dev* 2004;18:344–354.
89. van Kampen NG. Stochastic processes in physics and chemistry. Amsterdam: Elsevier; 1992.
90. Elowitz MB, Levine AJ, Siggia ED, Swain PS. Stochastic gene expression in a single cell. *Science* 2002;297:1183–1186.
91. Ao P, Yin L. Towards the understanding of stability puzzles in phage lambda. (2003; eprint: cond-mat/0307747: http://arxiv.org/PS_cache/cond-mat/pdf/0307/0307747.pdf)
92. Raser JM, O'Shea EK. Control of stochasticity in eukaryotic gene expression. *Science* 2004;304:1811–1814.
93. Kramers HA. Brownian motion in a field of force and the diffusion model of chemical reactions. *Physica* 1940;7:284–304.
94. Hanggi P, Talkner P, Borkevic M. Reaction-rate theory: Fifty years after Kramers. *Rev Mod Phys* 1990;62:251–341.
95. Dodd IB, Perkins AJ, Tsemitsidis DT, Egan JB. Octamerization of λ CI repressor is needed for effective repression of P_{RM} and efficient switching from lysogeny. *Genes Dev* 2001;15:3013–3022.
96. Revet B, von Wilcken-Bergmann B, Bessert H, et al. Four dimers of λ repressor bound to two suitably spaced pairs of λ operators form octamers and DNA loops over large distances. *Curr Biol* 1999;9:151–154.
97. Pray TR, Burz DS, Ackers GK. Cooperative non-specific DAN binding by octamerizing λ CI repressors: A site-specific thermodynamic analysis. *J Mol Biol* 1998;282:947–958.
98. Hochschild A, Douhan III J, Ptashne M. How λ repressor and λ Cro distinguish between O_{R1} and O_{R3}. *Cell* 1986;47:807–816.

99. Bremmer H, Dennis PP. Modulation of chemical composition and other parameters of the cell by growth rate. In: Neidhardt FC, ed. *Escherichia coli* and *Salmonella*. ASM Press; 1996:1553–1569.
100. Shean CS, Gottesman ME. Translation of the prophage lambda cI transcript. *Cell* 1992;70:513–522.
101. Ringquist S, Shinedling S, Barrick D, et al. Translation initiation in Escherichia coli: sequences within the ribosome-binding site. *Mol Microbiol* 1992;6:1219–1229.
102. Kennell D, Riezman H. Transcription and translation initiation frequencies of the Escherichia coli lac operon. *J Mol Biol* 1977;114:1–21.
103. Eisen H, Brachet P, Pereira da Silva L, Jacob F. Regulation of repressor expression in λ. *Proc Natl Acad Sci USA* 1970;66:855–862.
104. Calef E, Avitabile LDG, Marchelli C, et al. The genetics of the anti-immune phenotype of defective lambda lysogens. In: Hershey AD, ed. The Bacteriophage Lambda. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press; 1971:609–620.
105. Bialek W. Stability and noise in biochemical switches. In: Advances in Neural Information Processing 13. Leen TK, Dietterich TG, Tresp V, eds. Cambridge: MIT Press; 2001:103.
106. Paulsson J. Summing up the noise in gene networks. *Nature* 2004;427: 415–418.
107. Casci T. Systems biology—Noise is golden. *Nature Rev Genet* 2005;6: 346–346.
108. Hartwell LH, Hopfiedl JJ, Leibler S, Murray AW. From molecular to modular cell biology. *Nature* 1999;402:C47–C52.
109. Nijhout HF. The nature of robustness in development. *Bioessays* 2002;24: 553–563.
110. Kitano H. Biological robustness. *Nature Rev Genet* 2004;5:826–837.
111. Ma L, Iglesias PA. Quantifying robustness of biochemical network models. *BMC Bioinformatics* 2002;3:38.
112. Bluthgen N, Herzel H. How robust are switches in intracellular signaling cascades? *J Theor Biol* 2003;225:293–300.
113. Kerszberg M. Noise, delays, robustness, canalization and all that. *Curr Opin Genet Dev* 2004;14:440–445.
114. Goulian M. Robust control in bacterial regulatory circuits. *Curr Opin Microbiol* 2004;7:198–202.
115. Stelling J, Sauer U, Szallasi Z, et al. Robustness of cellular functions. *Cell* 2004;118:675–685.
116. Li FT, Long T, Lu Y, et al. The yeast cell-cycle network is robustly designed. *Proc Natl Acad Sci USA* 2004;101:4781–4786.
117. Flintoft L. Gene networks—The flexible network. *Nature Rev Genet* 2005;6:252–252.
118. Ao P. Laws in Darwinian evolutionary theory. *Phys Life Rev* 2005;2: 117–156.
119. Hill TL. Cooperativity Theory in Biochemistry: Steady State and Equilibrium Systems. New York: Springer; 1985.
120. Ben-Naim A. Cooperativity in binding of proteins to DNA. *J Chem Phys* 1997;107:10242–10252.
121. Glass L, Kauffman SA. The logical analysis of continuous, non-linear biochemical control networks. *J Theor Biol* 1973;39:103–129.
122. von Dassow G, Meir E, Munro EM, Odell GM. The segment polarity network is a robust developmental module. *Nature* 2000;406:188–192.
123. Vohradsky J. Neural network model of gene expression. *FASEB J* 2001;15: 846–854.

124. Tchuraev RN, Galimzyanov AV. Parametric stability evaluation in computer experiments on the mathematical model of *Drosophila* control gene subnetwork *In Silico*. *Biology* 2003;3:0100.
125. Waddington CH. Organisers and Genes. Cambridge: Cambridge University Press; 1940.
126. Monod J, Jacob F. General conclusions: teleonomic mechanisms in cellular metabolism, growth and differentiation. *Cold Spring Harbor Symp Quant Biol* 1961;26:389–401.
127. Savageau MA. Design principles for elementary gene circuits: Elements, methods, and examples. *Chaos* 2001;11:142–159.
128. Muller-Hill B. The *lac* Operon: a short history of a genetic paradigm. Berlin: Walter de Gruyter; 1996.
129. Hendrix RW, Roberts JW, Stahl FW, Weisberg RA. Lambda II. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory; 1983.
130. Birge EA. Bacterial and Bacteriophage Genetics. 4th edition. Berlin: Springer; 2000.
131. Wegrzyn G, Wegrzyn A. Genetic switches during bacteriophage lambda development. *Prog Nucl Acid Res Mol Biol* 2005;79:1–48.
132. Atsumi S, Little JW. Regulatory circuit design and evolution using phage lambda. *Gene Dev* 2004;18:2086–2094.
133. Svenningsen SL, Costanitno N, Court DL, Adhya S. On the role of Cro in lambda prophage induction. *Proc Natl Acad Sci USA* 2005;102:4465–4469.
134. Kobiler O, Rokney A, Friedman N, et al. Quantitative kinetic analysis of the bacteriophage lambda genetic network. *Proc Natl Acad Sci USA* 2005;102: 4470–4475.
135. Buchler NE, Gerland U, Hwa T. Nonlinear protein degradation and the function of genetic circuits. *Proc Natl Acad Sci USA* 2005;102:9559–9564.
136. Roma DM, O'Flanagan RA, Ruckenstein AE, Sengupta AM. Optimal path to epigenetic switching. *Phys Rev E* 2005;71:011902.
137. Baak K, Svenningsen S, Eisen H, et al. Single-cell analysis of lambda immunity regulation. *J Mol Biol* 2003;334:363–372.
138. Bintu L, Buchler NE, Garcia HG, et al. Transcriptional regulation by the numbers: application. *Curr Opin Genet Dev* 2005;15:125–135.
139. Tian T, Burrage K. Bistability and switching in the lysis/lysogeny genetic regulatory network of bacteriophage lambda. *J Theor Biol* 2004;227:229–237.
140. Warren PB, ten Wolde PR. Enhancement of the stability of genetic switches by overlapping upstream regulatory domains. *Phys Rev Lett* 2004;92:128101.
141. Walczak AM, Sasai M, Wolynes PG. Self-consistent proteomic field theory of stochastic gene switches. *Biophys J* 2005;88:828–850.
142. Campbell A. The future of bacteriophage biology. *Nature Rev Genet* 2003;4: 471–477.
143. Mann NH. The third age of phage. *PloS Biol* 2005;3:753–755.
144. Friedman DI, Court DL. Bacteriophage lambda: alive and well and still doing its thing. *Curr Opin Microbiol* 2001;4:201–207.
145. Wright S. The roles of mutation, inbreeding, crossbreeding and selection in evolution. Proceedings of the 6th International Congress of Genetics. 1932;1:356–366.
146. Delbruck M. Discussion in Unités biologiquées de continuité génétique (International Symposium CNRS *8*), Paris; 1949:33–35.
147. Thom R. Mathematical Models of Morphogenesis. New York: Wiley; 1983.
148. von Bertalanffy L. General System Theory: Foundations, Development, Applications. New York: G. Braziller; 1968.
149. Wiener N. Cybernetics: or, Control and Communication In the Animal and the Machine, 2nd edition. Cambridge: M.I.T. Press; 1961.

Introduction to Systems Biology

Edited by

Sangdun Choi, PhD
Department of Biological Sciences
Ajou University
Suwon, Korea

HUMANA PRESS TOTOWA, NEW JERSEY

[2007]