

# Probability landscape of heritable and robust epigenetic state of lysogeny in phage lambda

Youfang Cao<sup>\* †</sup>, Hsiao-Mei Lu<sup>†</sup> and Jie Liang<sup>† \*</sup>

<sup>\*</sup>Ministry of Education Key Laboratory of Systems Biomedicine, Shanghai Center for Systems Biomedicine, Shanghai Jiao Tong University, Shanghai 200240, China, and <sup>†</sup>Department of Bioengineering, University of Illinois at Chicago, Chicago, IL, 60607, USA

Submitted to Proceedings of the National Academy of Sciences of the United States of America

Computational studies of biological networks can help to identify components and wirings responsible for observed phenotypes. However, studying stochastic networks controlling many biological processes is challenging. Similar to Schrödinger's equation in quantum mechanics, the chemical master equation (CME) provides a basic framework for understanding stochastic networks. However, except for simple problems, the CME cannot be solved analytically. Here we use a method called direct chemical master equation (dCME) to compute directly the full steady-state probability landscape of the lysogeny maintenance network in phage lambda from its CME. Results show that wild type phage lambda can maintain a constant level of repressor over a wide range of repressor degradation rate, and is stable against UV irradiation, ensuring heritability of the lysogenic state. Furthermore, it can switch efficiently to the lytic state once repressor degradation increases past a high threshold by a small amount. We find that beyond bistability and nonlinear dimerization, cooperativity between repressors bound to  $O_R1$  and  $O_R2$  is required for stable and heritable epigenetic state of lysogeny that can switch efficiently. Mutants of phage lambda lack stability and do not possess a high threshold. Instead, they are leaky and respond to gradual changes in degradation rate. Our computation faithfully reproduces the hair triggers for UV-induced lysis observed in mutants and the limitation in robustness against mutations. The landscape approach computed from dCME is general and can be applied to study broad issues in systems biology.

**keywords:** phage lambda, stochasticity, epigenetic control, chemical master equation.

## Section 1: Introduction

Bacteriophage lambda is a virus that infects *E. coli* cells. It has served as a model system for studying regulatory networks and for engineering gene circuits [1–5]. Of central importance is the molecular circuitry that controls phage lambda to choose between two productive modes of development, namely, the lysogenic state and the lytic state (Fig 1a). In the lysogenic state, phage lambda represses its developmental function, integrates its DNA into the chromosome of the host *E. coli* bacterium, and is replicated in cell cycles for potentially many generations. When threatening DNA damage occurs, phage lambda switches from the epigenetic state of lysogeny to the lytic state and undergoes massive replications in a single cell cycle, releases 50–100 progeny phages upon lysis of the *E. coli* cell. This switching process is called *prophage induction* [5].

The molecular network that controls the choice between these two different physiological states has been studied extensively during the past 40 years [5–9]. All of the major molecular components of the network have been identified, binding constants and reaction rates characterized, and there is a good experimental understanding of the general mechanism of the molecular switch [5]. Theoretical studies have also contributed to the illumination of the central role of stochasticity [3] and the stability of lysogen against spontaneous switching [4, 10]. With the advent

---

## Reserved for Publication Footnotes

of systems biology, studying the switching network of phage lambda and lysogeny maintenance has gained added importance, as it provides an ideal ground for developing models and algorithms to study regulatory networks.

However, a general bottle-neck problem for computational studies of regulatory networks is the limitation of existing techniques for studying stochastic networks. Because reactions often involve only low copy numbers of molecules and have large separations in time scale, stochasticity has a strong influence on the behavior of molecular networks [3,4,10]. Deterministic models based on the principles of mass action are often incapable of capturing the multi-stable nature of the network when copy numbers are small [11]. Although the theory of the chemical master equation (CME) provides a general framework for studying stochastic networks [12,13], there are no analytical solutions to the CME except for simple toy problems [14].

One approach is to approximate the CME through various formulations of stochastic differential equations, including the Langevin and the Fokker-Planck formulation [12]. However, the consequences of such approximations for realistic problems involving many molecular species and complex reactions are unknown. Another approach is to carry out extensive stochastic simulations. Powerful simulation tools, including the Gillespie algorithm, have been developed [13,15]. With this approach, many trajectories of simulated reaction events are followed, which are analyzed to reconstruct a probabilistic picture of the stochastic network.

As the CME plays roles in systems biology equivalent to the roles the Schrödinger equation played in quantum mechanics, the development of computational solutions to the CME has important implications, just as the development of techniques for solving the Schrödinger equation for systems with many atoms is [16,17]. However, currently no numeric algorithms can solve the CME directly. Stochastic simulations can follow many cellular events, but have difficulty characterizing rare events that may be biologically critical. As the switching network in phage lambda is stable against random fluctuation, the transition from lysogeny to lysis occurs rarely under normal conditions. In this case, it is difficult to determine whether adequate sampling has been achieved, and for each trajectory whether the simulation time is sufficient.

In this paper, we describe a general approach to study molecular stochastic networks, called dCME, by directly obtaining accurate steady state solutions to the CME that underlies a molecular network. Using a model of the network for the maintenance of lysogeny in phage lambda, we calculate its steady state probability landscape, including those associated with the transition phase from lysogeny to lysis. Such a full stochastic characterization was previously computationally inaccessible. To understand the basic properties required for the maintenance of lysogeny, we characterize the probability landscape of the networks of wild type and mutant phage lambda at different physiological conditions and identify the molecular and architectural determinants of the observed biological behavior. We aim to understand the origin of the network's stability against environmental fluctuations in UV irradiation, the basis of its efficiency in switching, its robustness against changes in network components, as well as the mechanism of the heritability of the epigenetic state of lysogeny.

## Section 2: Results

**Subsection 2.1: Model of Epigenetic Switch of Phage Lambda.** To study how lysogeny is maintained and how it transitions to the lytic state, we assume that lysogeny has already been established. We use a simplified stochastic model for the molecular regulatory network that controls the epigenetic switch in phage lambda (Fig 1b). Using a total of 54 biochemical reactions involving 13 molecular species, our model explicitly includes key components, essential reactions, and cooperativities of the phage lambda decision circuitry. Details can be found in Methods and SI Appendix.

A stochastic biochemical network is characterized at any instant by the probability associated with each microstate of the network, namely, the probability of a specific combination of copy numbers of the molecular species. If the probabilities for all possible microstates at an instant are known, we have the *probability landscape* of the biochemical network for that instant. For a given initial condition, this landscape usually evolves with time. Our interest here is the steady state probabilistic landscape, which describes the overall behavior of phage lambda in the steady state. This landscape also governs the transient chemical reaction dynamics of the system. As the microstates of a biochemical network in general are

too numerous, the probabilistic landscape usually cannot be directly computed. Here we show the probability landscape of the decision circuit of phage lambda can be studied using the newly developed method for optimal enumeration of microstates and for exact calculation of the steady state probability landscape [18]. Briefly, this method slices through the high dimensional space of microstates following the submanifold of accessible microstates for a given initial condition, without visiting the vast space outside this submanifold. Technical details of this algorithm can be found in [18].

**Subsection 2.2: Probability Landscape of Phage Lambda in Lysogenic and Lytic States.** The epigenetic network model shown in Fig 1b can reach around 1.7 million microstates. We have calculated through dCME the steady state probability associated with each of these microstates. Since the space of the microstate is 13D for 13 molecular species, we project the landscape to the 2-dimensional subspace and record only the total copy numbers of  $CI_2$  dimer and  $Cro_2$  dimer molecules. With a high copy number of  $CI_2$  repressor, the lysogenic state of the phage lambda is maintained, whereas a high copy number of  $Cro_2$  protein signifies the lytic state [6]. The  $CI_2$  copy number therefore can be regarded as an indicator of the physiologic state of the phage. Fig 2a shows a probability landscape of the lysogenic state, which has one pronounced peak centered at around the location of 14 copies  $CI_2$  dimer and 0 copies of  $Cro_2$  dimer. As a range of  $CI_2$  values all have high probability in the lysogenic state, our results suggest that minor changes in  $CI_2$  concentration during cell growth and cell division will not affect this epigenetic state. This is important to ensure the heritability of the lysogenic state. Such a conclusion cannot be drawn from studies obtained using the ODE (ordinary differential equation) model, which would give only a fixed point of  $CI_2$  concentration for the lysogenic state, rather than a probability distribution.

We then examine the effects of accelerated CI degradation due to DNA damage from UV irradiation that leads to the activation of RecA-mediated CI cleavage [5, 19]. Fig 2c shows that when the CI degradation rate is raised from  $k_d = 7.0 \times 10^{-4}/s$  to  $k_d = 3.6 \times 10^{-3}/s$ , the probability landscape peaks at a different location, with about 4 copies of  $Cro_2$  dimer and 0 copies of  $CI_2$  dimer.

Our results show that the steady state probability landscapes change adaptively. In the lysogenic state, the probability landscape has one peak, which accounts for the vast majority of locales where there is a significant amount of  $CI_2$  molecules and few  $Cro_2$  molecules. The probability for the system to have high  $Cro$  copy numbers spontaneously is very small. In conditions that eventually lead to the lytic state, the probabilistic landscape adapts and changes, becoming dominated by a peak at different locations where there is a large amount of  $Cro$  and few  $CI$  molecules. At the same time, the peak for the lysogenic state disappears (Fig 2c).

**Subsection 2.3: High Threshold and Efficient Switching of Phage Lambda.** To model the effects of UV irradiation on RecA-mediated acceleration in CI degradation [5, 19], we have systematically examined the behavior of the network by computing its probability landscape at different CI degradation rates. To summarize our results, we calculated the mean copy numbers of  $CI_2$  and  $Cro_2$  of the network from the probability landscape. For example, for the probability landscape shown in Fig 2a, we integrate with properly weighted probabilities of the copy numbers of  $CI_2$  and  $Cro_2$  at each microstate, and obtain the mean copy numbers of 14.3 and 0.0. This is repeated for 36 different values of CI degradation rate. The results are summarized in computationally generated titration curves shown in Fig 3a (solid lines).

Our results show that the epigenetic network of phage lambda is very stable against changes in CI degradation rate due to environmental fluctuation of UV irradiation (Fig 3a, solid lines). Phage lambda stays firmly in the lysogenic state at normal condition. Even when UV irradiation is at a dosage that leads to doubling of the degradation rate of CI, the expected copy number of  $CI_2$  in the system changes very little. Over a wide range of CI degradation rates ( $1.0 \times 10^{-4}/s - 1.7 \times 10^{-3}/s$ ), the lysogenic state is maintained. The switching threshold for the lytic state is high.

Once the switching threshold of degradation rate  $k_d = 1.8 \times 10^{-3}/s$  is reached (Fig 3a, solid lines), a further small increase of  $0.4 \times 10^{-3}/s$  turns the expected copy number of  $CI_2$  from 88% of its maximum value at  $k_d = 1.8 \times 10^{-3}/s$  to 16%. The system therefore can be fully thrown efficiently to the lytic state with a small additional increase in CI degradation at the threshold.

**Subsection 2.4: Cooperativity between  $O_{R1}$  and  $O_{R2}$  Enables the Lysogenic State.** Next we studied the effects of cooperativity. It is well known that repressor dimer  $CI_2$  binds cooperatively to neighboring operator sites [20]. Similarly, Cro dimers also bind cooperatively, although to a lesser extent [21]. These cooperativities are fully incorporated in our model (see SI Appendix for details). In addition, we assume that the looping cooperativity always exists in the form of enhanced CI synthesis when  $O_{R2}$  is occupied by a  $CI_2$ .

We first studied the effects of removal of all cooperativities between neighboring repressor  $CI_2$  and between neighboring  $Cro_2$ . The results show that without these cooperativities, phage lambda cannot maintain the lysogenic state (Fig 3a, dashed lines). As soon as CI degradation becomes slightly above zero, CI repressors are depleted in the steady state, and Cro proteins accumulate.

The strength of the cooperativity between  $CI_2$  dimers and between  $Cro_2$  dimers at different operator sites is uneven. An interesting question is whether the coordination of multiple cooperativities is required, or whether one key cooperativity is sufficient to lead phage lambda to the lysogenic state. We tested different alternative possibilities. When all of the 5 cooperativities are removed ( $CI_2$  binding to  $O_{R1}$ - $O_{R2}$ , to  $O_{R2}$ - $O_{R3}$ ,  $Cro_2$  to  $O_{R1}$ - $O_{R2}$ , to  $O_{R2}$ - $O_{R3}$ , and to  $O_{R1}$ - $O_{R2}$ - $O_{R3}$ , respectively), phage lambda cannot lysogenize (Fig 3a, dashed lines, and details shown in Fig S4 in SI Appendix). Restoring the cooperativity between  $CI_2$  dimers binding to  $O_{R1}$  and  $O_{R2}$  alone can recover the lysogenic state of phage lambda (Fig 3b, solid lines, and Fig S4 in SI). In contrast, the stability and the high threshold are still missing when all cooperativities other than that between  $CI_2$ s binding to  $O_{R1}$ - $O_{R2}$  are restored (Fig 3b, dashed lines, and Fig S4 in SI). This shows the cooperativity of CI dimer binding to  $O_{R1}$  and  $O_{R2}$  is a key enabling factor for phage lambda to maintain the lysogenic state.

Our results are consistent with the experimental observation that the cooperativity between  $CI_2$  dimer binding to  $O_{R1}$  and  $O_{R2}$  plays the dominant role [20, 22, 23]. The stronger affinity and higher occupancy of  $O_{R1}$  bound by  $CI_2$  dimer leads to cooperative binding of  $CI_2$  to the neighboring operator site  $O_{R2}$ . This  $CI_2$  dimer precludes the alternative cooperative binding of  $CI_2$  dimers between  $O_{R2}$  and  $O_{R3}$  [5]. Our findings support the model that stability against prophage induction largely results from cooperative DNA binding by  $CI_2$  to the  $O_{R1}$  and  $O_{R2}$  sites [22, 23]. Zhu *et al.* also showed the importance of cooperativity between CI dimers in an earlier computational study [10]. The new finding emerging from this study is that the cooperativity of  $CI_2$  dimers between  $O_{R1}$  and  $O_{R2}$  is the dominant enabling factor and it alone is sufficient to endow phage lambda with the ability to adopt the lysogenic life style.

**Subsection 2.5: Effects of Altered Operators: Mechanism of Hair-Triggers of the Little Mutants.** To assess the robustness of phage lambda against changes in the molecular components of the epigenetic network, Little *et al* replaced the ordered operator sites of  $O_{R321}$  in the wild type phage lambda with the symmetric variants of  $O_{R323}$ ,  $O_{R121}$ , and  $O_{R3'23'}$ , respectively. In  $O_{R3'23'}$ , the  $O_{R3'}$  site has one of the  $O_{R3}$  nucleotides replaced by that of  $O_{R1}$  [8]. In this seminal work, all of the mutants are found to have functional epigenetic circuits, all have the ability to form lysogens, to grow lytically, and to undergo prophage induction upon UV irradiation. However, these mutants have markedly different tolerance to UV irradiation. The wild type lysogen has a high threshold and requires the highest level of UV irradiation for prophage induction to occur. In contrast, all mutant variants exhibit the behavior of a hair-trigger, and require much less UV irradiation for the onset of prophage induction.

To study the effects of different dosage of UV irradiation on the Little mutants, we calculated their probability landscapes at different CI degradation rates (Fig 4a-d and details shown in Fig S3 of SI Appendix). Overall, we find that the Little mutants all exhibit threshold behavior in our model, as was found experimentally [8]. However, these mutants are generally defective, with reduced thresholds for prophage induction, and are hair-triggered. As seen in Fig 3a (solid lines), wild type phage lambda has a deep threshold for prophage induction ( $k_d = 2.0 \times 10^{-3}/s$  at about 50% induction). In contrast, although all mutants have thresholds for the lytic response, these thresholds are much shallower ( $\leq 1.0 \times 10^{-3}/s$ ) (Fig 4a-d). This is consistent with the experimental finding that they are hair-triggered [8].

Our model can also reproduce detailed differences in UV responses among the mutants. Both  $O_{R121}$  (Fig 4a) and  $O_{R3'23'}$  (Fig 4b) mutants are experimentally found to have a higher threshold of UV irradiation than mutant  $O_{R323}$  (Fig 4c) for prophage induction [8], which we reproduced in our model. We find that the degradation rate at which the amount of  $Cro_2$  surpasses  $CI_2$  for both mutants  $O_{R121}$  and  $O_{R3'23'}$

is  $k_d = 1.6 - 1.7 \times 10^{-3}/s$ , while the  $O_R323$  mutant has a lower value of  $k_d = 4.0 - 5.0 \times 10^{-4}/s$ . There is also another subtle difference in the behavior among the Little mutants. Mutant  $O_R121$  is found experimentally to be slightly more stable than  $O_R3'23'$  according to Fig 3 in reference [8], which is again reflected in our results. When CI degradation is small, the mutant  $O_R121$  (Fig 4a) has a larger amount of CI repressor molecules than the mutant  $O_R3'23'$  (Fig 4b), hence it has a higher tolerance for UV irradiation.

We find that these mutants cannot maintain a constant level of  $CI_2$ . Although it has been suggested that the Little mutants have lower levels of CI, which contributes to the reduced threshold, the full mechanism of how the switching network functions in these mutants is not well understood. From our model, we find that although prophage inductions can occur in both  $O_R121$  and  $O_R3'23'$  mutants, their lysogenic states are different from that of wild type and can be characterized as “leaky”. There is a gradual reduction in the level of CI in response to a gradual increase in the CI degradation rate (Fig 4a and b), corresponding to graded amounts of DNA damage. When the degradation rate of CI increases, the level of CI also decreases, and this effect is cumulative, even though phage lambda remains overall in the lysogenic state. In contrast, wild type phage lambda exhibits a true high threshold behavior that is very stable: there is little change in the amount of CI, even when there is a large increase in the CI degradation rate. This state is maintained until the set-point of switching threshold is reached (Fig 3a, solid lines). Even the best mutants  $O_R121$  and  $O_R3'23'$  with relatively high thresholds respond to gradual changes in CI degradation rate and do not have a high threshold.

The epigenetic circuit of lambda phage is generally robust against system parameter perturbations [8]. However, there are limits to this robustness. As shown in reference [8], a nonsymmetric mutant  $O_R123$ , with the positions of  $O_R1$  and  $O_R3$  swapped could not lysogenize, a finding our results also reproduce. In our model (Fig 4d), this mutant has a severely impaired ability to generate CI repressor, even when there is little or no CI degradation due to UV damage.

**Subsection 2.6: Comparison with Other Methods.** To examine whether the same results can be obtained using stochastic simulation algorithm, we carried out simulation using the Gillespie algorithm [24]. For rare events such as transition from lysogenic to lytic state ( $k_d = 0.0020/s$ ), the algorithm failed to converge to the steady state probability distribution after  $> 6$  times of more computation time compared to the dCME method, and the results strongly depend on the initial conditions. There remain significant residual errors at each of the initial conditions tested. Conclusions drawn from non-converged simulations can give incorrect prediction that the system is still in the lysogenic state. Such failure of convergence can be difficult to detect. Residual error remains significant when the system is in the region of starting to enter the lytic state ( $k_d = 0.0022/s$ ). Furthermore, the expected copy number of  $CI_2$  can be over-estimated by 300% in this region, and the small amount of  $Cro_2$  calculated in the lysogenic region can be off by 3-order of magnitude with comparable computational time, which would lead to unreliable estimation of the frequency of very rare events such as spontaneous lysis in lysogenic state. Details of error analysis can be found in SI.

For this comparison, we assume that the steady state can be established during the life span of a cell. As the models of the reactions are Markovian, the probability landscape of the steady state computed by all stochastic methods should be independent of the initial condition. The issue of assessing differences in computed time-evolving probability landscape before reaching the steady state using different methods requires further investigations.

We also carried out calculation using both stochastic differential equation (SDE) and ordinary differential equation (ODE) models based on the studies of Santillán and Mackey [25] and Gillespie [26], with modifications so it is directly comparable to our model. There are significant qualitative differences in both cases. The SDE model failed to reach the correct steady state, with the landscape much further away than that from the stochastic simulation algorithm. In the ODE model, which can be regarded as the skeleton of the stochastic models, there is no stabilization of the  $CI_2$  concentration against fluctuations in the CI degradation rate, as the amount of  $CI_2$  rapidly decreases when the CI degradation rate increases. Wild type phage lambda would be hair-triggered by this model, which disagrees with experimental data. In addition, the transition from lysogeny to lysis in the 1-2-1 and 3'-2-3' mutants occurs at higher CI degradation rate than the wild type, which would lead to the erroneous conclusion that these mutants are more resistant to UV irradiation. Details of these comparison can be found in the SI.

### Section 3: Discussion

In this study, we characterized the properties of the epigenetic decision network of phage lambda through direct computation of its steady state probability landscape using the method of dCME, an approach previously unfeasible. We find that phage lambda is very stable and is strongly buffered with a high threshold against fluctuations in CI degradation rate due to environmental changes in UV irradiation, a behavior only observed in the wild type phage lambda. This high threshold against environmental fluctuations is important for the self-perpetuating nature of the epigenetic state of *E. coli* cells, allowing lysogeny to be passed on to its offspring. We also find that once the degradation rate of CI reaches a threshold, phage lambda switches very efficiently to the lytic state, and this efficiency is not built at the expense of stability against random fluctuation. Phage lambda can integrate signaling in the form of different CI degradation rates and can distinguish a true signal above the high threshold from random noise fluctuating below this threshold.

Our results further indicate that the cooperativity of CI<sub>2</sub> binding between O<sub>R</sub>1 and O<sub>R</sub>2 plays the key role in enabling the wild type behavior of phage lambda. Nevertheless, the phenotype of a high threshold and robustness against mutations are best viewed overall properties of the network. In addition to this cooperativity, disruptions of other structural and architectural features may also result in the loss of these phenotypes. The distributive nature of such network properties have been discussed in reference [27].

Our results are consistent with many experimental findings [8]. Our results point out that the Little mutants have a leaky switch, compromising stability against fluctuation in UV irradiation and reducing switching efficiency. Our finding suggests that the leaky response to UV-damage and the lack of a high threshold in mutants O<sub>R</sub>121 and O<sub>R</sub>3'23' are responsible for the hair-trigger of prophage induction upon UV damage.

There exist a number of theoretical studies of phage lambda and its mutants [4, 10, 25]. Here we discuss the main differences in biological findings. In a pioneering study, Aurell *et al.* investigated the effects of two free parameters on the stability of lysogenic state in a stochastic differential equation model of phage lambda based on measured lysis frequencies, with the goal of exploring unrecognized control mechanism for the stability of lysogeny. A major conclusion is that the total affinity of Cro for O<sub>R</sub>3 is a key factor in determining the stability of lysogenic state, which affects the transcription rate of derepressed P<sub>RM</sub>, and therefore the synthesis rate of CI. In addition to the differences in underlying model and computational techniques, a focus of our study is to explore the effects of cooperativity between CI dimers, with otherwise experimentally derived model parameters. We are able to pinpoint to the cooperativity between CI dimers on O<sub>R</sub>1 and O<sub>R</sub>2 as the key factor for the stability of phage lambda. An effect of this cooperativity is the increased repression of P<sub>R</sub>, which promotes the production of CI. Santillán and Mackey developed an ODE model [25]. Without stochasticity, this model shows neither stability against small increases of CI degradation rate, nor switching efficiency when CI degradation rate reaches the set-point threshold. The study of Zhu *et al* was based on a potential surface reconstructed from a stochastic differential equation model [10], with three free parameters adjusted for best fit of experimental data. All parameters in our model are derived from the literature. Although both studies show the importance of cooperativity between CI dimers, our results pinpoint to the important role of the cooperativity between CI dimers binding to O<sub>R</sub>1 and O<sub>R</sub>2 (see SI).

This work is also a significant improvement over a preliminary version of our model [28], which does not consider cooperativities between CI dimers or Cro dimers, and does not account for self-promoted synthesis of CI. Without such considerations, lysogeny can only be reached by increasing CI synthesis rate, which is unrealistic and cannot be used to model the effects of UV irradiation.

Overall, our study indicates that an important mechanistic understanding of the system behavior of a stochastic network can be gained through direct computing of the network probability landscape. The study of the effects of altered molecular components and altered wiring of the network further suggests that we could predict the outcome of manipulated phage lambda through computational studies, as elegantly laid out experimentally by Little *et al* [8, 29]. In addition, we can explore evolutionary history of how modern organism emerges. Our focus in this work is exploring the overall global behavior of the phage lambda switching network. There are many aspects of the model that can be further improved, for example, the non-specific binding of CI molecules to DNA and the effects of cooperative looping between O<sub>R</sub> and O<sub>L</sub> are currently only modeled implicitly. In addition, if discrepancies in reported and *in vivo* parameters can be reconciled and obtained, we expect that better quantitative predictions can be made.

Computational studies of biological systems allows exploration of many alternative hypotheses, and facilitate the identification of key elements responsible for phenotypic observations [3,4,10,30,31]. We show that the approach of dCME can offer fresh insight and generate testable predictions. An advantage of this approach is that rare events can be characterized, without difficulties associated with stochastic simulation in determining whether millions or billions of samples are required.

The approach of obtaining exact solutions of simplified models with enumerated microstates advocated has an analogy in previous studies of protein folding. Models such as lattice self-avoiding walks with only hydrophobic and polar (HP) interactions allow the complete enumeration of all feasible conformations and the calculation of exact thermodynamics as well as folding dynamics from the master equations for model molecules. Such studies have played important roles in elucidating the principles of protein folding [32–35].

It is likely that direct solutions to simplified but realistic stochastic networks can lead to the elucidation of the mechanisms of many biological processes. As the molecular mechanism underlying the control of phage lambda applies to many other biological regulatory processes as well, and as similar processes are likely underlie many developmental and epigenetic processes, including cooperative control of histone coding [27], the decision control of phage lambda has offered us a paradigm for studying broad issues in cell development and cell fate [4,5,10,36,37]. The approach described in this study is generally applicable and can be extended to these other systems as well.

## Materials and Methods

Our model includes the repressor protein CI, its dimer CI<sub>2</sub>, the Cro protein, and its dimer Cro<sub>2</sub>. The synthesis and degradation of CI monomer and Cro monomer, as well as their dimerizations are modeled explicitly. In addition, the three operator sites, O<sub>R</sub>1, O<sub>R</sub>2, and O<sub>R</sub>3, are modeled to bind to either CI<sub>2</sub> dimer or Cro<sub>2</sub> dimer with different affinities. These sites can also be unoccupied. The cooperativity between CI<sub>2</sub> dimers and between Cro<sub>2</sub> dimers on neighboring sites is included. In addition, the enhanced CI synthesis when O<sub>R</sub>2 is occupied by CI dimer (currently understood to be due to the looping effect between O<sub>L</sub> and O<sub>R</sub>) is also included [38,39]. The self repression of CI<sub>2</sub> at high concentration is also included. The values of protein-operator binding affinities, cooperativity, protein synthesis rate, and protein degradation rates are based on experimental measurement and are described in detail in the SI Appendix. In our model, we assume that there are a total of about 100 copies of CI repressors [4,40], of which about 50 copies are free in a cell volume available to bind to the operator sites in lysogen. The remaining 50 copies are assumed to bind to DNA non-specifically, as it is expected that a significant amount of CI repressors are bound to DNA in regions other than the operators, and the copy number of free CI repressor may be as few as 10–20 [40]. The rationale of this assumption is further described in the SI Appendix.

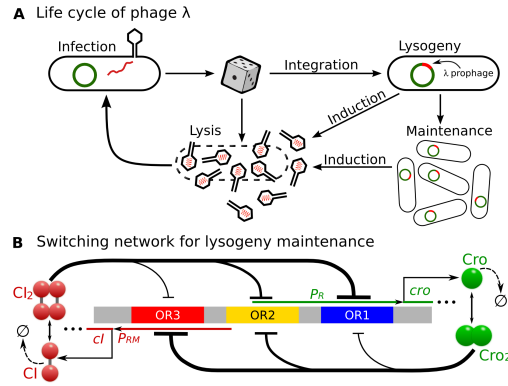
## Section 4: Acknowledgments

We thank Drs. Ping Ao, William Hendrickson, Linda Kenney, Konstantin Mischaikow, Shoudan Liang, Qing Nie, Garyk Papoian, Hong Qian, Michael Samoilov, and Kim Sneppen for helpful discussions, Dr. Zhu Chen and Zhifeng Shao for encouragement, computing center at SCSB for machine time. This work was supported by a phase II 985 Project, 973 grant 2007CB914703, NIH grants GM079804-01A1, GM081682, NSF grant DMS-0800257, and a grant from the Chicago Biomedical Consortium.

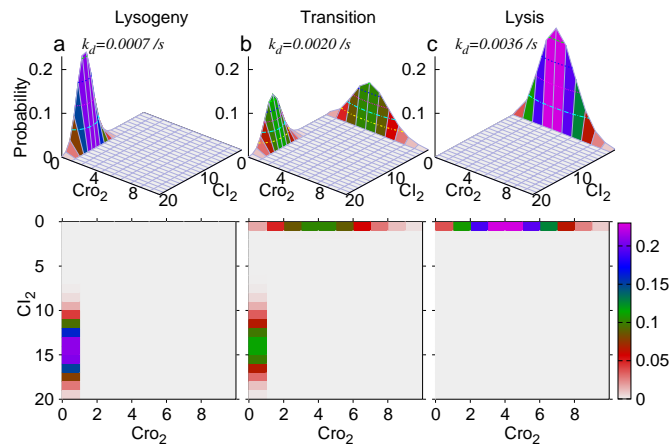
1. F. Jacob and J. Monod. Genetic regulatory mechanisms in the synthesis of proteins. *Journal of molecular biology*, 3(3):318–356, 1961.
2. M. A. Shea and G. K. Ackers. The OR control system of bacteriophage lambda a physical-chemical model for gene regulation. *Journal Molecular Biology*, 181(2):211–230, 1985.
3. A. Arkin, J. Ross, and H. H. McAdams. Stochastic kinetic analysis of developmental pathway bifurcation in phage lambda-infected Escherichia coli cells. *Genetics*, 149(4):1633–1648, 1998.
4. E. Aurell, S. Brown, J. Johanson, and K. Sneppen. Stability puzzles in phage  $\lambda$ . *Physical Review E*, 65(5):051914, 2002.
5. M. Ptashne. *A Genetic Switch: Phage Lambda Revisited*. Cold Spring Harbor Laboratory Press; 3 edition, 2004.
6. A.D. Johnson, et al. lambda repressor and cro—components of an efficient molecular switch. *Nature*, 294(5838):217–223, Nov 1981.
7. M. Ptashne, et al. Autoregulation and function of a repressor in bacteriophage lambda. *Science*, 194(4261):156–161, 1976.
8. J. W. Little, D. P. Shepley, and D. W. Wert. Robustness of a gene regulatory circuit. *The EMBO Journal*, 18(15):4299–4307, 1999.
9. L. M. Anderson and H. Yang. DNA looping can enhance lysogenic CI transcription in phage lambda. *Proceedings of the National Academy of Sciences*, 105(15):5827–5832, 2008.
10. X. M. Zhu, L. Yin, L. Hood, and P. Ao. Robustness, stability and efficiency of phage lambda genetic switch: dynamical structure analysis. *J Bioinform Comput Biol*, 2(4):785–817, 2004.
11. M.N. Artyomov, M. Mathur, M.S. Samoilov, and A.K. Chakraborty. Stochastic bimodalities in deterministically monostable reversible chemical networks due to network topology reduction. *The Journal of chemical physics*, 131:195103, 2009.
12. N. G. Van Kampen. *Stochastic processes in physics and chemistry*. North Holland, Amsterdam, 1992.
13. D. T. Gillespie. Exact stochastic simulation of coupled chemical reactions. *Journal of Physical Chemistry*, 81:2340–2361, 1977.
14. M. Vellela and H. Qian. Stochastic dynamics and non-equilibrium thermodynamics of a bistable chemical system: the schlogl model revisited. *J R Soc Interface*, 6(39):925–940, Oct 2009.
15. H. Li, Y. Cao, L.R. Petzold, and D.T. Gillespie. Algorithms and software for stochastic simulation of biochemical reacting systems. *Biotechnology progress*, 24(1):56, 2008.
16. W. Kohn, L.J. Sham, et al. Self-consistent equations including exchange and correlation effects. *Phys. Rev.*, 140(4A):A1133–A1138, 1965.
17. J. Kohanoff. *Electronic structure calculations for solids and molecules: theory and computational methods*. Cambridge Univ Pr, 2006.
18. Y. Cao and J. Liang. Optimal enumeration of state space of finitely buffered stochastic molecular networks and exact computation of steady state landscape probability. *BMC Systems Biology*, 2:30, 2008.
19. R.T. Sauer, M.J. Ross, and M. Ptashne. Cleavage of the lambda and P22 repressors by recA protein. *Journal of Biological Chemistry*, 257(8):4458, 1982.
20. K. S. Koblan and G. K. Ackers. Site-specific enthalpic regulation of DNA transcription at bacteriophage  $\lambda$  OR. *Biochemistry*, 31:57–65, 1992.
21. P. J. Darling, J. M. Holt, and G. K. Ackers. Coupled energetics of  $\lambda$  cro repressor self-assembly and site-specific DNA operator binding II: Cooperative interactions of cro dimers. *Journal of Molecular Biology*, 302(3):625–638, 2000.

22. A.D. Johnson, B.J. Meyer, and M. Ptashne. Interactions between DNA-bound repressors govern regulation by the  $\lambda$  phage repressor. *Proceedings of the National Academy of Sciences of the United States of America*, 76(10):5061–5065, 1979.
23. G.K. Ackers, A.D. Johnson, and M.A. Shea. Quantitative model for gene regulation by lambda phage repressor. *Proceedings of the National Academy of Sciences of the United States of America*, 79(4):1129, 1982.
24. H. Li, Y. Cao, L.R. Petzold, and D.T. Gillespie. Algorithms and software for stochastic simulation of biochemical reacting systems. *Biotechnology progress*, 24(1):56, 2008.
25. M. Santillán and M. C. Mackey. Why the Lysogenic State of Phage lambda Is So Stable: A Mathematical Modeling Approach. *Biophys. J.*, 86(1):75–84, 2004.
26. D.T. Gillespie. The chemical Langevin equation. *The Journal of Chemical Physics*, 113:297, 2000.
27. I.B. Dodd, M.A. Micheelsen, K. Sneppen, and G. Thon. Theoretical analysis of epigenetic cell memory by nucleosome modification. *Cell*, 129(4):813–822, 2007.
28. Y. Cao, H.M. Lu, and J. Liang. Stochastic probability landscape model for switching efficiency, robustness, and differential threshold for induction of genetic circuit in phage  $\lambda$ . In *IEEE EMBS*, pages 611–614, 2008.
29. S. Atsumi and J. W. Little. A synthetic phage  $\lambda$  regulatory circuit. *Proceedings of the National Academy of Sciences*, 103(50):19045–19050, 2006.
30. S. Lapidus, B. Han, and J. Wang. Intrinsic noise, dissipation cost, and robustness of cellular networks: The underlying energy landscape of MAPK signal transduction. *Proceedings of the National Academy of Sciences*, 105(16):6039, 2008.
31. J. Wang, L. Xu, and E. Wang. Potential landscape and flux framework of nonequilibrium networks: Robustness, dissipation, and coherence of biochemical oscillations. *Proceedings of the National Academy of Sciences*, 105(34):12271, 2008.
32. K. A. Dill, et al. Principles of protein-folding - a perspective from simple exact models. *Protein Science*, 4(4):561–602, 1995.
33. N. D. Socci and J. N. Onuchic. Folding kinetics of proteinlike heteropolymer. *J Chem Phys*, 101:1519–1528, 1994.
34. S. B. Ozkan, I. Bahar, and K. A. Dill. Transition states and the meaning of  $\phi$ -values in protein folding kinetics. *Folding & Design*, 3:R45–R58, 1998.
35. S. Kachalo, H. Lu, and J. Liang. Protein folding dynamics via quantification of kinematic energy landscape. *Phys Rev Lett*, 96(5):058106, 2006.
36. M. Ptashne. Binding reactions: epigenetic switches, signal transduction and cancer. *Current Biology*, 19(6):R234–R241, 2009.
37. J. Liang and H. Qian. Computational cellular dynamics based on the chemical master equation: A challenge for understanding complexity. *Journal of Computer Science and Technology*, 25(1):154–168, 2010.
38. B.J. Meyer and M. Ptashne. Gene regulation at the right operator (OR) of bacteriophage lambda. III lambda Repressor directly activates gene transcription. *Journal of Molecular biology*, 139(2):195–205, 1980.
39. B. Révet, et al. Four dimers of lambda repressor bound to two suitably spaced pairs of lambda operators form octamers and DNA loops over large distances. *Current Biology*, 9(3):151–154, 1999.
40. A. Bak and R. Metzler. Nonspecific binding of the OR repressors CI and Cro of bacteriophage lambda. *Journal of theoretical biology*, 231(4):525–533, 2004.

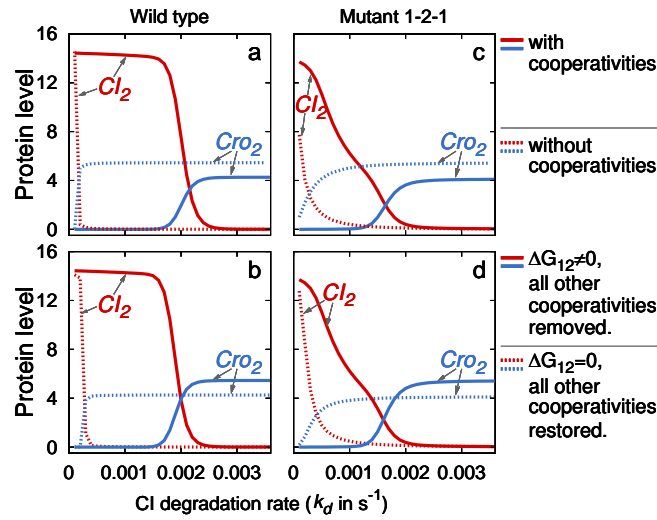




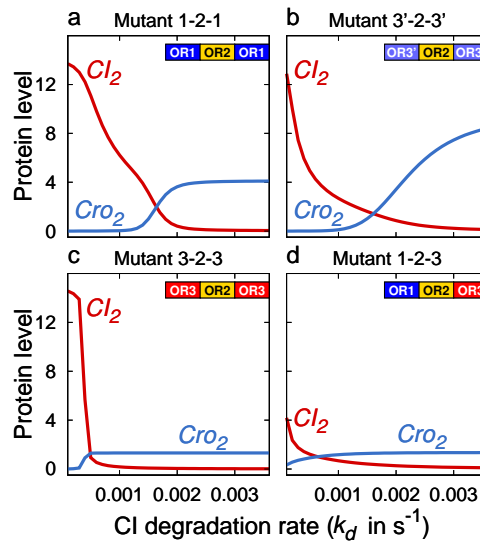
**Fig. 1.** Different selection of cell fate of *E. coli* infected by phage lambda and a model of the epigenetic circuit for lysogeny maintenance. (a) The lysogenic and lytic states of phage lambda. (b) A simplified model of the epigenetic switch for lysogeny maintenance.



**Fig. 2.** The probability landscape of the epigenetic circuits of lysogeny maintenance in phage lambda. (a) At the CI degradation rate of  $k_d = 7.0 \times 10^{-4}/s$ , probability landscape centers at locations with high copy numbers of CI and close to 0 copy of Cro. This corresponds to the lysogenic state of phage lambda. The landscape is shown both in three- and two-dimensional projections. (c) When CI is degraded at a faster rate of  $k_d = 3.6 \times 10^{-3}/s$ , the probability landscape centers at locations where there are higher copy numbers of Cro dimer and close to 0 copy of CI. This corresponds to the lytic state of phage lambda. (b) The probability landscape in transition with simultaneously two peaks when phage lambda is being induced to the lytic state. When  $k_d$  increases from  $k_d = 1.8 \times 10^{-3}/s$  to  $2.2 \times 10^{-3}/s$ , the peak located at lysogenic states gradually diminishes, whereas the peak located at lytic states gradually increases. At about  $k_d = 2.0 \times 10^{-3}/s$ , phage lambda has about equal probability to be in either lysogenic or lytic state (see SI Appendix for more information).



**Fig. 3.** Stability of the epigenetic network for lysogeny maintenance against fluctuation in UV irradiation and its switching efficiency in wild type phage lambda and key role of cooperativity between  $Cl_2$  bound on neighboring  $O_{R1}$  and  $O_{R2}$  sites in maintaining lysogeny. The mean copy number of  $Cl_2$  and  $Cro_2$  are plotted against the CI degradation rate  $k_d$  (in per second unit). (a) Solid lines: The lysogenic state in wild type phage lambda is stable against fluctuation of UV-induced CI degradation rate. The level of  $Cl_2$  is constant for a wide range of CI degradation rate. The threshold for switching to the lytic state is deep. Switching occurs efficiently once the threshold ( $k_d = 1.8 \times 10^{-3}/s$ ) is reached: a further small increase of  $0.4 \times 10^{-4}/s$  completely throws the phage lambda to the lytic state. (Dashed lines): When all cooperativities are removed, phage lambda cannot generate sufficient amount of  $Cl_2$  and therefore cannot maintain lysogenic state. (b) The stability and the deep threshold are missing if all other cooperativities except that between neighbor  $Cl_2$  dimers on  $O_{R1}$  and  $O_{R2}$  are restored ( $\Delta G_{12} = 0$ , dashed line). In contrast, if all other cooperativities are missing, but that between  $Cl_2$  dimers on  $O_{R1}$  and  $O_{R2}$  is restored ( $\Delta G_{12} \neq 0$ , solid line), phage lambda recovers to a large extent the stability and gained significant depth in its switching threshold. (c-d) The same phenomenon is observed in mutant 1-2-1 studied by Little *et al* [8]. With cooperativity between  $O_{R1}$  and  $O_{R2}$ , this mutant can still enter the lysogenic state, albeit without much stability and the threshold is shallow. Without this cooperativity ( $\Delta G_{12} = 0$ ), mutant 1-2-1 cannot maintain lysogeny (see Fig S4 in SI Appendix for more details).



**Fig. 4.** Instability, shallow threshold, and switching inefficiency of the network against fluctuation in UV irradiation in mutant phage lambda [8]. (a-b). In contrast to wild type (Fig 3a, dashed lines), mutant 1-2-1 and 3'-2-3' do not maintain a stable level of  $Cl_2$ . They are leaky and responds gradually to graded changes in  $k_d$ . Their thresholds and that of 3-2-3 (c) for lytic transition are much shallower. (d) Mutant 1-2-3 does not maintain a sufficient amount of  $Cl_2$ , and therefore cannot maintain lysogeny.