

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

Youfang Cao^a, Hsiao-Mei Lu^b, and Jie Liang^{a,b,1}

^aMinistry of Education Key Laboratory of Systems Biomedicine, Shanghai Center for Systems Biomedicine, Shanghai Jiao Tong University, Shanghai 200240, China; and ^bDepartment of Bioengineering, University of Illinois, Chicago, IL 60607

Edited* by Ken A. Dill, University of California, San Francisco, CA, and approved August 18, 2010 (received for review February 5, 2010)

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 discrete 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 non-linear 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.

bistable switch | discrete chemical master equation | epigenetic control | stochasticity | stochastic switch

Bacteriophage lambda is a virus that infects *Escherichia 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 y (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 of systems biology, studying the switching network of phage lambda and lysogeny maintenance has gained added importance, because

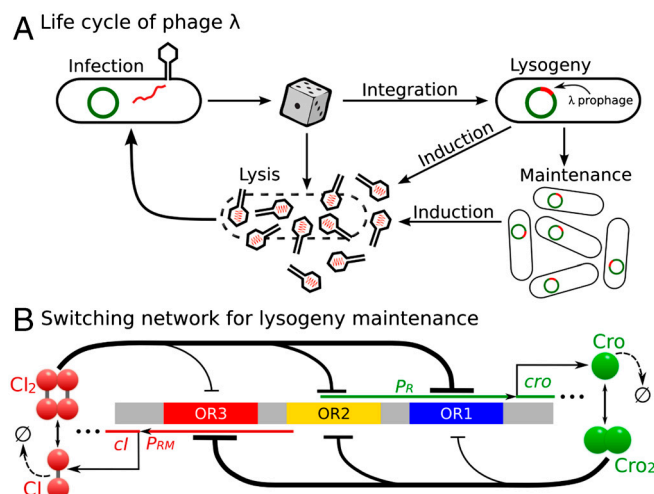


Fig. 1. Different selection of cell fate of *E. coli* infected by phage lambda and a model of the epigenetic circuit for lysogeny maintenance. (A) The lysogenic and lytic states of phage lambda. (B) A simplified model of the epigenetic switch for lysogeny maintenance.

it provides an ideal ground for developing models and algorithms to study regulatory networks.

However, a general bottleneck 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 timescale, 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 multistable 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,

Author contributions: Y.C., H.-M.L., and J.L. designed research; Y.C. and H.-M.L. performed research; Y.C., H.-M.L., and J.L. analyzed data; and Y.C., H.-M.L., and J.L. wrote the paper.

The authors declare no conflict of interest.

*This Direct Submission article had a prearranged editor.

¹To whom correspondence should be addressed. E-mail: jliang@uic.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1001455107/-DCSupplemental.

which are analyzed to reconstruct a probabilistic picture of the stochastic network.

Because 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 does (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 discrete CME (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.

Results

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 ref. 18.

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. Because the space of the microstate is 13D for 13 molecular species, we project the landscape to the 2D subspace and record only the total copy numbers of CI dimer (CI_2) and Cro dimer (Cro_2) molecules. With a high copy number of CI_2 , also known as 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 of CI_2 and 0 copies of Cro_2 . As a range of values of CI_2 copy numbers 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 insensitivity to fluctuation in CI degradation is important to ensure the heritability of the lysogenic state. Such a conclusion cannot be drawn from studies obtained using the ordinary differential equation (ODE) 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 four copies of Cro_2 and 0 copies of CI_2 .

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_2 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).

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 calculation is repeated for 36 different values of CI degradation rate representing different dosage levels of UV irradiation. 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 (1.0×10^{-4} – $1.7 \times 10^{-3}/s$), 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

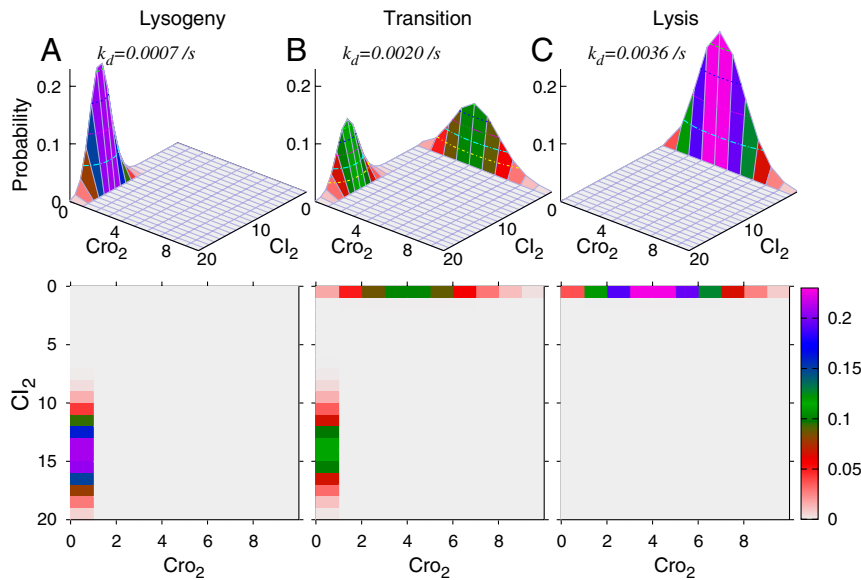


Fig. 2. The probability landscape of the epigenetic circuits of lysogeny maintenance in phage lambda. (A) At the CI degradation rate of $k_d = 7.0 \times 10^{-4}/s$, probability landscape centers at locations with high copy numbers of CI and close to 0 copy of Cro. This landscape corresponds to the lysogenic state of phage lambda. The landscape is shown both in three- and two-dimensional projections. (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). (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 landscape corresponds to the lytic state of phage lambda.

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.

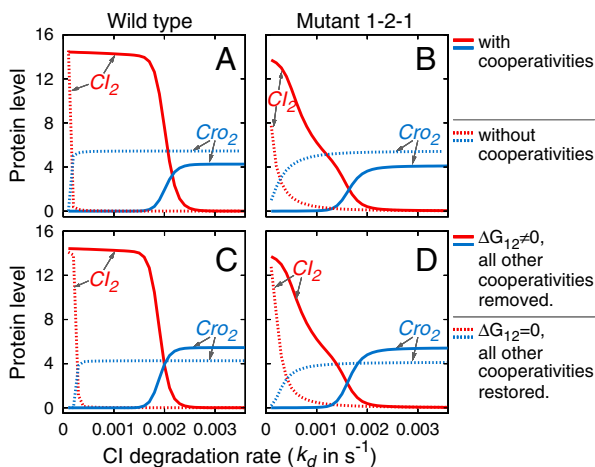


Fig. 3. Stability of the epigenetic network for lysogeny maintenance against fluctuation in UV irradiation and its switching efficiency in wild-type phage lambda and key role of cooperativity between CI_2 bound on neighboring O_R1 and O_R2 sites in maintaining lysogeny. The mean copy number of CI_2 and 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 CI_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 CI_2 and therefore cannot maintain lysogenic state. (B) The stability and the deep threshold are missing if all other cooperativities except that between neighbor CI_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 CI_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 and 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 [SI Appendix](#) for more details).

Cooperativity between O_R1 and O_R2 Enables the Lysogenic State. Next we studied the effects of cooperativity. It is well known that repressor 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 repressors (CI_2 dimers) and between neighboring Cro dimers. 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, repressors are depleted in the steady state, and Cro dimers 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 five 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 the [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 the [SI Appendix](#)). This finding shows the cooperativity of CI dimers 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 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.

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 et al. mutants (8), we calculated their probability landscapes at different CI degradation rates (Fig 4, and details shown in *SI Appendix*). Overall, we find that the Little et al. 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 4). These results are consistent with the experimental finding that they are hair-triggered (8).

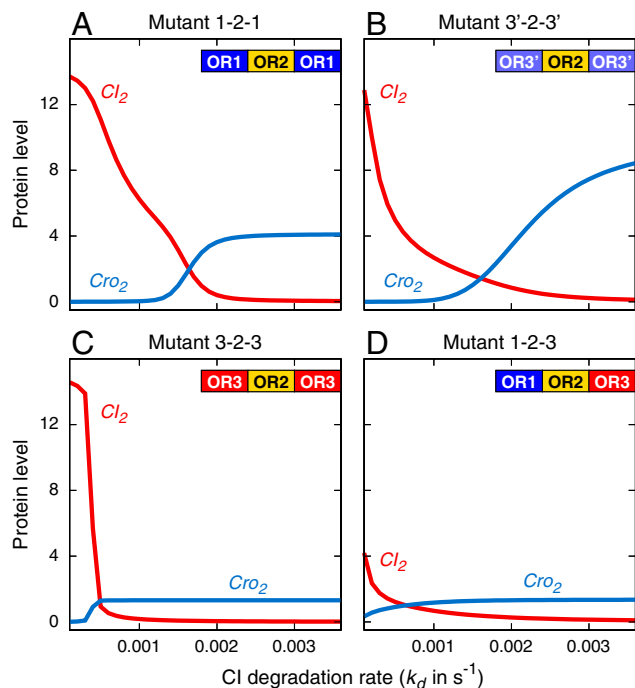


Fig. 4. Instability, shallow threshold, and switching inefficiency of the network against fluctuation in UV irradiation in mutant phage lambda (8). (A–C). In contrast to wild type (Fig 3A, solid lines), mutant 1-2-1 and 3'-2-3' do not maintain a stable level of CI_2 . They are leaky and responds gradually 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 CI_2 , and therefore cannot maintain lysogeny.

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, whereas 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 et al. mutants (8). Mutant O_R121 is found experimentally to be slightly more stable than $O_R3'23'$, according to figure 3 in ref. 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 et al. mutants (8) 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 ref. 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.

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 (15). 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 nonconverged 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 overestimated by 300% in this region, and the small amount of Cro_2 calculated in the lysogenic region can be off by three orders 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 the *SI Appendix*.

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 ODE models based on the studies of Santillán and Mackey (24) and Gillespie (25), 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 Appendix](#).

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 the switching network of 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_2 binding between O_R1 and O_R2 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 ref. 26.

Our results are consistent with many experimental findings (8). Our results point out that the Little et al. (8) 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_R121 and $O_R3'2'3'$ 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, 24). Here we discuss the main differences in biological findings. In a pioneering study, Aurell et al. (4) investigated the effects of two free parameters on the stability of lysogenic state in an SDE 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_R3 is a key factor in determining the stability of lysogenic state, which affects the tran-

scription rate of derepressed P_{RM} , 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 the cooperativity between CI dimers on O_R1 and O_R2 as the key factor for the stability of phage lambda. An effect of this cooperativity is the increased repression of P_R , which promotes the production of CI . Santillán and Mackey developed an ODE model (24). 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 an SDE 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_R1 and O_R2 (see [SI Appendix](#)).

This work is also a significant improvement over a preliminary version of our model (27), 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, 28). 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 nonspecific binding of CI molecules to DNA and the effects of cooperative looping between O_R and O_L 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 study of biological systems allows exploration of many alternative hypotheses and facilitates the identification of key elements responsible for phenotypic observations (3, 4, 10, 29, 30). 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 here has an analogy in previous studies of protein folding. Models such as lattice self-avoiding walks with only hydrophobic and polar interactions allow the complete enumeration of all feasible conformations and the calculation of exact thermodynamics as well as folding dynamics from the master equation for model molecules. Such studies have played important roles in elucidating the principles of protein folding (31–34).

It is likely that direct solutions to simplified but realistic stochastic networks can lead to the elucidation of the mechanisms of many biological processes. Because the molecular mechanism underlying the control of phage lambda applies to many other biological regulatory processes as well, and because similar processes likely underlie many developmental and epigenetic processes, including cooperative control of histone coding (26), the decision control of phage lambda has offered us a paradigm for studying broad issues in cell development and cell fate (4, 5,

10, 35, 36). 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_2 , the Cro protein, and its dimer Cro_2 . 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_R1 , O_R2 , and O_R3 , are modeled to bind to either CI_2 dimer or Cro_2 dimer with different affinities. These sites can also be unoccupied. The cooperativity between CI_2 dimers and between Cro_2 dimers on neighboring sites is included. In addition, the enhanced CI synthesis when O_R2 is occupied by CI dimer (currently understood to be due to the looping effect between O_L and O_R) is also included (37, 38). The self repression of CI_2 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, 39), 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 nonspecifically, because 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 (39). The rationale of this assumption is further described in the *SI Appendix*.

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; and the computing center at Shanghai Center for Systems Biomedicine for machine time. This work was supported by a phase II 985 Project, 973 Grant 2007CB914703, National Institutes of Health Grants GM079804-01A1 and GM081682, National Science Foundation Grant DMS-0800257, and a grant from the Chicago Biomedical Consortium.

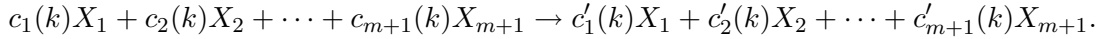
- Jacob F, Monod J (1961) Genetic regulatory mechanisms in the synthesis of proteins. *J Mol Biol* 3:318–356.
- Shea MA, Ackers GK (1985) The OR control system of bacteriophage lambda: a physical-chemical model for gene regulation. *J Mol Biol* 181:211–230.
- Arkin A, Ross J, McAdams HH (1998) Stochastic kinetic analysis of developmental pathway bifurcation in phage lambda-infected *Escherichia coli* cells. *Genetics* 149:1633–1648.
- Aurell E, Brown S, Johanson J, Sneppen K (2002) Stability puzzles in phage λ . *Phys Rev E* 65:051914.
- Ptashne M (2004) *A Genetic Switch: Phage Lambda Revisited* (Cold Spring Harbor Laboratory Press, Plainview, NY), 3 edition, pp 1–66.
- Johnson AD, et al. (1981) Lambda repressor and cro-components of an efficient molecular switch. *Nature* 294:217–223.
- Ptashne M, et al. (1976) Autoregulation and function of a repressor in bacteriophage lambda. *Science* 194:156–161.
- Little JW, Shepley DP, Wert DW (1999) Robustness of a gene regulatory circuit. *EMBO J* 18:4299–4307.
- Anderson LM, Yang H (2008) DNA looping can enhance lysogenic CI transcription in phage lambda. *Proc Natl Acad Sci USA* 105:5827–5832.
- Zhu XM, Yin L, Hood L, Ao P (2004) Robustness, stability and efficiency of phage lambda genetic switch: Dynamical structure analysis. *J Bioinform Comput Biol* 2:785–817.
- Artyomov MN, Mathur M, Samoilov MS, Chakraborty AK (2009) Stochastic bimodalities in deterministically monostable reversible chemical networks due to network topology reduction. *J Chem Phys* 131:195103.
- Van Kampen NG (2007) *Stochastic Processes in Physics and Chemistry* (North-Holland, Amsterdam), 3rd Ed, pp 96–133 pp 193–200.
- Gillespie DT (1977) Exact stochastic simulation of coupled chemical reactions. *J Phys Chem* 81:2340–2361.
- Vellela M, Qian H (2009) Stochastic dynamics and non-equilibrium thermodynamics of a bistable chemical system: The schlogl model revisited. *J R Soc Interface* 6:925–940.
- Li H, Cao Y, Petzold LR, Gillespie DT (2008) Algorithms and software for stochastic simulation of biochemical reacting systems. *Biotechnol Progr* 24:56–61.
- Kohn W, et al. (1965) Self-consistent equations including exchange and correlation effects. *Phys Rev* 140:A1133–A1138.
- Kohanoff J (2006) *Electronic Structure Calculations for Solids and Molecules: Theory and Computational Methods* (Cambridge Univ Press, Cambridge, UK), pp 51–69.
- Cao Y, Liang J (2008) Optimal enumeration of state space of finitely buffered stochastic molecular networks and exact computation of steady state landscape probability. *BMC Syst Biol* 2:30.
- Sauer RT, Ross MJ, Ptashne M (1982) Cleavage of the lambda and P22 repressors by recA protein. *J Biol Chem* 257:4458–4462.
- Koblan KS, Ackers GK (1992) Site-specific enthalpic regulation of DNA transcription at bacteriophage λ OR. *Biochemistry* 31:57–65.
- Darling PJ, Holt JM, Ackers GK (2000) Coupled energetics of λ cro repressor self-assembly and site-specific DNA operator binding II: Cooperative interactions of cro dimers. *J Mol Biol* 302:625–638.
- Johnson AD, Meyer BJ, Ptashne M (1979) Interactions between DNA-bound repressors govern regulation by the λ phage repressor. *Proc Natl Acad Sci USA* 76:5061–5065.
- Ackers GK, Johnson AD, Shea MA (1982) Quantitative model for gene regulation by lambda phage repressor. *Proc Natl Acad Sci USA* 79:1129–1133.
- Santillán M, Mackey MC (2004) Why the lysogenic state of phage lambda is so stable: A mathematical modeling approach. *Biophys J* 86:75–84.
- Gillespie DT (2000) The chemical Langevin equation. *J Chem Phys* 113:297–306.
- Dodd IB, Micheelsen MA, Sneppen K, Thon G (2007) Theoretical analysis of epigenetic cell memory by nucleosome modification. *Cell* 129:813–822.
- Cao Y, Lu HM, Liang J (2008) Stochastic probability landscape model for switching efficiency, robustness, and differential threshold for induction of genetic circuit in phage λ . *IEEE EMBS* 611–614.
- Atsumi S, Little JW (2006) A synthetic phage λ regulatory circuit. *Proc Natl Acad Sci USA* 103:19045–19050.
- Lapidus S, Han B, Wang J (2008) Intrinsic noise, dissipation cost, and robustness of cellular networks: The underlying energy landscape of MAPK signal transduction. *Proc Natl Acad Sci USA* 105:6039–6044.
- Wang J, Xu L, Wang E (2008) Potential landscape and flux framework of nonequilibrium networks: Robustness, dissipation, and coherence of biochemical oscillations. *Proc Natl Acad Sci USA* 105:12271–12276.
- Dill KA, et al. (1995) Principles of protein-folding—a perspective from simple exact models. *Protein Sci* 4:561–602.
- Socchi ND, Onuchic JN (1994) Folding kinetics of proteinlike heteropolymer. *J Chem Phys* 101:1519–1528.
- Ozkan SB, Bahar I, Dill KA (1998) Transition states and the meaning of ϕ -values in protein folding kinetics. *Fold Des* 3:R45–R58.
- Kachalo S, Lu H, Liang J (2006) Protein folding dynamics via quantification of kinematic energy landscape. *Phys Rev Lett* 96:058106.
- Ptashne M (2009) Binding reactions: Epigenetic switches, signal transduction and cancer. *Curr Biol* 19:R234–R241.
- Liang J, Qian H (2010) Computational cellular dynamics based on the chemical master equation: A challenge for understanding complexity. *J Comput Sci Technol* 25:154–168.
- Meyer BJ, Ptashne M (1980) Gene regulation at the right operator (OR) of bacteriophage lambda. III lambda Repressor directly activates gene transcription. *J Mol Biol* 139:195–205.
- Révet B, et al. (1999) Four dimers of lambda repressor bound to two suitably spaced pairs of lambda operators form octamers and DNA loops over large distances. *Curr Biol* 9:151–154.
- Bakk A, Metzler R (2004) Nonspecific binding of the OR repressors CI and Cro of bacteriophage lambda. *J Theor Biol* 231:525–533.

Supporting Information

S0. Discrete Chemical Master Equation

Here we first discuss the chemical master equation on a discrete state space. The CME describes the gain and loss of probability associated with each microstate due to chemical reactions. The chemical reactions can be thought as jump processes that bring the system from one combination of copy number of molecular species (micro state) to a different combination of copy number of molecular species once a reaction occurs. The CME describes the change of probability of different microstates connected by such jump processes due to reactions.

Specifically, we assume a system with $m + 1$ molecular species $\{X_1, \dots, X_{m+1}\}$, where X_i is the label of the i -th molecular species, and have n chemical reactions $\mathcal{R} = \{R_1, \dots, R_n\}$. We denote the copy number of the i -th molecular species as x_i . The combination of the copy numbers at time t is a vector of integers and is denoted as $\mathbf{x}(t) = (x_1(t), \dots, x_{m+1}(t)) \in \mathbb{N}^{m+1}$. We call $\mathbf{x}(t)$ the *microstate* of the system at time t . The probability for the system to be in state $\mathbf{x}(t)$ is $\mathbf{P}(\mathbf{x}, t)$. The set \mathcal{X} of all possible combinations of copy numbers $\mathcal{X} = \{\mathbf{x}(t) | t \in (0, \infty)\}$ is the *state space* of the system. Its size is denoted as $|\mathcal{X}|$. The collection of probabilities for each of the microstate at time t is the *probability landscape* $\mathbf{P}(t)$. Suppose a chemical reaction r_k has the form:



It brings the system from the microstate \mathbf{x}_j to \mathbf{x}_i . The difference between \mathbf{x}_j and \mathbf{x}_i is the stoichiometry vector \mathbf{r}_k of reaction k :

$$\mathbf{r}_k = \mathbf{x}_i - \mathbf{x}_j = (c_1(k) - c'_1(k), \dots, c_{m+1}(k) - c'_{m+1}(k)). \quad (1)$$

Here \mathbf{r}_k can admit 0 entries if a molecular species does not participate in the reaction, so \mathbf{r}_k has the same dimension as that of the microstate.

The rate of the k -th reaction that connects state \mathbf{x}_j to state \mathbf{x}_i is determined by the intrinsic reaction rate constant r_k , and the copy numbers of relevant reactants, which is given by the state \mathbf{x}_j :

$$A_k(\mathbf{x}_i, \mathbf{x}_j) = A_k(\cdot, \mathbf{x}_j) = A_k(\mathbf{x}_j) = r_k \prod_{l=1}^{m+1} \binom{x_l}{c_l}, \quad (2)$$

assuming the convention $\binom{0}{0} = 1$. If the k -th reaction can lead the system from state \mathbf{x}_j to state \mathbf{x}_i , we have $A_k(\mathbf{x}_i, \mathbf{x}_j) > 0$, otherwise $A_k(\mathbf{x}_i, \mathbf{x}_j) = 0$. In most cases, only one reaction connects two microstates. However, since in principle more than one reaction may connect state j to state i , we have the overall reaction rate that brings the system from \mathbf{x}_j to \mathbf{x}_i as:

$$A(\mathbf{x}_i, \mathbf{x}_j) = \sum_{k \in \mathcal{R}} A_k(\mathbf{x}_i, \mathbf{x}_j).$$

where $A(\mathbf{x}_i, \mathbf{x}_j)$ represents the transition probability function per unit time from \mathbf{x}_j to \mathbf{x}_i .

The discrete chemical master equation can then be written as:

$$\frac{d\mathbf{P}(\mathbf{x}, t)}{dt} = \sum_{\mathbf{x}'} [A(\mathbf{x}, \mathbf{x}')\mathbf{P}(\mathbf{x}', t) - A(\mathbf{x}', \mathbf{x})\mathbf{P}(\mathbf{x}, t)] \quad (3)$$

with $\mathbf{x}' \neq \mathbf{x}$. Note here we regard the probability $\mathbf{P}(\mathbf{x}, t)$ of a microstate is continuous in time, while the states are all discrete. The CME in this form fully account for the probabilities of jumps between states, regardless whether the copy number components of \mathbf{x}_i and \mathbf{x}_j are small or large. It gives a full account for the stochasticity due to small copy number events.

In Eqn. (3), the rate constant for leaving the current state $A(\mathbf{x}, \mathbf{x})$ does not appear. If we define

$$A(\mathbf{x}, \mathbf{x}) = - \sum_{\mathbf{x}', \mathbf{x}' \neq \mathbf{x}} A(\mathbf{x}', \mathbf{x}),$$

Eqn. (3) can be written in a more compact form:

$$\frac{d\mathbf{P}(\mathbf{x}, t)}{dt} = \mathbf{A}\mathbf{P}(\mathbf{x}, t), \quad (4)$$

where $\mathbf{A} \in \mathbb{R}^{|\mathcal{X}| \times |\mathcal{X}|}$ is the rate matrix formed by the collection of all $A(\mathbf{x}_i, \mathbf{x}_j)$ s:

$$\mathbf{A} = ||A(\mathbf{x}_i, \mathbf{x}_j)||, \quad \mathbf{x}_i, \mathbf{x}_j \in \mathcal{X}.$$

If we treat the state space as continuous, that is, we assume the amount of a molecular species x_i is measured by a real value (such as concentration) instead of an integer (copy numbers), the vector $\mathbf{x}(t)$ becomes a real-valued vector $\mathbf{x}(t) \in \mathbb{R}^{m+1}$. We have the chemical master equation equivalent to Eqn. (3) on continuous state space as:

$$\frac{\partial \mathbf{P}(\mathbf{x}, t)}{\partial t} = \int_{\mathbf{x}'} [A(\mathbf{x}, \mathbf{x}')\mathbf{P}(\mathbf{x}', t) - A(\mathbf{x}', \mathbf{x})\mathbf{P}(\mathbf{x}, t)]d\mathbf{x}' \quad (5)$$

where the kernel $A(\mathbf{x}, \mathbf{x}')$ represents the transition probability function per unit time from \mathbf{x}' to \mathbf{x} . The CME in this form is equivalent to the Chapman-Kolmogorov equation frequently used to describe continuous Markov processes.

Remark. The continuous state space version of the CME requires a strong assumption. It is only appropriate if one can assume that the difference in the amount of molecules in neighboring states is infinitesimally small, which is valid only if the copy number of the molecular species in the system are much larger than 1, and larger than the changes in the numbers of molecules when a reaction occurs. The continuous CME therefore cannot be used when the total amount of molecules involved is very small, for example, in systems of single or a handful of particles. In these cases, the discrete CME should be used, as it does not contain any intrinsic singularity difficulties.

S1. Enumeration of Microstates of Biological Network

A major hurdle in studying systems using discrete CME is the characterization of the space of microstates. Here we briefly summarize the key elements of our optimal algorithm for state enumeration. We use a slightly different notation to conform with conventions in computer science. Details can be found in [6]. For a network with m molecular species and n reactions, we calculate all microstates that the network can reach starting from a given initial condition. We denote a network as $\mathbf{N} = (\mathbf{M}, \mathbf{R})$, which has $m + 1$ number of molecular species: $\mathbf{M} = (M_1, \dots, M_{m+1})$, and n reactions: $\mathbf{R} = \{R_1, \dots, R_n\}$. A buffer of finite capacity is used from which synthesis reactions can generate new molecules, and to which degradation reactions can deposit molecules removed from the network. A

microstate is a specific combination of copy numbers of all molecular species $\mathbf{s} = (c_1, \dots, c_m, c_{m+1})$. Here c_1, \dots, c_m are the copy numbers of molecular species $1, \dots, m$. c_{m+1} is the number of net new molecules that can still be synthesized at this microstate. A reaction in principle can involve an arbitrary number (≥ 1 and $\leq m$) of molecular species as reactants and/or products, with any arbitrary positive integer coefficient (*i.e.*, arbitrary stoichiometry). Synthesis reaction is allowed to occur only if the buffer capacity is not exhausted, namely, only if $c_{m+1} > 0$. The set of all possible microstates \mathbf{s} that can be reached from an initial condition following these rules constitute the state space \mathcal{X} of the system: $\mathcal{X} = \{\mathbf{s}\}$. The set of allowed transitions is $\mathbf{T} = \{t_{ij}\}$, in which t_{ij} maps the microstate \mathbf{s}_j before the reaction to the microstate \mathbf{s}_i after the reaction. The initial condition is given as: $\mathbf{s}^{t=0} = (c_1^0, c_2^0, \dots, c_m^0, c_{m+1}^0)$, where c_i^0 is the initial copy number of the i -th molecular species at time $t = 0$, and $c_{m+1}^0 = B$ is the predefined buffer capacity.

The algorithm for enumerating the state space is summarized as Algorithm 1. After initialization, it starts with the given initial microstate $\mathbf{s}^{t=0}$. Each reaction is then examined in turn to determine if this reaction can occur for the current microstate. If so, and if the buffer is not used up, the state that this reaction leads to is generated. If the newly generated state was never encountered before, we add it to our collection of states for the state space, and declare it as a new state. We repeat this for all new states, which is maintained by a stack data structure. This terminates when all new states are exhausted. Details can be found in [6].

Following the approach outlined in references [9, 23, 26], the transition coefficient $\{a_{ij}\}$ between two different microstates \mathbf{s}_j and \mathbf{s}_i connected by a reaction is calculated by multiplying the intrinsic rate of this reaction with the reaction order dependent combination number of copies of reactants in the “before” state (see reference [6] for more details).

S2. Calculation of Steady State Probability Landscape

Following Kachalo *et al* [14], we obtain the Markovian state transition matrix \mathbf{M} from the reaction rate matrix \mathbf{A} : $\mathbf{M} = \mathbf{I} + \mathbf{A} \cdot \Delta t$, where \mathbf{I} is the identity matrix, and Δt is the discrete time unit and is chosen to be 1. The steady state probability landscape over the microstates, namely, the probability distribution function \mathbf{P} of the microstates can be obtained by solving the equation

$$\mathbf{P} = \mathbf{M}\mathbf{P}.$$

Since the steady state distribution also corresponds to the eigenvector of \mathbf{M} with eigenvalue of 1, one can also use the Arnoldi method as implemented in the software ARPACK to compute the steady state distribution \mathbf{P} [18]. An alternative is to solve $\mathbf{A}\mathbf{P} = 0$ directly.

Note this probabilistic landscape is different from that of a previous study, in which the landscape is that of a potential function, where the two peaks and the transition saddle point exist simultaneously, with transition between lysogenic and lytic states described by a Kramers process [30].

S3. Calculations of CI_2 and Cro_2 levels

The steady state protein concentration of CI_2 and Cro_2 dimer are calculated from the landscape as: $C_m = \sum_{\mathbf{s} \in \mathcal{X}} \mathbf{P}(\mathbf{s}) s_m$, $m \in \{\text{CI}_2, \text{Cro}_2\}$, where \mathcal{X} is the state space, \mathbf{s} is a microstate in \mathcal{X} , $\mathbf{P}(\mathbf{s})$ is the

Algorithm 1 State Enumerator(M, R, B)

Network model: $N \leftarrow \{M, R\}$;
Initial condition: $s^{t=0} \leftarrow \{c_1^0, c_2^0, \dots, c_m^0\}$; Set the value of buffer capacity: $c_{m+1}^0 \leftarrow B$;
Initialize the state space and the set of transitions: $\mathcal{X} \leftarrow \emptyset$; $T \leftarrow \emptyset$;
Stack $ST \leftarrow \emptyset$; Push($ST, s^{t=0}$); $StateGenerated \leftarrow \text{FALSE}$
while $ST \neq \emptyset$ **do**
 $s_j \leftarrow \text{Pop}(ST)$;
 for $k = 1$ to n **do**
 if reaction R_k occurs under condition s_j **then**
 if reaction R_k is a synthetic reaction and generates u_k new molecules **then**
 $c_{m+1} \leftarrow c_{m+1} - u_k$
 if $c_{m+1} \geq 0$ **then**
 Generate state $s(j, R_k)$ that is reached by following reaction R_k from s_j ;
 $StateGenerated \leftarrow \text{TRUE}$
 end if
 else
 if reaction R_k is a degradation reaction and breaks down u_k molecules **then**
 $c_{m+1} \leftarrow c_{m+1} + u_k$
 end if
 Generate state $s(j, R_k)$ that is reached by following reaction R_k from s_j ;
 $StateGenerated \leftarrow \text{TRUE}$
 end if
 if ($StateGenerated = \text{TRUE}$) and ($s(j, R_k) \notin \mathcal{X}$) **then**
 $\mathcal{X} \leftarrow \mathcal{X} \cup s(j, R_k)$;
 Push($ST, s(j, R_k)$);
 $T \leftarrow T \cup t_{s(j, R_k), s_j}$;
 $a_{i,j} \leftarrow \text{Transition Coefficient}(s(j, R_k), s_j, R_k)$
 end if
 end if
 end for
end while
Output \mathcal{X}, T and $A = \{a_{i,j}\}$.

probability of \mathbf{s} in the steady state probability distribution, and \mathbf{s}_m is the copy number of molecular species m in state \mathbf{s} .

S4. Stochastic Model of Epigenetic Switch of Phage Lambda

The architecture of our model of the epigenetic switch for prophage induction is shown in Fig S1. In this model, the O_R region contains the lysogenic promoter P_{RM} for the transcription of the cI gene and the lytic promoter P_R for transcription of the cro gene. Both CI and Cro proteins dimerize and self-regulate through positive feedback loops, at the same time suppresses the expression of the other protein. There are three operator sites in O_R that both CI_2 protein (also called the repressor) and Cro_2 protein bind, but with different affinities. CI_2 maintains the lysogenic state by blocking O_R sites and preventing transcription of lytic genes including cro . As a result, there is continued expression of the CI protein and the suppression of the cro gene. The switch from lysogenic to lytic state depends on reduction of CI_2 levels, which reflects the state of the bacterium.

When DNA is damaged, the protease RecA from the SOS system is activated, which mediates cleavage of CI repressor. When CI is below a certain level, P_R becomes depressed and transcription of lytic gene is activated, starting with cro . The switch to lytic state is then thrown [21, 24]. In the lytic state, Cro fully represses CI expression.

In this study, our interest is to model once lysogeny has been established, how phage lambda maintain lysogeny, and how it transits to lytic state at adverse environmental condition. The process of establishing lysogeny is not modeled explicitly. For clarity, we use the term of lysogenic state and lytic state instead of lysogenic pathway and lytic pathway, which are sometimes used in the literature. The rationale is that lysogeny and lysis are physiological conditions rather than a specific sequence of events, as implied by the term of *pathway*.

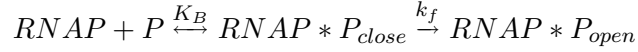
There are a total of 13 molecular species in our model: three empty operator sites ($OR1$, $OR2$ and $OR3$), three operator sites separately bound by CI dimer ($ROR1$, $ROR2$, and $ROR3$), and operator sites separately bound by Cro dimer ($COR1$, $COR2$, and $COR3$), protein monomers (CI and Cro), and protein dimers (CI_2 and Cro_2). Our model is an open system, *i.e.*, there exist synthesis and degradation reactions. A buffer is introduced to limit the state space [6]. The buffer size used in this study, namely, the maximum copy number of net molecules synthesized in the system is 50. This is sufficient to model phage lambda [4, 5]. We do not consider the detailed effects of promoter strengths.

S5. Reactions and Parameters

The 54 biochemical reactions can be classified as synthesis, degradations, dimerizations, binding and dissociation reactions.

Protein synthesis. We follow the approach introduced by Arkin *et al* [2] to model the protein synthesis process. In this approach, the time required for protein synthesis includes that required for the transcription process (transcription initiation and elongation) and the translational process. The rate limiting step in transcription initiation is taken to be the closed- to open-complex isomerization

reaction [12]. Therefore, the rate of the transcription initiation reaction



is taken as $k_f = 6.5 \times 10^4 / M \cdot sec$ (when OR2 is empty) or $9.5 \times 10^5 / M \cdot sec$ (when OR2 is occupied by a repressor) for CI protein and $k_f = 6.7 \times 10^6 / M \cdot sec$ for Cro protein [12, 13, 20]. Transcription elongation rate on average is $= 30$ nt/sec [2, 11, 15, 29], and the elongation time for CI and Cro is calculated based on the length of the coding DNA and is taken as 23.6 and 7.6 sec, respectively. An average of ten copies of proteins are assumed to be produced per transcript for CI and Cro [2, 27]. The translational time for 10 copies of CI and Cro is calculated based on the translational rate of 100 nt/sec [1, 2, 15, 28], and is taken as 7.1 and 2.3 sec for CI and Cro, respectively.

Based on these considerations, transcription and translation of a protein are combined into a single synthesis reaction for simplification. The synthesis reactions and rates are taken from [12, 13] according to [2], and are listed in Table S1. These reactions and parameters are used in both wild type and mutant phage lambda.

Protein degradation. CI and Cro protein degradation is modeled by the proteolysis of the monomeric form, a common degradation mode for multimeric proteins [2, 27]. The values of the rate parameters are taken from [2]. The degradation reactions and reaction rates are listed in Table S1.

Dimerization reactions. We follow [2] to model the dimerization reactions, in which the parameters were obtained from experimental measurements [2, 27]. The sources of the parameter values as documented in [2, 27] for individual reactions are also listed in Table S1.

Free energy and equilibrium constant of binding/dissociation reactions. We use ΔG to denote the Gibbs free energy of the binding/dissociation reactions. It is related to the equilibrium constant k_{eq} through the relationship:

$$k_{eq} = \exp\left(\frac{-\Delta G}{R \cdot T}\right).$$

It was found experimentally that CI binds to OR1 tightly and to OR3 weakly [16]. Conversely, Cro binds to OR1 and OR2 weakly, but to OR3 tightly. We use ΔG_1 to denote the binding free energy of CI_2 to OR1, ΔG_2 for CI_2 to OR2, and ΔG_3 for CI_2 to OR3, respectively. We use ΔG_{1*} , ΔG_{2*} , and ΔG_{3*} to represent binding free energy of Cro_2 to OR1, OR2, and OR3, respectively. The binding free energies of CI repressor to the operators are listed in Table S3 and are taken from Koblan and Ackers [16], those of Cro are from Darling *et al.* [7].

Free energy change due to cooperativities of CI_2 and Cro_2 bound to neighboring operator sites. Following the approach described in [7, 16], we explicitly incorporate cooperativity between repressor CI dimers binding on adjacent operator sites (Fig S2). That is, CI_2 dimer binding on OR1 promotes binding of another CI_2 on OR2. This is reflected by a modification to the free energy change for the dissociation interaction. Specifically, the reaction



has a free energy change $\Delta G_1 = -12.5$ kcal/mol for Cl_2 dissociating from OR1. When a Cl_2 dimer is already bound to the OR2 site, the dissociation reaction of another Cl_2 dimer from OR1



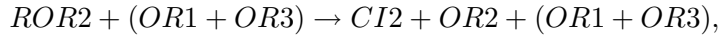
has a larger free energy change and Cl_2 dimers bound on adjacent OR1 and OR2 form a more stable complex, which is modeled by adding a correction term ΔG_{12} due to cooperativity to the original free energy change:

$$\Delta G_{1,coop} = \Delta G_1 + \Delta G_{12}.$$

Similarly, we modify the free energy ΔG_2 for Cl_2 dimer dissociating from OR2 if a Cl_2 dimer is already bound to OR1:

$$\Delta G_{2,coop} = \Delta G_2 + \Delta G_{12}.$$

The same approach is used for modeling cooperativity of Cl_2 dissociating from OR2 and OR3. The dissociation reaction



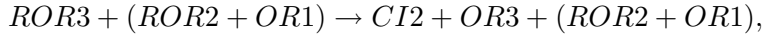
has a free energy change of ΔG_2 . When a Cl_2 dimer is already bound to the OR3 site, the dissociation reaction



has a modification of ΔG_{23} added to ΔG_2 :

$$\Delta G_{2,coop} = \Delta G_2 + \Delta G_{23}.$$

When a Cl_2 dimer is already bound to the OR2 site, the dissociation reaction



also has a modified amount of ΔG_{23} added to ΔG_3 :

$$\Delta G_{3,coop} = \Delta G_3 + \Delta G_{23}.$$

With the same approach, we also explicitly incorporate cooperativity between Cro_2 dimers binding on adjacent operator sites. For the binding cooperativity when Cro_2 dimer binds to OR1 and OR2, and when there is already a Cro_2 dimer bound to OR1 or OR2 site, we add a stabilizing modification term ΔG_{1*2*} to ΔG_{1*} or ΔG_{2*} , respectively. For the binding cooperativity when Cro_2 dimer binds to OR2 and OR3, and when there is already a Cro_2 dimer bound to either of these two sites, we add a stabilizing modification term ΔG_{2*3*} to ΔG_{2*} and ΔG_{3*} accordingly.

In addition, we introduce a cooperativity term ΔG_{1*2*3*} for Cro_2 dimer dissociating from OR1 when both OR2 and OR3 are occupied, and for Cro_2 dissociating from OR3 when both OR1 and OR2 are occupied. In this case, the final binding free energy is modeled as:

$$\Delta G_{1*,coop} = \Delta G_{1*} - \Delta G_{2*3*} + \Delta G_{1*2*3*}$$

and

$$\Delta G_{3*,coop} = \Delta G_{3*} - \Delta G_{1*2*} + \Delta G_{1*2*3*}$$

respectively.

The free energy changes due to these cooperativities for CI are taken from Koblan and Ackers [16], those of Cro are from Darling *et al.* [7], and are listed in Table S3.

Other cooperativity effects. Our model implicitly models the cooperative effect of looping between OR and OL. The cooperativity due to the looping effect stabilizes CI₂ binding to OR2, leading to increased CI synthesis [20]. This is reflected by about 10 fold increase in CI synthesis rate when OR2 is CI₂-bound as shown in Table S3.

In addition, the suppression of the P_{RM} of CI bound to OR3 is also included in our model, which would not be possible without the looping effects between OL and OR.

Association rates of protein binding to operators. We assume CI₂ and Cro₂ proteins have equal access to the operators on DNA. The association rate constant $k_a = 2.57 \times 10^7 / M \cdot sec$ is derived from experimentally measured protein diffusion coefficient $10^{-7} cm^2/s$ [8], which is within the *in vitro* measured range ($10^6 - 10^8 / M \cdot sec$) of CI binding to OR1, OR2, and OR3 [22]. In order to calculate the mesoscopic rate constant for binding reactions involving two reactants, we need to divide k_a by $A \cdot V$, in which $A = 6.023 \times 10^{23}$ is the Avogadro's number and $V = 2 \times 10^{-15} l$ is roughly the volume of *E. coli* cell [17]. We use the resulting rate constant $K_{b_DimerDNA} = k_a / A \cdot V = 0.021 / sec$,

Dissociation rates of CI₂ and Cro₂ from operator sites.

Wild type phage lambda. We follow the model in reference [2], and assume the binding of CI₂ and Cro₂ to the operator sites have the same association rates but different dissociation rates. The dissociation rate k_d is calculated from the equilibrium constant k_{eq} :

$$k_d = k_a / k_{eq},$$

where k_a is the association rate, and $k_{eq} = \exp(\frac{-\Delta G}{R \cdot T})$, ΔG is the Gibbs free energy of the binding reactions, R is the universal gas constant, and $T = 310.15$ Kelvin (*i.e.*, 37 Celsius) is the absolute temperature at which the experiments were performed. The resulting values of dissociation rates of CI₂ and Cro₂ from OR1, OR2, and OR3 calculated from Gibbs free energies are listed in Table S2.

Mutant 1-2-1. Following Little *et al*, we assume that mutations do not alter the properties of the P_{RM} and P_R promoters but affect only binding interactions of Cro₂ and CI₂ to OR [21]. In mutant 1-2-1, an OR1 operator replaces the original OR3. Based on parameters reported in [7, 16], we have now $\Delta G_3 = \Delta G_1 = -12.5$ kcal/mol, and $\Delta G_{3*} = \Delta G_{1*} = -12.0$ kcal/mol. In addition, we have $\Delta G_{23} = \Delta G_{12} = -2.7$ kcal/mol, and $\Delta G_{2*3*} = \Delta G_{1*2*} = -1.0$ kcal/mol. The dissociation rates of mutant 1-2-1 are then calculated using the appropriate modified ΔG values (Table S3).

Mutant 1-2-3. In this mutant, operators OR1 and OR3 exchange places. Based on parameters reported in [7, 16], we have now $\Delta G_1 = -9.5$ kcal/mol, $\Delta G_3 = -12.5$ kcal/mol, $\Delta G_{1*} = -13.4$ kcal/mol, and $\Delta G_{3*} = -12.0$ kcal/mol. In addition, we have $\Delta G_{12} = -2.9$ kcal/mol, $\Delta G_{23} = -2.7$ kcal/mol, $\Delta G_{1*2*} = -0.6$ kcal/mol, and $\Delta G_{2*3*} = -1.0$ kcal/mol. The dissociation rates of mutant 1-2-3 are then calculated accordingly (Table S3).

Mutant 3-2-3. In this mutant, operator OR3 replaces the original OR1. Based on parameters reported in [7, 16], we have now $\Delta G_1 = \Delta G_3 = -9.5$ kcal/mol, and $\Delta G_{1*} = \Delta G_{3*} = -13.4$ kcal/mol. In addition, we have $\Delta G_{12} = \Delta G_{23} = -2.9$ kcal/mol, and $\Delta G_{1*2*} = \Delta G_{2*3*} = -0.6$ kcal/mol. The dissociation rates of mutant 3-2-3 are then calculated accordingly (Table S3).

Mutant 3'-2-3'. This mutant is similar to mutant 3-2-3, with the difference of one nucleotide in OR3 replaced by one from OR1 [21]. Based on experimental data reported in [7, 21], we have the modification

of $\Delta\Delta G_3 = -0.8$ kcal/mol to the binding free energy ΔG_3 of CI_2 to OR3, and $\Delta\Delta G_{3^*} = 1.3$ kcal/mol to the binding free energy ΔG_{3^*} of Cro_2 to OR3 (Fig. 4 from [21]). From these considerations, we have $\Delta G_1 = \Delta G_3 = \Delta G_1 + \Delta\Delta G_3 = -13.3$ kcal/mol for CI_2 binding, and $\Delta G_{1^*} = \Delta G_{3^*} = \Delta G_{1^*} + \Delta\Delta G_{3^*} = -10.7$ kcal/mol for Cro_2 dimer binding.

Additionally, since we replaced the original OR1 with OR3', we have $\Delta G_{12} = \Delta G_{23} = -2.9$ kcal/mol, and $\Delta G_{1^*2^*} = \Delta G_{2^*3^*} = -0.6$ kcal/mol. As the cooperativities of CI_2 and Cro_2 between OR2 and OR3' have not been measured, we use the values of cooperativity between OR2 and OR3 instead. The dissociate rates of mutant 3'-2-3' are calculated accordingly (Table S3).

Free proteins and DNA bound proteins. Our goal is to study the stochastic effects of the phage lambda network, which are most prominent when the copy numbers of molecules are small. Here we make the assumption that only the free concentration of the regulatory proteins contributes to specific binding. The fluctuation of the number of DNA-bound CI dimer is not modeled explicitly. The total number of CI molecules is often thought to be in the range of 100 to 500 [3]. However, there exists significant amount of nonspecific binding of CI and Cro proteins to DNA at regions other than the operators, which may account for about 86% of the total of about 250 of CI monomers in lysogen. Since there is about only 10 copies of free CI dimers in a lysogenic cell [4, 5], we set the total maximum number of free CI monomers to be 50 copies. This choice of the upper limit offers the additional advantages in facilitating large scale computational investigations necessary when studying systematically the behavior of the phage lambda regulatory network at different conditions.

Probability Landscape of Wild Type and Mutant Phage Lambda. The probability landscape of wild type projected in the space of CI_2 and Cro_2 at five different CI degradation rates k_d (lysogeny, start of transition, transition, end of transition, and lysis) are plotted in Fig S3. Probability landscape of 4 mutant phage lambda at the same degradation rates are also plotted.

Effects of Cooperativity on Probability Landscape of Wild Type and Mutant Phage Lambda The probability landscapes of wild type phage lambda with all cooperativities intact, with all cooperativities removed, with only cooperativity between CI_2 binding to O_R1 and O_R2 restored, with all but that between CI_2 binding to O_R1 and O_R2 restored at different K_d are shown in in Fig S4. The probability landscapes of mutant 1-2-1 are shown in Fig S5.

S6. Comparisons to Other Methods

In this section, we compare results from our method with results from the stochastic simulation algorithm [9], a stochastic differential equation (SDE) model, and a deterministic ODE model.

Stochastic Simulation Algorithm. Stochastic simulation algorithm [9] is a widely used method for simulating biochemical reaction systems. However, it is inefficient in simulating rare events, such as the transition between lysogenic and lytic states in phage lambda. To study how this transition is affected by different parameters such as different UV irradiation, it is important to characterize the probability distribution of the switching network at the transition state.

We illustrate here how the stochastic simulation algorithm (SSA) and our direct chemical master equation (dCME) method differ in computing the steady state probability distribution of the phage lambda in transition phase. We use exactly the same model and parameters for both methods. With the wild-type phage lambda model, we set the CI degradation rate $k_d = 0.0020/s$ so the decision network is in the transition state. We start SSA simulation from three different initial conditions: (1) all protein monomer and dimer are initially zero concentration; (2) 10 copies of CI monomers and dimers each, and zero copies of Cro monomer and dimer; (3) 10 copies of Cro monomers and dimers each, but zero copies of CI monomer and dimer. Because of the Markovian properties of the system, all simulations should give the same steady state distribution. We run calculation and simulations in single thread mode on the same machine with 2GHz AMD Quad Core CPU. We use the STOCHKIT package for SSA simulations [19].

The exact solution of the chemical master equation using dCME method was obtained in < 8 hours. The steady state probability distribution is shown in Fig 2b in the main text. Fig S6 shows the probability distributions we obtained from stochastic simulation after 1, 12, 24, 36, and 48 hours of calculation, respectively. It can be seen that after 48 hour, the simulation had not yet converged, as the probability landscape distributions are different for simulations run at the three different initial conditions. In fact, some of these simulations would erroneously lead to the conclusion that the steady state is still predominantly lysogenic if certain initial conditions are chosen. We calculate the residual error between the probability landscape computed in 48-hour of stochastic simulations and that computed from dCME in < 8 hours of calculation (last row of Fig. S6). Here residual errors are obtained by subtracting the probability landscapes of stochastic simulation at 48 hours (Fig. S6) from the directly solved landscape of chemical master equation (Fig. 2b). After 48 hours of simulation, the differences between these probability landscapes are still large. This example indicates that the stochastic simulation algorithm in this case is neither accurate nor efficient for studying rare events.

We have also compared probability landscapes computed from dCME and from SSA for the system in other regions: in the lysogenic region ($k_d = 0.0007/s$), at the beginning of the transition region ($k_d = 0.0018/s$), at the end of the transition region ($k_d = 0.0022/s$), and in the lytic region ($k_d = 0.0036/s$), each with three different initial conditions. Fig S7 summarizes the residual errors in landscape from SSA simulations and Fig S8 shows the errors in estimated mean copies of CI_2 and Cro_2 after 8 hours of simulation, Fig S9 and Fig S10 show errors after 48 hours of simulation. Our results show that in all cases, the probability landscape obtained from SSA depend on initial condition. We find that error in landscape remains large when the system start to enter the lytic region ($k_d = 0.0022/s$), and the expected copy number of CI_2 can be over-estimated by 300% and 150% after 8 hours and 48 hours of computation, respectively. In addition, the small amount of Cro_2 calculated in the lysogenic condition can be off by 3-order of magnitude.

Deterministic ODE Model. We built an ordinary differential equation (ODE) model following the deterministic formulation of Santillán and Mackey [25]. To facilitate a direct comparison, the O_L operators are omitted, and the same parameter values as in the dCME model are used. Correspondingly, the transcription and translation of CI and Cro are model here as two synthesis reactions, one for each protein. The synthesis rate of CI monomer is calculated as the product of the probability of promoter P_{RM} being occupied by RNA polymerase ($f_{RM}^c([CI_2], [Cro_2])$ when CI_2 is bound on O_{R2} , $f_{RM}([CI_2], [Cro_2])$ when O_{R2} is unoccupied by CI_2), times the rate constant of CI synthesis ($k_{s,CI}^c$ when CI_2 is bound on O_{R2} , and $k_{s,CI}$ when O_{R2} is unoccupied by CI_2). The synthesis rate of Cro monomer is calculated as the the product of probability of promoter P_R being occupied by the RNA

polymerase ($f_R([CI_2], [Cro_2])$), times the rate constant of synthesis $k_{s,Cro}$.

The same model of thermodynamics of interactions between protein dimers and O_R operator sites of Shea and Ackers [27] is used to calculate the probabilities $f_{RM}^c([CI_2], [Cro_2])$, $f_{RM}([CI_2], [Cro_2])$, and $f_R([CI_2], [Cro_2])$. The rate constants of CI and Cro synthesis take the same values as in the dCME model (Table S1).

Following the formulation of Santillán and Mackey [25], we use the quasi-steady state assumption to approximately calculate the amount of CI and Cro dimers from monomers, although dimer association and dissociation reactions of CI and Cro are explicitly modeled in the dCME study. The equilibrium CI and Cro dissociation constants K_D^{CI} and K_D^{Cro} necessary for this approximation are calculated from the association and dissociation rates of the dimerization reactions listed in Table S1. We have: $K_D^{CI} = 0.5s^{-1}/0.05nM^{-1}s^{-1} = 10nM = 0.01\mu M$, and $K_D^{Cro} = 0.5s^{-1}/0.0307nM^{-1}s^{-1} \approx 16.3nM = 0.0163\mu M$. Our ODE model can be summarized as:

$$\frac{d[CI]}{dt} = k_{s,CI}^c f_{RM}^c([CI_2], [Cro_2]) + k_{s,CI} f_{RM}([CI_2], [Cro_2]) - k_{d,CI}[CI], \quad (6)$$

$$\frac{d[Cro]}{dt} = k_{s,Cro} f_R([CI_2], [Cro_2]) - k_{d,Cro}[Cro], \quad (7)$$

and

$$[CI_2] = \frac{1}{2}[CI] - \frac{K_D^{CI}}{8} \left[\sqrt{1 + 8 \frac{[CI]}{K_D^{CI}}} - 1 \right], \quad (8)$$

$$[Cro_2] = \frac{1}{2}[Cro] - \frac{K_D^{Cro}}{8} \left[\sqrt{1 + 8 \frac{[Cro]}{K_D^{Cro}}} - 1 \right] \quad (9)$$

where $k_{s,CI}^c = 0.066/s$ is the synthesis rate of CI monomer with a CI dimer bound on O_R2 , $k_{s,CI} = 0.0069/s$ is the CI synthesis rate without CI dimer bound on O_R2 , and $k_{s,Cro} = 0.0929/s$ is the synthesis rate of Cro. $k_{d,CI}$ and $k_{d,Cro}$ are degradation rates of CI and Cro. $k_{d,Cro}$ takes a fixed value of 0.0025/s, and the value of $k_{d,CI}$ is changed systematically from 0.0001/s to 0.0036/s.

The bifurcation curves of the steady state of the ODE model for the wild-type phage lambda and the Little's mutants are shown in Fig. S11. Red lines represent the concentrations of CI_2 at different CI degradation rate, and blue lines represent that of Cro_2 . Comparison with Fig. 3 and 4 in the main text shows that there are at least three major qualitative differences between these two approaches. First, results from the ODE model (Fig. S11a) lacks the ability in maintaining a relatively stable amount of CI dimer in wild-type phage lambda when CI degradation rate increases, unlike results from the chemical master equation model (Fig. 3a, solid line in the main text). Rather, the amount of CI dimer rapidly decreases when CI degradation rate increases. Second, the overall behavior of wild type and mutant 1-2-1 phage lambda do not show appreciable differences other than the transition point (Fig. S11a and b), whereas they are significantly different in results from the chemical master equation model (Fig 3a,c solid lines in the main text). Third, results from the ODE model show that the transition point from lysogenic to lytic state occurs at a later stage in mutants 3'-2-3' and 1-2-1 than in wild-type (Fig. S11a, b and c), which would suggest that that mutant 1-2-1 and 3'-2-3' would be more stable against the UV irradiation than the wild-type, whereas our results (Fig. 3a, solid line, Fig. 4a and b) show that these mutants have a hair trigger and lack the stability against UV irradiation, in agreement with experimental observations [21].

This comparison is important, as ODE model provides the “skeleton” of all stochastic models (both dCME and SDE models), and a full grasp of the corresponding deterministic ODE will help to gain a thorough understanding of the dCME model.

Model of Stochastic Differential Equation. We also compare our results with that of stochastic differential equation (SDE). Following the approach described in reference [10], we derive the SDE model by extending the ODE model discussed above. The model takes the following form:

$$d[\text{CI}] = S_{\text{CI}}dt - D_{\text{CI}}dt + \sqrt{S_{\text{CI}}dt} \cdot \mathcal{N}_1(0, 1) - \sqrt{D_{\text{CI}}dt} \cdot \mathcal{N}_2(0, 1), \quad (10)$$

$$d[\text{Cro}] = S_{\text{Cro}}dt - D_{\text{Cro}}dt + \sqrt{S_{\text{Cro}}dt} \cdot \mathcal{N}_3(0, 1) - \sqrt{D_{\text{Cro}}dt} \cdot \mathcal{N}_4(0, 1), \quad (11)$$

where

$$S_{\text{CI}} = k_{s,\text{CI}}^c f_{RM}^c([\text{CI}_2], [\text{Cro}_2]) + k_{s,\text{CI}} f_{RM}([\text{CI}_2], [\text{Cro}_2]), \quad (12)$$

$$D_{\text{CI}} = k_{d,\text{CI}}[\text{CI}], \quad (13)$$

$$S_{\text{Cro}} = k_{s,\text{Cro}} f_R([\text{CI}_2], [\text{Cro}_2]), \quad (14)$$

$$D_{\text{Cro}} = k_{d,\text{Cro}}[\text{Cro}] \quad (15)$$

and $\mathcal{N}_1(0, 1), \mathcal{N}_2(0, 1), \mathcal{N}_3(0, 1)$, and $\mathcal{N}_4(0, 1)$ are independent Gaussian noises with mean of 0 and variance of 1.

We set $k_{d,\text{CI}} = 0.0020/s$ so the system is in the transition phase, and monitor the rate of convergence of the SDE simulations when starting with different initial conditions. The initial conditions are: (1) both concentrations of CI and Cro monomers are $0 \mu M$: $[\text{CI}]_0 = [\text{Cro}]_0 = 0 \mu M$; (2) $[\text{CI}]_0 = 0.1 \mu M, [\text{Cro}]_0 = 0 \mu M$; and (3) $[\text{CI}]_0 = 0 \mu M, [\text{Cro}]_0 = 0.1 \mu M$. Here only concentrations of monomers are specified, dimer concentrations are calculated as in the above ODE model.

Our dCME model results show that the system is in the transition phase when $k_{d,\text{CI}} = 0.0020/s$, while SDE model shows the system is completely in lytic state. That is, the SDE model has not converged after 48 hours, 6 times of the 8 hours required by the dCME model (Fig. S12). This suggests that it is challenging to produce accurate results using SDE model for the phage lambda system. Further study is required to obtain a comprehensive understanding of the behavior and limitations of SDE models.

S7. Additional information on comparison with a previous theoretical study

Zhu *et al.* modeled the phage lambda switching network through a potential surface reconstructed from a stochastic differential equation model [30]. They found that the wild-type phage lambda was not significantly affected when the cooperative binding of CI dimers is reduced by half. Our results show that the titration curve of the CI concentration against increases in CI degradation rate depends significantly on the cooperative binding free energy between CI dimers on O_R1 and O_R2 (ΔG_{12}). Fig. S13 shows the titration curves at cooperative binding free energy (ΔG_{12}) multiplied by a factor of 0.0, 0.5, 0.8, and 1.2, respectively. When the binding free energy is reduced by a factor of 0.5, the switching threshold at about 50% induction is decreased from $k_d = 0.0020/s$ to $k_d = 0.0007/s$. If the cooperative binding free energy is increased by a factor 1.2, the switching threshold is raised to $k_d = 0.0028/s$. Therefore, we find that the resistance of wild type phage to the UV irradiation

is strongly affected by the cooperative binding: the larger the cooperativity the stronger resistance phage lambda is to the UV irradiation. In contrast, Zhu *et al*'s calculation suggested that there is little change in stability for wild type phage lambda [30].

References

- [1] S. Adhya and M. Gottesman, *Promoter occlusion: transcription through a promoter may inhibit its activity*, Cell **29** (1982), no. 3, 939–944.
- [2] 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) (1998), 1633–1648.
- [3] Erik Aurell, Stanley Brown, Johan Johanson, and Kim Sneppen, *Stability puzzles in phage λ* , Physical Review E **65** (2002), no. 5, 051914.
- [4] A. Bakk and R. Metzler, *In vivo non-specific binding of lambda CI and Cro repressors is significant*, FEBS Letters **563** (2004), no. 1-3, 66–68.
- [5] A. Bakk and R. Metzler, *Nonspecific binding of the OR repressors CI and Cro of bacteriophage lambda*, Journal of theoretical biology **231** (2004), no. 4, 525–533.
- [6] Youfang Cao and Jie Liang, *Optimal enumeration of state space of finitely buffered stochastic molecular networks and exact computation of steady state landscape probability*, BMC Systems Biology **2** (2008), 30.
- [7] Paul J. Darling, Jo M. Holt, and Gary K. Ackers, *Coupled energetics of λ cro repressor self-assembly and site-specific DNA operator binding II: Cooperative interactions of cro dimers*, Journal of Molecular Biology **302** (2000), no. 3, 625–638.
- [8] Michael B. Elowitz, Michael G. Surette, Pierre-Etienne Wolf, Jeffry B. Stock, and Stanislas Leibler, *Protein Mobility in the Cytoplasm of Escherichia coli*, J. Bacteriol. **181** (1999), no. 1, 197–203.
- [9] D. T. Gillespie, *Exact stochastic simulation of coupled chemical reactions*, Journal of Physical Chemistry **81** (1977), 2340–2361.
- [10] D.T. Gillespie, *The chemical Langevin equation*, The Journal of Chemical Physics **113** (2000), 297.
- [11] SL Gotta, OL Miller Jr, and SL French, *rRNA transcription rate in Escherichia coli.*, Journal of bacteriology **173** (1991), no. 20, 6647.
- [12] D. Hawley and W. McClure, *In vitro comparison of initiation properties of bacteriophage lambda wild-type pr and x3 mutant promoters.*, Proc Natl Acad Sci U S A **77**(11) (1980), 6381–6385.
- [13] D. Hawley and W. McClure, *Mechanism of activation of transcription initiation from the lambda prm promoter.*, J Mol Biol **157**(3) (1982), 493–525.
- [14] S. Kachalo, H. Lu, and J. Liang, *Protein folding dynamics via quantification of kinematic energy landscape.*, Phys Rev Lett **96**(5) (2006), 058106.

- [15] D. Kennell and H. Riezman, *Transcription and translation initiation frequencies of the Escherichia coli lac operon*, Journal of Molecular Biology **114** (1977), no. 1, 1–21.
- [16] Kenneth S. Koblan and Gary K. Ackers, *Site-specific enthalpic regulation of DNA transcription at bacteriophage λ OR*, Biochemistry **31** (1992), 57–65.
- [17] Céline Kuttler and Joachim Niehren, *Gene Regulation in the Pi Calculus: Simulating Cooperativity at the Lambda Switch*, Transactions on Computational Systems Biology VII **4230** (2006), 24–55.
- [18] R. Lehoucq, D. Sorensen, and C. Yang, *Arpack users’ guide: Solution of large scale eigenvalue problems with implicitly restarted arnoldi methods*, SIAM, Philadelphia, 1998.
- [19] H. Li, Y. Cao, L.R. Petzold, and D.T. Gillespie, *Algorithms and software for stochastic simulation of biochemical reacting systems*, Biotechnology progress **24** (2008), no. 1, 56.
- [20] M. Li, W. McClure, and M. Susskind, *Changing the mechanism of transcriptional activation by phage lambda repressor.*, Proc Natl Acad Sci U S A **94**(8) (1997), 3691–3696.
- [21] John W. Little, Donald P. Shepley, and David W. Wert, *Robustness of a gene regulatory circuit*, The EMBO Journal **18** (1999), no. 15, 4299–4307.
- [22] E. Merabet and G.K. Ackers, *Calorimetric analysis of lambda cI repressor binding to DNA operator sites*, Biochemistry **34** (1995), 8554–8563.
- [23] Brian Munsky and Mustafa Khammash, *The finite state projection algorithm for the solution of the chemical master equation*, The Journal of Chemical Physics **124** (2006), no. 4, 044104.
- [24] Mark Ptashne, *A genetic switch: Phage lambda revisited.*, Cold Spring Harbor Laboratory Press; 3 edition, 2004.
- [25] Moises Santillán and Michael C. Mackey, *Why the Lysogenic State of Phage lambda Is So Stable: A Mathematical Modeling Approach*, Biophys. J. **86** (2004), no. 1, 75–84.
- [26] D. Schultz, J. N. Onuchic, and P. G. Wolynes, *Understanding stochastic simulations of the smallest genetic networks.*, J Chem Phys **126**(24) (2007), 245102.
- [27] Madeline A. Shea and Gary K. Ackers, *The OR control system of bacteriophage lambda a physical-chemical model for gene regulation*, Journal Molecular Biology **181** (1985), no. 2, 211–230.
- [28] M.A. Sørensen and S. Pedersen, *Absolute in vivo translation rates of individual codons in Escherichia coli: The two glutamic acid codons GAA and GAG are translated with a threefold difference in rate*, Journal of molecular biology **222** (1991), no. 2, 265–280.
- [29] U. Vogel and KF Jensen, *The RNA chain elongation rate in Escherichia coli depends on the growth rate.*, Journal of bacteriology **176** (1994), no. 10, 2807.
- [30] X.-M. Zhu, L. Yin, L. Hood, and P. Ao, *Robustness, stability and efficiency of phage lambda genetic switch: dynamical structure analysis*, Journal of Bioinformatics and Computational Biology **2** (2004), 785–817.

Table S1: Reactions and associated parameters. These reactions and parameter values are common for wild type and all mutants. Here $COR2$ denotes Cro_2 dimer-bound OR2, $ROR2$ denotes CI_2 dimer-bound OR2. Note that molecular species enclosed in parenthesis are those whose presence is required for the specific reactions to occur, but their copy numbers do not influence the transition rates between microstates.

Reactions	Rates(k)
Synthesis reactions [2, 12, 13, 20]	
$\emptyset + (OR3 + OR2) \rightarrow CI + (OR3 + OR2)$	0.0069/s
$\emptyset + (OR3 + COR2) \rightarrow CI + (OR3 + COR2)$	0.0069/s
$\emptyset + (OR3 + ROR2) \rightarrow CI + (OR3 + ROR2)$	0.066/s
$\emptyset + (OR1 + OR2) \rightarrow Cro + (OR1 + OR2)$	0.0929/s
Degradation reactions [2, 27]	
$CI \rightarrow \emptyset$	0.0007/s
$Cro \rightarrow \emptyset$	0.0025/s
Dimerising reactions [2, 27]	
$2 \times CI \rightarrow CI_2$	$0.05/nM \cdot s$
$2 \times Cro \rightarrow Cro_2$	$0.0307/nM \cdot s$
$CI_2 \rightarrow 2 \times CI$	0.5/s
$Cro_2 \rightarrow 2 \times Cro$	0.5/s
Association rate of binding reactions [17]	
$CI_2 + OR1 \rightarrow ROR1$	0.021/s
$CI_2 + OR2 \rightarrow ROR2$	0.021/s
$CI_2 + OR3 \rightarrow ROR3$	0.021/s
$Cro_2 + OR1 \rightarrow COR1$	0.021/s
$Cro_2 + OR2 \rightarrow COR2$	0.021/s
$Cro_2 + OR3 \rightarrow COR3$	0.021/s

Table S2: Reactions and their parameters (continued). This table lists dissociation reactions and their rates in wild type phage lambda. Rates of dissociation for other mutants of phage lambda can be calculated by substituting corresponding free energies with those values listed in Table S3.

Reactions	ΔG_s (kcal/mol)	Rates(s^{-1})
Dissociation reactions - CI_2 dissociation from OR1		
$ROR1 + (OR2) \rightarrow CI2 + OR1 + (OR2)$	$\Delta G_1 = -12.5$	0.03998/s
$ROR1 + (ROR2 + OR3) \rightarrow CI2 + OR1 + (ROR2 + OR3)$	$\Delta G_1 + \Delta G_{12} = -15.2$	0.0005/s
$ROR1 + (ROR2 + ROR3) \rightarrow CI2 + OR1 + (ROR2 + ROR3)$	$\Delta G_1 + \Delta G_{12} - \Delta G_{23} = -12.3$	0.05531/s
$ROR1 + (ROR2 + COR3) \rightarrow CI2 + OR1 + (ROR2 + COR3)$	$\Delta G_1 + \Delta G_{12} = -15.2$	0.0005/s
$ROR1 + (COR2) \rightarrow CI2 + OR1 + (COR2)$	$\Delta G_1 = -12.5$	0.03998/s
Dissociation reactions - CI_2 dissociation from OR2		
$ROR2 + (OR1 + OR3) \rightarrow CI2 + OR2 + (OR1 + OR3)$	$\Delta G_2 = -10.5$	1.026/s
$ROR2 + (ROR1 + OR3) \rightarrow CI2 + OR2 + (ROR1 + OR3)$	$\Delta G_2 + \Delta G_{12} = -13.2$	0.01284/s
$ROR2 + (OR1 + ROR3) \rightarrow CI2 + OR2 + (OR1 + ROR3)$	$\Delta G_2 + \Delta G_{23} = -13.4$	0.00928/s
$ROR2 + (ROR1 + ROR3) \rightarrow CI2 + OR2 + (ROR1 + ROR3)$	$\Delta G_2 + \Delta G_{12} = -13.2$	0.01284/s
$ROR2 + (COR1 + OR3) \rightarrow CI2 + OR2 + (COR1 + OR3)$	$\Delta G_2 = -10.5$	1.026/s
$ROR2 + (OR1 + COR3) \rightarrow CI2 + OR2 + (OR1 + COR3)$	$\Delta G_2 = -10.5$	1.026/s
$ROR2 + (COR1 + COR3) \rightarrow CI2 + OR2 + (COR1 + COR3)$	$\Delta G_2 = -10.5$	1.026/s
$ROR2 + (ROR1 + COR3) \rightarrow CI2 + OR2 + (ROR1 + COR3)$	$\Delta G_2 + \Delta G_{12} = -13.2$	0.01284/s
$ROR2 + (COR1 + ROR3) \rightarrow CI2 + OR2 + (COR1 + ROR3)$	$\Delta G_2 + \Delta G_{23} = -13.4$	0.00928/s
Dissociation reactions - CI_2 dissociation from OR3		
$ROR3 + (OR2) \rightarrow CI2 + OR3 + (OR2)$	$\Delta G_3 = -9.5$	5.19753/s
$ROR3 + (ROR2 + OR1) \rightarrow CI2 + OR3 + (ROR2 + OR1)$	$\Delta G_3 + \Delta G_{23} = -12.4$	0.04702/s
$ROR3 + (ROR2 + ROR1) \rightarrow CI2 + OR3 + (ROR2 + ROR1)$	$\Delta G_3 = -9.5$	5.19753/s
$ROR3 + (ROR2 + COR1) \rightarrow CI2 + OR3 + (ROR2 + COR1)$	$\Delta G_3 + \Delta G_{23} = -12.4$	0.04702/s
$ROR3 + (COR2) \rightarrow CI2 + OR3 + (COR2)$	$\Delta G_3 = -9.5$	5.19753/s
Dissociation reactions - Cro_2 dissociation from OR1		
$COR1 + (OR2) \rightarrow Cro2 + OR1 + (OR2)$	$\Delta G_{1*} = -12.0$	0.08999/s
$COR1 + (ROR2) \rightarrow Cro2 + OR1 + (ROR2)$	$\Delta G_{1*} = -12.0$	0.08999/s
$COR1 + (COR2 + OR3) \rightarrow Cro2 + OR1 + (COR2 + OR3)$	$\Delta G_{1*} + \Delta G_{1*2*} = -13.0$	0.01776/s
$COR1 + (COR2 + ROR3) \rightarrow Cro2 + OR1 + (COR2 + ROR3)$	$\Delta G_{1*} + \Delta G_{1*2*} = -13.0$	0.01776/s
$COR1 + (COR2 + COR3) \rightarrow Cro2 + OR1 + (COR2 + COR3)$	$\Delta G_{1*} + \Delta G_{1*2*3*} - \Delta G_{2*3*} = -12.3$	0.05531/s
Dissociation reactions - Cro_2 dissociation from OR2		
$COR2 + (OR1 + OR3) \rightarrow Cro2 + OR2 + (OR1 + OR3)$	$\Delta G_{2*} = -10.8$	0.6306/s
$COR2 + (ROR1 + OR3) \rightarrow Cro2 + OR2 + (ROR1 + OR3)$	$\Delta G_{2*} = -10.8$	0.6306/s
$COR2 + (OR1 + ROR3) \rightarrow Cro2 + OR2 + (OR1 + ROR3)$	$\Delta G_{2*} = -10.8$	0.6306/s
$COR2 + (ROR1 + ROR3) \rightarrow Cro2 + OR2 + (ROR1 + ROR3)$	$\Delta G_{2*} = -10.8$	0.6306/s
$COR2 + (COR1 + OR3) \rightarrow Cro2 + OR2 + (COR1 + OR3)$	$\Delta G_{2*} + \Delta G_{1*2*} = -11.8$	0.12448/s
$COR2 + (OR1 + COR3) \rightarrow Cro2 + OR2 + (OR1 + COR3)$	$\Delta G_{2*} + \Delta G_{2*3*} = -11.4$	0.23822/s
$COR2 + (COR1 + COR3) \rightarrow Cro2 + OR2 + (COR1 + COR3)$	$\Delta G_{2*} + \Delta G_{1*2*3*} = -11.7$	0.14641/s
$COR2 + (ROR1 + COR3) \rightarrow Cro2 + OR2 + (ROR1 + COR3)$	$\Delta G_{2*} + \Delta G_{2*3*} = -11.4$	0.23822/s
$COR2 + (COR1 + ROR3) \rightarrow Cro2 + OR2 + (COR1 + ROR3)$	$\Delta G_{2*} + \Delta G_{1*2*} = -11.8$	0.12448/s
Dissociation reactions - Cro_2 dissociation from OR3		
$COR3 + (OR2) \rightarrow Cro2 + OR3 + (OR2)$	$\Delta G_{3*} = -13.4$	0.00928/s
$COR3 + (ROR2) \rightarrow Cro2 + OR3 + (ROR2)$	$\Delta G_{3*} = -13.4$	0.00928/s
$COR3 + (COR2 + OR1) \rightarrow Cro2 + OR3 + (COR2 + OR1)$	$\Delta G_{3*} + \Delta G_{2*3*} = -14.0$	0.00351/s
$COR3 + (COR2 + ROR1) \rightarrow Cro2 + OR3 + (COR2 + ROR1)$	$\Delta G_{3*} + \Delta G_{2*3*} = -14.0$	0.00351/s
$COR3 + (COR2 + COR1) \rightarrow Cro2 + OR3 + (COR2 + COR1)$	$\Delta G_{3*} + \Delta G_{1*2*3*} - \Delta G_{1*2*} = -13.3$	0.01092/s

Table S3: Free energies of wild type phage and mutants from [7, 16]. The subscripts of ΔG , 1, 2, and 3 stand for binding energies of Cl_2 to OR1, OR2 and OR3, respectively. Accordingly, 1*, 2* and 3* are for the binding Cro_2 to OR1, OR2 and OR3, respectively. The subscripts of ΔG , 12 and 23 stand for cooperatively binding energies of Cl_2 among OR1-OR2 and OR2-OR3, respectively. Accordingly, 1*2*, 2*3*, and 1*2*3* stand for cooperatively binding energies of Cro_2 among OR1-OR2, OR2-OR3, and OR1-OR2-OR3, respectively.

$\Delta G(\text{kcal/mol})$	WT	1-2-1	3-2-3	1-2-3	3'-2-3'
ΔG_1	-12.5	-12.5	-9.5	-9.5	-13.3
ΔG_2	-10.5	-10.5	-10.5	-10.5	-10.5
ΔG_3	-9.5	-12.5	-9.5	-12.5	-13.3
ΔG_{1*}	-12.0	-12.0	-13.4	-13.4	-10.7
ΔG_{2*}	-10.8	-10.8	-10.8	-10.8	-10.8
ΔG_{3*}	-13.4	-12.0	-13.4	-12.0	-10.7
ΔG_{12}	-2.7	-2.7	-2.9	-2.9	-2.9
ΔG_{23}	-2.9	-2.7	-2.9	-2.7	-2.9
ΔG_{1*2*}	-1.0	-1.0	-0.6	-0.6	-0.6
ΔG_{2*3*}	-0.6	-1.0	-0.6	-1.0	-0.6
ΔG_{1*2*3*}	-0.9	-0.9	-0.9	-0.9	-0.9

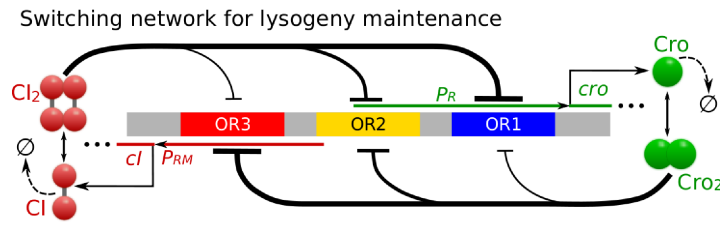


Figure S1: Model of the epigenetic circuit for lysogeny maintenance.

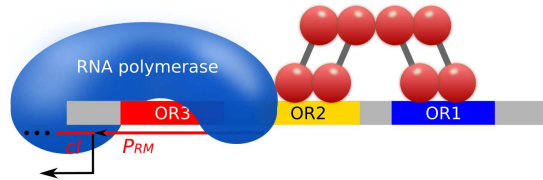


Figure S2: An example of cooperative binding in our model. When either of O_{R1} or O_{R2} site is occupied by CI_2 , the next CI_2 will bind to the remaining empty site with enhanced affinity.

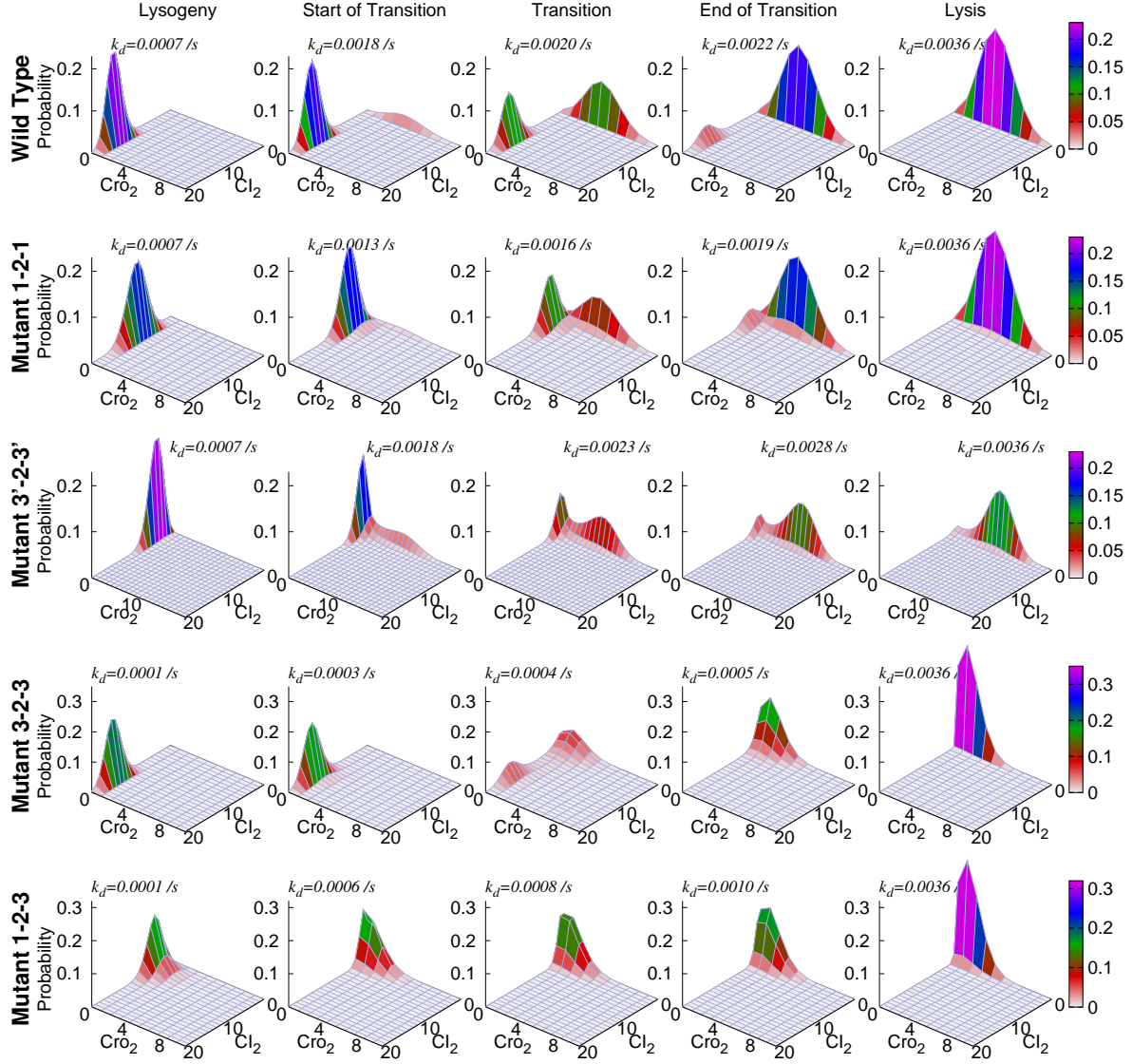


Figure S3: Probability landscapes of wild type, mutant 1-2-1, mutant 3'-2-3', mutant 3-2-3, and mutant 1-2-3 phage lambda at different CI degradation rate k_d .

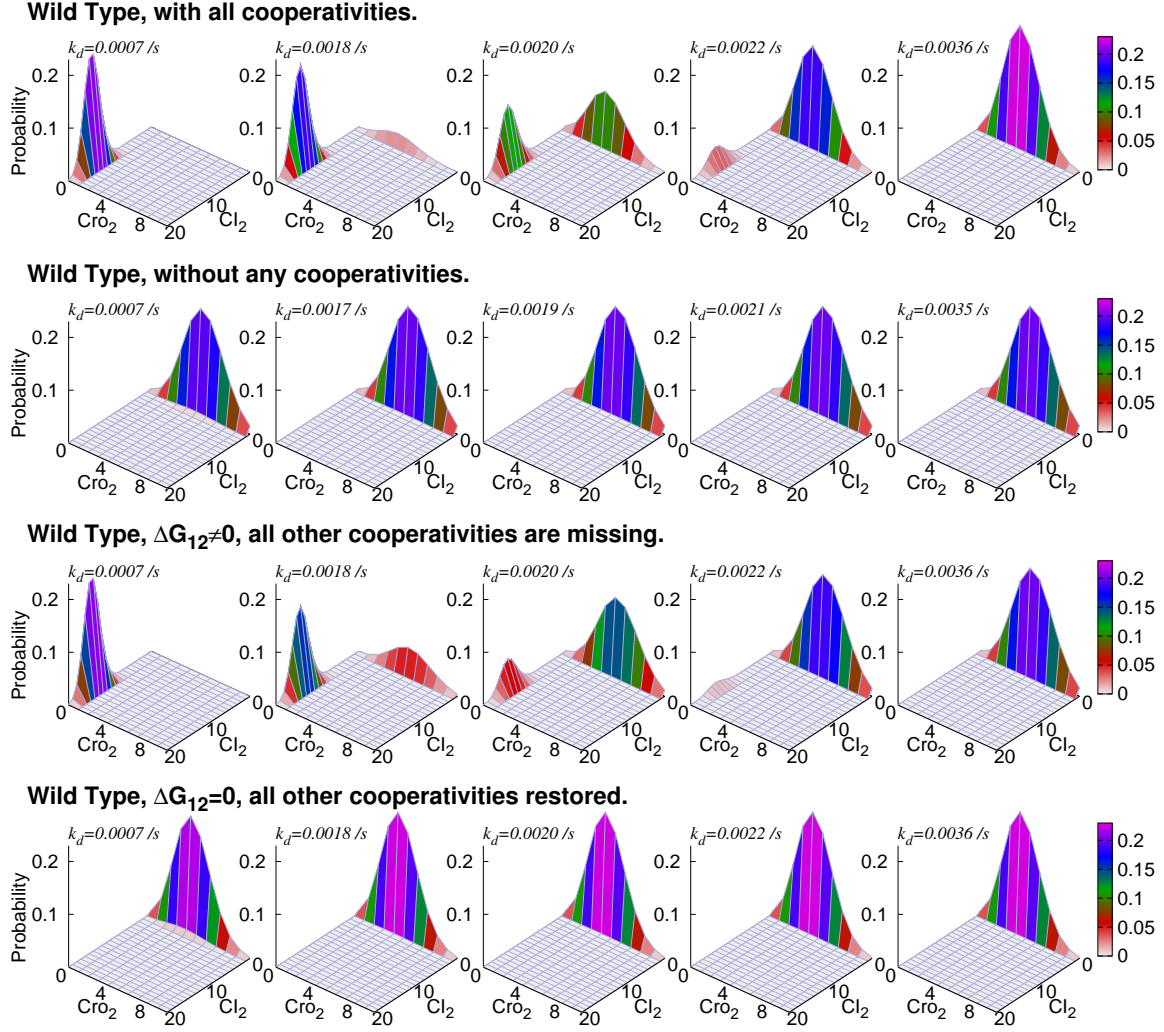


Figure S4: Probability landscapes of wild type phage lambda at different CI degradation rate k_d with all cooperativities intact, all cooperativities removed, with only cooperativity between Cl_2 binding to O_{R1} and O_{R2} restored, and with all but that between Cl_2 binding to O_{R1} and O_{R2} restored.

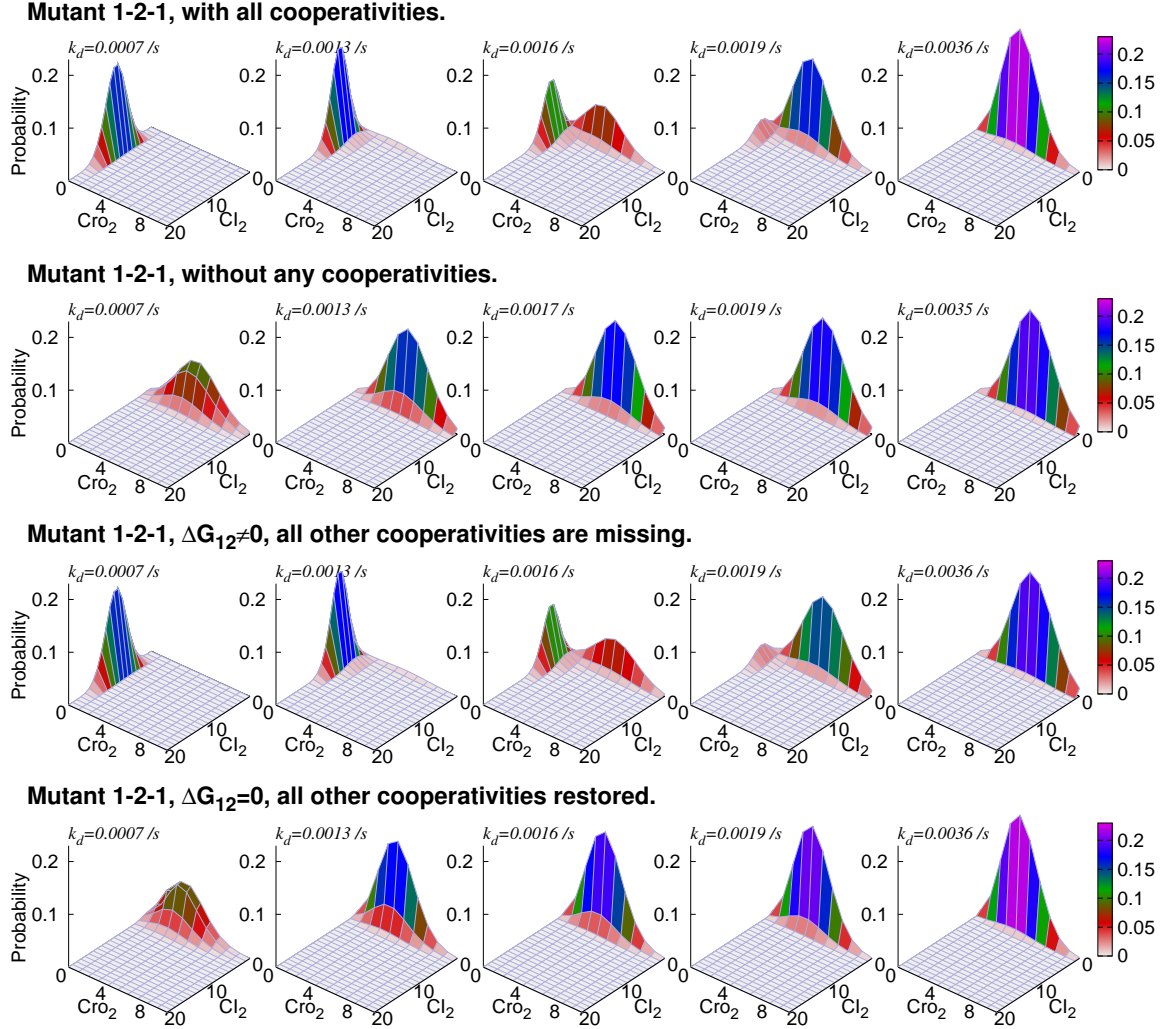


Figure S5: Probability landscapes of mutant 1-2-1 phage lambda at different CI degradation rate k_d with all cooperativities intact, all cooperativities removed, with only cooperativity between Cl_2 binding to O_{R1} and O_{R2} restored, and with all but that between Cl_2 binding to O_{R1} and O_{R2} restored.

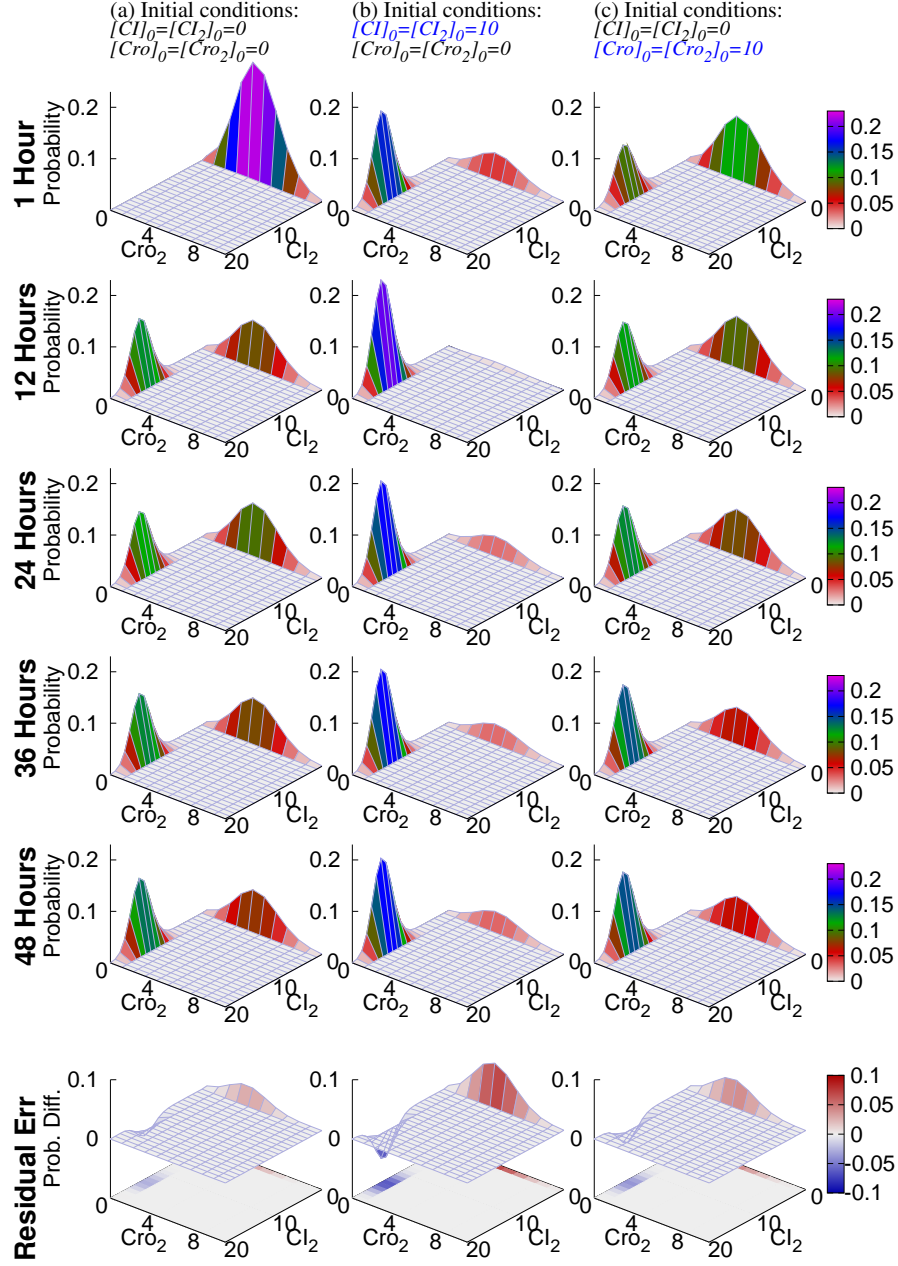


Figure S6: Probability landscape at lysogeny-lysis transition phase ($k_d = 0.0020/s$) computed using stochastic simulation algorithm have not converged after 48 hours when started at different initial conditions. Simulations are run using exactly the same model and parameters for wild-type phage lambda. The true steady state landscape can be seen in the 3rd plot in the first row of Fig. S3. The residual errors between probability landscape distributions obtained from stochastic simulations and the exact solution obtained from dCME are significant after 48 hours of computation (last row).

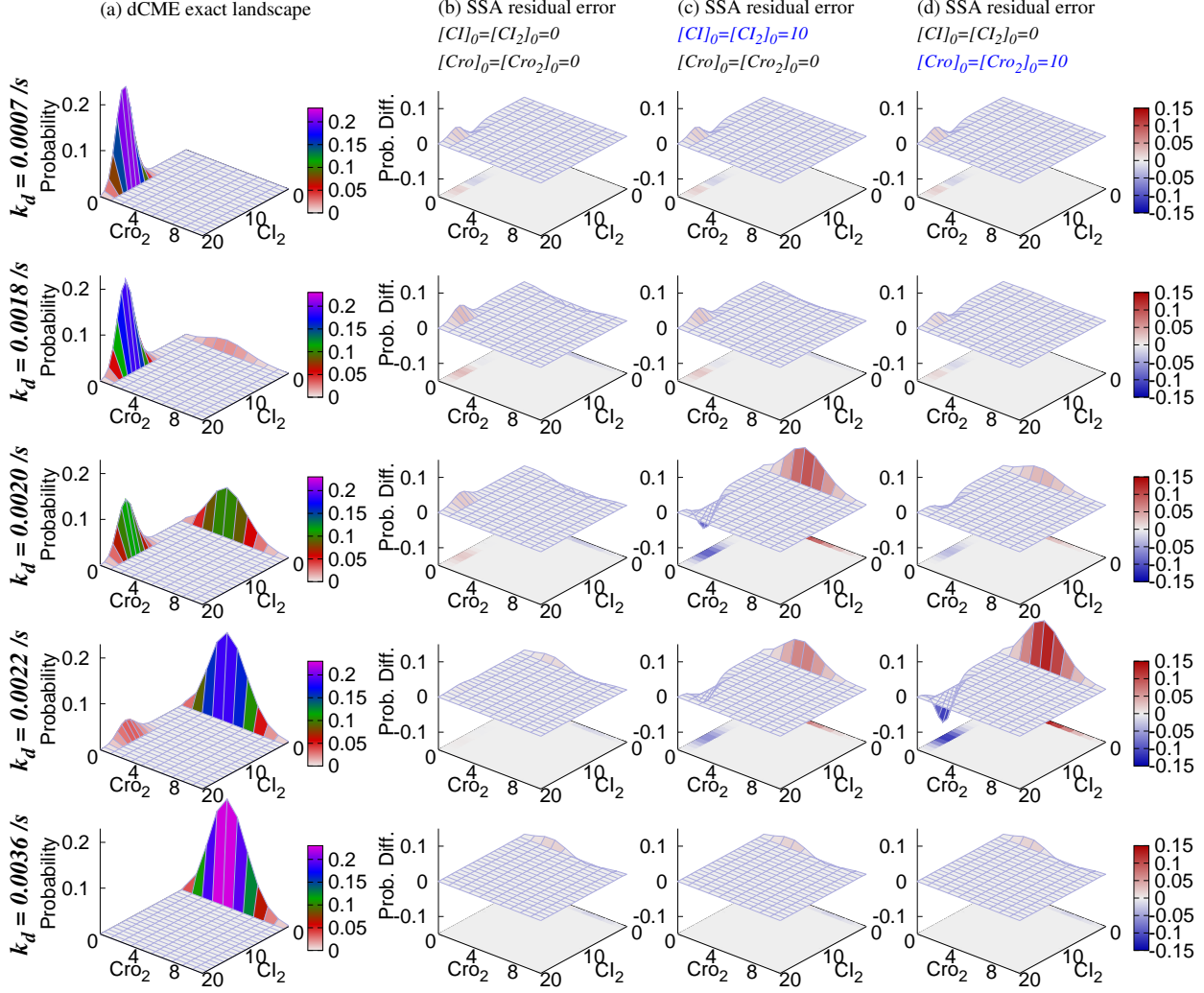


Figure S7: Probability landscape at different regions and corresponding residual errors from SSA simulations after 8 hours. (Left column): The probability landscape from dCME in the lysogenic region ($k_d = 0.0007/s$), at the beginning of the transition region ($k_d = 0.0018/s$), at the transition region ($k_d = 0.0020/s$), at the end of the transition region ($k_d = 0.0022/s$), and in the lytic region ($k_d = 0.0036/s$) are plotted. (Right 3 columns): Residual error from stochastic simulation algorithm with three different initial conditions. For ease of visualization, we plot $P_{\text{dCME}}(x) - P_{\text{SSA}}(x)$ projected onto the $[Cl_2]$ - $[Cro_2]$ space.

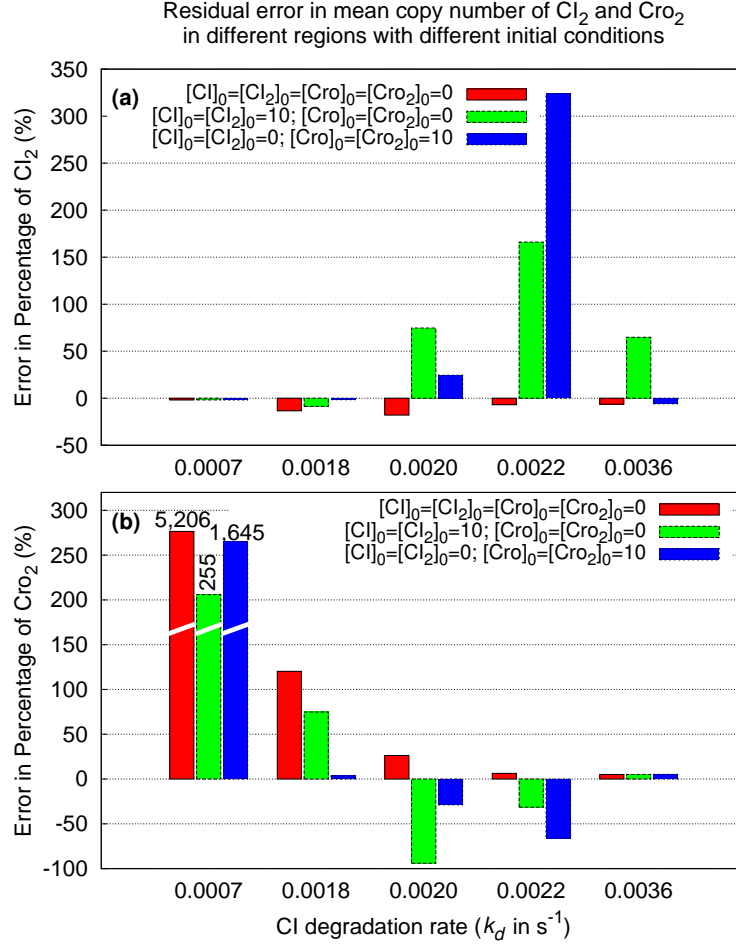


Figure S8: Errors in estimated mean copies of Cl_2 and Cro_2 from SSA simulations at different regions and different initial conditions. (a) Error in estimated copy numbers of Cl_2 ($\mathbb{E}[Cl_2]^{SSA} - \mathbb{E}[Cl_2]^{dCME}) / \mathbb{E}[Cl_2]^{dCME}$ in percentage. (b) Error in estimated copy numbers of Cro_2 ($\mathbb{E}[Cro_2]^{SSA} - \mathbb{E}[Cro_2]^{dCME}) / \mathbb{E}[Cro_2]^{dCME}$ in percentage.

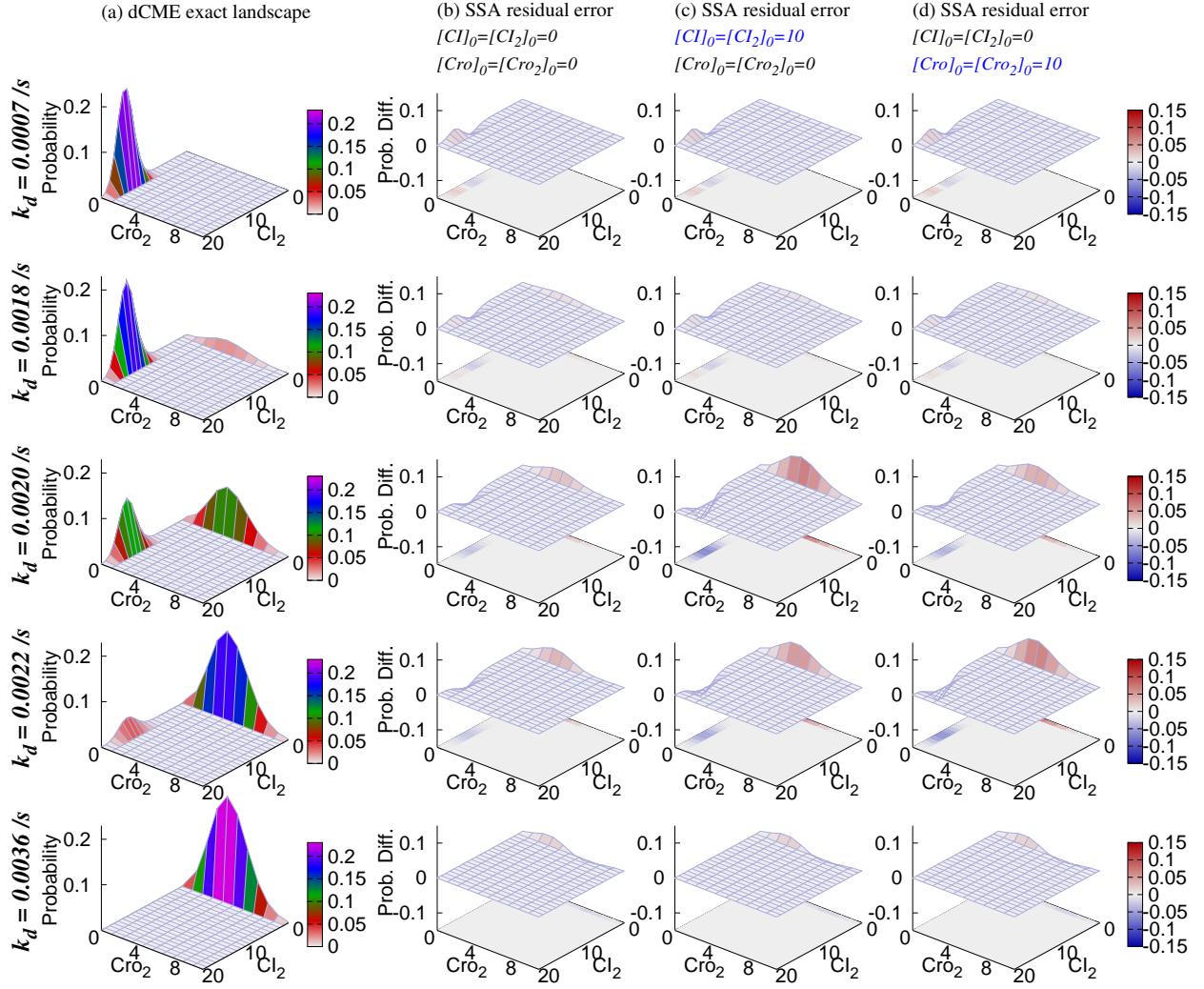


Figure S9: Same as Fig S7 but with SSA simulations after 48 hours.

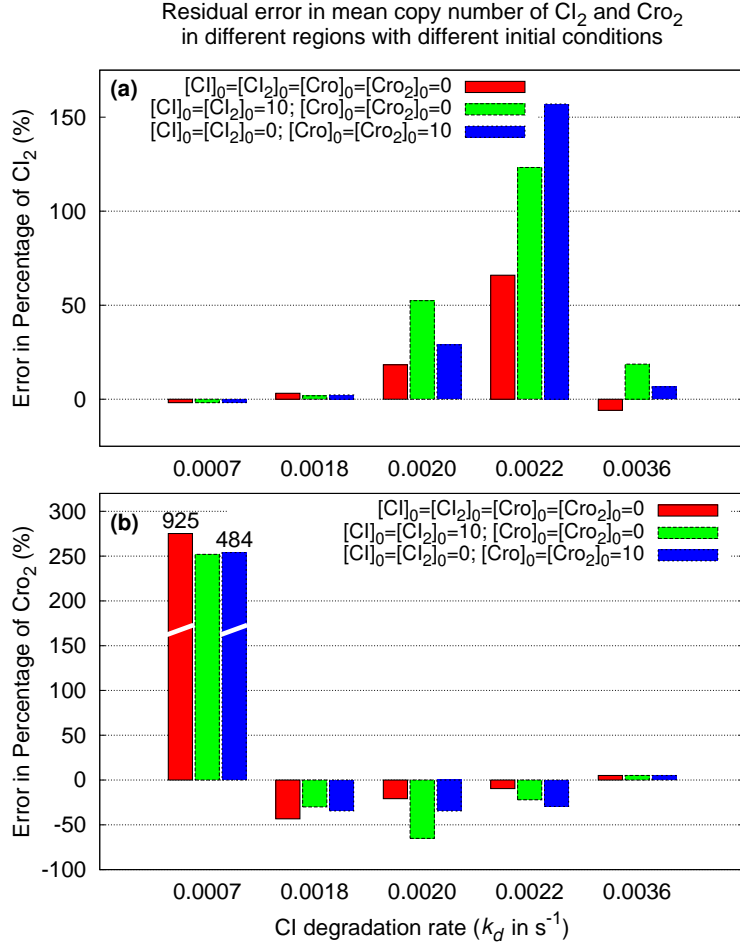


Figure S10: Same as Fig S8 but with SSA simulations after 48 hours.

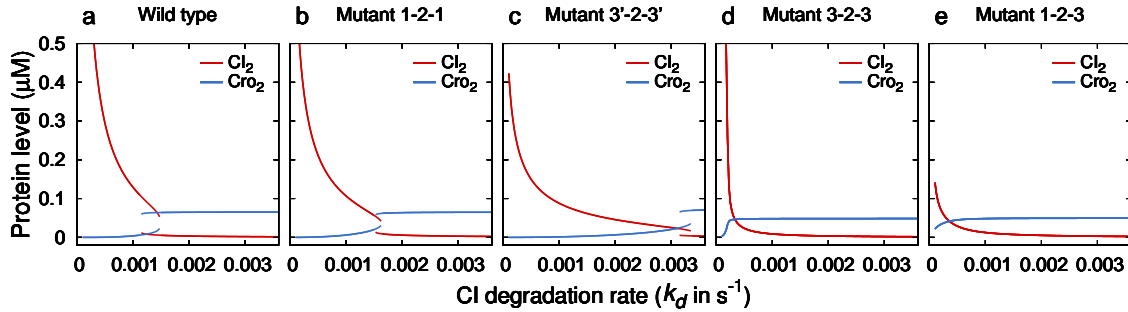


Figure S11: Bifurcation curves of the deterministic model for wild-type phage lambda and the Little's mutants. The deterministic model was based on the formulation of the model from [25], and was built using the same parameters as in the chemical master equation model [27]. Red lines represent the concentrations of Cl_2 in steady state for different CI degradation rates, and the blue lines represent that of Cro_2 in steady state.

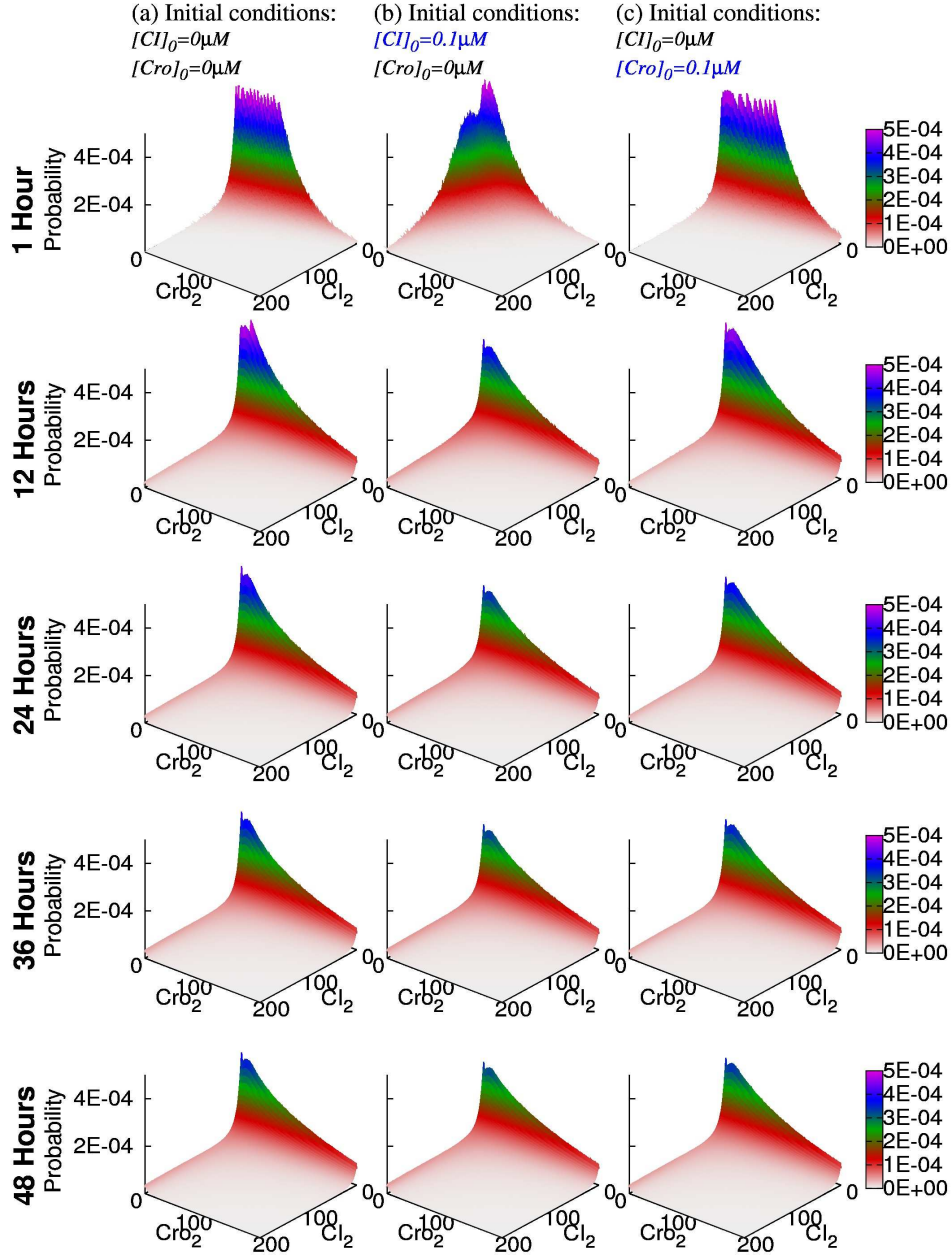


Figure S12: Probability landscape at lysogeny-lysis transition phase ($k_d = 0.0020/s$) computed using stochastic differential equations (SDE) have not converged after 48 hours when started at different initial conditions, and is further away from steady state than results from stochastic simulation. SDE model is built according to the convention of chemical Langevin equation [10] by adding Gaussian noises to ODE model.

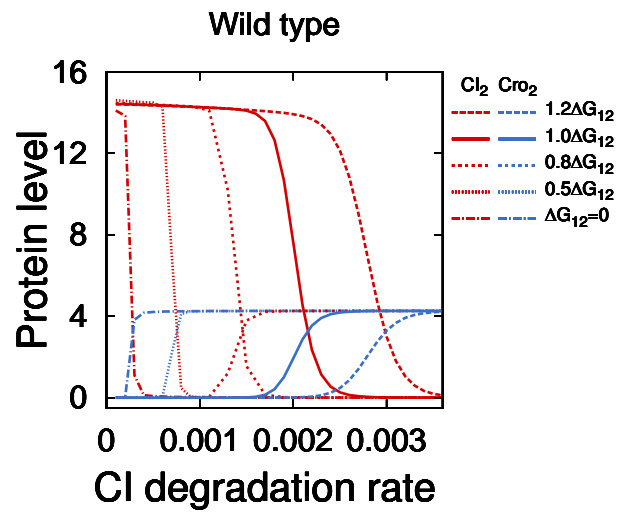


Figure S13: Changes in the binding cooperativity between Cl dimers on O_{R1} and O_{R2} shifts the switching threshold, while the lysogens remain stable. Red lines represent the amounts of Cl_2 , and blue lines the amount of Cro_2 . Solid lines, long dashed lines, short dashed lines, dotted lines, and dash-dot lines stand for the ΔG_{12} modified by a factor of 1.0, 1.2, 0.8, 0.5 and 0.0, respectively.