Design Principle of 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 possesses dual strategy of replication. Upon infecting its host, *Escherichia coli*, it can either choose lytic pathway, in which the host undergoes lysis, releasing hundreds of progeny viruses, or opt for lysogeny, in which the viral genome exists as part of bacterial chromosome. A classic and a molecular study have shown that the lysis/lysogeny decision depends upon the number of coinfecting phages, viz. the multiplicity of infection (MoI): lysis at low MoI; lysogeny at high MoI. Here, by constructing an expression for quality of the lysis/lysogeny minimal two-protein switch which, beside another thing, demands high equilibrium concentration of Cro-like protein (Lyt) and low equilibrium concentration of CI-like protein (Lys) - that is, lytic development - at MoI of 1, and vice versa - that is, lysogeny development - at MoI of 2, I demonstrate that positive feedback loop formed by activation of *cI*'s transcription by its own product in a cooperative manner underlies the switch's design. The minimal two-protein model is justified by showing its analogy with the GRN responsible for lysis/lysogeny decision. Existence of bistability at MoI of 1 is argued to be responsible for lysogen stability. By comparing the minimal model and its variants, possessing the positive feedback loop, with other models, without having the positive feedback loop, such as the mutual repression model, it is shown why lysis/lysogeny switch involving positive autoregulation of *cI* is evolved instead of one without it. A three-protein simplified version of lambda switch is shown to be equivalent to a close variant of the two-protein minimal switch. Only a fraction, if at all, of parameter sets that produced switch deterministically were able to do so in stochastic simulations more than 95% of the time. Bistability at MoI of 1 was not found during stochastic simulation.

Keywords: Bacteriophage λ, switch, positive feedback, bistablity

1. Introduction

Virulent bacteriophages possesses only one method of replication; that is, lytic strategy. However, other bacteriophages have a dual perpetuation strategy, viz. lytic and lysogeny. In lytic strategy, phage injects its genetic material into the host bacterium, viral genes are transcribed, m-RNAs, thus produced, are translated, and phage's genetic material is replicated. Finally, the host bacterium undergoes lysis, releasing progeny particles. In lysogeny, lytic pathway is repressed, the viral genome is integrated into that of the host bacterium, and thus, the virus exists in a latent state known as prophage. As the teleological explanation goes, lytic strategy leads to fast multiplication, but its risky, as 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 therefore, reproduces by a slower process as compared to lytic strategy, but this way phage safeguards its survival. Should a phage infect a bacterium containing lysogenized phage, lambda repressors (CI) present in the cytosol will not allow expression from *pR*. Thus, the newly entered phage would remain inert and, ultimately, get digested by the host's nucleases.

Being one of the simplest organisms to make a genetic decision, viz. to choose between lysis and lysogeny, bacteriophage lambda is a paradigm for gene regulation and developmental decision. Extensive study of phage's core genetic network involved in lysis/lysogeny decision has yielded a complex picture of interconnected feedback loops. A classic [1] and a molecular study [2] have shown that the lysis/lysogeny decision depends upon MoI. Avlund et al. analysed [3] Kourilsky's data [1,4] 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. This ability of a phage to choose between lysis and lysogeny based upon 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 was on a simplified model that possesses three critical proteins, viz. Cro, CI, and CII, and two promoters, viz. *pRM*/*pR* and *pRE* [5]. The work showed that the transient level of CII increases with MoI, and past a certain threshold, the nonlinearity of positive feedback

loop, formed by the self-activation of *cI*, gives rise to switch in terms of equilibrium values of the three proteins. Another way to understand design principles of a complex GRN is to obtain its minimal models, which lend themselves to deeper theoretical study than their original counterpart. In this study, I obtain a two-protein minimal model for lambda's GRN; and additionally, other two-protein models were constructed based upon certain features, such as the presence of feedback loop. It was shown that a positive feedback loop, formed by CI activating its own gene, is at the base of switchs design and lysogen stability. A three-protein simplified model of lambda's GRN, which differed slightly from the one modelled in aforementioned study [5], was shown to be reducible to a close variant of the two-protein minimal model in terms of equilibrium values of the proteins.

2. Result and Discussion

2.1. Minimal two-protein lysis/lysogeny switch

The promoter of *lyt* gene is constitutive; whereas, that of *lys* gene is positively regulated as they are in lambda phage's GRN. The role of Lys in the minimal two-protein model; that is, binding cooperatively to the intergenic region, activating transcription of its own gene, and inhibiting transcription of *lyt* gene, is identical to that of CI in lambda phage's GRN. The role of Lyt was conceptualized from first principle in the following way. At MoI of 2, equilibrium concentrations of Lyt and Lys should be much lower and much higher, respectively, as compared to those at MoI of 1. However, if Lyt did not affect *lys* promoter, assuming no basal expression of *lys* (which is weak promoter anyway), equilibrium concentration of Lyt at MoI 2 would be even higher, let alone much lower, than that at MoI of 1. And equilibrium concentration of Lys would be very low, instead of being high enough to repress *lyt*, at MoI of 2. Since the only protein present to actuate any process is Lyt, it was argued that Lyt should engender lysogeny and inhibit lytic pathway at MoI of 2.

Thus, Lyt activates transcription of *lys* (whose product causes lysogeny development), represses transcription of its own gene, thereby suppressing lytic development (though, as

shown below, the last interaction is dispensable), and activates imaginary downstream pathway which leads to lytic development. This seemingly paradoxical role of Lyt, as explained below, is due to it being proxy for CII, which causes lysogeny, and anti-termination factor Q, which enables transcription of lytic genes. The positive feedback loop constituted by transcriptional activation of *lys* by its own protein causes Lys to accumulate to low concentration at MoI of 1 and high concentration at MoI of 2. Thus, at MoI of 1, Lyt's equilibrium concentration is high because it is constitutively produced and Lys' equilibrium concentration is not high enough to repress its production. On the other hand, at MoI of 2, Lyt's equilibrium concentration is low because of repression by Lys, which is present in high concentration.

GRN underlying lysis/lysogeny decision is much more complex than the minimal two-protein model proposed here, because MoI is but one of many signals taken into account by the phage to decide between lysis and lysogeny. Since the expression for quality of lysis/lysogeny switch (the switching quotient) takes equilibrium values into account, the values of degradation constants of X (concentration of Lyt) and Y (concentration of Lys), viz. k_2 and k_5, respectively, can be subsumed into k_1, k_3, and k_4. Hence, they are taken to be unity for all two-protein models. This model would henceforth be referred to as 1A_Lyt_Lys.

1A_Cro_CI:

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

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

where, m is multiplicity of infection, k_1 is basal expression rate of *lys*, 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 ”combined” dissociation constants of Lyt and Lys, respectively (see Methods). In those

models where *lys* has basal expression, k_3 represents basal expression rate. Exponents a and b are Hill coefficients for binding of Lyt and Lys, respectively.

2.2. Analogy between the minimal two-protein model (1A_Lyt_Lys) and lambda phages GRN

Upon infection, RNA polymerase transcribes from the constitutive promoters, *pL* and *pR*, till it encounters transcription terminators *tL1* and *tR1*, respectively. *N* and *cro* genes are transcribed by *pL* and *pR*, respectively. The product of *N* is an anti-termination factor that modifies subsequent RNAPs initiating at *pL* and *pR* so that they move past their respective terminators and transcribe *cIII* and *cII* genes, respectively. Such an RNAP from *pR* is also able to transcribe through another terminator, *tR2*, present upstream of gene *Q* (see Figure 2). Up to this point, the pathway for lytic and lysogeny are identical. Lytic pathway is chosen when the extended transcription from *pR* also causes gene *Q* to be transcribed. Q, being an anti-termination factor, causes transcription of *pR'* to not terminate, as it would otherwise do, at *tR*', which is present at about 200 bases away from the beginning, thereby allowing transcription of the lytic genes downstream of *Q*. Once this happens, the cell is committed to lysis. CIII protein has an indirect role in establishing lysogeny. It prevents the degradation of CII by inhibiting bacterial protease HflB [6,7]. As the current paper focuses on the design principle of lysis/lysogeny switch, the (indrect) role of cIII will not be taken into consideration.

In lambda's GRN, *cII* and *Q* are under the control of promoter *pR*. Since in 1A_Lyt_Lys *lyt* is transcribed from *pR*, Lyt protein should be functionally equivalent to CII and Q. That is, on the whole, CII and Q should carry out three actions: activate transcription from *lys*, inhibit transcription from *lyt* gene, and engender lytic development. When CII accumulates in sufficient concentration, it activates transcription from three promoters: *pI*, *pRE*, and *pAQ* [10,11]. Promoter *pI* transcribes *int* gene, required for the integration of phage genome into that of the host bacterium. Transcript produced from *pRE* contains orf for *cI*; hence, activation of this promoter leads to production of CI. Thus, the action of CII on promoters *pI* and *pRE*

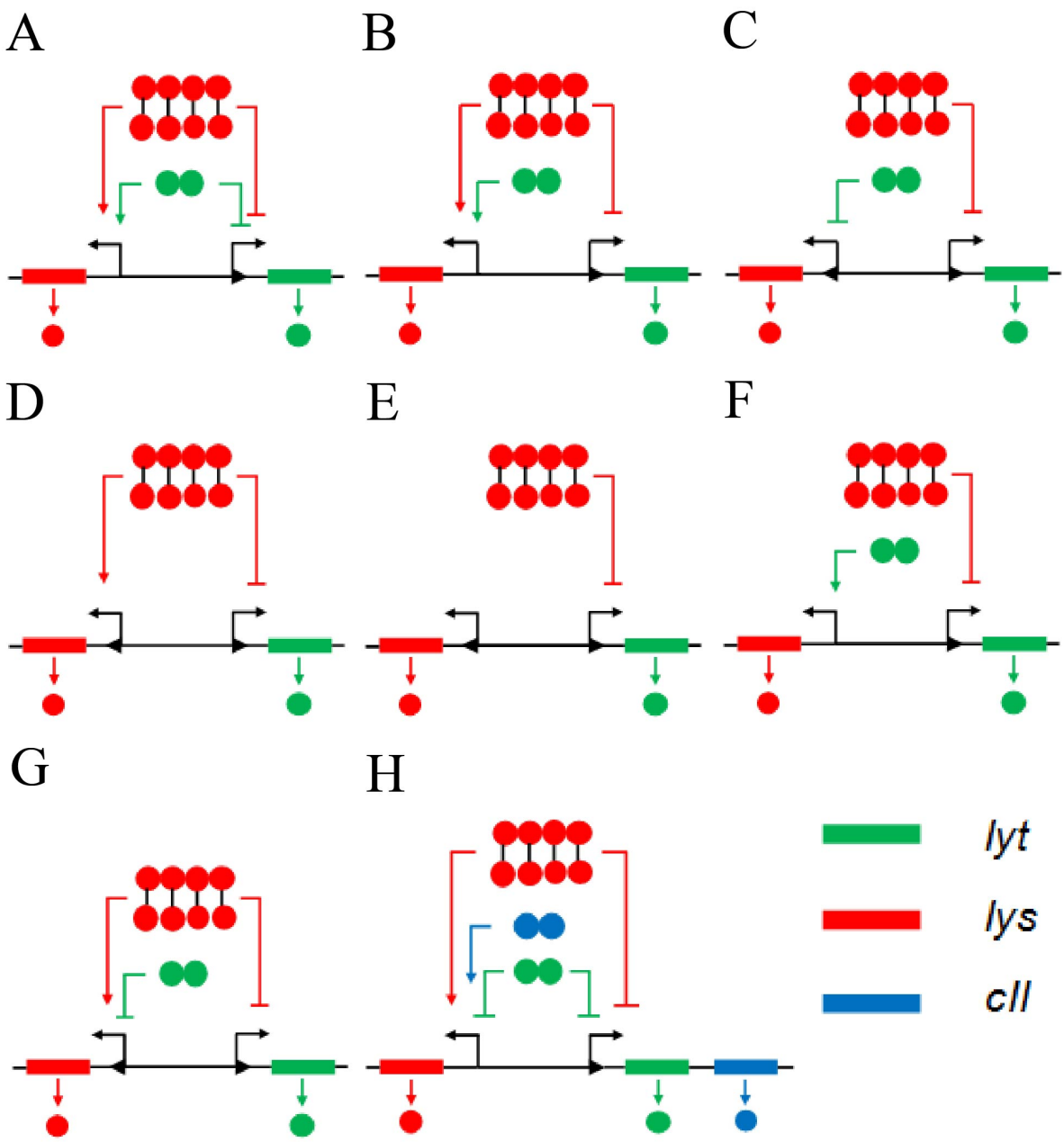

Figure 1: Various two-protein models, and three-protein model. Lower arrowhead represents basal expression. (**A**) The minimal model or 1A_Lyt_Lys. (**B**) Previous model with self-repression of *lyt* removed or 1B_Lyt_Lys. (**C**) Mutual repression or 2_Lyt_Lys. (**D**) 3_Lyt_Lys. (**E**) 4_Lyt_Lys.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

is functionally equivalent to Lyt protein activating transcription of *lys*. Notably, while the role of Cro in lambda's GRN is to inhibit the expression of *lys*, Cro-like protein (Lyt) activates the expression of *lys* in the 1A_Lyt_Lys.

CII inhibits lytic development by activating transcription from *pAQ*, which is located within *Q* gene in the opposite polarity. The transcript, thus produced, being antisense to (a part of) *Q* mRNA hybridizes with the latter, thereby preventing the translation of *Q* m-RNA, which is essential for lytic development [2]. Thus, the action of CII on promoter *pAQ* is functionally equivalent to Lyt protein inhibiting transcription of its own gene. If CII is not produced in sufficient amount, *Q* m-RNA is translated and anti-terminator Q, thus produced, causes lysis.

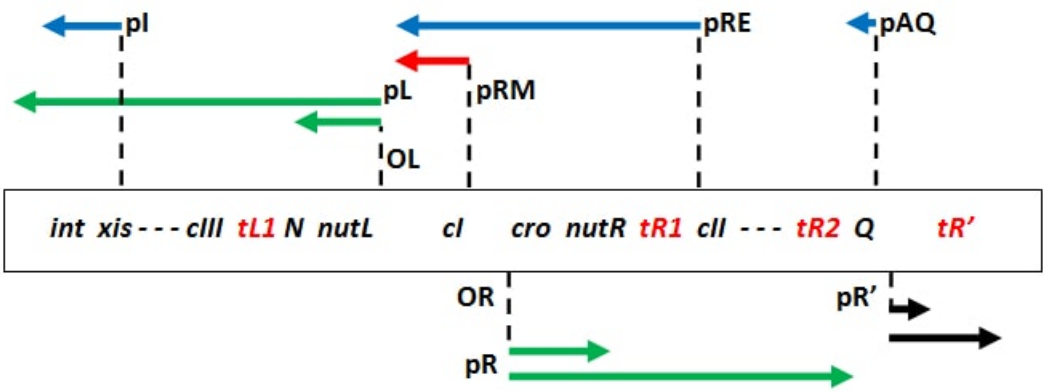

Figure 2: GRN and transcription map of lambda (adapted from Figure 1 of [8]). Transcripts that are produced earliest, viz. from *pL* and *pR* promoters, are depicted as green arrows. The late transcript, viz. from *pR'*, is a black arrow. Transcripts from CII-activated promoters, viz. *pI*, *pRE*, and *pAQ*, are shown as blue arrows. Transcript from *pRM*, which is activated by CI, is shown as 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 better demonstrate that the positive feedback underlies lysis/lysogeny switch, I considered variants of 1A_Lyt_Lys, mutual repression model, which doesn't have positive feedback loop, and its variants, and a model having the features of 1A_Lyt_Lys and mutual repression model. Since two features, viz. constitutive expression of *lyt* and its inhibition by Lys, are common, they would not be mentioned in the description of the models below. Since *cI* gene is positively regulated in lambda's GRN, *lys* has to have either basal expression or be activated

by Lyt. All of these models can be categorized in terms of three factors, as shown in the Table 1. First column shows whether *lys* possesses basal expression or is activated by Lyt. Second column shows if positive feedback, constituted by transcriptional activation of *lys* by its own product, is present. Third column shows if inhibition of *lys* by Lyt is present. Inhibition of *lys* by Lyt can only be present 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 self-inhibition of Lyt. The inhibition of *lyt*, required at MoI of 2, by its own product is dispensable, as Lys performs the same function, and more so, because at MoI of 2 Lyt's concentration is required to be much lower than that of Lys in order for switch to be of good quality. In terms of lambda's GRN, this would mean CI, instead of CII, activating transcription from *pAQ*.

$$\frac{dX}{dt} = \frac{mk_1(1 + \frac{X^a}{KD_1})}{1 + \frac{X^a}{K_{D1}} + \frac{Y^b}{K_{D2}}} - k_2 X \tag{3}$$

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

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.

$$\frac{dX}{dt} = \frac{mk_1(1 + \frac{X^a}{K_{D1}})}{1 + \frac{X^a}{K_{D1}} + \frac{Y^b}{K_{D2}}} - k_2 X \tag{13}$$

$$\frac{dY}{dt} = \frac{m(k_3 + k_4 \frac{Y^b}{K_{D2}})}{1 + \frac{X^a}{K_{D1}} + \frac{Y^b}{K_{D2}}} - k_5 Y \tag{14}$$

2.4. Deterministic simulation

The models were evaluated with respect to 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 dissociation constants would henceforth be referred to as a parameter set (That is, Hill coefficients are not part of parameter set). Since Cro forms dimer, Hill coefficient for Lyt's binding is taken to be 2; whereas, since CI forms tetramer, Hill coefficient for Lys' binding is taken to be 4. However, in the interest of completeness, another set of Hill coefficients, viz. a=2, b=2, was also considered. The rate constants and dissociation constants of equations defining a given model were searched (see Methods) in two stages: order search and linear search (as they are called here). Parameters search was guided by the quality of switch, viz. switch quotient (as it is called here), they generated. Switch quotient was initially considered to be determined by the expression

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

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

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

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

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

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

This expression (like the older one) varies between 0 and 1. Only those parameter sets were selected whose corresponding switch quotients (SQ) were positive.

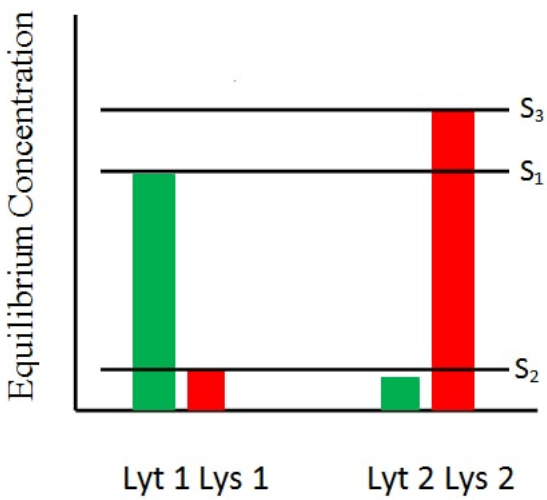

Figure 3: Schematic of switch's profile, viz. equilibrium concentrations of Lyt and Lys at the two MoIs.

As Table 2 shows, all of the models possessing the positive feedback loop have average deterministic switch quotient (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 have average DSQ of 0.5283 (highest DSQ was 0.6666); and, for that of a=2, b=4 all DSQs were more than 0.9 except for one parameter set, whose DSQ was 0.5. 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. Thus, if we exclude 2_Lyt_Lys for Hill coefficients' set of a=2, b=4 from the analysis, average DSQs of models with the positive feedback loop, viz. 1A_Lyt_Lys, 1A_Lyt(1)_Lys, 1B_Lyt_Lys, 3_Lyt_Lys, and 6_Lyt_Lys, were much higher than average DSQs of models without it, viz. 2_Lyt_Lys and 4_Lyt_Lys. Not just that, the lowest DSQ among the first set of models was much higher than the highest DSQ among the second set of models.

In the former set, Lys activating its own gene lets the value of Lys at MoI of 1 to be disproportionately lower for its desired particular value at MoI of 2. On the other hand, in 4_Lyt_Lys, since increase in genome copy number leads to proportional increase in the equilibrium activity of *lys* promoter, value of Lys at MoI of 1 would be half its value at MoI of 2. However, mutual repression model does generate many parameter sets with SQ greater than 0.9 for Hill coefficients' set of a=2, b=4. Since this model exhibits very different behaviour in the stochastic simulations, it will be discussed further in the section for stochastic simulations.

The model 5_Lyt_Lys did not generate any parameter set. The reason is that in the absence of the positive feedback loop, *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. Equivalently, the desired high concentration of Lyt at MoI of 1, also leads to excessive production of Lys at the same MoI. Thus, both proteins are present in similar amounts at both MoIs. In hindsight, one notes that the equations for Lyt and Lys are almost identical for this model.

In order to examine the significance of cooperativity in positive feedback here, another set of Hill coefficients, viz. a=2, b=1, was also considered for 1A_Lyt_Lys. However, parameter sets generated by this set gave DSQs which were almost equal to zero. For models having the positive feedback loop, average DSQ of parameter sets was very slightly, almost negligibly, greater for Hill coefficients' set of a=2, b=2 than that for set of a=2, b=4.

Table 2: Average and SD of SQs under deterministic and stochastic conditions for various thresholds of stochastic success rate.

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	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)
2_Lyt_Lys	0.5283 (0.0696)	0.8917 (0.1766)	none	0.5001 (N/A)	none	0.2725 (N/A)
3_Lyt_Lys	0.9938 (0.0078)	0.9873 (0.0157)	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	0.9988 (0.0004)	0.9876 (0.0135)	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

[a] SSR = Stochastic Success Rate

Table 3: Average and SD of SQs under deterministic and stochastic conditions for various thresholds of stochastic success rate.

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	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)
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

[a] SSR = Stochastic Success Rate

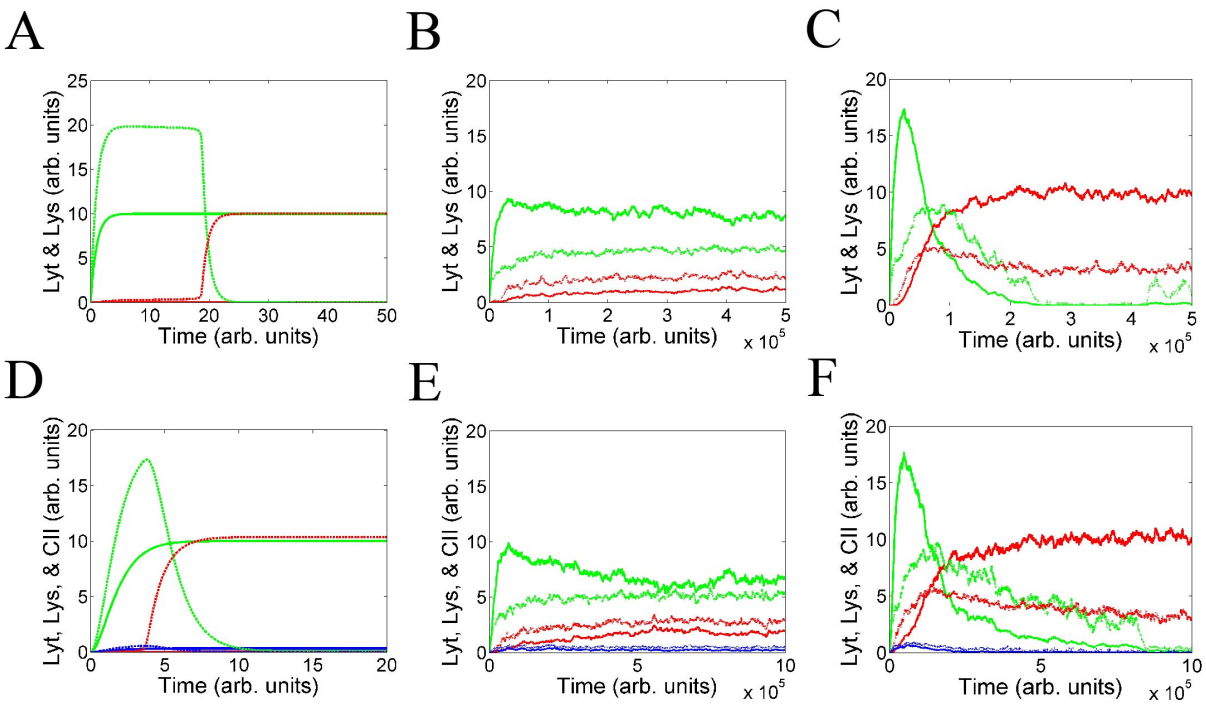

Figure 4: Deterministic and stochastic simulations of the minimal two-protein model (1A_Lyt_Lys) and three-protein model (Lyt_Lys_CII). Lyt, Lys, and CII are represented by green, red, and blue, respectively. For deterministic simulations, concentrations of proteins at MoI of 1 and 2 are depicted by solid curve and dashed curve, respectively. For 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 had maximum stochastic success rate was used for simulation. The stochastic simulation trajectories shown here are qualitatively similar to those of all other models for parameter sets with high stochastic success rate; whereas, the deterministic simulation trajectories were so, irrespective of stochastic success rate. In stochastic simulation graphs, the original abscissa, which had unequally spaced time intervals, was converted to one with equally spaced time intervals. Each (arb.) 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, initially, the concentration of Lyt becomes more than its equilibrium concentration at MoI of 1 but then comes back to very low level. It is due to double initial rate of production of Lyt at MoI of 2 as compared to that at MoI of 1; however, as Lyt's concentration increases, *lys*' transcription becomes stronger, leading to production of Lys, which in turn 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, initially, concentrations of Lyt and CII become more than their respective equilibrium concentrations at MoI of 1 but then come back to very low levels. This was also observed for a three-protein model, which is very similar to that of this paper, in a theoretical study [5]. Analogous to the two-protein model, it's due to heightened initial rate of production of CII at MoI of 2 as compared to that at MoI of 1; however, as CII's concentration increases, transcription of *lys* becomes stronger, leading to production of Lys, which represses *lyt* and *cI*. (**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 [2].

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

In order to further verify if 1A_Lyt_Lys represents reduced form of lambda's GRN, I consider a three-protein simplified version of lambda switch and show that it is equivalent to a two-protein model possessing the positive feedback loop: 1B_Lyt_Lys. A CII-like protein is added to 1A_Lyt_Lys beside extending the role of Lyt. Since genes *lyt* and *cII* are under the control of same promoter, in order to allow for potentially different rates of translation of their corresponding cistrons during stochastic simulations, their mRNAs are considered explicitly. The role of Lyt in this model is identical to that of Cro in lambda phage's GRN. That is, now Lyt represses transcription of *lys*, in addition to repressing that of its own gene. As in the two-protein models, the function of Lys here is identical to that of CI in the lambda phage's GRN. The role of CII in the minimal three-protein model is to activate transcription of *lys*. This corresponds to CII's activation of *pRE* promoter of lambda's GRN, leading to synthesis of mRNA which contains orf for *cI*. The three-protein model considered here is different from that in [5], in which CII activates transcription of *cI* from a distinct (*pRE*) promoter. Since in the three-protein model, CII has to compete with Lyt, which represses transcription of *lys*, for binding to the intergenic region, the demonstration of equivalence of the three-protein model (Lyt_Lys_CII) with 1A_Lyt_Lys, or any of its variants, gets more challenging. The degradation constants for xz (concentration of *lyt-cII* mRNA) , X (concentration of Lyt), Z (concentration of CII), and Y (concentration of Lys), viz. k_6, k_7, k_9, k_8, respectively, are taken to be unity for the same reason why degradation constants for two-protein models were set equal to 1. Since for 1A_Lyt_Lys SQs generated by Hill coefficients' set of a=2, b=2 were as high as SQs generated by that of a=2, b=4, applying occam's razor, Hill coefficients for binding of Lyt and Lys are taken to be 2 and 2, respectively, not 2 and 4. Further, taking lead from here, Hill coefficient for CII's binding is taken to be 2, even though it has been shown to exist as tetramer in solution [14] and in crystallized free and DNA-bound state [15].

Model equations for three-protein model are as follows.

Transcription of *lyt-cII* genes:

$$\frac{dxz}{dt} = \frac{mk_1(1 + \frac{Z^c}{K_{D3}})}{1 + \frac{X^a}{K_{D1}} + \frac{Y^b}{K_{D2}} + \frac{Z^c}{K_{D3}}} - k_6 xz \tag{15}$$

Translation of *lyt*:

$$\frac{dX}{dt} = k_2 xz - k_7 X \tag{16}$$

Translation of *cII*:

$$\frac{dZ}{dt} = k_4 xz - k_9 Z \tag{17}$$

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 the Hill coefficient of CII's binding, k_1 is basal expression rate of *lyt-cII* genes, K_{D3} is the ”combined” 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.

Equilibrium values of xz, X, Z, and Y are

$$k_6 \overline{xz} = \frac{mk_1(1 + \frac{\overline{Z}^a}{K_{D3}})}{1 + \frac{\overline{X}^a}{K_{D1}} + \frac{\overline{Y}^b}{K_{D2}} + \frac{\overline{Z}^a}{K_{D3}}} \tag{19}$$

$$k_7 \overline{X} = k_2 \overline{xz} \tag{20}$$

$$k_9 \overline{Z} = k_4 \overline{xz} \tag{21}$$

$$k_8\overline{Y} = \frac{m(k_5\frac{\overline{Y}^b}{K_{D2}} + k_3\frac{\overline{Z}^a}{K_{D3}})}{1 + \frac{\overline{X}^a}{K_{D1}} + \frac{\overline{Y}^b}{K_{D2}} + \frac{\overline{Z}^a}{K_{D3}}} \tag{22}$$

From (20) and (21), it can be seen that equilibrium value of CII is in constant proportion to that of Lyt. Hence, CII can be written in terms of Lyt

$$\overline{Z} = p\overline{X} \tag{23}$$

where

$$p = \frac{k_4k_7}{k_2k_9}$$

Using (20) and (23), (19) and (22) can be written as

$$\overline{X} = \frac{m\frac{k_1k_2}{k_6k_7}(1 + \frac{(p\overline{X})^a}{K_{D3}})}{1 + \overline{X}^a(\frac{1}{K_{D1}} + \frac{p^a}{K_{D3}}) + \frac{\overline{Y}^b}{K_{D2}}} \tag{24}$$

$$\overline{Y} = \frac{m\frac{1}{k_8}(k_5\frac{\overline{Y}^b}{K_{D2}} + k_3\frac{(p\overline{X})^a}{K_{D3}})}{1 + \overline{X}^a(\frac{1}{K_{D1}} + \frac{p^a}{K_{D3}}) + \frac{\overline{Y}^b}{K_{D2}}} \tag{25}$$

The equivalence of equations (24) and (25) to the defining equations of 1B_Lyt_Lys which have reached equilibrium validates two-protein model. Two-protein model being sufficient for producing lysis/lysogeny switch constitutes an argument that *cro* in lambda's GRN is expendable. Mathematically, the reason for Cro being expendable lies in its equilibrium concentration being proportional to that of CII.

Kobiler et al. [2] showed that infection with lambda lacking cro gene (λcro$^-$) leads to production of CII to level sufficient to cause lysogeny even at MoI of 1. This, however, does not mean that Cro, per se, is required to engender lytic development. Cro represses *pL* and

pR by fourfold and twofold, respectively [12]. Thus, the absence of Cro increases the level of CII in two ways: first, by allowing transcription of *cII*, which is under the control of *pR*, and *cIII*, which is under the control of *pL* and whose product prevents degradation of CII by protease HflB. In the wild type strain, parameters associated with transcription rates of *cII* and *cIII*, translation and degradation rates of their respective mRNAs, and degradation rates of CII and CIII are such that enough CII is produced, despite Cro's repression of *pL* and *pR*, at higher MoIs so as to sufficiently activate *pRE* promoter, leading to production of CI to level which is enough to cause lysogeny. However, when *cro* is deleted, CI produced even at MoI of 1 is enough to engender lysogeny. With appropriate changes in the aforementioned parameters, it would be possible to model λcro$^-$ strain which behaves like its wild type counterpart.

As stated above, there are experimental evidences for CII present as tetramer in solution [14] and in crystallized free and DNA-bound state [15]. Additionally, as Figure 4 in [10] shows, the binding curve of CII to *pAQ* has appreciable lag phase, indicating that it binds as a multimer. However, Figure 2c in [2] shows that curve of *pRE*'s activity with respect to CII levels is not sigmoidal as expected from multimeric binding, but hyperbolic as seen in monomeric binding. Therefore, another model was considered where Hill coefficient for CII binding was taken to be 1 (Lyt_Lys_CII(1)). Additionally, one more model was considered where Hill coefficient for Lyt too was taken to be 1 (Lyt(1)_Lys_CII(1)). This made the current author go back to two-protein models and consider 1A_Lyt_Lys model too with Hill coefficients' set of a=1, b=2 and a=1, b=4, named 1A_Lyt(1)_Lys. SQs generated by all new variants were similar in values to those generated from their counterparts, where Hill coefficient of either Lyt or CII, or both, were taken to be 2. Specifically, for 1A_Lyt(1)_Lys all SQs were more than 0.98 for both sets of Hill coefficients. For all of the three protein models, all SQs were greater than 0.95. Just like the Hill coefficients' set of a=2, b=1, parameter sets generated by the set of a=1, b=1 gave SQs which were almost equal to zero.

2.6. Stochastic simulation

Since gene expression is stochastic [17,18], the true validity of results obtained in the deterministic simulations lie in their being replicated in the stochastic simulations. Thus, stochastic simulations were performed, using Gillespie algorithm [19], for parameter sets obtained in the deterministic simulations.

For both two-protein and three-protein models, for any given parameter set, SQ generated in the stochastic simulation, or stochastic switch quotient (SSQ), was less than its deterministic counterpart, viz. DSQ. No parameter set was able to produce switch in every run during stochastic simulation. That is because either the SSQ was negative ($S_1 < S_2$) or, rarely, S_3 was zero. Percentage of runs that produce finite, positive SQs during stochastic simulation for a given parameter set and set of Hill coefficients would henceforth be referred to as stochastic success rate. For two-protein and three-protein models, 500 and 200 simulations, respectively, were performed, for a given parameter set and set of Hill coefficients.

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	21	17
1A_Lyt(1)_Lys	0	0	1	1	3	9	17	15
1B_Lyt_Lys	1	1	1	5	2	2	11	11
2_Lyt_Lys	0	1	0	0	0	0	6	6
3_Lyt_Lys	0	0	0	0	1	4	8	10
4_Lyt_Lys	0	2	0	0	0	0	2	2
6_Lyt_Lys	0	0	0	0	0	0	12	12
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

[a] SSR = Stochastic Success Rate

An interesting property was observed for mutual repression model for Hill coefficients' set of a=2, b=4. It was the only set of Hill coefficients for any model lacking the positive feedback that produced a DSQ more than 0.9 (highest SQ for the same model for Hill coefficients' set of

a=2, b=2 was 0.6666). As aforementioned, all of the parameter sets for Hill coefficients' set of a=2, b=4 produced DSQ of more than 0.9 except one, whose DSQ was 0.5. Notably, this is the parameter set which had very high stochastic success rate, viz. that of 97%; while, maximum stochastic success rate among other parameter sets was 50%. This peculiar result for mutual repression has been reported earlier also.

Avlund et al. showed that various two-protein models, based upon mutual repression model, which were able to produce switch in a noise-less environment, did not function when noise was introduced [9]. However, additional CII-like protein conferred robustness to noise in 8% of the parameter sets that produced switch deterministically. The different behaviour of mutual repression model in deterministic simulations with respect to stochastic simulations warrants theoretical investigation. Notably, one of their rare two-protein models (i.e., b of Figure 2) which did produce switch 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 current paper.

Thus, taking into account stochastic success rate of at least 95%, two-protein models can be divided into two sets based upon DSQs or SSQs. One set comprises of two models with the positive feedback loop, viz. 1A_Lyt_Lys and 1B_Lyt_Lys, and another without it, viz. 2_Lyt_Lys and 4_Lyt_Lys. The one with the positive feedback loop has appreciably higher DSQs and SSQs than the one without it. In fact, the lowest DSQ and SSQ among the first set of models were much higher than the highest DSQ and SSQ, respectively, among the second set of models. The comparison could not be made for stochastic success rate's range of less than 95% and greater than or equal to 90% because neither 2_Lyt_Lys nor 4_Lyt_Lys produced switch.

However, for the same two thresholds of stochastic success rate, average SSQs for parameter sets with Hill coefficients' set of a=2, b=2 were greater than average SSQs for those with Hill coefficients' set of a=2, b=4, for any given model (Table 2 and Table 3). Not just that, the lowest SSQ among parameter sets with Hill coefficients' set of a=2, b=2 was greater than the highest SSQ among those with Hill coefficients' set of a=2, b=4, for any given model. This result is against one's expectation: since Lys activating transcription of its own gene in a cooperative manner is crux of the switch, increasing Hill coefficient of Lys should have, if at all, increased

the SSQ. This comparison could not be made in models without the positive feedback loop because none of their parameter sets with Hill coefficients' set of a=2, b=2 had stochastic success rate of at least 90%.

Table 5: Maximum stochastic success rate.

Model	Hill coefficients' set	Maximum stochastic success rate
1A_Lyt_Lys	a=2, b=4	96.8
1A_Lyt(1)_Lys	a=1, b=2	91
1B_Lyt_Lys	a=2, b=4	97.2
2_Lyt_Lys	a=2, b=4	97
3_Lyt_Lys	a=2, b=4	87
4_Lyt_Lys	a=2, b=4	98.8
6_Lyt_Lys	a=2, b=4	73
Lyt_Lys_CII	a=2, b=2, c=2	95.5
Lyt_Lys_CII(1)	a=2, b=2, c=1	97
Lyt(1)_Lys_CII(1)	a=1, b=2, c=1	93.5

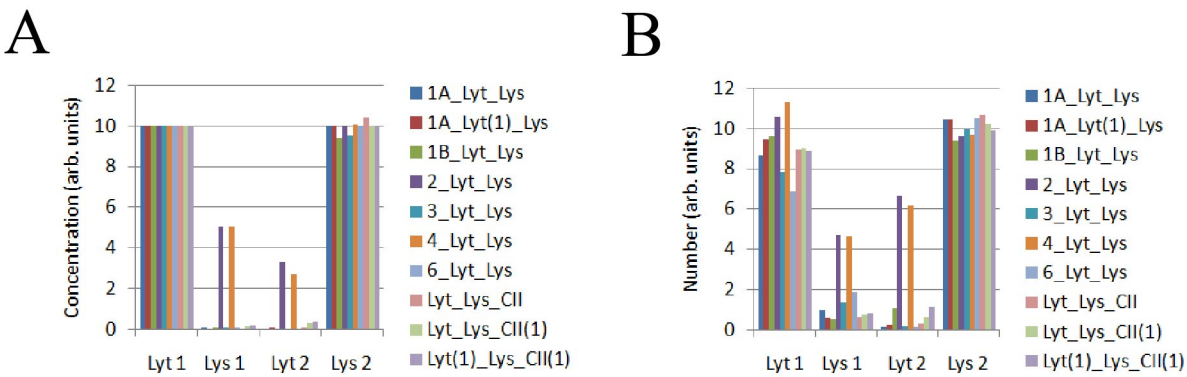

Figure 5: Equilibrium values correspond to those parameter sets which gave maximum stochastic success rate for their respective models (see Table 5). (**A**) Deterministic simulations. (**B**) Stochastic simulations. Note how the values of Lys at MoI of 1 and Lyt at MoI of 2 for 2_Lyt_Lys and 4_Lyt_Lys are much higher than those of the any other model.

2.7. Bistability at MoI of 1 and lysogen stability

In this study, parameter sets were searched for their ability to cause lysis at MoI of 1 and lysogeny at MoI of 2. However, if only one of the phage genomes gets integrated into the

bacterial chromosome, it would not be able to maintain lysogeny, and lysis would ensue, if only one stable state existed at MoI of 1. In the deterministic simulations, all of the two-protein models possessing the positive feedback exhibited bistability at MoI of 1 for all of the parameter sets, except one (for 1B_Lyt_Lys). In the other stable state, the concentration of Lyt is almost zero and that of Lys is about half of its concentration at MoI of 2. Arguably, in lambda's system, the level of Lys in the second stable state would be high enough to maintain lysogeny.

For 4_Lyt_Lys, none of the parameter sets produced bistability at MoI of 1. For 2_Lyt_Lys, for Hill coefficient's set of a=2, b=2 one parameter set generated bistability at MoI of 1, but its stochastic success rate was just 7.6% (Bistability exists for two more parameter sets, but their second stable states are at very high values of Lyt (>50) and very low values of Lys (<2); hence, inconsequential for lysogeny maintenance, and in any case, never reached by the phase point). For Hill coefficients set of a=2, b=4, the only parameter set which did not exhibit bistability at MoI of 1 had stochastic success rate of 97%, while maximum stochastic success rate among other parameter sets was 50% (as aforementioned in the section for stochastic simulations). All of the three-protein models exhibited bistability at MoI of 1. The Lyt and Lys values of second stable states at MoI of 1 in three-protein models are about same as those of second stable states in two-protein models at the said MoI.

DNA between OL and OR sites forms a loop that has been shown to be important for the stable maintenance of lysogeny [12]. The loop forms due to interaction between CI dimers bound at OL1 and OL2 with those bound at OR1 and OR2 [13]. Therefore, the contribution of OL-CI-OR complex to production of CI would be represented by adding a term proportional to [CI], raised to the power 8, to numerator and denominator. Since bistability at MoI of 1 in the two-protein models is the consequence of *lys*' transcription getting activated by its own product in a cooperative manner (i.e., by the binding of Lys dimer), in lambda's GRN, activation of *cI* promoter when present in looped DNA, stabilized by CI octamer, would either generate bistability or contribute to already existing bistability due to two CI dimers activating the transcription of *cI*. Thus, it is reasonable to propose that the role of OL_CI_OR loop formation is to produce or strengthen bistability at MoI of 1. This argument becomes stronger in the light

of the finding that looping also activates transcription from *pRM* by allowing the α-CTD of RNAP bound at *pRM* to contact UP element at OL [16]. The heightened rate of transcription from *pRM* when present in looped DNA would also lead to higher equilibrium concentration of CI in the other stable state, thereby enabling better maintenance of lysogeny.

During prophage induction, Lys undergoes autocleavage, facilitated by activated RecA coprotease [20], which results into removal of repression of pL and pR and lytic development ensues. As can be seen in the phase diagram (Figure 6), Lys' concentration should become extremely low for prophage induction to occur, viz. phase point reaching fixed point representing lysis, thereby making this process hard to achieve. This result was seen for all of the paramter sets exhibiting bistablilty. This could be explained away by simply noting that the criterion of parameter selection here does not include the process of prophage induction. That would have demanded the threshold of Lys' concentration for induction to be neither too high, so that lysogen becomes unstable, nor too low, so that induction becomes difficult. In the stochastic simulations, however, none of the two-protein and three-protein models produced bistability at MoI of 1.

At MoI of 2, only two models, viz. 2_Lyt_Lys and 6_Lyt_Lys, show bistability for about 80% and 60%, respectively, of the parameter sets. Notably, only these two models have Lyt repressing the transcription of *lys*. Since at the second stable state the concentration of Lyt is very high and that of Lys is very low, a parameter set would not, if at all, generate good quality switch if its phase point reached this stable fixed point. Hence, bistability at MoI of 2 is inconsequential.

2.8. Why positive feedback?

There can be two reasons why lysis/lysogeny switch is based upon the positive feedback: 1) biological properties of the switch, viz. a) highest switch quotient and b) presence of bistability at MoI of 1, and 2) quickest evolution of such a model. It should be noted, however, that speed of evolution would not matter if evolution is path-independent. That is, it's possible that nature

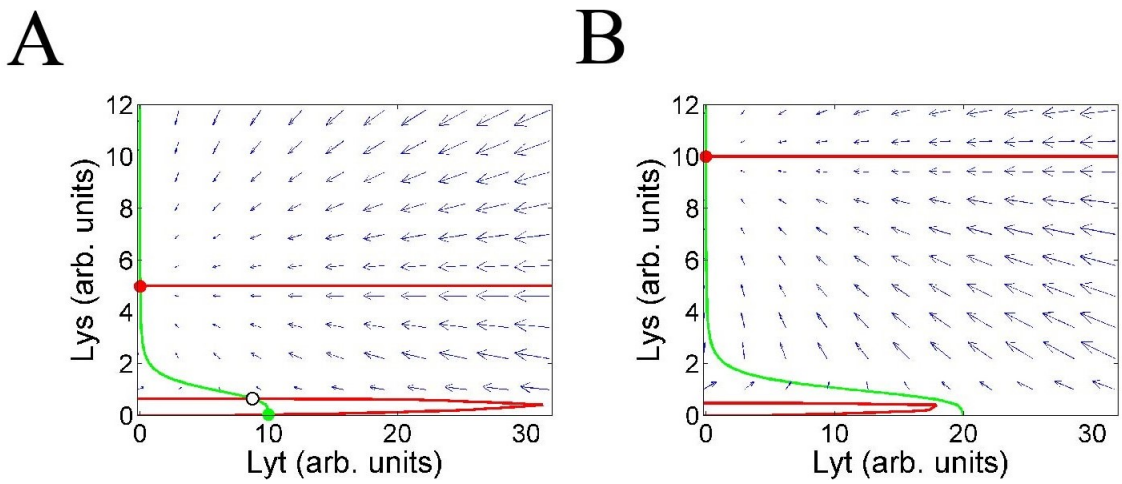

Figure 6: (**A-B**)Phase diagram of 1A_Lyt_Lys corrresponding to the parameter set that gave maximum stochastic success rate, at Mo1 of 1 and 2. Green and red full circles are stable fixed points, whereas empty black circle is unstable fixed point. Green stable point is where system reaches when a single phage infects a bacterium. Red stable point is where system reaches when lysogeny is established by two phages, but only one of them gets integrated into the host's genome.

initially evolves a sub-optimal design but which, given enough time, gets superseded by an optimal one.

1a) Switch quotient: As mentioned in the previous sections, SQs generated in the deterministic and the stochastic simulations, respectively, for models possessing the positive feedback are much greater than those of the models lacking positive feedback.

1b) Bistability at MoI of 1: As stated in the last section, for models not possessing the positive feedback loop, no parameter set, if at all, having sufficiently good stochastic success rate generated bistability. If one ignores the possibility of any other mechanism generating bistability, such as the formation of OL-CI-OR complex, this reason alone is sufficient for nature to choose models which possess the positive feedback loop over those which do not.

2) Speed of evolution: Even though the maximum stochastic success rate is very low for 3_Lyt_Lys and (especially) 6_Lyt_Lys, they are still compared with 4_Lyt_Lys and 2_Lyt_Lys, respectively, as these two are the only pairs within which mathematical comparison with regard to the positive feedback loop is possible. 2_Lyt_Lys and 4_Lyt_Lys differ from 6_Lyt_Lys and 3_Lyt_Lys, respectively, only in not having the positive feedback loop. Thus, model equations

of former two models differ from those of latter two only in the dynamics of Lys. In models with the positive feedback loop, the term representing binding of Lys to the intergenic region (i.e., Y^b/K_{D2}) is multiplied by rate constant for transcriptional activation of *lys* by Lys, k_4. On the other hand, in models without the positive feedback loop Y^b/K_{D2} is multiplied by k_3, the basal expression rate of *lys*. Thus, 2_Lyt_Lys and 4_Lyt_Lys can be thought of as being equivalent to 6_Lyt_Lys and 3_Lyt_Lys, respectively, whose k_4 is equal to k_3. That is, the former two models are those latter two models, respectively, whose rate constant for transcriptional activation of *lys* by Lys is equal to the basal expression rate of *lys*. This constrain of having $k_3 = k_4$ reduces the potential parameter space for 2_Lyt_Lys and 4_Lyt_Lys by one dimension. Hence, the two parameters being independent in 3_Lyt_Lys and 6_Lyt_Lys makes nature more likely to discover them. This explains why 2_Lyt_Lys (11, 11) and 4_Lyt_Lys (2, 2) produced fewer parameter sets than 6_Lyt_Lys (16, 16) and 3_Lyt_Lys (11, 13), respectively, for both sets of Hill coefficients during the order search (as shown in the parenthesis).

Now, qualitative equivalence of 3_Lyt_Lys and 6_Lyt_Lys with 1B_Lyt_Lys, which is equivalent to 1A_Lyt_Lys, is shown. 1B_Lyt_Lys is qualitatively equivalent to 3_Lyt_Lys for the reason that in the former, transcriptional activation of *lys*' is achieved by binding of Lyt to its promoter; whereas, in the latter, *lys* possesses basal expression. 6_Lyt_Lys differs from 3_Lyt_Lys in having Lyt as a repressor of *lys*. This interaction is expendable, as at MoI of 2, the concentration of Lyt is anyway very low, and, qualitatively speaking, at MoI of 1, the repression of *lys* by Lyt can be compensated by reducing basal expression of *lys*. As Table 6 shows, for a given set of Hill coefficients and range of stochastic success rate, average k_3 is higher (except being equal on one occasion) for 6_Lyt_Lys than that for 3_Lyt_Lys, for those cases where both models produce switch at least for one parameter set. It should be noted that the comparison is not mathematical but only qualitative.

Table 6: Average and SD of k_3 for various thresholds of stochastic success rate.

Design	AVG k_3 (SD) (SSR[a] $\geqslant$ 60)		AVG k_3 (SD) (60 >SSR $\geqslant$ 50)		AVG k_3 (SD) (50 >SSR $\geqslant$ 40)	
	a=2, b=2	a=2, b=4	a=2, b=2	a=2, b=4	a=2, b=2	a=2, b=4
3_Lyt_Lys	0.0074 (0.0026)	0.0478 (0.0234)	0.0037 (N/A)	none	0.0031 (0.0005)	none
6_Lyt_Lys	none	0.1165 (0.0882)	0.0037 ($\approx$ 0)	6.9641 (9.2816)	0.0711 (0.0678)	4.5459 (4.2041)

[a] SSR = Stochastic Success Rate

3. Methods

3.1. Derivation of model equations

The model, using the fact that binding of protein to itself or DNA is a much quicker process than transcription and translation, assumes quick equilibration for the processes of protein binding to itself or DNA, in order to calculate the ”combined” dissociation constants of proteins. In the expressions below, P, X, and Y are promoter, Lys, and Lyt, respectively.

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

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

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

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

$$P + X_2 \overset{K'_{D1}}{\rightleftharpoons} PX_2 \qquad \frac{[P][X^2]}{[PX_2]} = K'_{D1}K_{DX} = K_{D1}$$

$$P + Y_4 \overset{K'_{D2}}{\rightleftharpoons} PY_4 \qquad \frac{[P][Y^4]}{[PY_4]} = K'_{D2}K^2_{DY1}K_{DY2} = K_{D2}$$

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, viz. monomeric and dimeric Lys, and monomeric and dimeric CII.

Processes of transcription and translation are not considered explicity except for *lyt-CII* genes in the three-protein models. Hence, the model equations describe concentrations of proteins only. With expressions for concentrations of promoter-protein complexes, one can write generalized form of term representing protein production.

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

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

3.2. Parameter search

Parameter sets, viz. rate constants and dissociation constants, of model equations were searched deterministically in two stages, viz. order search and linear search (as they are named here). In the order search, rate constants and dissociation constants were searched as 3's exponent, which was varied between -5 and 5 with the difference of 1, in a nested fashion. Thus, the number of parameter sets searched was equal to the number of parameters raised to the power 11. Notably, switch quotients generated by this approach are unrefined because rate constants and dissociation constants were increased geometrically, thereby causing a lot of intervening values to remain unsampled. Therefore, parameter sets generated from order search were further refined by linear search, which searches the neighbourhood of parameter set arithmetically. It

was noted that those parameter sets generated in the order search whose SQs were too close to each other were either rescaled form of each other, or differed in those parameters to which SQ was resilient up to a certain range. Thus, in order to remove redundancy and in the interest of time, for linear search, the parameter sets were taken in such a way that the difference between consecutive SQs is at least 0.01.

Parameter sets, and thus accompanied SQs, generated through order search were refined by linear search in the following way. The value of each parameter (say, V) of a set was varied between -3*V/5 and 3*V/5 with the increment of V/5, in a nested fashion. Thus, the number of parameter sets searched was equal to the number of parameters raised to the power 7. However, for three-protein model, which had eight parameters, in the interest of saving time, each parameter was varied between -2*V/5 and 2*V/5 with the increment of V/5, in a nested fashion. Search was ended if the latest SQ was either lower than the previous one (which never happened) or if ((latest SQ - previous SQ)/previous SQ) was less than 0.01. Again, in the interest of saving time, for three-protein model, the search was ended if the SQ at the end of the last iteration was more than or equal to 0.95. It should be noted that linear search is path dependent: it may happen that a path which initially yields lower SQs leads to higher SQ in the end than a path which initially yields higher SQs, and thus, treaded by the search. For both order and linear search and for all of the models, in order to expedite search, those parameter sets were rejected whose accompanying SQ was lower than the SQ of the previous parameter set. The values of the parameters were normalized such that the Lyt's equilibrium concentration was 10 arb. units. This was done for two purposes: a) to ensure that lowest values of Lyt at MoI of 1 and Lys at MoI of 2 never drop to zero in the stochastic simulations; b) in order to make comparison of parameter sets and equilibrium values of proteins visually easier. For both order and linear search, simulations were carried for time 100 arb. units. Thus, there was a possibility of a system of equations, defining a particular model, not reaching equilibrium in 100 arb. units for a given parameter set. In order to eliminate such parameter sets, simulations were done for 10^5 arb. units. Only few parameter sets had not reached equilibrium, and all of such parameter sets produced negative SQ. In order to calculate stochastic switch quotie nt,

levels of proteins were averaged between 100 and 200 arb. units. The transient kinetics, viz. inital rise and plateauing at MoI of 1 and bell-shaped trajectory MoI 2, were completed at most by 50 arb. units.

References

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

[2] 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.

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

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

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

[6] 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.

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

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

[9] 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.

[10] 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.

[11] 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.

[12] 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.

[13] Révet B, von Wilcken-Bergmann B, Bessert H, Barker A, Mller-Hill B Four dimers of lambda repressor bound to two suitably spaced pairs of lambda operators form octamers and DNA loops over large distances. Current Biology. 1999; 9: 151–154.

[14] 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.

[15] 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.

[16] Cui L, Murchland I, Shearwin KE, Dodd IB Enhancer-like long-range transcriptional activation by CI-mediated DNA looping. Proc Natl Acad Sci U S A. 2013; 110: 2922–2927.

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

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

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

[20] Little JW, Mount DW The SOS regulatory system of Escherichia coli. Cell. 1982; 29: 11–22.