Binding cooperativity in phage λ is not sufficient to produce an effective switch

Tomáš Gedeon [*] [†‡] **, Konstantin Mischaikow** [§] [¶]**, Kate Patterson** [*] **, and Eliane Traldi** [¶]

[*] Department of Mathematics, Montana State University, Bozeman, MT 59715, [†] Center for Computational Biology, Montana State University, Bozeman, MT 59715, [§] Department of Mathematics, Rutgers, The State University of New Jersey, 110 Frelinghuysen Road, Piscataway, NJ 08854, and [¶] BioMaPS Institute, Rutgers, The State University of New Jersey, 110 Frelinghuysen Road, Piscataway, NJ 08854

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

In the wild type phage λ, binding of CI to O_R2 helps RNA polymerase bound to P_{RM} transition from a closed to open complex. Activators on other promoters increase the polymerase-DNA binding energy, or affect both the binding energy and the closed-open transition probability. Using a validated mathematical model we show that these two modes of upregulation have very different effects on the promoter function. We predict that if CI_2 bound to O_R2 produced equal increase in RNAP-DNA binding constant (compared to wild type increase in the closed-open transition probability), the lysogen would be significantly less stable.

phage λ | cooperativity | transcriptional activation | switch

Although it is premature to assert that we have entered the era of synthetic biology, the groundwork for targeted design of functioning living organisms is being laid. The manipulation of DNA within an organism is by now a standard laboratory practice. Recent work has shown the feasibility of complete genome transplantation [1]. Thus, the tools exist, but to use them effectively requires the ability to design elements of signal transduction and gene regulatory networks. While this has been done [2, 3] much remains to be understood both on the level of the construction of the individual components to the design of the networks themselves. The focus of this paper is on the former.

Transcriptional control plays a fundamental role in gene expression. The initiation of transcription involves a series of reactions which can be summarized into three steps[1]: (*binding*) RNA polymerase binds to promoter DNA yielding a closed RNA polymerase promoter complex, (*opening*) RNA polymerase unwinds a short segment of DNA yielding an open RNA polymerase promoter complex, (*escape*) after abortive cycles of synthesis and release of short RNA products, the RNA polymerase escapes the promoter and enters into productive synthesis of RNA. The activation and repression of transcription initiation is primarily caused by regulatory proteins and the structure of DNA. Regulated recruitment [4] provides a conceptual model for this process. Considerable progress has been made in understanding the biochemistry of the various reactions in the process [5, 6] and in particular, it is clear that while the three steps are physically coupled there is considerable freedom for varying the respective energy profiles. To model these steps in the simplest way we will treat opening and escape as a single chemical reaction with forward reaction rate k determined by the regulatory proteins and their interaction with the DNA. Binding will be treated as a reversible reaction with an equilibrium constant K_B.

This simplification of the biochemistry allows one to develop thermodynamic models to quantify the rates of transcription initiation [5, 7, 8] that can be validated against experimental data [9, 10]. However, the combination of activators, repressors, and the above mentioned steps implies that control of transcription initiation is a highly nonlinear process, which in turn suggests that systematic mathematical analysis may lead to a deeper understanding of this regulatory mechanism [11]. Given the goal of synthetic biology, claims based on the mathematical models must be experimentally verifiable.

More is known about the phage λ machinery than any other gene regulation mechanism [4, 12]. After infection of *E. coli* the phage λ follows one of two pathways: *lysis*, where it uses the bacterial molecular machinery to make many viral copies, kills the host bacterium and leaves to infect other cells; or *lysogeny*, where it integrates its DNA into the bacterial DNA and divides for generations with the bacterium. The lysogen exhibits great stability, yet it induces readily when the bacteria are irradiated with ultraviolet light.

The primary objective of this paper is to use the above mentioned mathematical models to demonstrate that in the context of the proper functioning of the phage λ induction the binding constant K_B plays a fundamentally different role from the opening and clearing constant k. In particular, they are *not* interchangeable; that is, modifications in K_B cannot be directly compensated for by modifications in k and vice versa. To make this argument we begin, in Section with a review of a simplified biological model of the phage λ switch and a precise statement of why increases in K_B are not equivalent to increases in k. In Section we recall and explain the associated mathematical model and in Section relate it back to the biology. We validate the model in Section by considering several mutants, where our model recovers experimental observations of the lysogen stability. With this justification, in Section we make several mathematical predictions concerning the unequal role played by RNAP binding versus closed-open complex transition in transcription initiation process. These predictions are in principle experimentally testable.

The phage λ switch

The central controlling region for the lysogen maintenance is the *right operator* O_R, even though the long range cooperative binding with the O_L operator plays a crucial role in stability of the lysogen. For a more complete description of the regulatory mechanisms the reader is referred to [4]. O_R has three subregions designated O_R1, O_R2 and O_R3 (see Figure 1). The O_R region also contains two disjoint promoters P_R (*R*ight promoter) and P_{RM} (*R*epression *M*aintanence promoter). The promoter P_R completely overlaps O_R1 and partially overlaps O_R2, while P_{RM} completely overlaps O_R3 and partially overlaps O_R2. The

Conflict of interest footnote placeholder

Insert 'This paper was submitted directly to the PNAS office.' when applicable.

‡ To whom correspondence should be addressed. E-mail: gedeon@math.montana.edu

© 2007 by The National Academy of Sciences of the USA

[1] There are additional controls which occur in later stages of the process of transcription, but are not considered in this paper.

gene cI, that codes for the repressor protein CI and the gene cro, that codes for Cro protein, flank the O_R region. Binding of either CI or Cro dimers (CI_2, Cro_2) to O_R2 prevents binding of RNA polymerase (RNAP) to P_R, but it does not prevent such binding to P_{RM}. The initiation of transcription of cro occurs only if RNAP binds to P_R. Similarly, the initiation of transcription of cI occurs only if RNAP binds to P_{RM}.

The lytic pathway corresponds to a state where Cro_2 protein is bound to O_R3, blocking the P_{RM} promoter and thus transcription of cI. At the same time RNAP is free to bind P_R, thus maintaining the transcription of cro. The lysogenic pathway corresponds to the state of O_R where CI_2 binds to both O_R2 and O_R1 blocking the P_R promoter and hence the transcription of cro. RNAP is free to bind P_{RM} and thus maintain the transcription of cI. Even though these pathways are stable, the change from lysogeny to lysis, called *induction*, is experimentally well documented. When the bacterial population is subject to irradiation by UV light, the phage λ starts to lyse the bacteria and emerge in about 45 minutes. The irradiation causes RecA protein-mediated cleavage of CI which lowers its effective concentration [4, 13, 14, 15]. There are several key features which makes lysogen very stable and the induction "switch-like" [4].

1. High level of cooperativity between CI molecules: CI forms dimers CI_2 in the solution; when bound to neighboring regions O_R2 and O_R1 (or O_R2 and O_R3) it forms tetramers, and as was described in [4], it forms octomers with CI_2 bound to the O_L operator, which is fairly distant, at 3.6kb, from O_R along the DNA strand.

2. Cooperative binding of CI_2 to O_R2 and O_R1: binding of CI_2 to O_R1 facilitates binding of another CI_2 molecule to O_R2.

3. Variable binding affinities of CI_2 and Cro_2 to different O_R regions: CI_2 has the highest affinity to O_R1, lower for O_R2 and lowest for O_R3, while Cro_2 has the highest affinity to O_R3, lower for O_R2 and O_R1.

4. Cooperative binding of CI_2 to O_R2 and RNAP at P_{RM}: that is, O_R2 bound CI_2 increases the forward rate constant k at P_{RM} about 10-fold without having any significant effect on the binding of the RNA polymerase to the DNA [16].

We refer to the cooperativity in 4 as k-cooperativity. In an intriguing paper Li *et. al.* [17] have shown that after an Arg to His change in the σ subunit of RNAP, the wild type CI activates mutant RNAP by increasing K_B. We will refer to this cooperativity as K_B-cooperativity. This suggests that mutations allowing for an increase in K_B were (and are) evolutionary accessible to the phage. It is therefore likely that k-cooperativity, as opposed to an increase in K_B, has been selected for functional reasons. Further support for this hypothesis is provided by the fact that not all activators increase k. In fact in phage λ the factor CII acting on P_{RE} promoter uses both the K_B- and k-cooperativity [18] and the CAP activation of the *lac* operon in *E. coli* uses K_B-cooperativity [19].

To investigate this hypothesis we model the dynamics of the entire switch and study the effect of the K_B- and k-cooperativity on the stability of the lysogenic state. We show that the stability of the lysogen depends crucially not only on the fact that CI_2 interacts cooperatively with RNAP, but also on the fact that this cooperativity increases k rather than K_B. In fact, our computations suggest that increasing K_B 100 fold while abolishing k-cooperativity yields phage with lysogen that is significantly less stable than the wild type.

The Mathematical Model

We make use of a delay differential equation model developed by Santillán and Mackey [20]:

$$\frac{d[M_{cI}]}{dt} = [O_R]f^c_{RM}([CI_2]_{\tau_M}, [Cro_2]_{\tau_M}) + [O_R]f_{RM}([CI_2]_{\tau_M}, [Cro_2]_{\tau_M}) - (\gamma_M + \mu)[M_{cI}] \quad [1]$$

$$\frac{d[M_{cro}]}{dt} = [O_R]f_R([CI_2]_{\tau_M}, [Cro_2]_{\tau_M}) - (\gamma_M + \mu)[M_{cro}] \quad [2]$$

$$\frac{d[CI]}{dt} = \nu_{cI}[M_{cI}]_{\tau_{cI}} - (\gamma_{cI} + \mu)[CI] \quad [3]$$

$$\frac{d[Cro]}{dt} = \nu_{cro}[M_{cro}]_{\tau_{cro}} - (\gamma_{cro} + \mu)[Cro] \quad [4]$$

which, as is explained below, tracks concentrations of cI mRNA, cro mRNA, CI protein and Cro protein. Concentrations are denoted by square brackets; that is $[CI]$ is the total concentration of CI protein while $[M_{cro}]$ is the concentration of cro mRNA.

We will use $[Cro_2]$ and $[CI_2]$ to denote the concentration of CI and Cro dimers and $[RNAP]$ to denote concentration of the RNA polymerase. The concentration of the right operator is $[O_R]$. The subscript notation $[M_{cro}]_{\tau_{cro}}$ indicates that the concentration of cro mRNA is evaluated at $t - \tau_{cro}$ where t is the present time. The time delays τ_{cI} and τ_{cro} are incorporated to take into account the fact that the production of the proteins from the associated mRNA and the actual process of transcription is not instantaneous.

Equations (3) and (4) are based on the assumption that the changes in protein concentrations are linear functions of the corresponding mRNA concentrations. There are two sets of positive decay constants. Since the volume of the growing bacteria increases, concentrations of all chemicals in a cell decrease. This is modeled by the decay constant δ which is the same in all equations. In addition, each chemical species experiences a specific degradation rate denoted by γ_*. Of particular interest is the constant γ_{cI}. We will model the effect of UV light, which, as is noted earlier lowers the effective concentration of CI dimers, by an increase in the degradation rate γ_{cI} of the CI protein.

The ν_* are positive translation initiation constants.

The change in concentration of mRNA is described by equations (1) and (2). The nonlinear function $f_R([CI_2]_{\tau_M}, [Cro_2]_{\tau_M})$ describes the rate of transcription initiation at the promoter P_R. For the sake of clarity the rate of transcription initiation at the promoter P_{RM} is expressed as the sum of two functions $f^c_{RM}([CI_2]_{\tau_M}, [Cro_2]_{\tau_M})$ and $f_{RM}([CI_2]_{\tau_M}, [Cro_2]_{\tau_M})$, where the first applies to the state of the operator in which CI_2 is bound to O_R2 and the second when it is not.

Santillán and Mackey's [20] construction of these functions is based on the work of Ackers *et. al.* [7] and begins with expressions of the probability of binding of RNAP to the promoter in the presence or absence of the regulatory proteins. The probability of a particular macroscopic state s of the operator takes the form

$$\mathbb{P}_s([CI_2], [Cro_2]) = \frac{K_B(s)[Cro_2]^{\alpha_s}[CI_2]^{\beta_s}[RNAP]^{\gamma_s}}{\sum_i K_B(s_i)[Cro_2]^{\alpha_i}[CI_2]^{\beta_i}[RNAP]^{\gamma_i}} \quad [5]$$

where

$$K_B(s) = e^{\frac{-\Delta G_s}{RT}} \quad [6]$$

and the summation in the denominator is taken over all possible states. Since ΔG_s denotes the binding energy of the state, $K_B(s)$ determines the equilibrium constant for the biochemical reaction that results in binding of the regulatory proteins and/or RNAP to the DNA in a closed form. The right (O_R), the left (O_L) operator (each of which has three subdomains) and the three promoters (P_R, P_{RM}, and P_L) are included

in the model of Santillán and Mackey [20]. Therefore the state s of the operator is a description of which of the nine sites are empty or occupied by which of the three possible molecules CI_2, Cro_2, or RNAP.

These probabilities need to be multiplied by an appropriate constant, $k(s)$, to incorporate the forward reaction rates of the opening and escape steps in order to obtain a rate of transcription initiation. Thus for each state, the transcription initiation rate has the form

$$f_s([CI_2],[Cro_2]) = k(s)\frac{K_B(s)[Cro_2]^{\alpha_s}[CI_2]^{\beta_s}[RNAP]^{\gamma_s}}{\sum_i K_B(s_i)[Cro_2]^{\alpha_i}[CI_2]^{\beta_i}[RNAP]^{\gamma_i}}. \quad [7]$$

Though clearly a simplification, we assume that the rate constants $k(s)$ take on three values: k_{cro} when RNAP is bound to P_R, k^c_{cI} when RNAP is bound to P_{RM} and CI_2 is bound to O_R2, and k_{cI} when RNAP is bound to P_{RM} and CI_2 is not bound to O_R2.

Finally, f_R is the sum of all combinations of (7) with the restriction that each state s has a RNAP bound to P_R, with O_R1 and O_R2 unbound. Similarly, f^c_{RM} is the sum of (7) for all states s which have RNAP bound to P_{RM} and CI_2 bound to O_R2, and f_{RM} the sum of (7) for all states s which have RNAP bound to P_{RM} but CI_2 is not bound to O_R2.

To compare this model against experimental data, requires knowledge of the above mentioned constants. The experimentally determined values are taken from [20] and presented in Tables 1 and 2 in the Appendix.

Interpreting the Model

Based on the biochemistry of the phage λ switch, the phenomenological state of lysogeny is associated with low levels of Cro and high levels of CI. Similarly, lysis is associated with low levels of CI and high levels of Cro. With this in mind, we look for equilibria of the system (1)-(4) and declare that an equilibrium for which $0 \approx$ [Cro] << [CI] is a lysogenic equilibrium and an equilibrium for which $0 \approx$ [CI] << [Cro] is a lytic equilibrium.

The equilibria of this system are steady (time independent) states of the system and thus are not dependent on delays. Notice that since both CI and Cro proteins form dimers, the right hand side of the equations depend on the concentration of dimers. The conversion formula for computing the concentration of dimers from total concentration of monomers is

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

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

and its derivation is presented in the Appendix.

Let

$$\phi := \frac{k^c_{cI}}{k_{cI}}.$$

Observe that this provides a measure of the effect of O_R2 bound CI_2 on the forward reaction rate associated with opening and escape. In particular, $\phi > 1$ implies that the rate of transcription initiation with O_R2 bound CI_2 is higher than that without. We refer to this as *k-cooperativity*.

As is indicated in Section , γ_{cI} indicates the degradation rate of $[CI]$, induced for example by exposure to UV radiation. Since this is known to trigger induction of phage, we study the equilibria as a function of γ_{cI}. Observe that the equilibria satisfy the two equations

$$\Phi([CI],[Cro],\gamma_{cI}) = 0 \qquad \text{and} \qquad \Theta([CI],[Cro]) = 0$$

where

$$\begin{aligned}\Phi([CI],[Cro],\gamma_{cI}) &= \frac{\nu_{cI}}{\gamma_M+\mu}[O_R]\Big(f^c_{RM}([CI_2],[Cro_2]) \\ &\quad + f_{RM}([CI_2],[Cro_2])\Big) - (\gamma_{cI}+\mu)[CI] \\ \Theta([CI],[Cro]) &= \frac{\nu_{cro}}{\gamma_M+\mu}[O_R]f_R([CI_2],[Cro_2]) \\ &\quad -(\gamma_{cro}+\mu)[Cro].\end{aligned}$$

The intersection of these two curves in the [CI], [Cro] plane determines two protein concentrations at a dynamical equilibrium; the remaining two concentrations $[M_{cI}]$ and $[M_{cro}]$ can be found from equations (3) and (4) with the left hand side set equal to zero.

Observe that Θ is independent of γ_{cI}. The set $\Theta([CI],[Cro]) = 0$ is given by the black curve in Figure 2.A. According to Table 1, for wild type phage in the absence of UV radiation, $\gamma_{cI} = 0$ min^{-1}. The set $\Phi([CI],[Cro],0) = 0$ is plotted in red dash in Figure 2.A. There is a unique equilibrium, i.e. intersection point of $\Theta([CI],[Cro]) = 0$ and $\Phi([CI],[Cro],0) = 0$, for which $[CI] = 0.528\ \mu M$ and $[Cro] = 1.04 \times 10^{-5}\mu M$. This is a *lysogenic equilibrium*.

As the parameter γ_{cI} increases the $\Phi = 0$ curve shifts its relative position relative to the $\Theta = 0$ curve. When γ_{cI} is 0.00039 min^{-1}, a pair of new intersections corresponding to new equilibria appear. Plotted in blue dots in Figure 2.A is $\Phi([CI],[Cro],0.05) = 0$. The equilibrium with high value of $[Cro]$ and low value of $[CI]$ corresponds to lytic state and we call it a *lytic equilibrium*. Observe that there are three equilibria: a lysogenic equilibrium, a lytic equilibrium, and an unstable intermediate equilibrium. Finally, the green dash-dot curve represents $\Phi([CI],[Cro],0.35) = 0$ which intersects $\Theta = 0$ in a single point corresponding to the lytic equilibrium.

Clearly, the set of equilibria changes as a function of γ_{cI}. This is indicated in the bifurcation diagram of Figure 2.B, where the equilibrium values of $[Cro]$ are plotted on the vertical axis as a function of γ_{cI}. This graph allows us to describe the induction process. When no UV radiation is applied to bacterial population, $\gamma_{cI} = 0$ min^{-1} and the phage occupies lysogenic equilibrium. As γ_{cI} is slowly increased, the lysogenic equilibrium moves and the phage state tracks this slowly moving equilibrium. Immediately after γ_{cI} crosses the value of 0.343 the lysogenic equilibrium disappears and the state rapidly approaches the lytic equilibrium.

Therefore we define the value $\gamma^*_{WT} := 0.343$ min^{-1} as the *wild type induction value*. The color code in Figure 2.B shows the values of γ_{cI} that correspond to the same color curves in Figure 2.A.

In Sections and we make use of bifurcation diagrams such as that of Figure 2.B, thus we point out some of the important features. For the parameter values 0.00039 min^{-1} $\leq \gamma_{cI} \leq$ 0.343 min^{-1} the wild type phage λ switch is *bistable*; that is there are two stable equilibria, the lysogenic equilibrium (corresponding to the lower branch) and the lytic equilibrium (corresponding to the upper branch), and furthermore, for some initial concentrations the state of the phage will evolve toward the lysogenic equilibrium and for other initial concentrations toward the lytic equilibrium.

We introduced the dimensionless parameter ϕ to have a measure of the change in the forward reaction rate associated with opening and escape. We wish to have a similar measure for the binding probabilities. When the binding of a transcription factor increases RNAP residence time on the promoter, it is reflected in the Ackers model in the cooperative increase of the binding energy of the transcription factor-RNAP pair. We denote the binding energy between CI_2 and O_R2 by $\Delta G^{CI_2}_{O_R2}$ and binding energy between RNAP and P_{RM} by

$\Delta G^{\text{RNAP}}_{\text{P}_{\text{RM}}}$. In the absence of binding cooperation, as is the case in the wild type phage λ, the binding energy contribution from O_R2-bound CI and P_{RM}-bound RNAP to any state s that contains them is

$$\Delta G_{\text{ind}}(s) := \Delta G^{\text{CI}_2}_{\text{O}_\text{R}2} + \Delta G^{\text{RNAP}}_{\text{P}_{\text{RM}}} + \Delta G_{\text{rest}}(s),$$

where subscript 'ind' stands for independent binding of the binding factors and $\Delta G_{\text{rest}}(s)$ is the binding energy of the other factors in state s.

The cooperative binding between CI_2 and RNAP is reflected in additional binding energy $\Delta G^{\text{CI}_2\text{RNAP}}_{\text{O}_\text{R}2\text{P}_{\text{RM}}}$. If this energy is positive we refer to this as *K_B-cooperativity*. We express the cooperativity in terms of the binding constant $K_B(s)$ (see (6))

$$K_B(s) := \beta K^{\text{ind}}_B(s)$$

where $K^{\text{ind}}_B(s) = exp(-\frac{1}{RT}(\Delta G_{\text{ind}}(s)))$ and the state s independent multiplicative factor

$$\beta := exp(-\frac{1}{RT}(\Delta G^{\text{CI}_2\text{RNAP}}_{\text{O}_\text{R}2\text{P}_{\text{RM}}})).$$

In this formulation $\beta > 1$ represents the cooperative binding.

In summary, the k-cooperativity is manifested by the constant $\phi > 1$ (see section) and K_B-cooperativity by $\beta > 1$.

Model validation

In order to validate our biological interpretation of the equilibria of equations (1)-(4) we model the induction scenarios for several different phage mutants which are described in the literature.

O_R323 mutant. Little *et. al.* [21] constructed a mutant O_R323 in which the O_R1 domain was replaced by O_R3 and reported the following results:

R1 O_R323 can lysogenize;

R2 O_R323 has a threshold response, but at lower doses of UV radiation and at a higher level of free phage in the lysogen than the wild type;

R3 in the lytic state the burst size i.e. the number of phages per infected cell, of O_R323 is lower than that of wild type.

This mutation is easily incorporated into the mathematical model. To replace the O_R1 binding site by the O_R3 binding site we set the binding energy of CI_2 to O_R1 to be that of CI_2 to O_R3 (-9.5 kcal/mol). Similarly, the binding energy of Cro to O_R1 is set to that of Cro to O_R3 (-12.0 kcal/mol).

The bifurcation curves for the O_R323 mutation as compared with the wild type are presented in Figure 3. The graph shows the concentration of Cro as a function of γ_{cI}. The blue solid curve represents the wild type phage, while the black dashed curve represents the O_R323. The lower branch on both curves corresponds to the lysogenic equilibrium and the upper branch to the lytic equilibrium.

The existence of the lower branch in Figure 3 implies that O_R323 can lysogenize (compare **R1**). However, the induction value for the O_R323 mutant is $\gamma^*_{O_R323} = 0.09\ \text{min}^{-1} < 0.34\ \text{min}^{-1} = \gamma^*_{WT}$, which suggests that a lower level of UV radiation is required to induce lytic growth (compare **R2**). Observe that when $\gamma_{cI} = 0\ \text{min}^{-1}$ there are three equilibria in the system describing O_R323. Thus a stable lytic equilibrium is present even in the absence of UV radiation and thus in the presence of noise some phages can spontaneously induce and switch to lytic state. This would manifest itself experimentally in increased number of free phages (compare **R2**).

Finally, it is possible that the burst size (number of phages per infected cell) is proportional to the transcription level of the lytic pathway in phage's genome, which in turn may be proportional to the level of Cro production in the lytic state. This theory is in agreement with Figure 3 in which the Cro production in the lytic state for O_R323 (the upper black dashed branch) is significantly lower than in the wild-type lytic state (the upper blue solid branch) (compare **R3**). Of course, the burst size can also be determined by energetics of the cell or by available resources, and therefore the suggested relationship between Cro production and the burst size is, at best, speculative.

P_{RM} mutant. Michalowski and Little [22] (see also [23]) obtained multiple mutants of phage λ by subjecting the P_{RM} binding site to mutagenesis. These were then compared to wild type by three criteria: the ability to grow lytically, the ability to establish and maintain a stable lysogenic state, and the ability to undergo prophage induction. In the experiments they were particularly careful not to affect the O_R2 and O_R3 binding sites. Of these isolates they further analyzed nine which were selected because they were comparable to or more difficult to induce than the wild type. When compared to wild type these nine strains seem to share three properties: they had an equal or increased P_{RM} binding affinity, a decreased P_R binding affinity, and an increase in the k-cooperativity between CI_2 and RNA polymerase. To model such mutant we set $P_{RM}= -12.5$ kcal/mol, $P_R= -10.5$ kcal/mol, and $\phi = 4.5/.35$, which should be compared to wild type values $P_{RM}= -11.5$ kcal/mol, $P_R= -12.5$ kcal/mol and $\phi = 4.29/.35$. The resulting bifurcation diagrams are presented in Figure 4 The induction parameter $\gamma^*_{\text{P}_{\text{RM}}} \simeq 0.85\ \text{min}^{-1}$ for the mutation is much higher then the wild type $\gamma^*_{WT} \simeq 0.35\ \text{min}^{-1}$ implying greater stability of the lysogen.

cI-pc mutant. When a *pc* mutation is introduced to CI it eliminates the k-cooperativity between CI_2 protein bound to O_R2 and RNAP [4, 25]. This mutant forms lysogen in a wild-type bacteria, but suffers from high rate of spontaneous induction and induction at a very low levels of UV light.

To model this mutant we replace the k^c_{cI} in the function f^c_{RM} (see equation (2)) by k_{cI}. This implies $\phi = 1$. The associated bifurcation curves are indicated in Figure 5. Observe that our model predicts that the induction value is dramatically lower ($\gamma^*_{WT} = 0.34\ \text{min}^{-1}$ in wild type, $\gamma^*_{CIpc} = 0.01\ \text{min}^{-1}$ in the mutant). In the noisy environment of a cell we expect that this low stability threshold will yield a high spontaneous induction rate.

K_B- and k-cooperativity are not interchangeable

Our most significant prediction is that K_B- and k-cooperativity affect the stability of the lysogen differently, and thus are not interchangeable. To demonstrate this we compare the stability of the lysogen under k-cooperativity, $\beta = 1, \phi = \alpha > 1$, against K_B-cooperativity, $\phi = 1, \beta = \alpha > 1$, for different values of α. The analysis of the stability of the cI-pc mutant in Section provides the first step of this analysis. In this mutant both $\phi = 1$ and $\beta = 1$, thus all cooperation is abolished and our model predicts that the induction value is dramatically lower.

To test the ability of K_B-cooperativity to restore the lysogen stability, we fix $\phi = 1$ and solve for the equilibria at $\beta = 10$ and $\beta = 100$. The bifurcation diagrams are presented in Figure 6 where they can be compared against the cI-pc mutant and the wild type (recall that for the wild type $\phi \approx 12$ and $\beta = 1$). Observe that when $\beta = 10$, the induction value is $\gamma^*_{\beta=10} = 0.04\ \text{min}^{-1}$ which is much lower than $\gamma^*_{WT} = 0.34\ \text{min}^{-1}$. We predict that this produces a very unstable lysogen. Even in the case of unrealistically strong K_B-cooperativity, $\beta = 100$, the induction value is only $\gamma^*_{\beta=100} = 0.07\ \text{min}^{-1}$.

Figure 6 clearly indicates that K_B- and k-cooperativity are not equivalent. This difference is highlighted in Figure 7 where isoclines

of the induction value γ_* are plotted as a function of β and ϕ. The deviation of symmetry across the diagonal $\beta = \phi$ indicates the extent to which K_B- and k-cooperativity fail to be equivalent in maintaining the stability of the lysogenic state.

While Figures 6 and 7 clearly indicate that there is a difference between K_B- and k-cooperativity, they provide no explanation for this difference. Since the interactions between the binding factors are mediated through nonlinear functions we do not expect there to be a simple, but complete quantitative description of this difference. However, there are two mathematical results that provide a partial explanation.

The first has to do with the rate of production of CI. Let

$$f_{RM}^{\beta,\phi}([CI_2],[Cro_2]) := f_{RM}^{c}([CI_2],[Cro_2]) + f_{RM}([CI_2],[Cro_2])$$

for fixed values of β and ϕ. By [11, Theorem 4.8], if $\alpha > 1$, then

$$f_{RM}^{1,\alpha}([CI_2],[Cro_2]) > f_{RM}^{\alpha,1}([CI_2],[Cro_2]).$$

This means that the rate of transcription of cI mRNA is greater under k-cooperativity than under an equal amount of K_B-cooperativity.

The second has do with the biological fact that at low concentrations CI_2 up regulates its own transcription, while at high concentrations is down regulates its own transcription [4]. In the lysogen O_R1 is almost always bound by CI_2 protein and thus the production of Cro is very low. To produce a simple model that can be easily analyzed we assume CI_2 is always bound to O_R1, and thus the states of interest involve the binding of CI_2 to O_R2 and O_R3. In [11, Example 4.11] it is proven that under these assumptions there exists a unique critical concentration κ, such that if $[CI_2] < \kappa$, then CI_2 is an activator and if $[CI_2] > \kappa$, then CI_2 is a repressor. This implies that the maximal production rate of CI mRNA occurs at $[CI]_2 = \kappa$. As is shown in [11, Example 4.13] κ is larger under k-cooperativity than under an equal amount of K_B-cooperativity. In particular, the critical concentration for the wild type is greater than the critical concentration for the cI-pc mutant.

Conclusions

One of the common features of transcriptional control in bacteria and eukaryotes is "activation by recruitment", where subtle interactions between the transcription factors and RNAP control the rate of transcription. The three essential steps in this process (binding, opening and escape) coalesce in the Ackers modeling framework into two sets of constants. One set captures binding energies, while the other models the transcription initiation process which includes both opening and escape. If for some state of the operator the binding of a factor increases the binding probability of RNAP we call it K_B-cooperativity. If on the other hand the factor increases the probability of transcription initiation we call it k-cooperativity.

At the first glance it may appear that these two types of activation are interchangeable. We have shown using an experimentally validated dynamic model of phage λ that with respect to induction of the lysogenic state k- and K_B-cooperativity are not substitutable. Without k-cooperativity the lysogenic state of the phage λ switch is quite unstable and comparable to some known mutants like O_R323 [21].

Our model produced experimentally verifiable predictions and can serve to test hypothesis about induction of phage λ various mutants before they are constructed in the lab. Furthermore, the mathematical techniques and arguments used to obtain these predictions are quite general and thus in the long run we believe that this type of analysis will prove useful for bioengineers who are trying to design novel genetic control units.

Appendix

Computation of binding energies . For each state of the promoter P_R and P_L the transcription initiation rate is

$$f_s([CI_2],[Cro_2]) = k(s)\frac{K_B(s)[Cro_2]^{\alpha_s}[CI_2]^{\beta_s}[RNAP]^{\gamma_s}}{\sum_i K_B(s_i)[Cro_2]^{\alpha_i}[CI_2]^{\beta_i}[RNAP]^{\gamma_i}}$$

where

$$K_B(s) = e^{\frac{-\Delta G_s}{RT}}$$

and ΔG_s is a binding state energy. We calculate these energies using the following formula

$$\begin{aligned}\Delta G_s = &\sum_{X=R,L}\sum_{Y=CI_2,Cro_2}\sum_{\nu=1}^{3}\Delta G^Y_{O_{X\nu}}\Gamma^Y_{O_{X\nu}}(s) \quad + \\ &\sum_{X=R,L}\sum_{Y=CI_2,Cro_2}\sum_{\nu=1}^{2}\Delta G^Y_{O_{X\nu\nu+1}}\Gamma^Y_{O_{X\nu}}(s)\Gamma^Y_{O_{X\nu+1}}(s)\Gamma^{Cro_2}_{O_{X123}}(s) \\ + &\sum_{X=R,L}\Delta G^{Cro_2}_{O_{X123}}\Gamma^{Cro_2}_{O_{X^1}}(s)\Gamma^{Cro_2}_{O_{X^2}}(s)\Gamma^{Cro_2}_{O_{X^3}}(s) \quad + \\ &\sum_{X=RM,R,L}\Delta G^{RNAP}_{P_X}\Gamma^{RNAP}_{P_X}(s) + \sum_{\nu=1}^{3}\Delta G_{RL}\Gamma^{CI_2}_{O_{R^\nu}}(s)\Gamma^{CI_2}_{O_{L^\nu}}(s)\end{aligned}$$

where

$$\Gamma^Y_X(k) = \begin{cases}1, & \text{if molecule } Y \text{ is bound to site } X;\\ 0, & \text{otherwise}\end{cases}$$

and

$$\Gamma^{Cro_2}_{O_{X123}}(s) = \begin{cases}0, & \text{if Cro}_2 \text{ is bound to } O_R1,\ O_R2, \text{ and } O_R3\\ 1, & \text{otherwise}\end{cases}$$

All ΔG^*_* values in Table 2 are computed from [24]. The detailed explanation of how these energies have been computed can be found in [20]. The first sum includes all binding energies of transcription factors to the six binding sites on both left and right operator. The second sum includes all cooperation energies between any two adjacent factors and the third takes into account cooperativity that results from having Cro bound to all three binding sites on either O_R or O_L. It should be noted that in the measurements by Darling *et. al.* [24], the cooperative binding energies when Cro is bound to all three subdomains of O_R or O_L are not equal to the sum of the cooperative binding energies $\Delta G^{\text{Cro}_2}_{\text{O}_\text{X}12}$ and $\Delta G^{\text{Cro}_2}_{\text{O}_\text{X}23}$ (see Table 2). The term $\Gamma^{Cro_2}_{O_{X123}}(s)$ in the second sum guarantees that when Cro occupies all three subdomains in O_R or O_L, the cooperative energies $\Delta G^{\text{Cro}_2}_{\text{O}_\text{X}12}$ and $\Delta G^{\text{Cro}_2}_{\text{O}_\text{X}23}$ are not included in this sum. The energies $\Delta G^{\text{Cro}_2}_{\text{O}_\text{X}123}$ are then added in the third sum. The fourth sum adds the RNAP binding energy for the state, and the last one contributes any cross cooperation between CI_2 molecules bound to P_R and P_L.

The Differential Equation Model. In the differential equation model (1)-(4) the concentrations on the left hand side denote total monomer concentration, while on the right had side we have dimer concentrations $[CI_2]$ and $[Cro_2]$. To accurately represent this, the equations (8) and (9) represent this dimerization. As demonstrated in [20], these equations arose from the following chemical reaction:

$$2a_1 \underset{k_-}{\overset{k_+}{\longleftrightarrow}} a_2$$

where a_1 is a free monomer form of the protein and a_2 represents a dimer of protein a, k_+ and k_- are the forward and backward rate constants respectively.

In chemical equilibrium with, $K_D = k_-/k_+$, we have the following relation:

$$[a_1]^2 = K_D[a_2]. \qquad \textbf{[10]}$$

K_D is the dissociation constant and $[\cdot]$ represents concentration. In addition, if $[a]$ is the total monomer concentration,

$$[a] = [a_1] + 2[a_2]. \qquad \textbf{[11]}$$

The equations (10)-(11) can be used to solve for $[a_2]$ leading to

$$[a_2] = \frac{[a]}{2} - \frac{K_D}{8}\left(\sqrt{1+8\frac{[a]}{K_D}} - 1\right)$$

from which (8)-(9) follow.

We would like to thank R. Ebright for valuable discussions about prokaryotic transcription initiation. T. G. was partially supported by NSF/NIH grant 0443863, NIH-NCRR grant PR16445 and NSF-CRCNS grant 0515290. K.M. was partially supported by NSF DMS 0443827 and grants from D.O.E. and DARPA. K.P. was supported by NSF/NIH grant 0443863. E.T. was partially supported by NSF DMS 0443827 and CAPES, Brazil.

1. C. Lartigue, J.I. Glass, N. Alperovich, R. Pieper, P.P. Parmar, C.A. Hutchison 3rd, H.O. Smith, and J.C. Venter (2007) *Science* **317**, 632-638.
2. M.B. Elowitz and S. Leibler (2000) *Nature* **403**, 335-338.
3. T.S. Gardner, C.R. Cantor, and J.J. Collins (2000) *Nature* **403**, 339-342.
4. M. Ptashne (2004) in *A Genetic Switch:Phage Lambda Revisited* (Cold Spring Harbor Laboratory Press).
5. S. Roy, S. Garges, and S. Adhya (1988) *J Biol Chem* **273**, 14059-14062.
6. A.N. Kapanidis, E. Margeat, S. Ho, E. Kortkhonjia, S. Weiss, and R.H. Ebright (2006) *Science* **314**, 1144-1147.
7. G.K. Ackers, A.D. Johnson, and M.A. Shea (1982) *Proc Natl Acad Sci USA* **79**, 1129-1133.
8. L. Bintu, N.E. Buchler, H.G. Garcia, U. Gerland, T. Hwa, J. Kondev, and R. Phillips (2005) *Curr Opin Genet Dev* **15**, 116-124.
9. M. Santillán and M.C. Mackey (2001) *Proc Natl Acad Sci USA* **98**, 1364-1369.
10. M. Santillán and M.C. Mackey (2004) *Biophys J* **86**, 1282-1292.
11. T. Gedeon, K. Mischaikow, K. Patterson, and E. Traldi (2007) *preprint*.
12. I.B. Dodd, K.E. Shearwin, and J.B. Egan (2005) *Curr Opin Genet Dev* **15**, 145-152.
13. J.W. Roberts, C.W. Roberts, and D.W. Mount (1977) *Proc Natl Acad Sci USA* **74**, 2283-2287.
14. A. Bailone, A. Levine, and R. Devoret (1979) *J Mol Biol* **131**, 553-572.
15. R.T. Sauer, M.J. Ross, and M. Ptashne (1982) *J Biol Chem* **257**, 4458-4462.
16. D.K. Hawley and W.R. McClure (1982) *J Mol Biol* **157**, 493-525.
17. M. Li, W.R. McClure, and M.M. Susskind (1997) *Proc Natl Acad Sci USA* **94**, 3691-3696.
18. M.T. Marr, J.W. Roberts, S.E. Brown, M. Klee, and G.N. Gussin (2004) *Nucleic Acids Res* **32**, 1083-1090.
19. M. Ptashne and A. Gann (2002) in *Genes & Signals* (Cold Spring Harbor Laboratory Press).
20. M. Santillán and M.C. Mackey (2004) *Biophys J* **86**, 75-84.
21. J.W. Little, D.P. Shepley, and D.W. Wert (1999) *EMBO J* **18**, 4299-4307.
22. C.B. Michalowski and J.W. Little (2005) *J Bacteriol* **187**, 6430-6442.
23. C.B. Michalowski, M.D. Short, and J.W. Little (2004) *J Bacteriol* **186**, 7988-7999.
24. P.J. Darling, J.M. Holt, and G.K. Ackers (2000) *J Mol Biol* **302**, 625-638.
25. L. Guarente, J.S. Nye, A. Hochschild, and M. Ptashne (1982) *Proc Natl Acad Sci USA* **79**, 2236-2239.

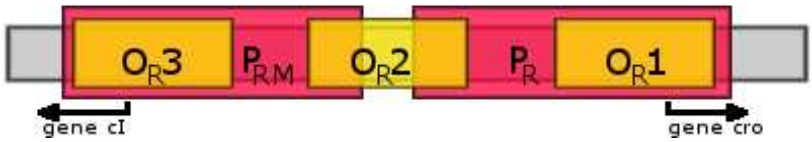
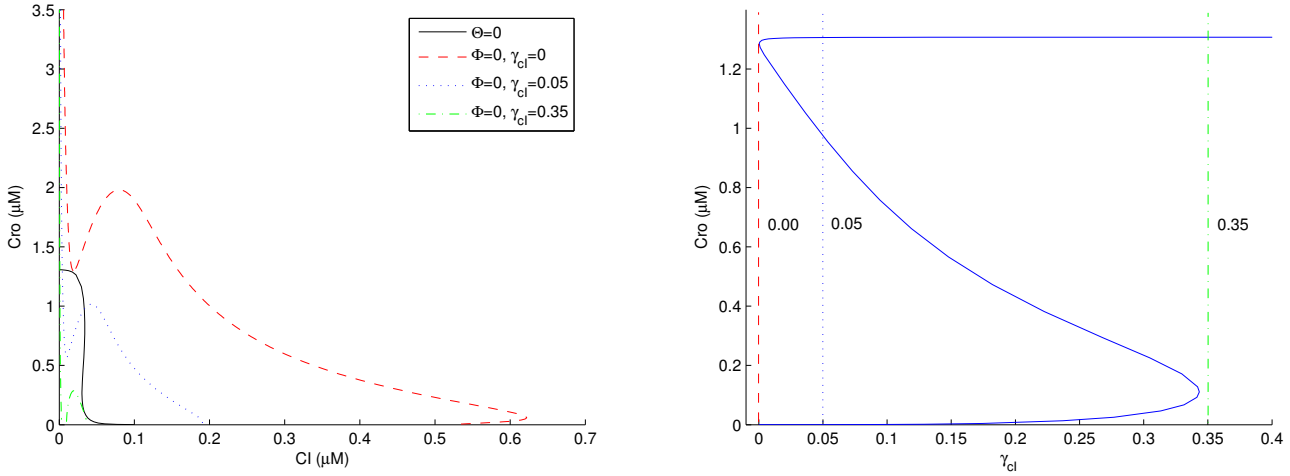


Fig. 1. O_R region



(a) (b)

Fig. 2. A. Nullclines for $\Theta = 0$ (black solid) and $\Phi = 0$ with $\gamma_{cI} = 0\,\mathrm{min}^{-1}$ (red dash), $\gamma_{cI} = 0.05\,\mathrm{min}^{-1}$ (blue dots) and $\gamma_{cI} = 0.35\,\mathrm{min}^{-1}$ (green dash-dot). B. Bifurcation diagram of γ_{cI} versus [Cro].

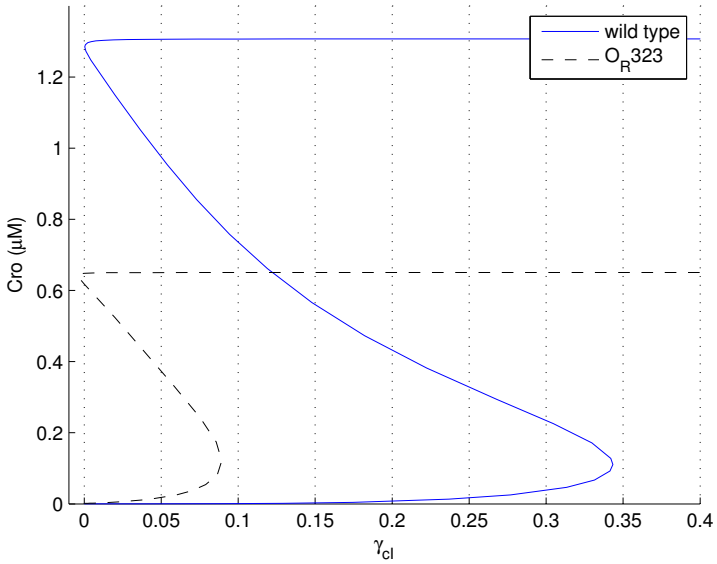


Fig. 3. Bifurcation diagrams for wild type and O_R323. The concentration of Cro is graphed as a function of γ_{cI}. The blue solid curve represents the wild type phage, while the black dashed curve represents the mutant.

Table 1. Estimated parameter values from [20] (with the addition of ϕ) for equations (1)-(4).

$\mu \simeq 2.0 \times 10^{-2}$ min^{-1}	$k_{cro} \simeq 2.76$ min^{-1}
$k_{cI}^c \simeq 4.29$ min^{-1}	$k_{cI} \simeq 0.35$ min^{-1}
$\gamma_M \simeq 0.12$ min^{-1}	$\gamma_{cI} \simeq 0.0$ min^{-1}
$\gamma_{cro} \simeq 1.6 \times 10^{-2}$ min^{-1}	$\nu_{cI} \simeq 0.09$ min^{-1}
$\nu_{cro} \simeq 3.2$ min^{-1}	$\tau_{cI} \simeq 0.24$ min
$\tau_{cro} \simeq 6.6 \times 10^{-2}$ min	$\tau_M \simeq 5.1 \times 10^{-3}$ min
$K_D^{cI} \simeq 5.56 \times 10^{-3} \mu$M	$K_D^{cro} \simeq 3.26 \times 10^{-1} \mu$M
$[O_R] \simeq 5.0 \times 10^{-3} \mu$M	[RNAP] $\simeq 3.0$ μM
$\Delta G_{RL} \simeq -3.1$ kcal/mol	$\phi \simeq 4.29/.35 = 12.26$

Table 2. Estimated binding energies from [20].

$\Delta G^{\mathrm{CI}_2}_{\mathrm{O_R}1} \simeq -12.5$ kcal/mol	$\Delta G^{\mathrm{CI}_2}_{\mathrm{O_L}1} \simeq -11.5$ kcal/mol
$\Delta G^{\mathrm{CI}_2}_{\mathrm{O_R}2} \simeq -10.5$ kcal/mol	$\Delta G^{\mathrm{CI}_2}_{\mathrm{O_L}2} \simeq -9.7$ kcal/mol
$\Delta G^{\mathrm{CI}_2}_{\mathrm{O_R}3} \simeq -9.5$ kcal/mol	$\Delta G^{\mathrm{CI}_2}_{\mathrm{O_L}3} \simeq -9.7$ kcal/mol
$\Delta G^{\mathrm{CI}_2}_{\mathrm{O_R}12} \simeq -2.7$ kcal/mol	$\Delta G^{\mathrm{CI}_2}_{\mathrm{O_L}12} \simeq -2.7$ kcal/mol
$\Delta G^{\mathrm{CI}_2}_{\mathrm{O_R}23} \simeq -2.9$ kcal/mol	$\Delta G^{\mathrm{CI}_2}_{\mathrm{O_L}23} \simeq -2.9$ kcal/mol
$\Delta G^{\mathrm{Cro}_2}_{\mathrm{O_R}1} \simeq -12.0$ kcal/mol	$\Delta G^{\mathrm{Cro}_2}_{\mathrm{O_L}1} \simeq -12.0$ kcal/mol
$\Delta G^{\mathrm{Cro}_2}_{\mathrm{O_R}2} \simeq -10.8$ kcal/mol	$\Delta G^{\mathrm{Cro}_2}_{\mathrm{O_L}2} \simeq -10.8$ kcal/mol
$\Delta G^{\mathrm{Cro}_2}_{\mathrm{O_R}3} \simeq -13.4$ kcal/mol	$\Delta G^{\mathrm{Cro}_2}_{\mathrm{O_L}3} \simeq -13.4$ kcal/mol
$\Delta G^{\mathrm{Cro}_2}_{\mathrm{O_R}12} \simeq -1.0$ kcal/mol	$\Delta G^{\mathrm{Cro}_2}_{\mathrm{O_L}12} \simeq -1.0$ kcal/mol
$\Delta G^{\mathrm{Cro}_2}_{\mathrm{O_R}23} \simeq -0.6$ kcal/mol	$\Delta G^{\mathrm{Cro}_2}_{\mathrm{O_L}23} \simeq -0.6$ kcal/mol
$\Delta G^{\mathrm{Cro}_2}_{\mathrm{O_R}123} \simeq -0.9$ kcal/mol	$\Delta G^{\mathrm{Cro}_2}_{\mathrm{O_L}123} \simeq -0.9$ kcal/mol
$\Delta G^{\mathrm{RNAP}}_{\mathrm{P_R}} \simeq -12.5$ kcal/mol	$\Delta G^{\mathrm{RNAP}}_{\mathrm{P_L}} \simeq -11.3$ kcal/mol
$\Delta G^{\mathrm{RNAP}}_{\mathrm{P_{RM}}} \simeq -11.5$ kcal/mol	

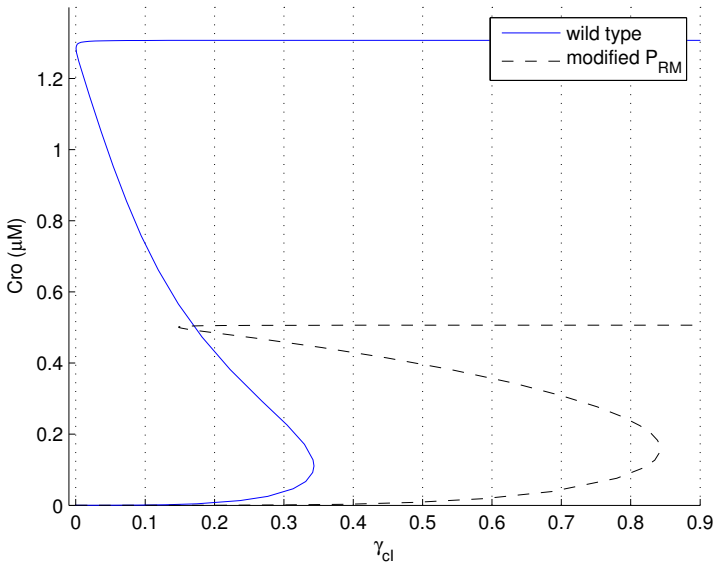


Fig. 4. Bifurcation diagram of wild type vs. a phage with mutated P_{RM} binding site which resulted in having $P_{RM} = -12.5$ kcal/mol, $\phi = 4.5/.35$, and $P_R = -10.5$ kcal/mol. For comparison, the wild type values were $P_{RM} = -11.5$kcal/mol, $\phi = 4.29/.35$, and $P_R = -12.5$kcal/mol.

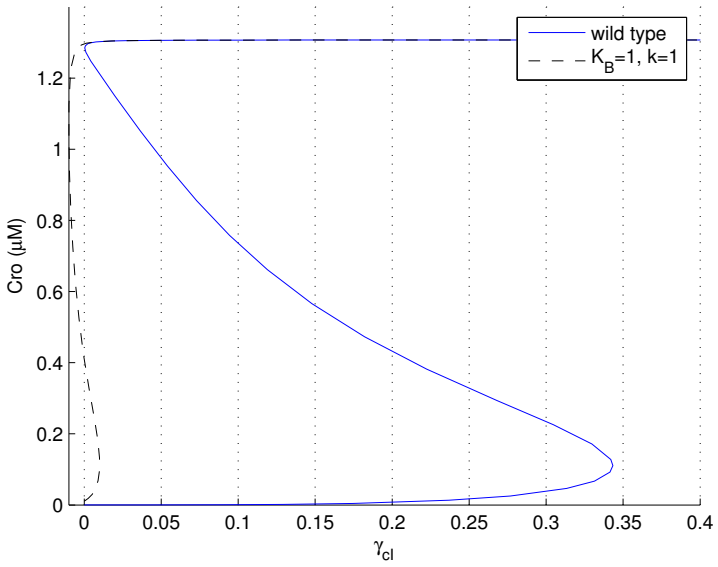


Fig. 5. Bifurcation diagram of wild type vs. cI-pc mutant where all positive control between CI_2 and RNAP has been eliminated ($\beta = 1, \phi = 1$).

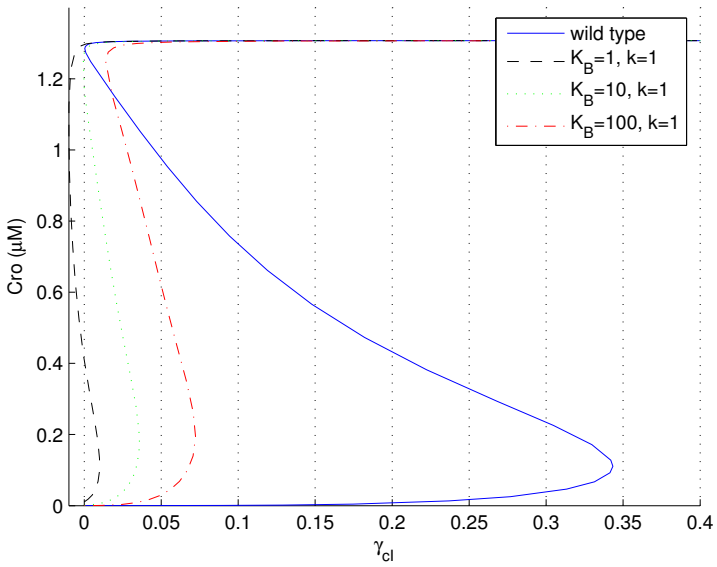


Fig. 6. Results from eliminating positive control ($\phi = 1$) with values of $\beta = 1$ (black dashed curve), $\beta = 10$ (green dotted curve) and $\beta = 100$ (red dash-dot curve). We graph concentration of Cro as a function of γ_{cI}.

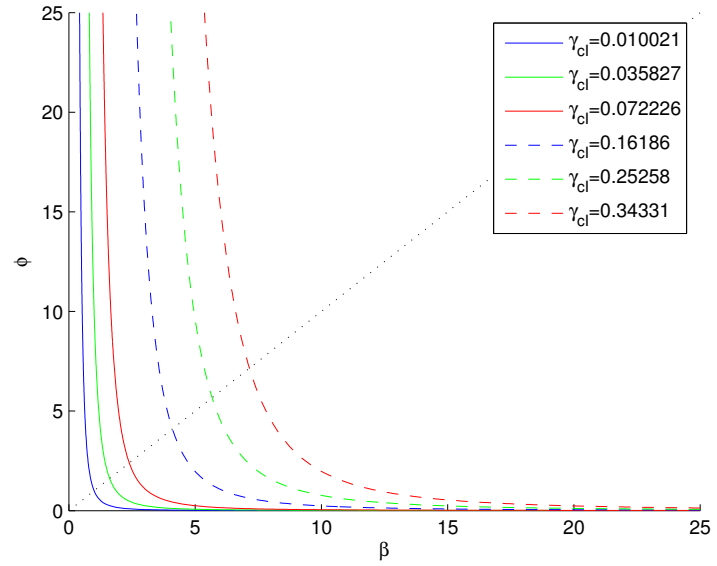


Fig. 7. Level curves of the induction value γ_* as a function of both β and ϕ. Here $\beta > 1$ represents K_B-cooperativity and $\phi > 1$ represents k-cooperativity.