Design Principles of Lambda's Lysis/lysogeny Decision vis-a-vis Multiplicity of Infection

Dinkar Wadhwa*

Department of Chemical Engineering

Indian Institute of Technology Bombay, Mumbai - 400 076, India

dinkar.wadhwa84@gmail.com

*The author is not currently affiliated with the institution.

Abstract

Bacteriophage lambda makes lysis/lysogeny decision based on the number of coinfecting phages, namely the multiplicity of infection (MoI): lysis at low MoI; lysogeny at high MoI. Here, by evaluating various rationally designed models on their ability a) to make lytic decision at MoI of 1 and lysogeny decision at MoI of 2, b) to exhibit bistability at the two MoIs, and c) to perform accurately in the presence of noise, it is demonstrated that lambda's lysis/lysogeny decision is based on three features, namely a) mutual repression, b) cooperative positive autoregulation of *cI*, and c) cooperative binding of the activator protein, and not basal expression, triggering the positive autoregulatory loop. Cro and CI are sufficient to acquire the first two features. CII is required to acquire the third feature. The quasi-minimal two-protein model for the switch is justified by showing its qualitative equivalence, except for the Cro repression of *pRM*, to lambda's gene regulatory network responsible for lysis/lysogeny decision. A three-protein simplified version of lambda's switch is shown to possess all the three design features. Existence of bistability at MoI of 1 and 2 confer lysogen stability and lytic development's stability, respectively.

Keywords: Bacteriophage λ, switch, mutual repression, positive feedback, bistability

1. Introduction

While virulent bacteriophages possess only one method of reproduction, namely lytic strategy, temperate bacteriophages have a dual perpetuation strategy, namely lytic and lysogeny. In

lytic strategy, the phage injects its genetic material into the bacterial host, the viral genes are transcribed, mRNAs thus produced are translated, and the phage's genetic material is replicated. Finally, the bacterial cell is lysed, releasing the progeny particles. In lysogeny, lytic pathway is repressed, the viral genome is integrated into the bacterial chromosome, and thereby the virus exists in a latent state known as prophage. As the teleological explanation goes, lytic strategy leads to fast multiplication but is risky because the viral progenies have to find new hosts which don't already contain lysogenized phages. On the other hand, a lysogenized phage replicates along with its host and, thus, reproduces slowly as compared to its growth by lytic strategy, but this way the phage safeguards its survival. Should a phage infect a bacterium containing lysogenized phage, lambda repressors (CI) present in the cytosol will not allow the expression from its promoter *pR*.

Being one of the simplest organisms to make a genetic decision, that is, between lysis and lysogeny, bacteriophage lambda is a paradigm for gene regulation and developmental decision. Extensive study of the phage's core genetic network involved in the lysis/lysogeny decision has yielded a complex picture of interconnected feedback loops [1]. A classic [2] and a molecular study [3] have shown that the lysis/lysogeny decision depends on multiplicity of infection (MoI). Avlund et al. [4] analysed Kourilsky's data [2,5] and determined the probability of lysogeny at MoI of 1 to be almost zero, at MoI of 2 to be around 0.6960, and at all higher MoIs to be around 0.9886. In abstract terms, this ability of the phage to choose between lysis and lysogeny based on the number of coinfecting phages is but a form of quorum sensing occurring inside a bacterium.

One of the first theoretical studies of lambda's decision with respect to multiplicity of infection considered a model which possessed two proteins, namely Cro and CI, which are produced from their respective promoters, namely *pR* and *pRM*, and a simplified model of the phage's gene regulatory network (GRN) which possessed three critical proteins, namely Cro, CI, and CII. Cro and CI had the same promoters as in the two-protein model. CII is under the control of *pR* and activates the third promoter *pRE*, resulting in the production of CI [6]. For the two-protein model, it was shown that if basal expression of *cro-like* gene exceeds that of

cI-like gene, one stable state exists at the lower and higher MoIs, dominated by Cro-like protein and CI-like protein, respectively, whereas bistability is present for the intermediate MoIs. Additionally, it was demonstrated that positive autoregulation of *lys* and protein dimerization are necessary for generating the switching behaviour. For the three-protein model, the work showed that the transient level of CII increases with MoI, and past a certain threshold, the nonlinearity of positive autoregulation of *cI* gives rise to the switching behaviour in terms of equilibrium values of the three proteins. Although the study demonstrated that a two-protein and three-protein reduced models of lambda's lysis/lysogeny switch are able to generate the switching behaviour deterministically, it did not determine how the switching behaviour is rooted in the features of the models. The study, for example, did not address the question as to what advantage can the additional protein, namely CII, confer to the two-protein model, which also is able to produce the switching behaviour.

Avlund et al. constructed all possible motifs comprised of two proteins, namely Cro-like protein (Lyt) and CI-like protein (Lys) [7]. The networks were selected on their ability to stably commit to lysis and lysogeny at MoI of 1 and MoI of 2, respectively, by solving their corresponding dynamical equations deterministically (the “deterministic counting task”). 96% of the networks successful in the deterministic counting task possessed mutual repression motif. In the subsequent work, Avlund et al. tested the two-protein networks successful in the deterministic counting task for their ability to carry out the same task in the presence of noise (the “stochastic counting task”) [8]. It was found out that almost all of the two-protein networks did not pass the stochastic counting task. However, additional CII-like protein significantly improved robustness to noise. It was argued that the reason for the improved robustness was CII relieving Lys of its role of executing commitment to lysogeny, thereby allowing Lys to be independently optimized for maintenance of lysogeny.

The mean lifetime of CII in the networks passing the deterministic counting task was 0.8 phage replication time (τ_{rep}). The stochastic counting task selects for CII with shorter half-life: the mean lifetime of CII in the networks carrying out the stochastic counting task reduced to 0.4 τ_{rep}. The mean lifetime of Lys in the two-protein networks successful in the deterministic

counting task was 1.2 τ_{rep}, and this increased to 4.6 τ_{rep} in the presence of CII. Thus, the authors argued that Lys, because of having longer mean half-life, reaches its steady state slowly and, as a result, interferes less with CII, which has much shorter half-life, during the commitment phase.

In this study, conceptualizing the role of Cro-like protein from first principle, a two-protein quasi-minimal model for lambda's GRN is obtained; and additionally, other two-protein models based on certain departures from the quasi-minimal model are constructed. Together with the constraint of possessing bistability at both the MoIs and relating the accuracy with which the switching behaviour was produced in stochastic simulations by the two-protein models and a three-protein simplified model of lambda's GRN with their respective features reveals design principles of the lysis/lysogeny switch.

2. Result and Discussion

2.1. Quasi-minimal two-protein lysis/lysogeny switch

In the quasi-minimal two-protein model, promoter of *lyt* gene is constitutive and that of *lys* gene is positively autoregulated as they are in lambda's GRN. The roles of CI-like protein (Lys) in the quasi-minimal two-protein model; that is, binding cooperatively to the intergenic region, activating transcription of its own gene and inhibiting transcription of *lyt* gene, are identical to the roles of CI in lambda's GRN; that is, binding cooperatively to the intergenic region, activating promoter *pRM* and repressing promoter *pR*. Additionally, Lys activates the imaginary downstream pathway required for lysogeny.

The role of Cro-like protein (Lyt) was conceptualized from first principle in the following way. At MoI of 1, constitutive expression of *lyt* leads to a high equilibrium value of Lyt. On the other hand, the level of Lys remains low because *lys* is a very weak promoter [9,10]. Although the analysis of Avlund et al. [4] showed that lysogeny occurs in around 70% of the cases at MoI of 2, here, for the sake of simplicity, lysogeny is expected to always occur at MoI of 2. Thus, at MoI of 2, the situation should be reversed: the equilibrium level of Lyt should be much

lower than that at MoI of 1, and the equilibrium level of Lys should be much higher than that at MoI of 1. However, that is not likely to happen because *lys*, being a weak promoter, cannot produce Lys to a level sufficient to trigger the autoregulatory loop of *lys*. Because the only protein present to actuate any process at MoI of 2 is Lyt, it is argued here from first principle that Lyt should activate promoter *lys* and suppress its own promoter.

At MoI of 1, Lyt is produced constitutively but its level is not high enough to trigger the positive feedback loop, causing the level of Lys to remain low. On the other hand, at MoI of 2, Lyt is initially produced at a higher rate, triggering the positive feedback loop, resulting in the production of Lys. Lys, thus present at a high level, represses *lyt*, causing Lyt's level to come down to a low value. Thus, Lyt 1) activates transcription of *lys*, whose product, Lys, engenders lysogeny and suppresses lytic pathway, 2) represses transcription of its own gene, thereby suppressing lytic pathway (though, as argued below, this interaction is dispensable when the kinetic details are ignored), and 3) activates the imaginary downstream pathway which leads to lytic development. This seemingly paradoxical role of Lyt; that is, causing lysogeny as well as lysis, is, as explained below, due to it being proxy for CII, which causes lysogeny, and anti-termination factor Q, which enables transcription of lytic genes.

Because the expression for quality of lysis/lysogeny switch (the "switching quotient") takes equilibrium values into account, the degradation rate of Lyt (X) can be subsumed into k_1, and the degradation rate of Lys (Y) can be subsumed into k_3 and, if present, k_4. Hence, k_2 and k_5 are taken to be unity for all of the two-protein models. This model would henceforth be referred to as 1A_Lyt_Lys.

1A_Lyt_Lys:

$$\frac{dX}{dt} = \frac{mk_1}{1 + \frac{X^a}{K_{D1}} + \frac{Y^b}{K_{D2}}} - k_2 X \tag{1}$$

$$\frac{dY}{dt} = \frac{m(k_3\frac{X^a}{K_{D1}} + k_4\frac{Y^b}{K_{D2}})}{1 + \frac{X^a}{K_{D1}} + \frac{Y^b}{K_{D2}}} - k_5 Y \tag{2}$$

where, m is multiplicity of infection; k_1 is basal expression rate of *lyt*; k_3 and k_4 are rate constants for transcriptional activation of *lys* by Lyt and Lys, respectively. K_{D1} and K_{D2} are the effective equilibrium dissociation constants of Lyt and Lys, respectively (see Methods). In those models in which *lys* has basal expression, k_3 represents the basal expression rate. Exponents a and b are Hill coefficients for binding of Lyt and Lys, respectively.

2.2. Equivalence of the quasi-minimal two-protein model (1A_Lyt_Lys) with lambda's GRN

Upon the infection, RNA polymerase transcribes from the early promoters *pL* and *pR* until it encounters the transcription terminators *tL1* and *tR1*, respectively [11]. This leads to transcription of *N* gene, present between promoter *pR* and the terminator *tL1*.The product of the *N* gene is an anti-termination factor that modifies subsequent RNAPs initiating at *pL* and *pR* so that they are able to transcribe genes that are beyond their terminators, such as *cIII* and *cII* genes, respectively. [12,13]. Such a modified RNAP from promoter *pR* is also able to transcribe through another terminator, namely *tR2*, present upstream of gene *Q* (see figure 2). Up to this point, the lytic and lysogenic pathway are identical. Commitment to lytic occurs in the following way: extended transcription from *pR* results in transcription of gene *Q*. Q, being an anti-termination factor, causes transcription from the strong constitutive promoter *pR′* to not terminate, as it would otherwise do, at *tR*', which is present at about 200 bases away from *pR′*. This results in transcription of the rest of the lytic genes downstream of *Q* [14], causing the cell to commit to lysis. CIII protein has an indirect role in causing lysogeny, that is, by preventing degradation of CII by the bacterial protease HflB [15,16]. The (indirect) role of cIII will not be taken into consideration in the present work, as it investigates the decision-making process of

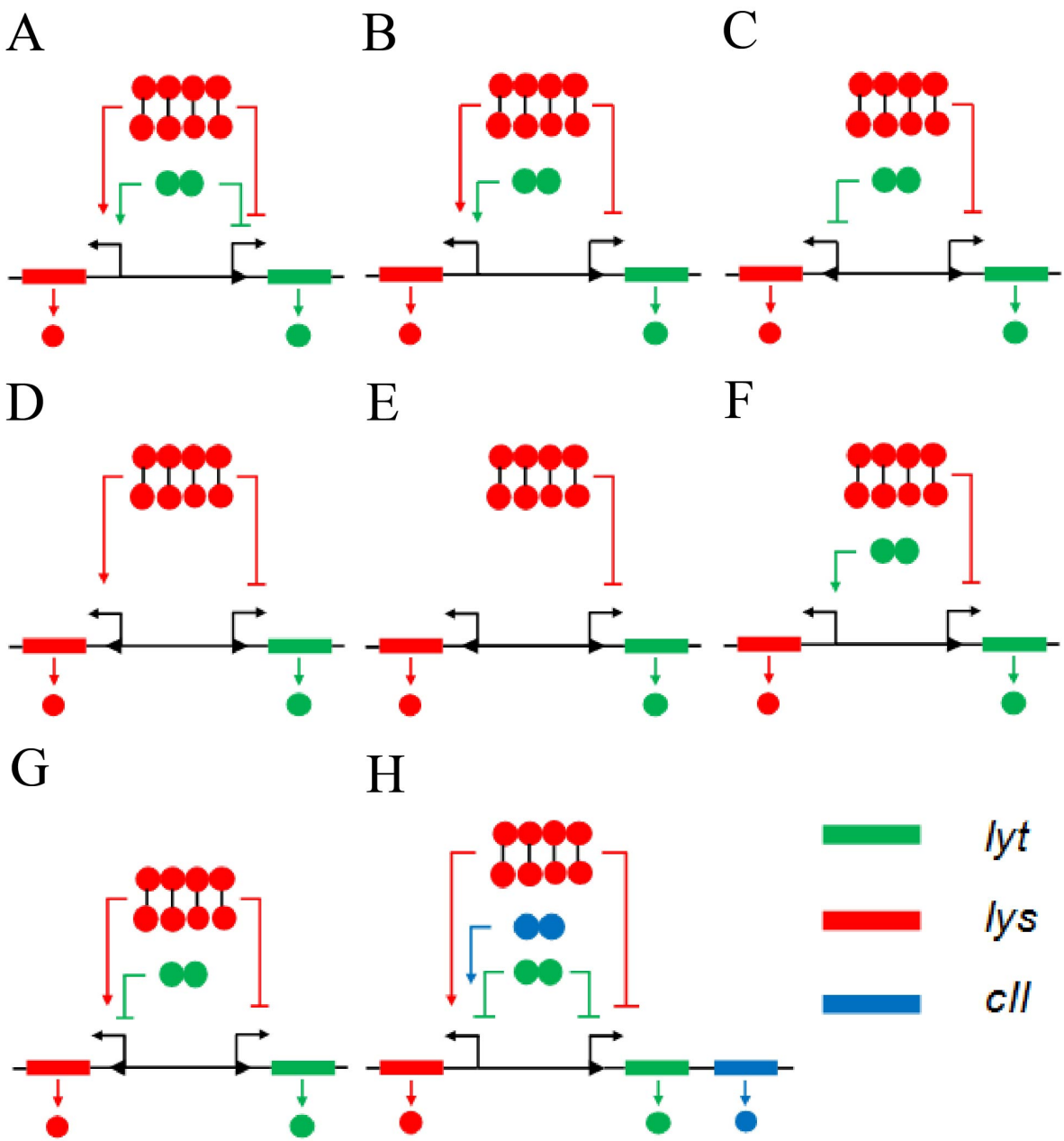

Figure 1: Various two-protein models, and the three-protein model. Lower arrowhead represents basal expression. (**A**) The quasi-minimal model or 1A_Lyt_Lys. (**B**) 1A_Lyt_Lys with self-repression of *lyt* removed, or 1B_Lyt_Lys. (**C**) Mutual repression or 2_Lyt_Lys. (**D**) 3_Lyt_Lys. (**E**) 4_Lyt_Lys. (**F**) 5_Lyt_Lys. (**G**) 6_Lyt_Lys. (**H**) A three-protein simplified version of lambda's switch or Lyt_Lys_CII.

lambda through minimalist models.

Commitment to lysogeny occurs when CII is produced to a concentration which is sufficient to activate transcription from three promoters, namely *pI*, *pRE*, and *pAQ*. Promoter *pI* controls the transcription of *int* gene, which codes for Int protein, required for the integration of the phage genome into that of the host bacterium [17]. The transcripts produced from *pRE* contain orf for *cI*, therefore activation of this promoter leads to the production of CI [18]. CII inhibits commitment to lytic pathway by activating transcription from *pAQ*, which is located within *Q* gene in the opposite polarity [19,20]. Being antisense to (a part of) *Q* mRNA, the transcripts produced from *pAQ* hybridize with them. This prevents translation of *Q* mRNAs, whose translated product is essential for commitment to lytic [3].

In lambda's GRN, *cro*, *cII*, and *Q* are under the control of promoter *pR*. Because in 1A_Lyt_Lys *lyt* is transcribed from *pR*, Lyt protein should be functionally equivalent to Cro, CII, and Q. Lyt's negative autoregulation is equivalent to Cro repression of *pR* and CII repression of Q's production, that is, by activating transcription from *pAQ*. Lyt activation of *lys* is equivalent to CII activation of *pRE*, which results in synthesis of CI. CII activation of *pI*, whose product, Int, is required for lysogeny, is modeled by Lyt engendering the production of Lys, which then activates imaginary downstream pathway which leads to lysogeny. In other words, promoter *lys* is equivalent to promoters *pRE* and *pI*, and Lys is equivalent to CI and Int. If CII is not produced in sufficient amount, *Q* mRNA is translated and anti-terminator Q thus produced causes lysis. The quasi-minimal model incorporates this process in this way: if Lyt's concentration doesn't become high enough to be able to trigger the positive feedback loop, which leads to production of Lys and consequent inhibition of *lyt*, Lyt is produced to an equilibrium concentration sufficient to result in commitment to lytic pathway.

Notably, while Cro in lambda's GRN inhibits transcription from promoter *pRM*, Cro-like protein (Lyt) activates the expression of *lys* gene in 1A_Lyt_Lys. It will be shown later that inhibition of *lys* (or *pRM*) by Lyt (or Cro) is required for the stability of lytic development; that is, to prevent lysogeny which would otherwise occur as a result of increase in MoI due to genome replication soon after commitment to lysis. Hence, 1A_Lyt_Lys is a quasi-minimal,

and not minimal, model. The promoter *pRM* possesses weak basal expression, whereas *lys* in 1A_Lyt_Lys lacks basal expression and is activated by Lyt. This difference can be neglected because transcription from *pRE*, when activated by CII, is much higher than the basal expression of *pRM*. Schmeissner et al. showed that transcription from *pRE* promoter, when activated by CII, was roughly 3 times higher than the basal expression of *pR* [18], which, as mentioned earlier, is a much stronger promoter than *pRM*.

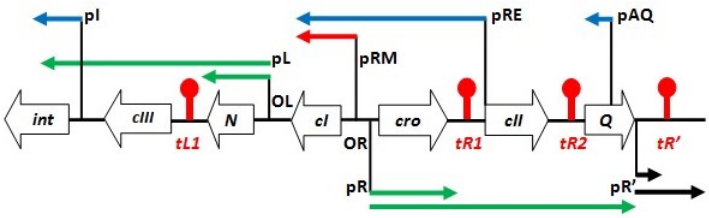

Figure 2: Gene and transcription map of lambda. The transcripts that are produced earliest, that is, from *pL* and *pR* promoters, are depicted as green arrows. The black arrow is the late transcript, that is, from *pR'*. The transcripts from CII-activated promoters, that is, *pI*, *pRE*, and *pAQ*, are shown as blue arrows. The transcript from *pRM*, which is activated by CI, is shown as a red arrow. Transcription terminators, namely *tL1*, *tR1*, and *tR2*, are depicted in red.

2.3. Variants of 1A_Lyt_Lys and mutual repression model

In order to further investigate design principles of the lysis/lysogeny switch, variants of 1A_Lyt_Lys and mutual repression model, and a model having features of 1A_Lyt_Lys and mutual repression model were considered. Because two features, namely constitutive expression of *lyt* and its inhibition by Lys, are present in all of them, they will not be mentioned in the description of the models below. A model possessed either Lyt activation of *lys*, a feature equivalent to positive regulation of *cI*, or basal expression of *lys*, a feature of mutual repression motif. All of these models can be categorized in terms of three factors, as shown in Table 1. The first column shows whether *lys* possesses basal expression or is activated by Lyt. The second column shows if positive feedback, formed by transcriptional activation of *lys* by Lys, is present. The third column shows if inhibition of *lys* by Lyt, ultimately forming mutual repression motif, is present. The inhibition of *lys* by Lyt can be present only when *lys* possesses basal expression. Thus, for

lys having basal expression, there are four models; and where it gets activated by Lyt, there are two models.

Table 1: Classification of additional two-protein models.

Model	Basal expression of *lys*/ Activation of *lys* by Lyt	Activation of *lys* by Lys	Inhibition of *lys* by Lyt
1B_Lyt_Lys	Activation	Yes	N/A
5_Lyt_Lys	Activation	No	N/A
3_Lyt_Lys	Basal	Yes	No
6_Lyt_Lys	Basal	Yes	Yes
4_Lyt_Lys	Basal	No	No
2_Lyt_Lys	Basal	No	Yes

1B_Lyt_Lys: This model differs from 1A_Lyt_Lys only in not having negative autoregulation of *lyt*, which in lambda represents repression of *pR* by Cro and inhibition of Q's production by CII. Lyt's negative autoregulation will be dispensable if these two processes can be shown to be redundant. Given that the present study takes into account equilibrium concentrations of proteins, arguably, kinetic details of Cro repression of *pR* may be ignored. Because Cro repression of *pR* reduces equilibrium concentrations of proteins whose genes are under the control of *pR*, the repression may be done away with at MoI of 1 while ensuring that equilibrium concentrations of proteins whose genes are under the control of *pR* remain unaffected. That can be achieved by readjusting the kinetic parameters associated with transcription from *pR*, translation and degradation of the resulting mRNAs, and degradation of proteins produced from those mRNAs. In the same way, Cro repression of *pR* at MoI of 2 is dispensable because CI performs the same function and much more strongly than Cro because the former is present in much higher concentration and the latter, being under the control of *pR* itself, cannot repress the promoter fully. As explained below, Cro repression of *pR* (and also *pL*) contributes to bistability at MoI of 2, but even this function of Cro can be subsumed into Cro repression of *pRM*.

It can be argued that CII inhibition of Q's production can be subsumed into Lys inhibition

of *lyt*. CII inhibition of Q's production is required at MoI of 2, where lysogeny is constrained to occur in the present study. As previously, if the kinetic detail of CII inhibition of Q's production is ignored, the process can be subsumed into repression of *pR* (*lyt*) by CI (Lys), whose production is triggered by CII, in the following way. CII activates transcription from textitpRE, leading to production of CI, which represses *pR*. This results into repression of *Q*, as its transcription is under the control of *pR*. Thus, CII −| Q is replaced by CII → CI (Lys) −|pR → Q (Lyt). This assumption becomes stronger in the light of the experimental result that, as compared to CI, CII has very low impact in reducing the production of Q [3]. In a strain lacking *cI* gene (λcI⁻), it was demonstrated that the level of Q protein, reported by GFP expressed from pR'-tR'-*gfp* promoter-terminator fusion, shows almost no change as a function of MoI (see figure 3 C and D of ref. 3).

$$\frac{dX}{dt} = \frac{mk_1(1 + \frac{X^a}{K_{D1}})}{1 + \frac{X^a}{K_{D1}} + \frac{Y^b}{K_{D2}}} - k_2 X \tag{3}$$

$$\frac{dY}{dt} = \frac{m(k_3 \frac{X^a}{K_{D1}} + k_4 \frac{Y^b}{K_{D2}})}{1 + \frac{X^a}{K_{D1}} + \frac{Y^b}{K_{D2}}} - k_5 Y \tag{4}$$

2_Lyt_Lys (Mutual repression): Lyt represses *lys*, which has basal expression.

$$\frac{dX}{dt} = \frac{mk_1(1 + \frac{X^a}{K_{D1}})}{1 + \frac{X^a}{K_{D1}} + \frac{Y^b}{K_{D2}}} - k_2 X \tag{5}$$

$$\frac{dY}{dt} = \frac{mk_3(1 + \frac{Y^b}{K_{D2}})}{1 + \frac{X^a}{K_{D1}} + \frac{Y^b}{K_{D2}}} - k_5 Y \tag{6}$$

3_Lyt_Lys: *lys* has basal expression and is activated by Lys cooperatively.

$$\frac{dX}{dt} = \frac{mk_1}{1 + \frac{Y^b}{K_{D2}}} - k_2 X \tag{7}$$

$$\frac{dY}{dt} = \frac{m(k_3 + k_4 \frac{Y^b}{K_{D2}})}{1 + \frac{Y^b}{K_{D2}}} - k_5 Y \tag{8}$$

4_Lyt_Lys: *lys* has basal expression.

$$\frac{dX}{dt} = \frac{mk_1}{1 + \frac{Y^b}{K_{D2}}} - k_2 X \tag{9}$$

$$\frac{dY}{dt} = mk_3 - k_5 Y \tag{10}$$

5_Lyt_Lys: *lys* is activated by Lyt.

$$\frac{dX}{dt} = \frac{mk_1(1 + \frac{X^a}{K_{D1}})}{1 + \frac{X^a}{K_{D1}} + \frac{Y^b}{K_{D2}}} - k_2 X \tag{11}$$

$$\frac{dY}{dt} = \frac{mk_3 \frac{X^a}{K_{D1}}}{1 + \frac{X^a}{K_{D1}} + \frac{Y^b}{K_{D2}}} - k_5 Y \tag{12}$$

6_Lyt_Lys: *lys* has basal expression, is activated by Lys and inhibited by Lyt. Notably, adding self-repression of *lyt* to this model makes it identical to the two-protein model of Weitz et al. [6].

$$\frac{dX}{dt} = \frac{mk_1(1 + \frac{X^a}{K_{D1}})}{1 + \frac{X^a}{K_{D1}} + \frac{Y^b}{K_{D2}}} - k_2X \tag{13}$$

$$\frac{dY}{dt} = \frac{m(k_3 + k_4\frac{Y^b}{K_{D2}})}{1 + \frac{X^a}{K_{D1}} + \frac{Y^b}{K_{D2}}} - k_5Y \tag{14}$$

2.4. Deterministic simulation

The models were evaluated in terms of the quality of switch they generated, by solving their defining equations. For a given set of Hill coefficients (a and b), a set of rate constants and equilibrium dissociation constants would henceforth be referred to as parameter set (that is, Hill coefficients are not a part of parameter set here). Because Cro forms dimer, Hill coefficient for Lyt's binding was taken to be 2, whereas because CI forms tetramer, Hill coefficient for Lys' binding was taken to be 4. However, in the interest of completeness, another set of Hill coefficients, namely a=2, b=2, was also considered. The rate constants and equilibrium dissociation constants of the equations defining a given model were searched (see Methods) in two stages: order search and linear search (as they are called here). The parameter search was guided by the quality of the switch (named here: switch quotient) generated by them. The switch quotient was initially considered to be determined by the expression

$$\mathrm{SQ} = \frac{(\mathrm{S}_1 - \mathrm{S}_2)}{\mathrm{S}_1}$$

S_1 = min{Lyt at MoI of 1, Lys at MoI of 2}

S_2 = max{Lys at MoI of 1, Lyt at MoI of 2}

The expression, however, selected parameter sets which gave equilibrium values of Lyt at MoI of 1 unequal to those of Lys at MoI of 2. From the perspective of simplicity, I believe that the difference between the two values should be minimum. Therefore, the previous expression is multiplied by the ratio of S_1 and S_3 in order to penalize any difference between S_3 and S_1.

$$SQ = \frac{(S_1 - S_2)}{S_1} \times \frac{S_1}{S_3} = \frac{(S_1 - S_2)}{S_3}$$

$$S_3 = \max\{\text{Lyt at MoI of 1, Lys at MoI of 2}\}$$

This expression, like the previous one, varies between 0 and 1. Only those parameter sets were considered that gave positive switch quotients (SQ). A switch quotient generated in the deterministic simulations is called deterministic switch quotient (DSQ). The reason why the present study does not take the ratio of the concentrations of the two proteins as the measure of lysis/lysogeny decision (as, for example, has been done by Avlund et al. [7]) is following. Suppose in a simulation, the concentrations of Lyt and Lys at MoI of 1 turn out to be 10 arb. units and 1 arb. units, respectively; and at MoI of 2, they turn out to be 1 arb. units and 10 arb. units, respectively. The ratio of Lyt/Lys at MoI of 1 is 10 arb. units, whereas at MoI of 2 it is 0.1. arb. units. Further, suppose in another simulation, the concentrations of Lyt and Lys at MoI of 1 turn out to be, as previously, 10 arb. units and 1 arb. units, respectively; but at MoI of 2, they turn out to be 10 arb. units and 100 arb. units, respectively. The ratio of Lyt/Lys at MoI of 1 and MoI of 2 are still the same. However, biologically the two scenarios are not identical, because if Lyt at the equilibrium concentration of 10 arb. units can activate lytic pathway in the presence of 1 arb. unit of Lys, there is no reason why it will not do so when present at the same concentration even in the presence of 100 arb. units of Lys. This is because Lys represses the production of Lyt but has no impact on the imaginary downstream lysis-causing pathway activated by Lyt. In the lambda's GRN, it will translate to CI having no impact on lytic development once the protein Q is made to a sufficient level. Thus a ratio-based measure is not likely to be a realistic representative of lysis and lysogeny states.

Parameter sets, though not all of them, of only 2_Lyt_Lys and 6_Lyt_Lys exhibited bistability at both the MoIs. Therefore, for these two models, those parameter sets which did not exhibit bistability at both the MoIs were discarded for all purposes. For rest of the models, those parameter sets which did not exhibit bistability at MoI of 1 were discarded for all purposes (this way only one parameter set, that is, of 1B_Lyt_Lys, was discarded). As Table 2 shows, all

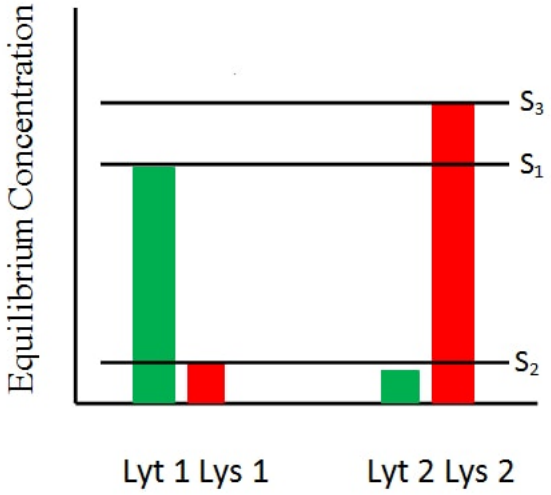

Figure 3: Schematic of the switching behaviour.

of the models possessing the positive feedback loop had average DSQ of more than 0.97 for both sets of Hill coefficients (lowest DSQ among all the models in this category was 0.9270). Mutual repression model for Hill coefficients' set of a=2, b=2 has only one parameter set, with DSQ of 0.6666; and, for that of a=2, b=4 average DSQ was 0.9746 (lowest DSQ was 0.9362). 4_Lyt_Lys for Hill coefficients' set of a=2, b=2 gives DSQs of 0.4794 and 0.4707; and, for that of a=2, b=4 both DSQs were almost 0.5. DSQs of 4_Lyt_Lys being around 0.5 can be explained by the fact that increase in genome copy number for this model leads to proportional increase in equilibrium activity of *lys* promoter, thereby causing the value of Lys at MoI of 1 to be half its value at MoI of 2.

The model 5_Lyt_Lys did not generate any parameter set. That's because, in the absence of positive autoregulation of *lys*, *lyt* needs to have strong basal expression in order to sustain high concentration of Lys, whose gene is activated by Lyt, at MoI of 2. That is, Lyt, too, is expressed in high concentration at MoI of 2. Equivalently, the desired high concentration of Lyt at MoI of 1 also leads to excessive production of Lys at the same MoI. Thus, the two proteins are present in similar amounts at both the MoIs.

In order to examine the significance of cooperativity in positive feedback here, another set of Hill coefficients, namely a=2, b=1, was considered for all models possessing positive

Table 2: Average and standard deviation of deterministic and stochastic switch quotients for a given range of stochastic success rate for time length of 100 arb. units

Model	AVG Deterministic SQ (SD)		AVG Deterministic SQ (SD) (SSR[a] ⩾ 95)		AVG Stochastic SQ (SD) (SSR ⩾ 95)	
	a=2, b=2	a=2, b=4	a=2, b=2	a=2, b=4	a=2, b=2	a=2, b=4
1A_Lyt_Lys	0.9917 (0.0106)	0.9896 (0.0049)	0.9898 (N/A)	0.9905 (0.0043)	0.7997 (N/A)	0.6204 (0.0624)
1A_Lyt(1)_Lys[b]	0.9950 (0.0053)	0.9923 (0.0045)	none	none	none	none
1B_Lyt_Lys	0.9806 (0.0236)	0.9769 (0.0277)	0.9971 (N/A)	0.9270 (N/A)	0.7995 (N/A)	0.7679 (N/A)
1B_Lyt(1)_Lys	0.9780 (0.0252)	0.9869 (0.0199)	none	0.9304 (N/A)	none	0.7520 (N/A)
2_Lyt_Lys	0.6666 (N/A)	0.9746 (0.0248)	none	none	none	none
3_Lyt_Lys	0.9950 (0.0073)	0.9893 (0.0150)	none	none	none	none
4_Lyt_Lys	0.4751 (0.0043)	0.4956 (0.0006)	none	0.4956 (0.0006)	none	0.2983 (0.0102)
6_Lyt_Lys	0.9984 (N/A)	0.9845 (0.0166)	none	none	none	none
Lyt_Lys_CII	0.9855 (0.0155)	N/A	0.9573 (N/A)	N/A	0.7526 (N/A)	N/A
Lyt_Lys_CII(1)	0.9801 (0.0151)	N/A	0.9718 (N/A)	N/A	0.7595 (N/A)	N/A
Lyt(1)_Lys_CII(1)	0.9894 (0.0134)	N/A	none	N/A	none	N/A
Lyt_Lys_CII′[c]	0.9840 (0.0126)	N/A	none	N/A	none	N/A

[a] SSR = Stochastic success rate
[b] Value in parenthesis in a model name is Hill coefficient of protein binding
[c] Parameter sets selected for bistability at MoI of 2.
N/A is used for cases with only one data point.

Table 3: Average and standard deviation of deterministic and stochastic switch quotients for a range of stochastic success rate for time length of 100 arb. units

Model	**AVG Deterministic SQ (SD) (95 >SSR[a] ⩾ 90)**		**AVG Stochastic SQ (SD) (95 >SSR ⩾ 90)**	
	a=2, b=2	a=2, b=4	a=2, b=2	a=2, b=4
1A_Lyt_Lys	0.9937 (0.0025)	0.9883 (0.0068)	0.8087 (0.0382)	0.5728 (0.1003)
1A_Lyt(1)_Lys[b]	0.9945 (N/A)	0.9930 (N/A)	0.8084 (N/A)	0.5825 (N/A)
1B_Lyt_Lys	0.9950 (N/A)	0.9972 (0.0020)	0.7891 (N/A)	0.6691 (0.0855)
1B_Lyt(1)_Lys	none	none	none	none
2_Lyt_Lys	none	none	none	none
3_Lyt_Lys	none	none	none	none
4_Lyt_Lys	none	none	none	none
6_Lyt_Lys	none	none	none	none
Lyt_Lys_CII	0.9808 (0.0008)	N/A	0.7614 (0.0211)	N/A
Lyt_Lys_CII(1)	0.9593 (N/A)	N/A	0.7551 (N/A)	N/A
Lyt(1)_Lys_CII(1)	0.9647 (N/A)	N/A	0.7178 (N/A)	N/A
Lyt_Lys_CII'[c]	0.9791 (0.0106)	N/A	0.6705 (0.0466)	N/A

[a] SSR = Stochastic Success Rate
[b] Value in the parenthesis is Hill coefficient of protein binding
[c] Parameter sets selected for bistability at MoI of 2.
N/A is used for cases with only one data point.

feedback loop. Highest DSQ for 1A_Lyt_Lys: 0.1262; 1B_Lyt_Lys: 0.1150; 3_Lyt_Lys: 0.0001; 6_Lyt_Lys: ≈ 0. Extremely low values of highest DSQs generated by this set of Hill coefficients underscores the absolute necessity of cooperative binding of Lys.

2.5. Closer to lambda's GRN: the minimal three-protein model

In order to further understand design principles of the switch, a three-protein simplified version of lambda's switch is considered. A CII-like protein is added to 1A_Lyt_Lys besides modifying the role of Lyt. Thus, the role of Lyt in this model is identical to that of Cro in lambda's GRN. That is, in addition to repressing transcription of its own gene and, as a result, the gene co-transcribed with *lyt*, namely *cII*, now Lyt represses transcription of *lys* too. As in the two-protein models, the function of Lys here is identical to that of CI in lambda's GRN. The role of CII in the three-protein model is to activate transcription of *lys*. This corresponds to CII activation of *pRE* promoter of lambda's GRN, leading to synthesis of mRNA which contains orf for *cI*. Like in 1A_Lyt_Lys, CII inhibition of Q's production is subsumed into Lys inhibition of transcription of *lyt-cII*. Three-protein model considered here is different from that in [6], in which CII activates transcription of *cI* from a distinct (*pRE*) promoter. Because in the three-protein model considered here, CII has to compete with Lyt, which represses transcription of *lys*, for binding to the intergenic region and, thus, causing transcription of *lys*, engendering lysogeny becomes more challenging than in the model in which CII activated the transcription from a distinct promoter.

Model equations for the three-protein model are as follows.

Transcription of *lyt-cII* genes:
$$\frac{dxz_{mRNA}}{dt} = \frac{mk_1(1 + \frac{Z^c}{K_{D3}})}{1 + \frac{X^a}{K_{D1}} + \frac{Y^b}{K_{D2}} + \frac{Z^c}{K_{D3}}} - k_6 xz_{mRNA} \tag{15}$$

Translation of *lyt*:
$$\frac{dX}{dt} = k_2 xz_{mRNA} - k_7 X \tag{16}$$

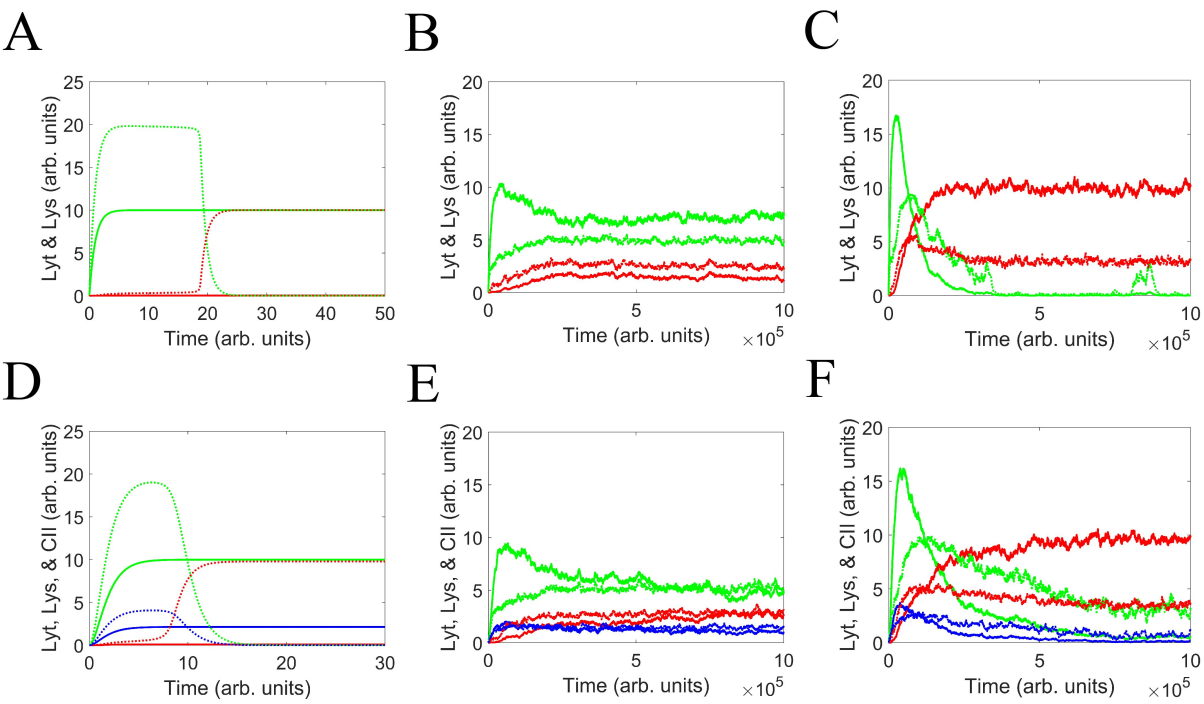

Figure 4: Deterministic and stochastic simulations of the quasi-minimal two-protein model (1A_Lyt_Lys) and the three-protein model (Lyt_Lys_CII). Lyt, Lys, and CII are represented by green, red, and blue, respectively. For the deterministic simulations, concentrations of proteins at MoI of 1 and 2 are depicted by solid curve and dashed curve, respectively. For the stochastic simulations, solid curve and dotted curve, respectively, represent average and standard deviation of number of protein molecules from 100 simulations. For a given model, the parameter set which generated highest stochastic success rate (SSR) for time length of 100 arb. units was used for simulations. Average stochastic simulation trajectories shown here are qualitatively similar to those of all of the other models for parameter sets which generated high values of SSR (data not shown), whereas the deterministic simulation trajectories were so, irrespective of the value of SSR (data not shown). In the stochastic simulation plots, the original abscissa, which had unequally-spaced time intervals, was converted to one with equally-spaced time intervals. Each unit of abscissa was divided into 10000 intervals. For the tiny fraction of intervals which still contained more than one event, their last events were defined to be their only events. (**A**) Deterministic simulations of 1A_Lyt_Lys. At MoI of 2, concentration of Lyt initially becomes more than its equilibrium concentration at MoI of 1 but then comes back to very low level. The bell-shaped kinetics occurs because the initial rate of production of Lyt at MoI of 2 (i.e., when Y remains very low) is higher than that at MoI of 1; however, as Lyt's concentration increases, positive autoregulatory loop of *lys* gets triggered, leading to production of Lys in high amount, which represses *lyt*. (**B-C**) Stochastic simulations of 1A_Lyt_Lys for MoI of 1 and 2, respectively. (**D**) Deterministic simulations of Lyt_Lys_CII. At MoI of 2, concentrations of Lyt and CII initially become more than their respective equilibrium concentrations at MoI of 1 but then come back to very low levels. A theoretical study of the lysis/lysogeny decision using a three-protein model, which is very similar to that of the present paper, also reported bell-shaped kinetics for CII [6]. Analogous to the two-protein model, the reason for bell-shaped kinetics for the three-protein model is that the initial rate of production of CII at MoI of 2 (i.e., when Y and Z remain very low) is higher than that at MoI of 1; however, as CII's concentration increases, positive autoregulatory loop of *lys* gets triggered, leading to production of Lys in high amount, which represses the promoter controlling the expression of *lyt* and *cII*. (**E-F**) Stochastic simulations of Lyt_Lys_CII for MoI of 1 and 2, respectively. Bell-shaped curve for CII at MoI of 6 was reported by an experimental study [3]. Curiously, while equilibrium concentration of CII is higher than that of CI in the deterministic simulation, CII settles down to lower level than CI in the stochastic simulation.

Translation of *cII*:

$$\frac{dZ}{dt} = k_4 x z_{mRNA} - k_9 Z \tag{17}$$

Production of Lys:

$$\frac{dY}{dt} = \frac{m(k_5 \frac{Y^b}{K_{D2}} + k_3 \frac{Z^c}{K_{D3}})}{1 + \frac{X^a}{K_{D1}} + \frac{Y^b}{K_{D2}} + \frac{Z^c}{K_{D3}}} - k_8 Y \tag{18}$$

where, c is Hill coefficient of CII binding; k_1 is basal expression rate of transcription of *lyt-cII* mRNA; k_6 is the degradation rate of *lyt-cII* mRNA; K_{D3} is the effective equilibrium dissociation constant of CII (see Methods); k_2 and k_4 are translation rates of *lyt* and *cII*, respectively. k_5 and k_3 are rate constants for transcriptional activation of *lys* by Lys and CII, respectively. The degradation rate of *lyt-cII* mRNA (xz_{mRNA}), namely k_6, was arbitrarily chosen to be unity so as not to increase the computation time. The degradation rate of Lyt (X), CII (Z), and Lys (Y), namely k_7, k_9, k_8, respectively, were taken to be unity for the same reason that the degradation rates in the two-protein models were set equal to 1. Because for 1A_Lyt_Lys, DSQs generated by Hill coefficients' set of a=2, b=2 were as high as DSQs generated by that of a=2, b=4, applying occam's razor, Hill coefficient for Lyt and Lys binding was taken to be 2 and 2, respectively, and not 2 and 4. Further, taking lead from here, Hill coefficient for CII binding was taken to be 2, even though it has been shown to exist as tetramer in solution [21] and in crystallized free and DNA-bound state [22].

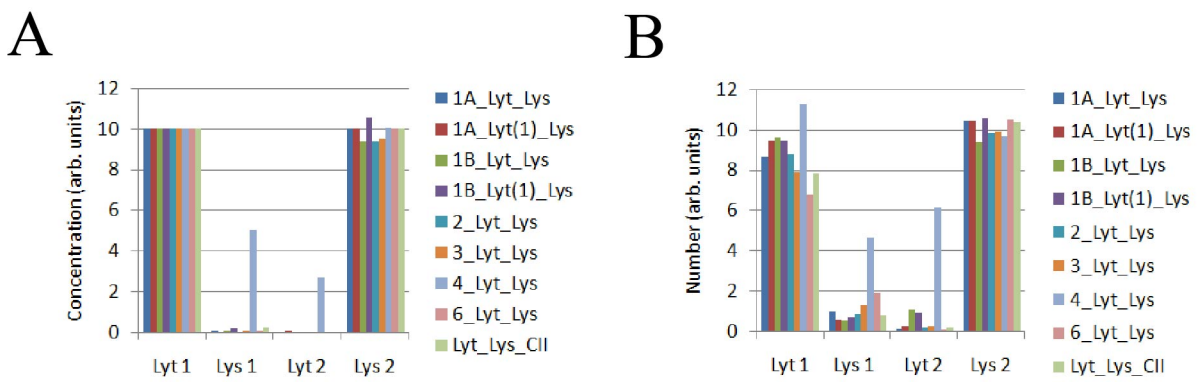

Figure 5: Equilibrium values corresponding to those parameter sets which gave maximum SSR for time length of 100 arb. units (see Table 5). (**A**) Deterministic simulations. (**B**) Stochastic simulations. Notably, value of Lys at MoI of 1 and that of Lyt at MoI of 2 for 4_Lyt_Lys are much higher than those of the any other model.

2.6. Stochastic simulation

Because gene expression is stochastic [23,24], it is crucial to simulate the model equations in the presence of noise. Therefore, stochastic simulations were performed by implementing Gillespie algorithm [25], using parameter sets obtained from the deterministic simulations. The transient kinetics of Lyt, namely the initial rise, followed by a partial drop, and plateauing of the trajectories at MoI of 1 and bell-shaped trajectories at MoI 2, were completed at most by 100 arb. units. In order to calculate the stochastic switch quotient (SSQ), the protein levels were averaged for 25, 50, 100, and 200 arb. units of time, keeping the starting time point at 100 arb. unit. That is, for averaging time=25, averaging was done between 100 and 125 arb. units; for averaging time=50, averaging was done between 100 and 150 arb. units; and so forth. All stochastic data, unless specified otherwise, presented in this paper is for time=100 arb. units (for other time lengths, see Supplementary Information: Table 1-9).

Barring a few parameter sets of two models for simulation time length of 200 arb. units, no parameter set was able to produce the switching behaviour in every run of the stochastic simulations. That was because either the SSQ was negative ($S_1 < S_2$) or, rarely, S_3 was zero. The percentage of runs that produced finite, positive SQ during stochastic simulation for a given parameter set and set of Hill coefficients would henceforth be referred to as stochastic success rate (SSR). Only those runs that produced finite, positive SSQ were taken into account in calculating average and standard deviation of SSQ. For the two-protein and the three-protein models, respectively, 500 and 200 simulations were performed for a given parameter set and a set of Hill coefficients.

For all the models, for any given parameter set, SSQ was less than its deterministic counterpart (DSQ). As Table 2 and 3, respectively, show for SSR $\geqslant$ 90 and 95 >SSR $\geqslant$ 90, for all the models, average SSQ for Hill coefficients' set of a=2, b=2 was higher than that for Hill coefficients' set of a=2, b=4. For all the models and for either of the ranges of SSR, lowest SSQ for the former pair of Hill coefficients was higher than highest SSQ for the latter pair of Hill coefficients. As Table 5 shows, maximum SSR across all the models, barring

Table 4: Number of parameter sets for various ranges of stochastic success rate.

Model	SSR[a] ⩾ 95		95 >SSR ⩾ 90		90 >SSR ⩾ 80		Total no. of parameter sets	
	a=2 b=2	a=2 b=4	a=2 b=2	a=2 b=4	a=2 b=2	a=2 b=4	a=2 b=2	a=2 b=4
1A_Lyt_Lys	1	4	5	3	10	3	20	17
1A_Lyt(1)_Lys	0	0	1	1	3	9	17	15
1B_Lyt_Lys	1	1	1	5	2	2	6	11
1B_Lyt(1)_Lys	0	1	0	0	8	6	14	10
2_Lyt_Lys	0	0	0	0	0	0	1	6
3_Lyt_Lys	0	0	0	0	1	6	10	12
4_Lyt_Lys	0	2	0	0	0	0	2	2
6_Lyt_Lys	0	0	0	0	0	0	1	13
Lyt_Lys_CII	1	N/A	2	N/A	1	N/A	9	N/A
Lyt_Lys_CII(1)	1	N/A	1	N/A	5	N/A	9	N/A
Lyt(1)_Lys_CII(1)	0	N/A	1	N/A	1	N/A	9	N/A
Lyt_Lys_CII′	0	N/A	2	N/A	4	N/A	9	N/A

[a] SSR = Stochastic Success Rate

a few cases, increases with time length of simulation. Also, the order of models arranged according to maximum SSR remains more or less same across all the four time lengths of the simulations. For a given model, for those time lengths which generated at least one parameter set with SSR being within at least one of the two ranges, (average) DSQ and (average) SSQ were similar to those of the time length of 100 arb. units for the respective range of SSR (see Supplementary information Table 1-6). 4_Lyt_Lys has highest maximum SSR, closely followed by 1B_Lyt(1)_Lys, 1B_Lyt_Lys, 1A_Lyt_Lys, and 1A_Lyt(1)_Lys, in that order. Maximum SSR of 3_Lyt_Lys is around halfway between that of 1B_Lyt(1)_Lys and 6_Lyt_Lys, which has second lowest maximum SSR. 2_Lyt_Lys has lowest maximum SSR. As Table 4 shows, maximum SSR of 6_Lyt_Lys and 2_Lyt_Lys are lower than 80. The last two models possess double-negative feedback, or mutual repression motif. This result that lambda-like switches based on mutual repression motif work poorly in the presence of noise has been reported earlier also. Avlund et al. showed that various two-protein minimal models, vast majority of which were based on mutual repression motif, of lambda's switch, which were able to carry out the counting task (as aforementioned) deterministically, did not function when noise was introduced [8]. Notably,

one of their two two-protein models (i.e., motif b of figure 2 of Avlund et al. [8]) which could pass the counting task even in the presence of noise (though with much lower success as compared to their three-protein models) is model 6_Lyt_Lys in the present paper.

Table 5: Highest stochastic success rate for the four time lengths

Model	Highest stochastic success rate			
	T = 25	**T = 50**	**T = 100**	**T = 200**
1A_Lyt_Lys	89.4	93.6	96.8	100
1A_Lyt(1)_Lys	83.8	84	91	97.2
1B_Lyt_Lys	90.8	93.4	97.2	99.2
1B_Lyt(1)_Lys	96	96.6	99.2	100
2_Lyt_Lys	39.4	40.2	46.4	50.6
3_Lyt_Lys	74.8	81.2	86	93.8
4_Lyt_Lys	87.4	93	98.8	99.8
6_Lyt_Lys	57	64.8	71.2	83.8
Lyt_Lys_CII	92.5	94.5	95	96.5
Lyt_Lys_CII(1)	90	96	95	97
Lyt(1)_Lys_CII(1)	88	87	94.5	97.5
Lyt_Lys_CII'[a]	83	85.5	92.5	96.5
Lyt_Lys_CII(1)'[a]	32.5	35	39	43

[a] Parameter sets selected for bistability at MoI of 2.

2.7. Bistability at MoI of 1 and lysogen stability

After commitment to lysogeny (here: at MoI of 2), if only one of the phage genomes gets integrated into the bacterial chromosome, thereby bringing MoI down to 1, maintenance of lysogeny would not be possible, and lysis would ensue, in case only one stable state existed at MoI of 1. Therefore, in order to assess lysogen stability in the given scenario, the model equations were solved at MoI of 1 with the initial values of Lyt and Lys being their equilibrium values at MoI of 2 (when the model equations were solved with the initial values of Lyt and Lys being equal to zero). That is, it was checked whether at MoI of 1 Lyt and Lys, starting from their equilibrium values at MoI of 2, could still settle down to values which are representative of lysogeny (i.e., low value of Lyt and high value of Lys). It should be noted that this method finds

out only those second stable states which represent lysogeny maintenance. In deterministic simulations, all of the two-protein models possessing the positive feedback (except 1B_Lyt_Lys for one parameter set) exhibited bistability at MoI of 1 for all of the parameter sets. In the other stable state, concentration of Lyt is almost zero and that of Lys is about half of its concentration at MoI of 2.

For 4_Lyt_Lys, none of the parameter sets produced bistability at MoI of 1. That's because the nullcline of Lys is a flat line and that of Lyt is a monotonically decreasing function. Therefore, 4_Lyt_Lys cannot produce bistability at either of the MoIs. For 2_Lyt_Lys, for Hill coefficient's set of a=2, b=2 , only one parameter set generated bistability at MoI of 1. For Hill coefficient's set of a=2, b=4, all parameter sets except one exhibited bistability at MoI of 1. It turned out that the parameter set which did not produce bistability has K_{D1} and K_{D2} about 5 orders of magnitude higher than average K_{D1} and K_{D2}, respectively, of rest of the parameter sets of 2_Lyt_Lys (Supplementary Information: Table 10). Because in 4_Lyt_Lys Lyt does not bind to any promoter, it can be thought to be having extremely large equilibrium dissociation constant (i.e., very large K_{D1}). Further, K_{D2} of the parameter set is within an order of average K_{D2} of parameter sets of 4_Lyt_Lys for the given pair of Hill coefficients. Therefore, it can be argued that the parameter set actually belonged to 4_Lyt_Lys.

All of the three-protein models for every parameter set exhibited bistability at MoI of 1. The values of Lyt and Lys of second stable state at MoI of 1 in the three-protein models were about the same as those of second stable states in the two-protein models at MoI of 1. Stochastic bistability was searched for using the same method employed to search for deterministic bistability. However, stochastic bistability could not be detected for any of the two-protein models and the three-protein model at MoI of 1. This may be due to noise driving trajectories out of the basin of attraction of one stable state to that of the other [26]. Stochastic bistability was checked only for those parameter sets whose SSR was more than 95%. This was done to avoid mistaking failure to generate the switching behaviour for bistability.

2.8. Bistability at MoI of 2 and lytic development's stability

Because replication of viral genome after committment to lysis causes MoI to become more than 1, (here: 2), the process of lytic development will not remain stable, and lysogeny will ensue, in case only one stable state existed at MoI of 2. Therefore, in order to assess the stability of lytic development, the model equations were solved at MoI of 2 with the initial values of Lyt and Lys being their equilibrium values at MoI of 1 (when the model equations were solved with the initial values of Lyt and Lys being equal to zero). That is, it was checked whether at MoI of 2 Lyt and Lys, starting from their equilibrium values at MoI of 1, could still settle down to values which are representative of lysis (i.e., high value of Lyt and low value of Lys). As previously, the method finds out only those second stable states which are relevant for the stability of lytic development. In deterministic simulations, only two models, namely 2_Lyt_Lys and 6_Lyt_Lys, exhibited bistability at MoI of 2. Notably, among all of the two-protein models, only these two models possess Lyt repression of *lys*. In the other stable state, concentration of Lyt is double of its concentration at MoI of 1 and that of Lys is almost zero. For 2_Lyt_Lys, for Hill coefficient's set of a=2, b=2, the parameter set that produced bistability at MoI of 1 is the only parameter set that exhibited bistability at MoI of 2. For Hill coefficient's set of a=2, b=4, all parameter sets (except the one which actually belonged to 4_Lyt_Lys) exhibited bistability at MoI of 2. For 6_Lyt_Lys, for Hill coefficients' set of a=2, b=2, one parameter set generated bistability at MoI of 2. For Hill coefficients' set a=2, b=4, all parameter sets except one exhibited bistability at MoI of 2.

None of the three-protein models for any parameter set exhibited bistability at MoI of 2. However, given that the three-protein models possess the repression of *pR* by Lyt, a feature associated with bistability at MoI of 2 for the two two-protein models, it was suspected that it, too, might exhibit bistability at MoI of 2 if this feature was actively selected for. Therefore, the parameter search was carried out by adding another selective filter. That is, Lyt at MoI of 1 (when the model equations were solved with the initial values of Lyt and Lys being equal to zero) ¡ Lyt at MoI of 2 (when the model equations were solved with the initial values of

Lyt and Lys being their equilibrium values at MoI of 1). This method generated 9 parameter sets for Lyt_Lys_CII (named Lyt_Lys_CII') and 3 for Lyt_Lys_CII(1) (named Lyt_Lys_CII(1)') but none for Lyt(1)_Lys_CII(1) (named Lyt(1)_Lys_CII(1)'). Further, while maximum SSR of Lyt_Lys_CII' was very close to that of Lyt_Lys_CII, maximum SSR of Lyt_Lys_CII(1)' was less than half of that of Lyt_Lys_CII(1). This result underscores the importance of cooperative binding for all the three key proteins of the switch. The most appreciable change between Lyt_Lys_CII and Lyt_Lys_CII') was in the equilibrium dissociation constants. Average K_{D1} and K_{D2} of Lyt_Lys_CII') were about an order of magnitude lower than those of Lyt_Lys_CII, whereas average K_{D3} was about two orders of magnitude lower for the same comparison (Supplementary Information: Table 11). Notably, parameter search with the additional selective filter did not produce any parameter set (i.e., which could exhibit bistability at MoI of 2) for the two-protein models other than 2_Lyt_Lys and 6_Lyt_Lys. This further confirms that Lyt (or Cro) repression of *lys* (or *pRM*) is required to produce bistability at MoI of 2.

A study demonstrated that average number of phages produced per infection (burst size) of a mutant defective in Cro repression of *pRM* was 75% of that of the wild type, whereas burst size produced after prophage induction triggered by ultra violet (UV) at 10 J/m^2 UV and at 5 J/m^2 UV in mutants were, respectively, 1/3 and 1/8 times of that of the wild type [27]. The lack of Cro repression of *pRM* having much larger impact on burst size in prophage induction as compared to infection can be explained in the following way. Replication of the viral genome post-induction leads to lysogeny-like situation (i.e., high MoI), thus residual CI and CI being produced by existing *cI* mRNAs which remain unaffected by UV irradiation, activate multiple *pRM* promoters which are defective in binding Cro, and trigger the positive feedback loop. CI, now present in higher concentration, represses promoter *pR*, under whose control transcription of *Q* gene lies. Therefore, lytical processes, namely viral genome replication and the associated proteins production, occur to a lesser extent, hence lower burst size.

On the other hand, for CI to affect lytic development post-infection in phages lacking Cro repression of *pRM*, CI has to be made to the level sufficient to activate *pRM*. But for that to happen, transcription from promoter *pRE*, activated by CII, and translation of the resulting

476 transcripts has to take place first. This extra time allows more rounds of replication of viral
477 genome and additional production of viral proteins, therefore the effect of defective Cro repression
478 of *pRM* on burst size during lytic development post-infection is lesser. Although before CI
479 could repress *pR*, CII will impede production of Q protein by activating transcription from
480 *paQ*, but as aforementioned, this process is much weaker than CI repression of *pR* in inhibiting
481 production of Q protein. It should be noted that modeling scheme in the present work is unable
482 to embody the fact that even without Lyt repression of *lys* lytic development does occur, albeit
483 to a lesser extent, namely burst size is lower than when Lyt repression of *lys* is intact. That
484 is, in the (quasi-) minimal models which do not exhibit bistability at MoI of 2 increase in
485 MoI due to genome replication after commitment to lytic will reverse the commitment and
486 lytic development (zero burst size), and consequently lysogeny would ensue. This inability
487 of the present modeling scheme stems from not considering lytical processes, initiated by Q,
488 explicitly.

489 The frequency of lysogeny (proportion of infected cells that formed lysogens) post-infection
490 in the mutant was 30% higher than that in the wild type, and the fraction of cells induced at
491 10 J/m^2 UV and at 5 J/m^2 UV was 32% and 48% lower for the same comparisions. The
492 increase in the lysogenization frequency post-infection and decrease in the fraction committing
493 to induction is due to defective Cro repression of *pRM* tilting the decision towards lysogenization.
494 However, for the latter case there exists an additional factor: activation of *pRM* by residual CI
495 and CI being produced by existing *cI* mRNA. It should be noted that the protein-only approach
496 of the present study cannot take into account production of CI by the already existing *cI* mRNAs
497 during the process of prophage induction.

498 Cro represses *pL* and *pR* by fourfold and twofold, respectively [28]. Therefore, the absence
499 of Cro increases the level of CII by allowing transcription of *cII*, which is under the control of
500 *pR*, and *cIII*, which is under the control of *pL* and whose product prevents degradation of CII
501 by protease HflB. CII engenders the production of CI by activating *pRE*, and CI activates *pRM*,
502 therefore like Cro repression of *pRM*, inhibition of *pL* and *pR*, too, by Cro should contribute to
503 bistability at MoI of 2. These processes are not modeled in the present study. However, because

the end result of production of CII is activation of *pRM*; qualitatively, the Cro inhibition of *pL* and *pR* can be subsumed into Cro inhibition of *pRM*. As is the case at MoI of 1, stochastic bistability could not be detected for any of the two-protein and the three-protein models at MoI of 2. Because stochastic bistability could be checked only for those parameter sets whose SSR was more than 95%, most models, including the two two-protein models which exhibited deterministic bistability at MoI of 2 (i.e., 2_Lyt_Lys and 6_Lyt_Lys), could not be checked for their potential ability to produce stochastic bistability at the MoI of 2.

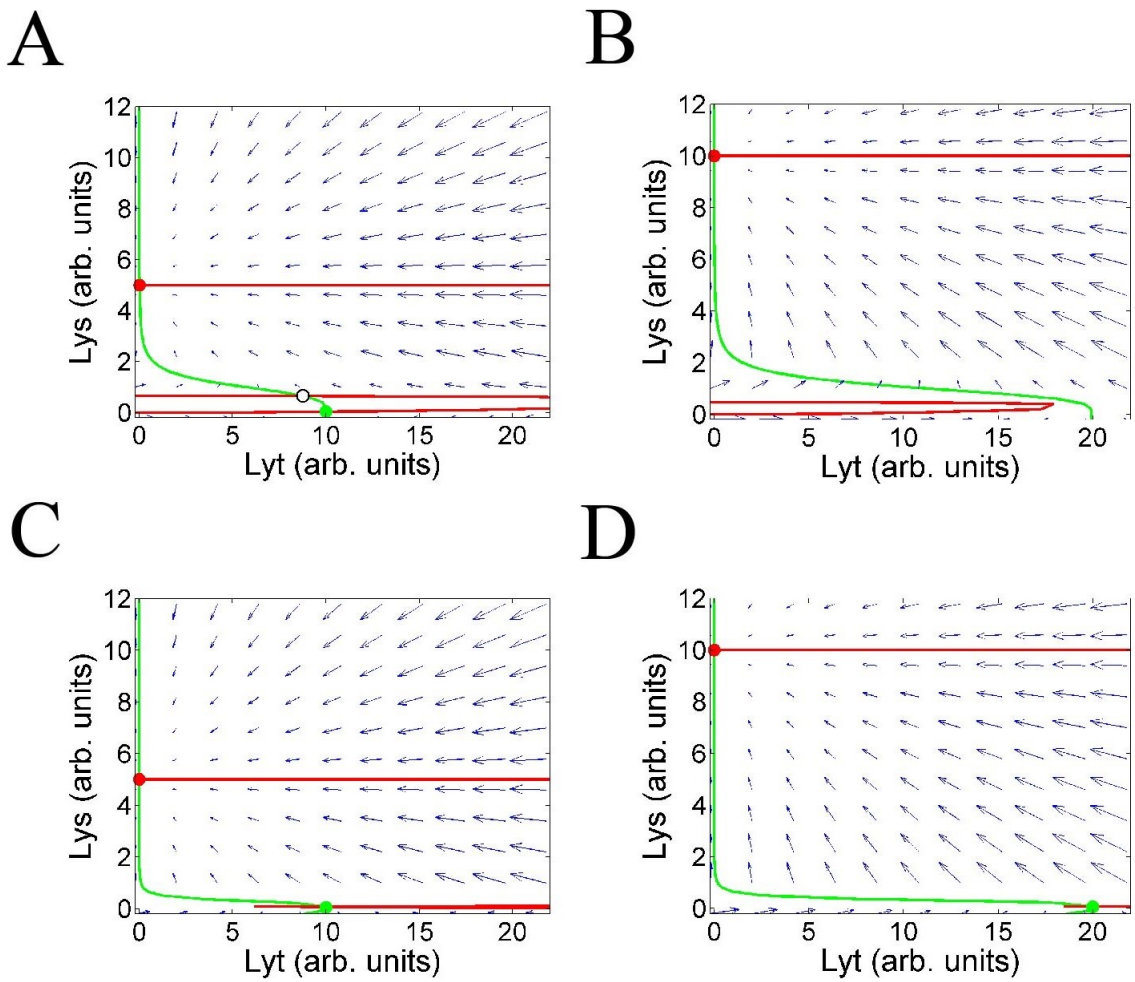

Figure 6: (**A-B**) Phase diagram of 1A_Lyt_Lys and 6_Lyt_Lys for MoI of 1 and 2 corresponding to their respective parameter set which gave maximum SSR Green and red full circles are stable fixed points, whereas empty black circle is unstable fixed point (not marked for 6_Lyt_Lys at both the MoIs because of being very close to the stable fixed point). At MoI of 1, the green dot shows anti-immune state, whereas red dot shows immune state; that is, lysogeny established by two phages, genome of one of which gets integrated into the host's chromosome. At MoI of 2, red dot shows immune state established by two phages, whereas green dot shows anti-immune state established at MoI of 1 but which persists even after genome replication

3. Design principles of lambda's switch

Because, barring 2_Lyt_Lys for a=2, b=2 and 4_Lyt_Lys for both sets of Hill coefficients, DSQ generated by every model for every parameter set is more than 0.9, design principles of lambda's switch cannot be based on the value of DSQ. One noticeable feature that the models clearly differ in is maximum SSR they generate. Also, because bistability at both the MoIs is an absolute requirement of the switch, among the two-protein models, only 2_Lyt_Lys and 6_Lyt_Lys (2_Lyt_Lys plus self-activation of *lys*) qualify as possible switch designs. Of these two, the latter is a better model than the former because of producing higher maximum SSR. 6_Lyt_Lys (3_Lyt_Lys plus Lyt repression of *lys*) has lower maximum SSR than 3_Lyt_Lys but higher SSR than 2_Lyt_Lys. Hence, it can intuitively be argued that 6_Lyt_Lys is a compromise between 3_Lyt_Lys (for its high maximum SSR) and 2_Lyt_Lys (for its bistability at MoI of 2). In other words, 6_Lyt_Lys incurs a drop in maximum SSR as compared to 3_Lyt_Lys in lieu of gaining bistability at MoI of 2.

Further, from the comparison of 1B_Lyt_Lys with 3_Lyt_Lys, it is inferred that activating promoter of *lys* in order to trigger the positive feedback loop at MoI of 2, by Lyt binding gives higher maximum SSR than having basal expression of *lys* carry out that function. Applying this cause-effect relationship in an additive fashion to 6_Lyt_Lys, it is hereby posited that maximum SSR of 6_Lyt_Lys is improved by replacing basal expression of *lys* with Lyt activation of *lys*. However, because Lyt's function in 6_Lyt_Lys is to repress *lys*, another protein, namely CII, whose gene is co-transcribed with *lyt*, is employed to activate *lys* (in a cooperative manner). The three-protein model thus generated is minimal model for lambda's switch. Thus, lambda's lysis/lysogeny switch is based on three features: a) mutual repression, b) cooperative positive autoregulation of *cI*, and c) cooperative binding of the activator protein, and not basal expression, triggering positive feedback loop formed by autoregulation of *cI*.

A speculative explanation for why 1B_Lyt_Lys has higher maximum SSR than 3_Lyt_Lys is proposed now. 1B_Lyt_Lys differs from 3_Lyt_Lys, in that in the former, Lyt activates transcription of *lys*, whereas in the latter, *lys* possesses basal expression. At MoI of 2, in both

the models, concentration of Lyt initially becomes more than its equilibrium concentration at MoI of 1 but then comes back to very low level, as shown in the qualitatively representative figure 4A. As figure 4A shows, until Lyt's trajectory begins to steeply fall back to low level, concentration of Lys remains very low; and as the concentration of Lys rises, Lyt's level drops. This indicates that the time since beginning of the simulation and until the steep fall of the Lyt's trajectory is the period when the positive autoregulatory loop is triggered. Additionally, as also shown in the figure 4A, for the two models, concentration of Lys at MoI of 1 remains very low as compared to that of Lyt. It may be argued that these observations justify approximating Y = 0 in the numerator and denominator of the first term on the right-hand side of equality of equations 4 and 8 for time interval between beginning of simulation and until the point when Lyt's trajectory begins to drop sharply. Thus one gets dY/dt equal to $m(k_3X^a/K_{D1})/(1+X^a/K_{D1})$ - k_5Y and mk_3 - k_5Y, respectively, - an approximation considered to be valid for the said time interval.

1B_Lyt_Lys and 3_Lyt_Lys can fail in two ways: either the positive feedback loop, formed by self-activation of *lys*, gets triggered at MoI of 1 or fails to get triggered at MoI of 2. The first term of the approximated form of dY/dt for the two models represents triggering of the positive feedback loop, which occurs during the said time interval. For 3_Lyt_Lys, it is equal to mk_3 and, as a result, increases proportionally to increase in MoI. For 1B_Lyt_Lys, the effect of increase in MoI on the first term of the approximated form of dY/dt depends on the value of X^a relative to K_{D1}. For $X^a/K_{D1} >> 1$, the term becomes approximately equal to that for 3_Lyt_Lys. On the other hand, for $X^a/K_{D1} << 1$, the term becomes approximately equal to $m(k_3X^a/K_{D1})$. Because during the said time interval, Lyt's concentration at MoI of 2 remains higher than its concentration at MoI of 1, increase in MoI will lead to more than proportional increase in the term. This greater separation between propensity to trigger the positive feedback loop at MoI of 1 and 2 for 1B_Lyt_Lys as compared to 3_Lyt_Lys is posited here to be responsible for the former model having higher maximum SSR than the latter. Another, perhaps less likely, reason could be that, while gene expression in 3_Lyt_Lys is ultimately driven by two promoters having independent (basal) expressions, that in 1B_Lyt_Lys (and also 1A_Lyt_Lys) is ultimately driven

by only one promoter, thereby lending tighter control over overall expression for the latter model. Because, as argued above, 1B_Lyt_Lys is qualitatively similar to 1A_Lyt_Lys – differing from the latter in only not possessing self-inhibition of textitlyt and, therefore, having identical equation for the dynamics of Lys – the argument in the previous paragraph also explains why 1A_Lyt_Lys has higher maximum SSR than 3_Lyt_Lys.

4. Methods

4.1. Derivation of the model equations

In order to calculate the effective equilibrium dissociation constants of proteins, the present work uses the fact that binding of protein to itself or DNA are much quicker processes than transcription and translation. Hence, the former processes are assumed to have reached quasi-steady states. In the expressions below, P, X, and Y are promoter, Lys, and Lyt, respectively.

$$\mathrm{X} + \mathrm{X} \overset{K_{DX}}{\rightleftharpoons} \mathrm{X}_2 \qquad \frac{[X]^2}{[X_2]} = K_{DX}$$

$$\mathrm{Y} + \mathrm{Y} \overset{K_{DY1}}{\rightleftharpoons} \mathrm{Y}_2 \qquad \frac{[Y]^2}{[Y_2]} = K_{DY1}$$

$$\mathrm{Y}_2 + \mathrm{Y}_2 \overset{K_{DY2}}{\rightleftharpoons} \mathrm{Y}_4 \qquad \frac{[Y_2]^2}{[Y_4]} = K_{DY2}$$

$$\mathrm{P} + \mathrm{X} \overset{K'_{D0}}{\rightleftharpoons} \mathrm{PX} \qquad \frac{[P][X]}{[PX]} = K'_{D0} = K_{D0}$$

$$\mathrm{P} + \mathrm{X}_2 \overset{K'_{D1}}{\rightleftharpoons} \mathrm{PX}_2$$

$$\frac{[P][X^2]}{[PX_2]} = K'_{D1}K_{DX} = K_{D1}$$

$$\mathrm{P} + \mathrm{Y}_4 \overset{K'_{D2}}{\rightleftharpoons} \mathrm{PY}_4 \qquad \frac{[P][Y^4]}{[PY_4]} = K'_{D2}K^2_{DY1}K_{DY2} = K_{D2}$$

The above expressions for concentrations of promoter-protein complexes are for cases where a) Lyt binds as monomer, b) Lyt binds as dimer, and c) Lys binds as tetramer. They exhaust all other cases, that is, of monomeric and dimeric Lys and CII. Further, the transcription and mRNA degradation are assumed to be much faster processes than the translation and protein degradation, hence mRNAs also are assumed to have reached quasi-steady states. Because genes *lyt* and *cII* are under the control of the same promoter, in order to allow for potentially different rates of translation of their corresponding cistrons during the stochastic simulations, the dynamics of *lyt-cII* mRNA is considered explicitly. Hence, except that for *lyt-CII* mRNA, all the equations describe the dynamics of only the proteins. With the expressions for concentrations of promoter-protein complexes, one can write generalized form of the term representing the rate of protein production.

$$\frac{b + \sum_i k_i.[DNA - Prot_i]}{[Unbound\ DNA] + \sum_i [DNA - Prot_i]}$$

where b, if present, is basal expression and k_i is rate constant for transcriptional activation by i_{th} protein.

4.2. Parameter search

Parameter sets, namely rate constants and equilibrium dissociation constants, of the model equations were searched deterministically in two stages, namely order search and linear search (as they are named here). In the order search, rate constants and equilibrium dissociation constants were searched as 3's exponent, which was varied between -5 and 5 with the difference of 1 (1/243, 1/81, ... 1 ..., 81, 243), in a nested fashion. Thus, number of parameter sets

searched for a given model was equal to the number of parameters raised to the power 11. For the three-protein models, in the interest of saving time, basal expression of *pR* and equilibrium dissociation constant of Lyt and Lys were taken from the parameter set of 1A_Lyt_Lys (for Lyt_Lys_CII and Lyt_Lys_CII(1)) and 1A_Lyt(1)_Lys (for Lyt(1)_Lys_CII(1)) which gave maximum SSR. In the interest of saving time, a parameter set was rejected if its DSQ was lower than that of the previous parameter set. Switch quotients generated by the order search are coarse because rate constants and equilibrium dissociation constants were increased geometrically, thereby causing a lot of intervening values to remain unsampled. Therefore, parameter sets generated by the order search were improved in terms their DSQ, by the linear search, which searches the neighbourhood of a parameter set arithmetically. It was noted that those parameter sets, generated by the order search, whose DSQs were too close to each other were either rescaled form of each other or differed in those parameters to which DSQ was resilient up to a certain range. Hence, in order to remove redundancy, the starting parameters for the linear search were taken in such a way that the difference between their consecutive DSQs is at least 0.001. However, because 1A_Lyt_Lys, 1A_Lyt(1)_Lys, 1B_Lyt_Lys, 1B_Lyt(1)_Lys, and all of the three-protein models generated large number of parameter sets, the difference between the consecutive DSQs for these four models is kept to be at least 0.01, in the interest of saving time. This way, Lyt_Lys_CII, Lyt_Lys_CII(1), Lyt(1)_Lys_CII(1) generated 15, 15, and 17 parameters, respectively. However, in order to save time, only their first 9 parameters were considered for the linear search.

The linear search was carried out in the following way. Value of each parameter (say, V) of a set was varied between V-3*V/5 and V+3*V/5 with increment of V/5, in a nested fashion. That is, suppose value of a parameter was 10, it was varied between 4 and 16 with increment of 2. Thus, number of parameter sets searched for one round was equal to the number of parameters raised to the power 7. However, for three-protein model, which had eight parameters, in order to save time, each parameter was varied between V-2*V/5 and V+2*V/5 with the increment of V/5, in a nested fashion. A parameter set was rejected if its DSQ was lower than that of the previous parameter set. The linear search was ended if the latest DSQ

was either lower than the previous one (which never happened) or if ((latest DSQ - previous DSQ)/previous DSQ) was less than 0.01. Again, in the interest of saving time, for three-protein model, the linear search was ended if DSQ at the end of last round was at least 0.95. It should be noted that the linear search is path dependent: a linear search may end up at a local maximum because it treaded a path which yielded higher DSQ initially instead of another path which, though, had yielded lower DSQ initially but would have led to the global maximum.

Parameter sets were normalized such that Lyt's equilibrium concentration at MoI of 1 becomes 10 arb. units. This was done for two purposes: a) to reduce the chances of the values of Lyt at MoI of 1 and those of Lys at MoI of 2 dropping to zero in the stochastic simulations; and b) in order to make the comparison of the parameter sets and equilibrium values of the proteins visually feasible. For both the types of searches, simulations were carried out for time 100 arb. units. There was a possibility of a system of equations, defining a particular model, not reaching equilibrium by 100 arb. units for a given parameter set. In order to eliminate such parameter sets, a final simulation was done for 10^5 arb. units. Only few parameter sets turned out to not had reached equilibrium, and all of such parameter sets produced negative DSQ.

References

[1] Court DL, Oppenheim AB, Adhya SL A new look at bacteriophage lambda genetic networks. J Bacteriol. 2007; 189: 298–304.

[2] Kourilsky P Lysogenization by bacteriophage lambda. I. Multiple infection and the lysogenic response. Mol Gen Genet. 1973; 122: 183–195.

[3] Kobiler O, Rokney A, Friedman N, Court DL, Stavans J, Oppenheim AB Quantitative kinetic analysis of the bacteriophage lambda genetic network. Proc Natl Acad Sci U S A. 2005; 102: 4470–4475.

[4] Avlund M, Dodd IB, Semsey S, Sneppen K, Krishna S Why do phage play dice? J Virol. 2009; 83: 11416–11420.

[5] Kourilsky P Lysogenization by bacteriophage lambda. II. Identification of genes involved in the multiplicity dependent processes. Biochimie. 1974; 56: 1517–1523.

[6] Weitz JS, Mileyko Y, Joh RI, Voit EO Collective decision making in bacterial viruses. Biophys J. 2008; 95: 2673–2680.

[7] Avlund M, Dodd IB, Sneppen K, Krishna S Minimal Gene Regulatory Circuits that Can Count like Bacteriophage Lambda. J. Mol. Biol. 2009; 394: 681–693.

[8] Avlund M, Krishna S, Semsey S, Dodd IB, Sneppen K Minimal gene regulatory circuits for a lysis-lysogeny choice in the presence of noise. PLoS One. 2010; 5:e15037.

[9] Hawley D, McClure WR Mechanism of activation of transcription initiation from the lambda P_{RM} promoter. J. Mol. Biol. 1982; 157: 493–525.

[10] Shih MC, Gussin GN Mutations affecting two different steps in transcription initiation at the phage λ *PRM* promoter. Proc Natl Acad Sci USA. 1983; 80: 496–500.

[11] Rosenberg M, Court D, Shimatake H, Brady C, and Wulff DL The relationship between function and DNA sequence in an intercistronic regulatory region in phage lambda. 1978 Nature 272:414–423.

[12] Cheng SW, Court DL, Friedman DI Transcription termination signals in the nin region of bacteriophage λ: identification of Rho-dependent termination regions. 1995 Genetics 140:875–887.

[13] Costantino N, Zuber M, Court D Analysis of mutations in the *ninR* region of bacteriophage lambda that bypass a requirement for lambda N antitermination. 1990 J. Bacteriol. 172:4610–4615.

[14] Roberts JW, Yarnell W, Bartlett E, Guo J, Marr M, Ko DC, Sun H, Roberts CW Antitermination by bacteriophage lambda Q protein. Cold Spring Harbor Symp. Quant. Biol. 1998 63:319–325.

[15] Banuett F, Hoyt MA, McFarlane L, Echols H, Herskowitz I hflB, a new Escherichia coli locus regulating lysogeny and the level of bacteriophage lambda cII protein. J Mol Biol. 1986; 187: 213–224.

[16] Kobiler O, Rokney A, Oppenheim AB Phage lambda CIII: a protease inhibitor regulating the lysis-lysogeny decision. PLoS One. 2007; 2:e363.

[17] Schmeissner U, Court D, McKenney K, Rosenberg M Positively activated transcription of lambda integrase gene initiates with UTP in vivo. 1981 Nature 292:173–175.

[18] Schmeissner U, Court D, Shimatake H, Rosenberg M Promoter for the establishment of repressor synthesis in bacteriophage λ. Proc Natl Acad Sci USA. 1980; 77: 3191–3195.

[19] Hoopes BC, McClure WR A cII-dependent promoter is located within the Q gene of bacteriophage lambda. Proc Natl Acad Sci U S A. 1985; 82: 3134–3138.

[20] Ho YS, Rosenberg M Characterization of a third, cII-dependent, coordinately activated promoter on phage lambda involved in lysogenic development. J Biol Chem. 1985; 260: 11838–11844.

[21] Ho Y, Lewis M, Rosenberg M Purification and properties of a transcriptional activator. The cII protein of phage lambda. J Biol Chem. 1982; 257: 9128–9134.

[22] Jain D, Kim Y, Maxwell KL, Beasley S, Zhang R, Gussin GN, Edwards AM, Darst SA Crystal structure of bacteriophage lambda cII and its DNA complex. Molecular Cell. 2005; 19: 259–269.

[23] Elowitz MB, Levine AJ, Siggia ED, Swain PS Stochastic gene expression in a single cell. Science. 2002; 297: 1183–1186.

[24] Golding I, Paulsson J, Zawilski SM, Cox EC Real-time kinetics of gene activity in individual bacteria. Cell. 2005; 123: 1025–1036.

[25] Gillespie DT Exact stochastic simulation of coupled chemical reactions. J Phys Chem. 1977; 81: 2340–2361.

[26] Samad HE, Khammash M, Petzold L, Gillespie D Stochastic modelling of gene regulatory networks. Int. J. Robust Nonlinear Control 2005; 15:691–711.

[27] Schubert RA, Dodd IB, Egan JB, Shearwin KE Cro's role in the CI–Cro bistable switch is critical for λ''s transition from lysogeny to lytic development. GENES & DEVELOPMENT. 2007; 21:2461–2472.

[28] Svenningsen SL, Costantino N, Court DL, Adhya S On the role of Cro in lambda prophage induction. Proc Natl Acad Sci U S A. 2005; 102: 4465–4469.