Lysogen stability is determined by the frequency of activity bursts from the fate-determining gene

Chenghang Zong1,5,*, Lok-hang So1,2, Leonardo A Sepúlveda2,3, Samuel O Skinner1,2 and Ido Golding1,2,3,4,*

1 Department of Physics, University of Illinois, Urbana, IL, USA, 2 Verna and Marrs McLean Department of Biochemistry and Molecular Biology, Baylor College of Medicine, Houston, TX, USA, 3 Center for Biophysics and Computational Biology, University of Illinois, Urbana, IL, USA and 4 Center for the Physics of Living Cells, University of Illinois, Urbana, IL, USA
5 Present address: Department of Chemistry and Chemical Biology, Harvard University, Cambridge, MA 02138, USA

* Corresponding authors. C Zong, Department of Chemistry and Chemical Biology, Harvard University, Cambridge, MA 02138, USA. Tel.: +1 617 496 8654; Fax: +1 617 496 8709; E-mail: czong@fas.harvard.edu or I Golding, Verna and Marrs McLean Department of Biochemistry and Molecular Biology, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA. Tel.: +1 713 798 4198; Fax: +1 713 796 9438; E-mail: igolding@illinois.edu

Received 6.7.10; accepted 19.10.10

The ability of living cells to maintain an inheritable memory of their gene-expression state is key to cellular differentiation. Bacterial lysogeny serves as a simple paradigm for long-term cellular memory. In this study, we address the following question: in the absence of external perturbation, how long will a cell stay in the lysogenic state before spontaneously switching away from that state? We show by direct measurement that lysogen stability exhibits a simple exponential dependence on the frequency of activity bursts from the fate-determining gene, cI. We quantify these gene-activity bursts using single-molecule-resolution mRNA measurements in individual cells, analyzed using a stochastic mathematical model of the gene-network kinetics. The quantitative relation between stability and gene activity is independent of the fine details of gene regulation, suggesting that a quantitative prediction of cell-state stability may also be possible in more complex systems. Molecular Systems Biology 6: 440; published online 30 November 2010; doi:10.1038/msb.2010.96

Subject Categories: simulation and data analysis; chromatin & transcription

Keywords: epigenetic stability; lysis; lysogeny; phage lambda; stochastic gene expression

This is an open-access article distributed under the terms of the Creative Commons Attribution Noncommercial No Derivative Works 3.0 Unported License, which permits distribution and reproduction in any medium, provided the original author and source are credited. This license does not permit commercial exploitation or the creation of derivative works without specific permission.

Introduction

The ability of living cells to maintain an inheritable memory of their gene-expression state is key to cellular differentiation (Monod and Jacob, 1961; Slack, 1991; Lawrence, 1992). A differentiated cellular state may be maintained for a long time, while at the same time allowing efficient state-switching (‘reprogramming’) in response to the proper stimulus (Gurdon and Melton, 2008). However, even in the absence of external perturbation, a cell’s gene-expression state may not be ‘infinitely stable’ (irreversible; Lawrence, 1992). This is a consequence of the stochastic nature of all cellular reactions (Acar et al., 2005; Maheshri and O’Shea, 2007; Raj et al., 2008), which shift individual cells away from the ‘average state’, and in particular may switch a cell from one state to another. A natural question then arises: how stable is a cell’s gene-expression state, in the absence of an external perturbation? In other words, how long will a differentiated cell stay in the same state before spontaneously switching to an alternative one? What features of the underlying gene-regulatory network determine this stability?

The lysogeny state of an Escherichia coli cell harboring a dormant bacteriophage (prophage) lambda serves as one of the simplest examples for a stable cellular state (Ptashne, 2004, 2007; Oppenheim et al., 2005). Lysogeny is maintained by the activity of a single protein species, the lambda repressor (CI), which acts as a transcription factor to repress all lytic functions from the prophage in the E. coli cell, as well as to regulate its own production (Figure 1A–C; Ptashne, 2004). This feature of auto-regulation by the fate-determining proteins is commonly observed in systems displaying long-term cellular memory (Lawrence, 1992; Gurdon and Melton, 2008; Crews and Pearson, 2009). The lambda lysogeny system exhibits extremely high stability: spontaneous switching events occur less than once per 108 cell generations in the absence of cellular RecA activity (Little et al., 1999). At the same time, this genetic switch also exhibits fast and efficient switching in response to the appropriate stimulus, for example, damage to the bacterial genome (Oppenheim et al., 2005).

The lambda system has been well characterized in terms of the regulatory circuitry that creates the stable lysogenic state. Specifically, the regulation of the two key promoters, P<sub>RAM</sub> and P<sub>RM</sub>,
The results of our stochastic simulation (red curve) are also shown for comparison. For experimental details see Materials and methods.

The system is wild-type lysogen MG1655. The strain is wild-type lysogen MG1655, which maintains the lysogenic state. The scale bar is 2 μm. (E) cI mRNA number distribution in lysogenic cells. Images containing ~500 cells were collected and analyzed to build the distribution of mRNA copy-number per cell. This experimental histogram was fitted to a negative binomial distribution (blue curve), parameters of which were used to calculate the transcriptional burst frequency r and burst size d_{off} (r=1.4±0.2, d_{off}=4.3±0.4, six independent experiments). The results of our stochastic simulation (red curve) are also shown for comparison. For experimental details see Materials and methods section.

To predict the stability of the lysogenic state, characterization of the steady state has to be accompanied by quantification of the stochastic dynamics of gene activity. Recent studies have demonstrated that the production of both mRNA (Golding et al., 2005) and proteins (Cai et al., 2006; Yu et al., 2006) exhibit intermittent, non-Poissonian kinetics. Such ‘bursty’ gene activity has been previously suggested to affect the switching of cellular states (Kaufmann et al., 2007; Schultz et al., 2007; Choi et al., 2008; Mehta et al., 2008; Gordon et al., 2009). Below, we characterize in detail the stochastic kinetics of gene activity in our system, in particular the frequency of activity bursts from the promoter P_{RM}, which maintains the lysogenic state.

Producing CI and P_{R} (which initiates the lytic cascade at low repressor levels) has been mapped as a function of CI and Cro (the ‘anti-repressor’) concentrations (Dodd et al., 2001; Ptashne, 2004; Figure 1A–C). A thermodynamic model using grand-canonical ensemble has been used to describe the occupancy of the operator sites controlling promoter activities and the corresponding protein levels (Shea and Ackers, 1985; Darling et al., 2000; Dodd et al., 2004; Anderson and Yang, 2008).

To predict the stability of the lysogenic state, characterization of the steady state has to be accompanied by quantification of the stochastic dynamics of gene activity. Recent studies have demonstrated that the production of both mRNA (Golding et al., 2005) and proteins (Cai et al., 2006; Yu et al., 2006) exhibit intermittent, non-Poissonian kinetics. Such ‘bursty’ gene activity has been previously suggested to affect the switching of cellular states (Kaufmann et al., 2007; Schultz et al., 2007; Choi et al., 2008; Mehta et al., 2008; Gordon et al., 2009). Below, we characterize in detail the stochastic kinetics of gene activity in our system, in particular the frequency of activity bursts from the promoter P_{RM}, which maintains the lysogenic state.
Knowing this fact allows us, in turn, to make a direct prediction of the stability of the lysogenic state.

Results

Single-molecule-resolution characterization of gene activity in a lysogen

Gene activity in individual cells was characterized using single-molecule fluorescence in situ hybridization (smFISH; Raj et al., 2008). We first quantified the statistics of cl mRNA numbers in a stable lysogen (MG1655(k_{ret}) at 37°C, Figure 1D and E). The observed mRNA statistics displayed a variance-to-mean ratio larger than 1 (\(\sigma^2/\langle n \rangle = 5.3 \pm 0.4\), six independent experiments; \(~500\) cells per experiment), indicating non-Poissonian kinetics for mRNA production.

mRNA number statistics were analyzed in the framework of a two-state model for transcription (Golding et al., 2005; Raj et al., 2006; Shahrezaei and Swain, 2008; Zenklusen et al., 2008; Figure 1C; Supplementary Figure 3). The gene is assumed to switch stochastically between ‘on’ and ‘off’ states, and mRNA is produced only in the ‘on’ state. The resulting time-series of mRNA production is intermittent or ‘bursty’ (Golding et al., 2005; Chubb et al., 2006; Raj et al., 2006; Zenklusen et al., 2008). The measured mRNA copy-number distribution allowed us to estimate the average transcriptional burst size \(b_{RX}\) (number of mRNA molecules produced at each bursting event) and the average number of burst events \(r\) per mRNA lifetime \(t_{mRNA}\). The lifetime of cl mRNA (and similarly cro mRNA, see Materials and methods section) was measured using quantitative RT–PCR after inhibition of transcription with rifampicin (Bernstein et al., 2002). Together, these measurements allowed us to estimate \(k_{on}=r/t_{mRNA}\), the rate of switching the gene ‘on’ in the two-state model (see Materials and methods section for detailed derivations). Thus, based on the combined smFISH and mRNA lifetime experiments, we were able to estimate the average burst size and the burst frequency (i.e. frequency of activity events) of cl transcription. In the case of MG1655(k_{ret}) at 37°C, we found a frequency of 1.4 ± 0.2 events per min with an average burst size of 4.3 ± 0.4 transcripts per event (six independent experiments).

Next, we extended the survey of system behavior by quantifying gene activity in the reporter strain NC416 (Svenningsen et al., 2005). The reporter strain carries a temperature-sensitive allele, cl857 (Hershey, 1971; Hecht et al., 1983). In this allele, a single mutation in the cl gene leads to decreased structural stability of the repressor protein at higher temperatures, and thus to a temperature-sensitive phenotype of the lysogenic state. The reporter strain contains the complete PRM/PR circuitry, but not the lytic genes. Therefore, cells do not die after switching occurs; instead, switched cells enter a Cro-dominated state (Svenningsen et al., 2005; Figure 1A and B).

We measured the copy-number distribution of cl and cro mRNA at different temperatures between 30 and 40°C (Figure 2A and B; two independent experiments; \(~500\) cells per experiment). One can observe the expected transition from cl dominance (lysogeny) at low temperatures to cro dominance at higher temperatures. Both mRNA species exhibited the typical negative binomial statistics, indicating a bursty mode of transcription from both \(P_{RM}\) and \(P_R\) promoters throughout the temperature range. Each of the promoters maintained an approximately constant burst size when active (4.1 ± 0.5 for cl and 1.7 ± 0.5 for cro).

We note that, at the population level, mRNA numbers were directly reflected in the protein level. We measured the Cl protein level at different temperatures using quantitative immunofluorescence (see Materials and methods section). As shown in Figure 2C, the population-averaged levels of cl mRNA and Cl protein were proportional to each other in the range: 30–40°C (\(R^2=0.99\)).

Measurement of lysogenic stability

We quantified stability using the ‘switching rate’ \(S\), the probability of switching from lysogeny to lysis in one cell generation \(S\) is actually the switching rate per \(~1.4\) cell generations; see Materials and methods section). \(S\) was measured experimentally in a fully functional lysogen (Figure 3). As host, we used a RecA-deficient strain, JLS902 (Little et al., 1999), because in wild-type RecA background the stability is masked by frequent spontaneous activation of the cell’s SOS response (Little et al., 1999; see Figure 3A). We estimated \(S\) based on the number of free phages in an exponentially growing culture of lysogens (Little et al., 1999). Specifically, \(S \approx \phi/BM\), where \(\phi\) is the number of free phages in the culture, \(B\) is the number of bacterial cells and \(M\) is the average number of phages released per cell lysis \(~200\) at 30°C and 40°C; see Materials and methods section). It is noteworthy that a constant switching rate \(S\) implies a constant ratio of free phages to bacteria during cell growth. Our data suggests that this is indeed the case (see Supplementary Figure 9). We measured \(S\) values for the temperature-sensitive prophage (cl857) in the temperature range 28–36°C (Figure 3A). The observed \(S\) values covered approximately eight orders of magnitude. We also conducted measurements for wild-type prophage, in both RecA+ and RecA- backgrounds, and observed very little change in \(S\) over the temperature range (Figure 3A), suggesting that the changes to repressor activity in the cl857 allele dominate over all other temperature-dependent effects (Ryals et al., 1982; Farewell and Neidhardt, 1998).

Stability is determined by the frequency of activity bursts from \(P_{RM}\)

When examining the relation between gene activity and lysogen stability (Figure 3B), we observed a simple exponential dependence of the switching rate \(S\) on the frequency of activity bursts from the \(P_{RM}\) promoter. Specifically, \(S\) was well-described by the expression \(S = \exp(-k_{on}\tau) = \exp(-k_{on}/\ln 2)\), where \(k_{on}\) is the rate of transcription bursts and \(\tau\) is the cell doubling time. Both parameters were measured in experiment. For the temperature-sensitive allele, \(R\) is further multiplied by a factor \(\mu(T)\), which describes the decreased fraction of active Cl proteins at increased temperatures. The value for \(\mu(T)\) was calculated using a comparison of the measured mRNA levels to the predictions of the stochastic simulation (see Materials and methods section; Supplemen-
Predicting the stability of a cellular state
C Zong et al

Figure 2 Tuning the state of the lysogeny system. (A) Two-color smFISH was used for counting cl (red) and cro (green) mRNA in the reporter strain NC416. Sample images of cells grown at different temperatures are shown in the top row. Images containing ~ 500 cells at each temperature were processed to yield the mRNA number distributions in the population (bottom two rows). The observed distributions (red/green circles) were fitted with negative binominal distributions (red/green lines) and compared to results of the stochastic simulation (blue lines). For details of smFISH and the stochastic simulation see Materials and methods section. (B) mRNA numbers as a function of temperature. Diamonds represent experimental data; dashed lines represent simulation data. cl and cro mRNA in NC416 strain (cro') are plotted against temperature. The measured cl mRNA level for wild-type lysogen at 37°C is also plotted. (C) Comparison of cl mRNA and protein levels. Protein levels at each temperature were quantified using immunofluorescence (see Materials and methods section). Each protein data point represents the mean of ~ 100 cells. mRNA levels are from smFISH data. Both mRNA and protein levels were normalized by the value at 30°C. Error bars correspond to the s.e.m. values obtained from two independent experiments. The population-averaged mRNA and protein levels were found to be proportional to each other (correlation coefficient is 0.99).

latory Figure 4). Our estimation of $\mu(T)$ also agrees well with previous experimental data (Villaverde et al., 1993; Isaacs et al., 2003).

The exponential dependence found above can be intuitively understood using the following simple model: we assume that CI molecules are produced from the PRM promoter following discrete bursts of $cl$ mRNA, and that the occurrence of the transcription-burst events obeys Poissonian statistics (Golding et al., 2005; Friedman et al., 2006). The average frequency of events is $k_{\text{on}}$. Thus, the probability of the cell NOT producing any $cl$ mRNA (and therefore repressor proteins) for a duration $t$ is $P_0(t) = \exp(-k_{\text{on}}t)$. Next, we note that the mean lifetime of CI proteins, due to cell growth and division, is $\tau/\ln 2$ (where $\tau$ is the cell-doubling time). This timescale is much longer than the lifetime of mRNA, as measured directly (see Supplementary Figure 1). By plugging the protein lifetime into the expression for $P_0(t)$, we obtain the probability of not producing new CI for the whole lifetime of the protein: $\exp(-k_{\text{on}}\tau/\ln 2) = \exp(-R)$. This is just the behavior observed in experiment for the spontaneous switching rate. Thus, the loss of lysogeny (switching) occurs if no new $cl$ mRNA (and the downstream proteins) is produced for the mean lifetime of CI protein, which is ~ 1.4 cell generations.

The expression $R = k_{\text{on}}\tau/\ln 2$ can also be written as $R = \langle CI \rangle / b_{\text{CI}}$, where $\langle CI \rangle$ is the average number of repressor protein per cell and $b_{\text{CI}}$ is the average number of proteins produced after one transcription burst. $b_{\text{CI}} = b_{\text{TX}}b_{\text{TL}}$, where $b_{\text{TX}}$ is the average number of $cl$ mRNA's produced per transcription burst and $b_{\text{TL}}$ is the average number of CI proteins made from one mRNA molecule during its lifetime. For CI, we estimated $b_{\text{CI}}$ to be ~ 20 (see Materials and methods section).

As an additional test for the agreement between the predicted switching rate and experiments, we plot in Figure 3B the predicted switching rate under two alternative hypotheses (while maintaining $\langle CI \rangle$ unchanged): the first one is that proteins are produced in a Poissonian manner, one-at-a-time (no burstiness), in which case one expects $S = \exp(-\langle CI \rangle)$. As shown in Figure 3B, the predicted switching rate in this case is orders of magnitude lower than the experimental data. The strong discrepancy demonstrates...
the critical role of bursty mRNA and protein production in limiting the stability of the lysogenic state (Mehta et al., 2008). A second case we test is what would happen if the cl gene had an efficient ribosome-binding site, for example, as that of the lacZ gene (Shean and Gottesman, 1992). In such a case, virulence. As seen in Figure 3B, the predicted switching rates in this case are considerably higher than those found in experiment. Thus, the relatively inefficient translation of the cl transcript increases the stability of the lysogenic state.

Finally, we asked whether the quantitative relation between the number of activity bursts and the spontaneous switching rate can be demonstrated in other alleles of PRM–cl beyond the cl857 case. The lambda lysogeny system has been studied for many years, and many mutations have been created, targeting multiple features of the lysis/lysogeny switch (see e.g. Little et al., 1999; Atsumi and Little, 2004, 2006a, b; Michalowski et al., 2004; Michalowski and Little, 2005). However, using PRM–cl mutants to test the gene-activity/stability relation presents the following challenge: unlike cl857, most mutants do not provide a continuous spectrum of lysogen stability but instead only a single stability phenotype. In most cases, the stability phenotype obtained is one of two: close to the wild-type value (and to the sensitivity limit of measuring lysogen stability, due to the appearance of unstable mutants (Little et al., 1999; Aurell and Sneppen, 2002; see Figure 3B) or the lysogen is ‘too unstable’, such that lysogenization of the host cell fails in the first place. As seen in Table 1 below, out of 18 alleles examined, 12 exhibited a switching rate very close to wild-type and four were ‘too unstable’. Only two gave an intermediate switching rate (see Figure 3B). Despite these limitations, it can be seen in Figure 3B that the measured stability of the mutant lysogens was consistent with the theoretical prediction. Taken together with our previous data set, these results further support our key observation that the stability of the lysogenic state depends in a simple manner on the number of activity bursts from the fate-determining gene, and that this relation is general, that is, it holds even when the promoter and gene-coding region are modified, as demonstrated by the mutants.

**Discussion**

We have shown that the stability of a bacterial lysogen is given by the simple expression $\exp\{\text{number of gene-activity events in } t\}$, where $\tau$ is the relevant time-scale for maintaining the lysogenic state. We note that exponential switching probabilities have been previously predicted in theoretical models of gene regulatory circuits, assuming weak-noise conditions (Bialek, 2001; Roma et al., 2005; Mehta et al., 2008). In particular, a few past studies applied thermal barrier-crossing (‘Kramers problem’) approaches to the problem of cellular state-switching (Bialek, 2001; Aurell and Sneppen, 2002; Mehta et al., 2008). However, it is important to note that when a system is controlled by pure dynamical rules (rather than thermal fluctuations), converting the master equation into a stochastic differential equation (e.g. Langevin) becomes challenging, especially when large fluctuations affect the outcome—as in the case of the lysogen. Barrier-crossing analysis is also restricted by whether an effective potential and an appropriate reaction coordinate can be identified or not, and by the question of how to define the effective ‘temperature’ (Lu et al., 2006).

In our study, we avoided these theoretical challenges by taking the following approach: the rate of switching was first calculated by Gillespie simulation of the master equation describing the gene network. An intuitive understanding of the
Table 1 $P_{\text{Ram-Cl}}$ mutants used for analyzing lysogen stability

| Name | Genotype | $P_{\text{Ram-Cl}}$ phenotype | Reference | Source | Burst size, $b$ | $S$ (generations$^{-1}$)$^a$ | $R$ (generations$^{-1}$)$^b$ |
|------|----------|-----------------------------|----------|--------|----------------|----------------|----------------|
| $\lambda_{GC31}$ | Wild type | Wild type | Sauer et al., 1990 | Lab stocks | 4.3 ± 0.4 | 9.0 ± 2.0 | 16 ± 1 |
| $\lambda_{GC2004}$ | cI TAT $\rightarrow$ TGG | Covalent dimerization | Sauer et al., 1990 | Lab stocks | 4.3 ± 0.5 | 5.8 ± 2.0 | 19 ± 3 |
| $\lambda_{GC1008}$ | 2× TAG (amber) at AA102 | N-terminus only | Reiderhagen and Sauer, 1988; Sauer et al., 1979 | Lab stocks | NL$^c$ | NL | NL |
| $\lambda_{GC4006}$ | cI E34K | Tight operator binding | Sauer et al., 1990 | Lab stocks | 0.3 ± 0.3 | 2.6 ± 1.0 | 3.3 ± 3.3 |
| $\lambda_{GC1503}$ | cI L18F | Misfolding | Lim and Sauer, 1989 | Lab stocks | NL | NL | NL |
| $\lambda_{GC4002}$ | cI Q44Y | Weak operator binding | Sauer et al., 1990 | Lab stocks | NL | NL | NL |
| $\lambda_{GC1501}$ | cI Q44L | Weak operator binding | Sauer et al., 1990 | Lab stocks | NL | NL | NL |
| $\lambda_{GC2001}$ | cI V36I | Misfolding | Lim and Sauer, 1989 | Lab stocks | 3.6 ± 0.6 | 1.4 ± 1.0 | 12 ± 4 |
| $\lambda_{GC2002}$ | cI L18V | Misfolding | Lim and Sauer, 1989 | Lab stocks | 4.0 ± 0.6 | 7.5 ± 3.0 | 16 ± 2 |
| $\lambda_{L5625}$ | cI D38N | No positive autoregulation; reduced PRM activity | Michalowski and Little, 2005 | Lab stocks | 4.0 ± 0.5 | 3.7 ± 1.0 | 9 ± 1 |
| $\lambda_{NP2}$ | cI D38N $\rightarrow$ GAC $\rightarrow$ GTG | No positive autoregulation; reduced PRM activity | Michalowski and Little, 2005 | Lab stocks | 4.5 ± 0.6 | 2.0 ± 1.0 | 18 ± 4 |
| $\lambda_{NP3}$ | cI D38N $\rightarrow$ GATG $\rightarrow$ GTAT | No positive autoregulation; reduced PRM activity | Michalowski and Little, 2005 | Lab stocks | 0.9 ± 0.4 | 1.0 ± 2.0 | 14 ± 1 |
| $\lambda_{NP4}$ | cI D38N $\rightarrow$ CATT $\rightarrow$ CAT | No positive autoregulation; reduced PRM activity | Michalowski and Little, 2005 | Lab stocks | 1.4 ± 0.5 | 1.7 ± 2.0 | 16 ± 1 |
| $\lambda_{NP5}$ | cI D38N $\rightarrow$ CATT $\rightarrow$ CAT | No positive autoregulation; reduced PRM activity | Michalowski and Little, 2005 | Lab stocks | 2.1 ± 0.5 | 3.0 ± 1.0 | 15 ± 4 |
| $\lambda_{NP6}$ | cI D38N $\rightarrow$ GAGT $\rightarrow$ GTAATT | No positive autoregulation; reduced PRM activity | Michalowski and Little, 2005 | Lab stocks | 1.8 ± 0.3 | 1.5 ± 2.0 | 15 ± 2 |
| $\lambda_{NP7}$ | cI D38N $\rightarrow$ CCA $\rightarrow$ ACC | No positive autoregulation; reduced PRM activity | Michalowski and Little, 2005 | Lab stocks | 0.5 ± 0.1 | 2.3 ± 1.0 | 25 ± 7 |
| $\lambda_{NP8}$ | cI D38N $\rightarrow$ CCA $\rightarrow$ ACC | No positive autoregulation; reduced PRM activity | Michalowski and Little, 2005 | Lab stocks | 2.8 ± 0.6 | 2.8 ± 1.0 | 14 ± 3 |
| $\lambda_{NP2}$ | cI D38N $\rightarrow$ TACC $\rightarrow$ TCT | No positive autoregulation; reduced PRM activity | Michalowski and Little, 2005 | Lab stocks | 1.1 ± 0.4 | 1.5 ± 2.0 | 16 ± 3 |
| $\lambda_{NP11}$ | cI D38N $\rightarrow$ TACC $\rightarrow$ TCT | No positive autoregulation; reduced PRM activity | Michalowski and Little, 2005 | Lab stocks | 1.4 ± 0.7 | 3.1 ± 1.0 | 21 ± 3 |
| $\lambda_{GC2903}$ | cI A66T $\rightarrow$ GCA $\rightarrow$ ACA | Temperature-sensitive | Hecht et al., 1983 | Lab stocks | 4.7 ± 0.6 | 5.8 ± 1.0 | 17 ± 3 |
| | 30°C | | | | 4.8 ± 0.7 | 1.1 ± 0.0 | 16 ± 2 |
| | 32°C | | | | 4.1 ± 0.7 | 2.1 ± 0.0 | 10 ± 2 |
| | 36°C | | | | 6.5 ± 0.3 | 1.7 ± 0.0 | 4 ± 1 |

$^a$Switching rate (per 1.4 cell generation), measured in JL5902 (RecA$^+$). Mean and s.e. for each strain were calculated from two to eight independent experiments.

$^b$Frequency of transcription bursts (per 1.4 cell generation), calculated using the formula $R=r/t_{\text{mRNA}}$ in 2. $t$ is the cell doubling time at 37°C, $t_{\text{mRNA}}$ is the mRNA lifetime (~4 min for NC416 strain, see Materials and methods section) and $r$ is the frequency of transcription bursts (per mRNA lifetime), obtained from the single-cell mRNA data. $R$, mean and s.e. for each strain were calculated from two to four independent smFISH experiments, except those of NP7, NP8 and NP11 (one smFISH experiment only).

$^c$NL, non-lysogenicizable.

$^d$For the temperature-sensitive allele cI D38S, switching rate data is from the $\lambda_{GC2001}$ lysogen, whereas burst size is from the NC416 reporter strain (Svenningsen et al., 2005).

Results from both experiments and simulation was obtained using the argument of survival probability; switching occurs if there are not enough CI molecules to maintain lysogeny. This event only happens at the rare events that no CI is made for a specific period of time. The ‘survival’ probability naturally explains the exponential behavior seen in experiment and simulations, with the key parameters being the promoter burst frequency and the protein lifetime.

At first glance, it may seem surprising that a simple expression captures the behavior of a real-life, naturally evolved system in which the stability is believed to be an important phenotype (Little et al., 1999). Specifically, the
The lambda lysogeny circuit has long served as a paradigm for the intricacy and precision of gene regulation (Pташаня, 2006; Court et al., 2007), in which the proper state of the system depends on the finely-tuned balance between the affinities of CI and Cro to their six DNA targets (O6l-1–3 and O3l-1–3). In contrast, our findings suggest that the stability of a genetic switch can be estimated simply based on the rate of gene activity, thus the intricacy is absent in the expression describing the stability of the switch.

We note that in line with the observation that lysogen stability is insensitive to many system parameters, there is a body of work from the last decade, mainly from the Little lab (Little et al., 1999; Atsumi and Little, 2004, 2006a,b; Michalowski et al., 2004; Michalowski and Little, 2005), pointing to the robust performance of the lambda lysogeny switch even when the underlying gene circuit is modified. For example, it has been shown that a stable lysogenic state could be maintained when the relative affinities of the operator sites to CI and Cro were reversed (Little et al., 1999), when the positive autoregulation by CI was deleted (Michalowski and Little, 2005), when P857 was made stronger or weaker (Michalowski et al., 2004), and even when Cro and CI were replaced by the Lac and Tet repressors, respectively (Atsumi and Little, 2006b). Although only semi-quantitative in nature, these studies suggest that the genetic circuitry found in lambda is not unique, and many alternative systems can maintain a stable lysogenic state. According to our findings here, the critical element is whether the new design can produce the proper rate of gene activity.

Finally, it is tempting to contemplate the possible relevance of our results concerning the stability of cellular states to higher systems, in which the ability of cells to maintain an inheritable memory of their gene-expression state is key to cellular differentiation (Monod and Jacob, 1961; Gurdon and Melton, 2008). Admittedly, the maintenance of bacterial lysogeny does not exhibit the complexity of cell differentiation in higher eukaryotic systems. However, even though a range of additional mechanisms have a role in cellular memory in the higher systems (Burrill and Silver, 2010), the fundamental feature of autoregulation by the fate-determining protein seems to be a central element (Slack, 1991; Lawrence, 1992; Gurdon and Melton, 2008; Crews and Pearson, 2009). We thus look forward to investigating the stability of epigenetic states in higher systems.

Materials and methods

Strains and media

Phage and bacterial strains used in this work are listed in Supplementary Table 1. Cell growth and manipulation followed standard protocols (Sambrook and Russell, 2001).

Two main phage strains, λac823 (λatt') and λac2293 (λ2ts), were used in this study. Both were kanamycin-resistant due to a cassette inserted at the bor gene (a gift from Barry Egan, Adelaide University). This antibiotic marker was used to select lysogens after infection. λatt' carried the wild-type cl allele, whereas λts carried the temperature-sensitive allele, cl857. λts was constructed following the method of Dodd et al. (2001). Briefly, mutations were introduced using the Quickchange kit (Stratagene) into a plasmid carrying the required fragment of the lambda genome. Cells carrying the mutated plasmid were infected with λts and the mutant sequences were crossed onto the lambda chromosome. The imm434 recombinants were selected among the progeny as plaques on λts lysogens. Recombination led to the replacement of imm434 region with lambda sequences.

We infected RecA+ strain MG1655 and RecA- strain JL5902 (a gift from John Little, University of Arizona) with λatt and λts to form four types of lysogens: (RecA+, clac), (RecA-, clac), (RecA+, clts), and (RecA-, clts). Two reporter strains were used, NC416 and NC417 (a gift from Don Court, NCI Fredrick) (Svenningsen et al., 2005). They were both kanamycin-resistant. The immunity region of the reporter strains is shown in Figure 1A.

Single-molecule fluorescence in situ hybridization (smFISH)

Our protocol is based on the ‘Star-FISH’ method developed by Raj et al. (2008). Modifications were made to adapt the protocol to E. coli.

Sample preparation

An overnight culture was diluted 104-fold into 10 ml LB, and incubated on a shaker. When the optical density (OD600) reached ~0.4, cell culture was cooled in an ice-water bath and the cells were collected by centrifugation at 4 °C. The cells were washed by resuspending in 1 ml ice-cold 1× PBS. The cell suspension was centrifuged again and the supernatant was removed. The pellet was gently resuspended in 1 ml 1× PBS with 3.7% formaldehyde. The solution was gently mixed at room temperature for 15 min. Cells were pelleted by centrifugation and washed twice with 1 ml 1× PBS. The pellet was resuspended in 1 ml 70% ethanol and mixed gently on a rotator at room temperature for at least 1 h to permeabilize cell membrane.

Hybridization procedure

The cells in 70% ethanol were pelleted by centrifugation. The pellet was resuspended in 1 ml wash solution (40% formaldehyde, 2× SSC), then set aside for a few minutes. Meanwhile, 2 μl of probed and 2 μl cro probes were added to 50 μl hybridization solution (10% dextran sulfate, 2 mM vanadyl-ribonucleoside complex, 0.02% RNase-free BSA, 50 μg E. coli tRNA, 2× SSC, 40% formaldehyde). Probes were mixed well with the viscous hybridization solution by vortexing. Next, the cells in wash solution were pelleted by centrifugation. After the supernatant was removed, the pellet was resuspended in the hybridization solution, and then incubated at 30 °C overnight. After hybridization, 5 μl of the hybridization mix was washed with 100 μl wash solution. Cells were washed twice with 100 μl wash solution. The pellet was then resuspended in 20 μl 2× SSC solution for imaging under the microscope.

DNA oligonucleotide probes were designed using the program (www.singlemoleculefish.com) from Arjun Raj (MIT), and were ordered with 5'-end amine modification from Biosearch Technologies. The probes for hybridizing cl mRNA were labeled with TAMRA (Invitrogen), and the probes for hybridizing cro mRNA were labeled with Cy5 (GE Health Life Science). The labeling process followed the procedures described previously (Raj et al., 2008). The sequences of the probes are given in Supplementary information.

Microscopy and image analysis

After hybridization and washing (above), 1–2 μl sample was placed between a piece of 1.5% agarose gel (~1 mm thick) and a glass coverslip. Still images of cells were acquired using an inverted epifluorescence microscope (Eclipse TE2000-E, Nikon) and a cooled CCD camera (Cascade 512, Photometrics). A ×100 oil immersion phase objective was used in conjunction with a ×2.5 lens in front of the camera. The microscope and camera were controlled by the Metamorph software (Molecular Devices). Phase contrast images were taken with 100-ms exposure time. TexasRed filter set (Nikon) was used for fluorescence imaging of cl mRNA, with 400 ms exposure time and nine-slice z-stacks with 120-nm spacing. Cy5 filter set (Nikon) was used for fluorescence imaging of cro mRNA, with 8-s exposure time.
For each sample, multiple images were taken and ~500 cells were analyzed from those images.

The images were automatically analyzed using home-made MATLAB programs. Cell recognition was done using the Schnitzcell program (a gift from Michael Elowitz (Caltech)). Spot recognition was performed with a home-made program developed in our lab, ‘to localize the foci, we scanned for the local maxima in both x- and y-directions. The spot at the local maximum in both directions was taken as the center of one fluorescent spot. As multiple mRNA can be in proximity of one fluorescent spot, the fluorescence intensity of the spot was measured. The histogram of the spot fluorescence intensity was plotted and we identified the position of the primary peak as the intensity of a single mRNA. The total mRNA in a cell was estimated by summing the estimated number of mRNA from all the spots in the cell.

We then fit the total mRNA per cell data to a negative binomial distribution with parameters $r$ and $p_0$ to obtain estimates of the transcription kinetics parameters: $r$ (burst frequency) and $b_{TX}$ (burst size that is obtained from $p_0$, see below). We report the mean and s.e.m. for these parameters, from two to six independent experiments in each case.

### Extraction of transcription parameters from smFISH data

Transcription is described using a phenomenological two-state model for gene activity (Supplementary Figure 3; Peccoud and Ycart, 1995; Golding et al., 2005; Golding and Cox, 2006; Raj et al., 2006; Shraerzea and Swain, 2008). In this model, the gene stochastically switches between ‘on’ and ‘off’ states, and mRNA is produced only when the gene is ‘on’. The mRNA distribution at the steady state can be analytically solved for this model (Raj et al., 2006; Shraerzea and Swain, 2008). When $k_{on} > k_{off}$, that is, when the dwell time at the ‘on’ state is much shorter than the mRNA lifetime, the mRNA statistics follows the negative binomial (NB) distribution:

$$P(n, r, p_0) = \frac{r}{n} (1 - p_0)^n$$

where $n$ is the number of mRNA molecules per cell. The parameter $p_0$ equals $k_{off} / (k_{on} + k_{off})$, and describes the probability of the gene turning from the ‘on’ state to the ‘off’ state. The parameter $r$ denotes the average number of gene activity events (i.e. the number of times that the gene has been ‘on’) during one mRNA lifetime. The mRNA molecules in a cell in general come from multiple gene-activity events, and the time window in which we can detect them is the lifetime of mRNA. In the special case that $r$ is equal to 1, we get $P(n, r=1, p_0) = p_0 (1 - p_0)^n$, which is simply the geometric distribution and it describes the probability of transcribing $n$ mRNAs when the gene has been ‘on’ once.

The measured mRNA distribution allowed us to estimate the average transcriptional burst size $b_{TX}$ (number of mRNA molecules produced at each bursting event): $b_{TX} = k_{off} / k_{on}$, as well as the number of average burst events per mRNA lifetime $\tau_{mRNA}$.

The expression for $k_{on}$ is derived as follows. The average number of mRNA per cell follows

$$d(mRNA) = (P_{mRNA}) k_{TX} - k_{off} (mRNA),$$

where $P_{mRNA} = k_{off} / (k_{on} + k_{off}) \approx k_{off} / k_{on}$ and $k_{off}$ is the decay rate of mRNA. At the steady state,

$$mRNA = \frac{k_{off} k_{TX}}{k_{off} k_{on}} = \frac{k_{off} k_{TX}}{k_{on}}$$

therefore the switching-on rate of the gene is

$$k_{on} = \frac{mRNA}{b_{TX}} = \frac{r k_{off}}{\tau_{mRNA}},$$

where we used the relation $\langle mRNA \rangle = b_{TX} r$. As $r$ is known from smFISH data, we can calculate the rate $k_{on}$ once we know mRNA lifetime $\tau_{mRNA}$. Standard RNA lifetime assay (Bernstein et al., 2002) was carried out as described below. The decay curves of cl mRNA and cro mRNA are shown in Supplementary Figure 1. Thus, both $k_{on}$ and the burst statistics: $b_{TX}$ for transcription are obtained. The transcription rate $k_{TX}$ is set at $b_{TX} k_{off}$.

### Immunofluorescence

Reagents, incubation times and temperatures were based on established immunofluorescence protocols in bacteria (Maddock and Shapiro, 1993; Harry et al., 1995; Raj et al., 2008). Unless otherwise stated, all steps were performed at room temperature. Cells were grown overnight at $30^\circ$C, diluted 1:1000 in 15 ml of LB medium, and grown at the experimentally desired temperature $(30-40^\circ$C) until cultures reached OD$_{600}$~0.4. Cells were collected by centrifugation at 4°C and resuspended in 1 ml of fixation solution (1 × PBS, 3.7% formaldehyde), transferred to an Eppendorf tube, and mixed gently for 30 min. Cells were then washed twice with 1 × PBS, resuspended in 300 μl water, ethanol added to reach a final concentration of 70% and mixed gently for 1 h. The suspension was centrifuged and permeabilized by adding 100 μl lysozyme solution (10 μg/ml lysozyme, 50 mM glucose, 25 mM Tris–HCl, 10 mM EDTA (pH 8.0)) and left for 10 min. 1 ml of 1 × PBS was then added, the cells collected, and washed twice in 1 × PBS. Before antibody labeling, two blocking steps were performed to decrease non-specific binding. Cells were incubated in 100 μl blocking solution (1 × PBS, 2% BSA, 0.05% Tween 20) for 30 min, and washed twice with 1 × PBS. Then, cells were resuspended in 100 μl of blocking solution containing 1:100 rabbit anti-CI serum (a gift from Ian Dodd, University of Adelaide; Dodd et al., 2001) and incubated for 1 h. Cells were collected by centrifugation, washed with 1 × PBS, blocking solution added and incubated for 30 min. Cells were collected by centrifugation and resuspended in 100 μl of blocking solution containing 1:100 mouse anti-rabbit secondary antibody labeled with Cy3 dye (Jackson Immunochemicals) and incubated for 1 h in the dark. The suspension was then diluted in 1 ml 1 × PBS, kept for 5 min and then washed twice in 1 × PBS. The immunolabeled cells were resuspended in 1 × PBS (5–10 μl), placed on a #1 slide, covered by a thin 1.5% agarose gel pad in 1 × PBS and imaged under the microscope.

### Quantification of protein levels

To quantify the relative levels of CI in the NC416 reporter strain, a fiducial marker (FM) strain was constructed by transforming NC416 with a pTcYe9A plasmid expressing GFP. In each experiment, three different cultures of NC416 were grown at 30, 32, 34, 36, 38 and 40°C. Cloning strains: MG1655 (wt), FM and NC416 were grown as described above. FM was grown overnight with 0.1 mM IPTG to induce the production of GFP. FM and MG1655 were grown overnight at 30°C, while independent cultures of NC416 were grown at 30, 32, 34, 36, 38 and 40°C. Before fixation, ~7.5 ml of MG1655 or NC416 cultures were mixed with the same volume of FM and the immunofluorescence protocol performed. After imaging, cells were recognized using automated procedures in MATLAB (see Microscopy and Image Analysis). At each temperature, the mean CI expression level was calculated as the ratio of the mean cell intensity of the Cy3 signal of NC416 (or MG1655) to the one from FM. The Cy3 level obtained from MG1655 was used as baseline and subtracted from the NC416 levels.

### RNA lifetime measurement using quantitative RT–PCR (qRT–PCR)

RNA decay experiments (Bernstein et al., 2002) were carried out as follows. A volume of 20 ml culture was grown to OD$_{600}$~0.4, and divided into two flasks at t=0. In one of the flasks, rifampicin was added to a final concentration of 500 μg/ml to inhibit transcription. The culture in the other flask, without rifampicin, served as the reference. We continued to incubate the cultures with shaking. A volume of 1.5 ml of culture was taken from each flask every 3 min and immediately mixed with RNAProtect Bacterial Reagent (Qiagen) to stabilize cellular RNA. This was continued until 12 min after adding rifampicin. RNA was then extracted from the RNA-stabilized samples using RNeasy Mini Kit (Qiagen). After RNA was extracted, the total RNA concentration and the purity were estimated from the UV absorption spectrum (Cary 100, Varian). All RNA samples were diluted
to a final concentration of 50 μg/ml for the qRT–PCR experiment. qRT–PCR was performed on a MiniOpticon System (Bio-Rad) using HotStart-IT SYBR Green One-step qRT–PCR Master Mix kit (USB). The primers for cI mRNA were 5′-AGAAGCAGCAGCTTAAAG-3′ and 5′-GCTTGGCCAACCGCTAA3′. The primers for cro mRNA were 5′-GCATTAAAATAAACGCAAGGCGGTCG-3′. The relative amount of target mRNA was measured for samples at different times and the decay rates were extracted from the exponential fit of experimental data (Supplementary Figure 1). The decay rates of cI mRNA in the lysogen with temperature-sensitive CI and the reporter system NC416 were measured at 30°C; the decay rate of cro mRNA in the reporter system NC416 was measured at 40°C. The lifetime values of cI mRNA and cro mRNA are 4.0 ± 0.8 and 2.8 ± 0.4 min, respectively.

Measuring spontaneous lysis rate

The measurement of lysogen stability is based on the protocol of Little et al. (1999). Lysogen culture was grown overnight in LBGM (LB with 1 mM MgSO4 and 0.2% glucose). The overnight culture was centrifuged and resuspended in 1 ml LBGM. The washed culture was diluted 103-fold in 15 ml LBGM and grown at the desired temperature (between 28 and 36°C). During the exponential growth phase (OD600 = 0.1–0.5), 1 ml culture was taken. OD600 value was immediately measured and 15 μl chloroform was quickly added to the sample and mixed by inverting 10 times. The samples were then stored at 4°C for phase titration later.

Phages were titrated using the following procedures (Little et al., 1999). Indicator strain (LE392) was grown overnight in LBMM (LB with 1 mM MgSO4 and 0.2% maltose). The overnight culture was diluted 100-fold in LBMM, and then grown to OD600 ~0.5. The cells were pelleted by centrifugation at 4°C, 10000 g for 10 min. The pellet was resuspended to one-tenth original volume in 1× PBS with 10 mM MgSO4 and stored at 4°C. A volume of 10 or 100 μl of the chloroform-treated solution or its dilution series was added to 100 μl host culture and mixed well. The tube was incubated at 37°C for 15 min, to allow phages to attach to cells. Meanwhile, NZYM (1% NZ amine, 0.5% yeast extract, 0.5% NaCl, 0.2% MgSO4, 0.75% agar (pH 7.0)) top agar was melted and kept at 48°C. After infection, the bacteria/phage solution was mixed with 3 ml NZYM top agar and plated onto dry and pre-warmed NZYM plates. Plaques were visible after ~12-h incubation at 37°C. The number of bacterial cells, B, was estimated based on OD600 value. The average number of phages released per lysis, M, was measured separately, as described in the next section. The spontaneous induction rate was calculated based on the formula derived below.

Derivation of the spontaneous induction rate

For the case of exponentially growing culture, the number of cells B at time t is:

\[ B(t) = B_0 e^{ln(2)/t}, \]

where t is the generation time and B0 is a constant reflecting the initial conditions. The rate of increase in phage concentration \( \phi(t) \) is the product of the three factors: the spontaneous induction rate \( I \), the number of phages released at each lytic event \( M \), and the bacterial concentration \( B \).

\[ \phi(t) = IMB(t - T) \]

Note that \( B \) is calculated at time \( t - T \), where \( T \) (the ‘latent period’ Little et al., 1999) is the time interval between induction and lysis. We can then write the expression for the free phage concentration in the lysogen culture at time t:

\[ \phi(t) = \int_0^{ln(2)/t} IMB_0 e^{(t - T)ln(2)/t} dt \]

\[ \phi(t) = \frac{IMB_0 T}{ln(2)} e^{-Tln(2)/t} e^{(t - T)ln(2)/t} - 1) \]

The ratio of free phage to bacteria in culture can be written as:

\[ \frac{\phi(t)}{B(t)} = \frac{IMT}{ln(2)} e^{-Tln(2)/t} (1 - e^{-Tln(2)/t}) \]

This expression converges within a few cell generations to a constant value. In addition, both the generation time and the latent period are of the order ~1 h, and thus \( e^{-Tln(2)/t} \) and \( 1 - e^{-Tln(2)/t} \) are close to ~1. We then find that the ratio of free phage to bacteria in a growing culture of lysogens is approximately given by the product of the spontaneous switching rate per generation and the number of phages released at lysis:

\[ \frac{\phi(t)}{B(t)} \approx IMT/ln(2). \]

Note that this simple model predicts that the free phage-to-bacteria ratio will be approximately constant during exponential growth. As seen in Supplementary Figure 9, we found this to be the case. In our study, we used this measured ratio to calculate the spontaneous induction rate, using: \( S = r/T \ln 2 \approx \phi/BM \) (per 1.4 cell generation).

Measuring the number of phages released at cell lysis

We used the following protocol (Little et al., 1999) to measure the average number of phages released per cell lysis. Bacterial culture (lysogen with temperature-sensitive CI) was grown in LBGM overnight at 30°C. The overnight culture was diluted 104-fold in 20 ml LB and incubated at 30°C. At OD600~0.4, 10 μl culture was taken for titering bacterial concentration. The rest of the culture was immediately transferred to 40°C and incubated for 15 min to induce lysis. The culture was then divided and moved to the studied temperatures (30°C and 40°C in this case). A volume of 1 ml culture was taken out every 15 min, gently mixed with 15 μl chloroform to kill the bacteria, and stored at 4°C. This procedure was continued until the culture became clear, indicating the complete lysis of bacterial cells. The phage concentration of each sample was determined using the titering assay described above. As shown in Supplementary Figure 2, the phage concentration reached plateau after all bacterial cells had lysed. The number of phages released per lysis was calculated as the ratio between the phage concentrations at the plateau to the bacterial cell concentration determined from colony counting.

Stochastic simulation of the lysis/lysogeny switch

Transcription parameters estimated from smFISH data, combined with known parameters of protein translation (Kennell and Riezman, 1977; Shea and Gottesman, 1992), allowed us to construct a fully calibrated stochastic model for the maintenance of lysogeny. We used the Gillespie algorithm (Gillespie, 1977) to simulate the stochastic kinetics of the system shown in Figure 1C. In the algorithm, the reaction probabilities are obtained from the kinetic rate parameters (summarized in Supplementary Table 2). Each gene under consideration (cI or cro, chemical rates for each species are denoted by superscripts) can be in the inactive (‘off’) or active (‘on’) state. When the gene is in the inactive state, it switches to the active state at a rate \( f_{on} = P_{RNAP} \cdot f_{on} \). In this study, \( P_{RNAP} \) is the probability that an RNA polymerase is bound to the promoter, calculated from a grand canonical ensemble that takes into account all occupancy configurations of CI and Cro dimers at the operator sites (Ackers et al., 1982; Shea and Ackers, 1985; Anderson and Yang, 2008). \( f_{on} \) is defined as the rate of transcription initiation when CI or Cro are bound at the proper operator to enhance transcription. We assume \( f_{on} \) is constant. At the level of ensemble average, we expect that \( \langle P_{RNAP} \cdot f_{on} \rangle \) as calculated from simulation to equal \( k_{on} \) as measured from smFISH experiment. We estimated \( f_{on} \) using the following procedure: we first use the experimentally derived \( k_{on} \) value as the ansatz for \( f_{on} \) used in the first round of simulation. We then calculate \( \langle P_{RNAP} \rangle \) and adjust \( f_{on} \) so that \( \langle P_{RNAP} \rangle \cdot f_{on} \) becomes closer to the experimental \( k_{on} \) value. We repeat this procedure till \( \langle P_{RNAP} \rangle \cdot f_{on} \) agrees well with the experimental \( k_{on} \) value. By the above procedure, we are able to tune our simulation model so that its mRNA profile matches with experimentally measured mRNA profile. For the
temperature sensitive strain NC416, we performed the above procedure for cl using the experimental mRNA profile at 30 °C and for cro using the mRNA profile at 40 °C.

In the simulation run, we update the value of P_RNAp after each time step, and calculate the new k_mRNA for each time step. When the gene is in the active state, it can switch back to the inactive state at rate k_off, or be transcribed to give a new mRNA at rate k_TX. To estimate k_off and k_TX, we note that k_TX = k_TXoff + k_TXon. k_TX on is the transcriptional burst size, estimated from smFISH experiments for each gene. We choose k_off to be an arbitrary fast rate, assuming the dwell time in the active state is very short. Each existing mRNA is translated into protein monomers at rate k_mRNA and is degraded at rate k_dsRNA. The translation rate of cl mRNA is taken to be six-fold lower than that of lacZ mRNA (Kennell and Riezman, 1977; Shean and Gottesman, 1992), or 0.02 s⁻¹. The translation rate of cro mRNA into Cro monomers is assumed to be the same as that of lacZ mRNA. The degradation rates of cl and cro mRNA were estimated from qRT–PCR as described above. CI dimerization was assumed to be a fast process, such that CI monomers and dimers are in equilibrium at all times (Johnson et al., 1980). Cro dimers association and dissociation rates were k_cro2cro and k_cro2pop, respectively (Jia et al., 2005). The resulting number of CI dimers discounted by μ(T) (see below) and the number of Cro dimers were fed into the calculation of P_RNAp to update the promoter activity. Cell growth rate k_g was measured directly. All mRNA and protein species were binomially partitioned upon cell division.

As a first indication of its validity, the average CI protein level predicted by the model (~300 molecules per cell in a lysogen) is in good agreement with the known value (Ackers et al., 1982). Note that no parameters were adjusted to obtain this estimate. Beyond previous models, our model accurately captures the fluctuations of gene activity that set a limit on the stability of the lysogenic state because the underlying burst processes are described in the model and calibrated by experimental data.

### Modeling the cl857 allele
To capture the effect of temperature in our model, we introduced the parameter μ(T), which denotes the fraction of active repressor proteins at a given temperature. For wild-type CI, μ(T) = 1. For cl857, μ(T) decreases as temperature is increased, indicating only partial functionality of repressor (Supplementary Figure 4). The exact molecular mechanism underlying the reduced activity of repressor at higher temperatures is unknown: CI may be degraded at a higher rate (Isacs et al., 2003); alternatively, only a fraction of CI molecules may remain active (Oppenheim and Noff, 1975). Importantly, however, these details are not critical for our analysis. We identified a simple sigmoidal shape for μ(T) based on the average number of cl and cro mRNA over the complete range of temperatures examined (Supplementary Figure 4). μ(T) relation thus allows a mapping between the simulation and experiments. As a test for the validity of this description, the same curve for μ(T) can be used to predict cl activity in a cro + reporter system (NC417; Svenningsen et al., 2005). This was done by eliminating Cro in the theoretical model. The good agreement between theory and experiment (Supplementary Figure 5) again supports the validity of our model.

### Estimating the switching rate
In the simulation, the rate of switching events was determined by enumerating the events in which CI numbers reach zero for the first time. ‘Brute-force’ sampling was performed when the rate was fast enough (> 0.0001 per generation). However, when the rate becomes smaller and the switching events become extremely rare, brute-force sampling is inefficient for generating reliable statistics. Instead, we used the Forward Flux Sampling (FFS) method (Allen et al., 2005). In this method, the frequency of rare events is sampled based on the initial flux and the probabilities of the flux reaching the successive interfaces along the reaction coordinate. We compared the results from FFS with those from brute-force sampling under the same conditions, and they agree with each other very well.

### Estimating the number of CI proteins generated from each transcription burst, b_CI
We can write b_CI = b_TXb_TL, in which b_TX is the average number of cl mRNAs produced per transcription burst and b_TL is the average number of CI proteins produced from one mRNA molecule during its lifetime. b_TX was estimated from the single-cell mRNA data and found to be ~4 (see Figure 2A). To estimate b_TL, we note that the translation rate of cl mRNA is about six-fold lower than that of lacZ mRNA (Shean and Gottesman, 1992), or 0.02 s⁻¹ (Kennell and Riezman, 1977). With cl mRNA lifetime measured to be ~4 min (see Supplementary Figure 1), we obtain b_TL = 5 and thus b_CI = 20.

### Gene activity in a Cro + reporter strain
We used a Cro + reporter strain, NC417 (Svenningsen et al., 2005) as a control for the validity of our stochastic model, as the difference between Cro + and Cro - has been measured before (Schubert et al., 2007). We performed the same smFISH experiments, as well as QRT–PCR, for this strain. The strain carries the cro27 allele, bearing a missense mutation. It also carries the same temperature-sensitive cl allele as in the Cro + reporter strain (NC416). The experiments were carried out between 30 and 40 °C in 2 °C intervals. Our stochastic model was extended to describe the Cro - strain, by simply setting the initial number of Cro proteins to zero and the transcription rate from promoter P_R to zero. The mRNA distributions and the averaged level from experiments and simulations are summarized in Supplementary Figure 5. The good agreement between theory and experiment in the Cro + strain further strengthens our confidence in the validity of the stochastic model, and in particular the mapping between physical temperature and CI activity μ. It is noteworthy that we used μ(T) from the results of the Cro - strain (Supplementary Figure 4).

### Comparing results from smFISH, qRT–PCR and LacZ activity assay
We compared cl mRNA levels determined using smFISH with the results from qRT–PCR. The comparison between the two methods is shown in Supplementary Figure 6. The good agreement between the smFISH and qRT–PCR data (correlation coefficient ~0.95) indicates that the quantitative measurement of mRNA using smFISH is reliable.

We also compared the mean level of cl mRNA obtained from smFISH experiments with the promoter activity reported in the literature, measured using β-galactosidase (LacZ) assay (Schubert et al., 2007). In those experiments, CI was expressed from an inducible promoter and the activity of P_R was assayed using LacZ as a reporter. The comparison is shown in Supplementary Figure 7 and our smFISH data is consistent with the LacZ assay results.

### Temperature does not significantly affect the number of phages released at cell lysis
Figure 3A shows that temperature has a very small effect on the spontaneous lysis rate. We also measured the number of phages released per lysis at 30 and 40 °C for both RecA + strain (JL5902) and RecA – strain (MG1655). As shown in Supplementary Figure 2, there is no significant difference in the number of phages released per lysis between 30 and 40 °C for both strains.

### ‘Continuous phase transition’-like behavior
As shown in the right panel of Supplementary Figure 8, there is no bimodality in the mRNA distribution at the transition temperature 36.5 °C. Sampling suggests a continuous phase transition-like behavior. As the temperature increases, the CI-dominant state (left panel of Supplementary Figure 8) shifts gradually toward the direction of Cro-dominant state. At the transition temperature, the different cells ‘spread’ across the region between the CI-dominant state and the Cro-dominant state.
Supplementary information
Supplementary information is available at the Molecular Systems Biology website (www.nature.com/msb).

Acknowledgements
We thank B Egan, A Raj, A van Oudenaarden, J Little, D Court, I Dodd, M Elowitz, P Ge, G Altan-Bonnet, T Gregor and all members of the Golding lab for supplying reagents and for their advice. We thank U Alon for commenting on an earlier version of this paper. Some of the CI mutants were constructed by IG while working at the lab of Ted Cox (Princeton). Work in the Golding lab is supported by NIH grant R01GM082837, HFSP grant RGY 70/2008 and NSF Grant 082265 (PFC: Center for the Physics of Living Cells).

Author contributions: IG and CZ conceived the stability measurement project. CZ performed the majority of experiments and developed the theoretical model for lysogen stability. LHs, LAS and SOS performed additional experiments and developed analysis tools. IG, CZ, LHs, LAS and SOS wrote the paper.

Conflict of interest
The authors declare that they have no conflict of interest.

References
Acar M, Becskei A, van Oudenaarden A (2005) Enhancement of cellular memory by reducing stochastic transitions. Nature 435: 228–232
Akers GK, Johnson AD, Shea MA (1982) Quantitative model for gene regulation by lambda phage repressor. Proc Natl Acad Sci USA 79: 1129–1133
Allen RJ, Warren PB, Ten Wolde PR (2005) Sampling rare switching events in biochemical networks. Phys Rev Lett 94: 018104
Anderson LM, Yang H (2008) DNA looping can enhance lysogenic CI transcription in phage lambda. Proc Natl Acad Sci USA 105: 5827–5832
Atsumi S, Little JW (2004) Regulatory circuit design and evolution using phage lambda. Genes Dev 18: 2086–2094
Atsumi S, Little JW (2006a) Role of the lytic repressor in prophage induction of phage lambda as analyzed by a module-replacement approach. Proc Natl Acad Sci USA 103: 4558–4563
Atsumi S, Little JW (2006b) A synthetic phage lambda regulatory circuit. Proc Natl Acad Sci USA 103: 19045–19050
Aurell E, Sneppen K (2002) Epigenetics as a first exit problem. Phys Rev Lett 88: 048101
Bernstein JA, Khodursky AB, Lin P-H, Lin-Chao S, Cohen SN (2002) Global analysis of mRNA decay and abundance in Escherichia coli at single-gene resolution using two-color fluorescent DNA microarrays. Proc Natl Acad Sci USA 99: 9697–9702
Biakle W (2001) Stability and noise in biochemical switches. Adv Neural Info Process 13: 103–109
Burrill DR, Silver PA (2010) Making cellular memories. Cell 140: 13–18
Cai L, Friedman N, Xie XS (2006) Stochastic protein expression in individual cells at the single molecule level. Nature 440: 358–362
Choi PJ, Cai L, Frieda K, Xie XS (2006) A stochastic single-molecule event triggers phenotype switching of a bacterial cell. Science 322: 442–446
Chubb JR, Trek T, Shenoy SM, Singer RH (2006) Transcriptional pulsing of a developmental gene. Curr Biol 16: 1018–1025
Court DL, Oppenheim AB, Adhya SL (2007) A new look at bacteriophage lambda genetic networks. J Bacteriol 189: 298–304
Crowe ST, Pearson JC (2009) Transcriptional autoregulation in development. Curr Biol 19: R241–R246
Darling PJ, Holt JM, Ackers GK (2000) Coupled energetics of lambda cro repressor self-assembly and site-specific DNA operator binding II: cooperative interactions of cro dimers. J Mol Biol 302: 625–638
Dodd IB, Perkins AJ, Tsemitsidis D, Egan JB (2001) Octamerization of lambda CI repressor is needed for effective repression of P(RM) and efficient switching from lysogenic. Genes Dev 15: 3013–3022
Dodd IB, Shearin KE, Perkins AJ, Burr T, Hochschild A, Egan JB (2004) Cooperativity in long-range gene regulation by the CI repressor. Genes Dev 18: 344–354
Farewell A, Neidhardt FC (1998) Effect of temperature on in vivo protein synthetic capacity in Escherichia coli. J Bacteriol 180: 4704–4710
Friedman N, Cai L, Xie XS (2006) Linking stochastic dynamics to population distribution: an analytical framework of gene expression. Phys Rev Lett 97: 168302
Gillespie DT (1977) Exact stochastic simulation of coupled chemical reactions. J Phys Chem 81: 2340–2361
Golding I, Cox EC (2006) Eukaryotic transcription: what does it mean for a gene to be ‘on’? Curr Biol 16: R371–R373
Golding I, Paulsson J, Zawilski SM, Cox EC (2005) Real-time kinetics of gene activity in individual bacteria. Cell 123: 1025–1036
Gordon AJ, Halliday JA, Blankschien MD, Burns PA, Yatagai F, Herman C (2009) Transcriptional infidelity promotes heritable phenotypic change in a bistable gene network. PLoS Biol 7: e44
Gurdon JB, Melton DA (2008) Nuclear reprogramming in cells. Science 322: 1811–1815
Harry EJ, Pogliano K, Losick R (1995) Use of immunofluorescence to visualize cell-specific gene expression during sporulation in Bacillus subtilis. J Bacteriol 177: 3386–3393
Hecht MH, Nelson HC, Sauer RT (1983) Mutations in lambda repressor’s amino-terminal domain: implications for protein stability and DNA binding. Proc Natl Acad Sci USA 80: 2676–2680
Hershey AD (1971) The Bacteriophage Lambda. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press
Isaacs FJ, Hasty J, Cantor CR, Collins JJ (2003) Prediction and measurement of an autoregulatory genetic module. Proc Natl Acad Sci USA 100: 7714–7719
Jia H, Satumbi WJ, Bidwell III GL, Mossing MC (2005) Slow assembly and disassembly of lambda Cro repressor dimers. J Mol Biol 350: 919–929
Johnson AD, Palbo CO, Sauer K (1980) Bacteriophage lambda repressor and cro protein: interaction with operator DNA. Methods Enzymol 65: 839–856
Kaufmann BB, Yang Q, Mettetal JT, van Oudenaarden A (2007) Heritable stochastic switching revealed by single-cell genealogy. PLoS Biol 5: e239
Kennon D, Riezman H (1977) Transcription and translation initiation frequencies of the Escherichia coli lac operon. J Mol Biol 114: 1–21
Lawrence PA (1992) The Making of a Fly: the Genetics of Animal Design. Oxford, Boston: Blackwell Scientific Publications
Lim WA, Sauer RT (1989) Alternative packing arrangements in the hydrophobic core of lambda repressor. Nature 339: 31–36
Little JW, Shepley DP, Wert DW (1999) Robustness of a gene regulatory circuit. EMBO J 18: 4299–4307
Lu T, Hasty J, Wolynes PG (2006) Effective temperature in stochastic kinetics and gene networks. Biophys J 91: 84–94
Maddock JR, Shapiro L (1993) Polar location of the chemoreceptor complex in the Escherichia coli cell. Science 259: 1717–1723
Maheshri N, O’Shea EK (2007) Living with noisy genes: how cells function reliably with inherent variability in gene expression. Annu Rev Biophys Biomol Struct 36: 413–434
Mehta P, Mukhopadhyay R, Wingreen NS (2008) Exponential sensitivity of noise-driven switching in genetic networks. Phys Biol 5: 26005
Michalowski CB, Little JW (2005) Positive autoregulation of cl is a dispensable feature of the phage lambda gene regulatory circuitry. J Bacteriol 187: 6430–6442

© 2010 EMBO and Macmillan Publishers Limited Molecular Systems Biology 2010 11
Michalowski CB, Short MD, Little JW (2004) Sequence tolerance of the phage lambda PRM promoter: implications for evolution of gene regulatory circuitry. *J Bacteriol* **186**: 7988–7999

Monod J, Jacob F (1961) General conclusions: teleonomic mechanisms in cellular metabolism, growth, and differentiation. *Cold Spring Harb Symp Quant Biol* **26**: 389–401

Oppenheim AB, Kobiler O, Stavans J, Court DL, Adhya S (2005) Switches in bacteriophage lambda development. *Annu Rev Genet* **39**: 409–429

Oppenheim AB, Noff D (1975) Deletion mapping of trans dominant mutations in the lambda repressor gene. *Virology* **64**: 553–556

Pecoul J, Ycart B (1995) Markovian modeling of gene-product synthesis. *Theor Popul Biol* **48**: 222–234

Ptashne M (2004) *Genetic Switch: Phage Lambda Revisited*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press

Ptashne M (2006) Lambda’s switch: lessons from a module swap. *Curr Biol* **16**: R459–R462

Ptashne M (2007) On the use of the word ‘epigenetic’. *Curr Biol* **17**: R233–R236

Raj A, Peskin CS, Tranchina D, Vargas DY, Tyagi S (2006) Stochastic mRNA synthesis in mammalian cells. *PLoS Biol* **4**: e309

Raj A, van den Bogaard P, Rifkin SA, van Oudenaarden A, Tyagi S (2008) Imaging individual mRNA molecules using multiple singly labeled probes. *Nat Methods* **5**: 877–879

Reidhaar-Olson JF, Sauer RT (1988) Combinatorial cassette mutagenesis as a probe of the informational content of protein sequences. *Science* **241**: 53–57

Rome DM, O’Flanagan RA, Ruckenstein AE, Sengupta AM, Mukhopadhyay R (2005) Optimal path to epigenetic switching. *Phys Rev E Stat Nonlin Soft Matter Phys* **71**: 011902

Ryals J, Little R, Bremer H (1982) Temperature dependence of RNA synthesis parameters in Escherichia coli. *J Bacteriol* **151**: 879–887

Sambrook J, Russell DW (2001) *Molecular Cloning: A Laboratory Manual*. New York: Cold Spring Harbor Laboratory Press

Sauer RT, Jordan SR, Pabo CO (1990) Lambda repressor: a model system for understanding protein–DNA interactions and protein stability. *Adv Protein Chem* **40**: 1–61

Sauer RT, Pabo CO, Meyer BJ, Ptashne M, Backman KC (1979) Regulatory functions of the lambda repressor reside in the amineterminal domain. *Nature* **279**: 396–400

Schubert RA, Dodd IB, Egan JB, Shearwin KE (2007) Cro’s role in the CI Cro bistable switch is critical for λ’s transition from lysogeny to lytic development. *Genes Dev* **21**: 2461–2472

Schultz D, Onuchic JN, Wolynes PG (2007) Understanding stochastic simulations of the smallest genetic networks. *J Chem Phys* **126**: 245102

Shahrezaei V, Swain PS (2008) Analytical distributions for stochastic gene expression. *Proc Natl Acad Sci USA* **105**: 17256–17261

Shen CS, Gottesman ME (1992) Translation of the prophage lambda cl transcript. *Cell* **70**: S13–S22

Slack JMW (1991) *From Egg to Embryo: Regional Specification in Early Development*, 2nd edn. Cambridge

Svensningsen SL, Costantino N, Court DL, Adhya S (2005) On the role of Cro in lambda prophage induction. *Proc Natl Acad Sci USA* **102**: 4465–4469

Villaverde A, Benito A, Viaplana E, Cubarsi R (1993) Fine regulation of cl857-controlled gene expression in continuous culture of recombinant *Escherichia coli* by temperature. *Appl Environ Microbiol* **59**: 3485–3487

Yu J, Xiao J, Ren X, Lao K, Xie XS (2006) Probing gene expression in live cells, one protein molecule at a time. *Science* **311**: 1600–1603

Zenklausen D, Larson DR, Singer RH (2008) Single-RNA counting reveals alternative modes of gene expression in yeast. *Nat Struct Mol Biol* **15**: 1263–1271

Molecular Systems Biology is an open-access journal published by European Molecular Biology Organization and Nature Publishing Group. This work is licensed under a Creative Commons Attribution-Noncommercial-No Derivative Works 3.0 Unported License.