nature genetics

A multistep epigenetic switch enables the stable inheritance of DNA methylation states

Han N Lim[1,2] & Alexander van Oudenaarden[1]

In many prokaryotes and eukaryotes, DNA methylation at *cis*-regulatory sequences determines whether gene expression is on or off. Stable inheritance of these expression states is required in bacterial pathogenesis, cancer and developmental pathways[1,2]. Here we delineate the factors that control the stability of these states by using the *agn43* gene in *Escherichia coli* as a model system. Systematic disruption of this system shows that a functional switch requires the presence of several, rarely occupied, intermediate states that separate the 'on' and 'off' states. Cells that leave the on and off state enter different intermediate states, where there is a strong bias that drives cells back to their original state. The intermediate states therefore act as buffers that prevent back and forth switching. This mechanism of generating multiple states is an alternative to feedback regulation[3–5], and its general principle should be applicable to the analysis of other epigenetic switches and the design of synthetic circuits.

Methylation of DNA at the *cis*-regulatory region of many prokaryotic and eukaryotic genes acts as an epigenetic switch to turn gene expression on or off[1,2]. The 'on' and 'off' expression states can be inherited for many generations, allowing the creation and maintenance of differentiated cell lineages. The factors involved in determining whether gene expression is on or off have been studied in detail for many epigenetic switches[2,6]; however, the dynamic process of transitioning between the on and off states, which determines the stability of these states, is not well characterized. To gain insight into the switching process, we have used the *agn43* epigenetic switch in *Escherichia coli* as a model system because it does not involve feedback regulation[7], which by itself can generate stable on and off expression states[3–5].

The *agn43* gene encodes an outer membrane protein, antigen 43 (refs. 8–10), that is involved in biofilm formation. The methylation state of three GATC sequences located in a switch region downstream of the promoter[7,10–13] determines whether *agn43* expression is on (methylated) or off (unmethylated). Switching between the methylated and unmethylated states occurs at rates of 10^{-3} to 10^{-4} cells per generation[13–16]. The methylation state of the GATC sites is determined by competitive binding between OxyR, a global oxidative stress protein, and DNA adenine methylase (Dam). DNA methylation by Dam is a two-step reaction that first generates hemimethylated DNA (M_H). The hemimethylated DNA can then become fully methylated (M_F) when Dam methylates the complementary strand. There is no DNA demethylation reaction; therefore, methylated DNA can be only diluted out of the population over many generations by the synthesis of unmethylated DNA at DNA replication, resulting in the conversion of fully methylated to hemimethylated DNA. Half of the hemimethylated DNA that is not methylated by the next round of DNA replication becomes unmethylated. The unmethylated switch region can be re-methylated or bound by OxyR as a tetramer[17,18], which blocks Dam and results in 'off' expression.

To characterize *agn43* switching, single-cell measurements were necessary to distinguish changes in the expression of *agn43* from changes in the fraction of cells in different states. A genetic amplifier was developed to facilitate single-cell measurements of transcription from weak bacterial promoters in their native positions (**Fig. 1a**). The amplification system replaced the coding sequence of *agn43* in the bacterial chromosome with the coding sequence of T7 RNA polymerase (T7 RNAP), resulting in deletion of the Ag43 protein. The expression of T7 RNAP, which is under the control of the native *agn43* regulatory region (which includes the first 48 bp of the coding sequence[7]), activates a T7 RNAP–specific promoter that regulates expression of the green fluorescent protein (*GFP*) gene located on a plasmid. The creation of multiple *GFP* transcripts by each T7 RNAP molecule and the placement of *GFP* on a multicopy plasmid increased the fluorescence signal and improved the signal-to-noise ratio. The amplification system enabled colonies and single cells in the on and off states of the wild-type *agn43* strain (cHNL135) to be distinguished by fluorescence microscopy and flow cytometry (**Fig. 1b**,**c**).

Overexpression of Dam (cHNL140) or unopposed Dam activity owing to the deletion of OxyR (cHNL26) shifted the system towards the methylated states (M_H and M_F). The distribution was unimodal with full transcription ('on'), indicating the M_H and M_F states have the same expression level and/or one of the states predominates. In support of the latter, most hemimethylated GATC sites become fully methylated soon after passage of the DNA replication machinery[19,20], suggesting that the on cells are predominately in the M_F state. The overexpression of OxyR (cHNL149) or unopposed OxyR binding

[1]Department of Physics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA. [2]Department of Integrative Biology, University of California, Berkeley, California 94720, USA. Correspondence should be addressed to H.N.L. (hanlim@berkeley.edu).

Received 31 August 2006; accepted 4 December 2006; published online 14 January 2007; doi:10.1038/ng1956

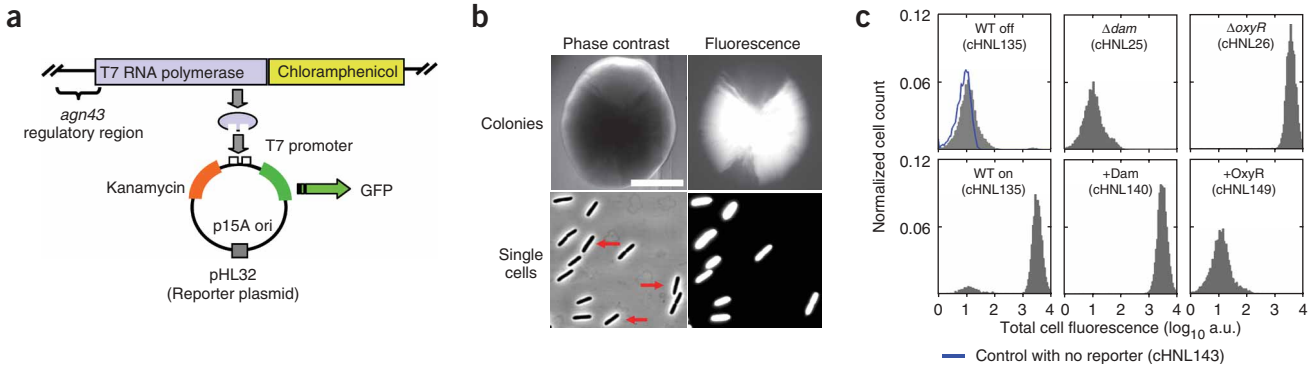

Figure 1 Genetic amplification system for measuring *agn43* expression and switching. (**a**) Genetic amplification system. (**b**) Phase-contrast and fluorescence microscopy images showing regions of 'on' and 'off' cells in a single wild-type colony (cHNL135; top) and individual on and off cells in liquid culture (bottom). Scale bar, ~1 mm; arrows indicate cells that are off and therefore not visible in the fluorescence image. (**c**) Histograms of single-cell fluorescence for each strain (all carrying the genetic amplification system) in M9 media supplemented with 0.4% glycerol and measured by flow cytometry. Left, wild-type (WT) background strain derived from two separate colonies, one of which is mainly off and the other of which is mainly on. Blue line indicates the control strain (cHNL143) with the pHL32 reporter plasmid but no T7 RNA polymerase. Middle and right, Dam and OxyR deletion strains (top) and overexpression strains (bottom), respectively.

owing to the deletion of Dam (cHNL25) prevented methylation and caused inactive transcription ('off') as expected. The observed behavior of *agn43* using the genetic amplifier is consistent with previous LacZ studies[11,13,18].

The deletion of both Dam and OxyR (cHNL172) trapped cells in an unmethylated state without OxyR bound to the switch region. This strain had a 'partial' expression state that was in between the on and off states (**Fig. 2a**,**b**). Partial expression in the absence of OxyR and Dam has also been reported *in vivo* with the LacZ reporter[13] and with an *in vitro* transcription assay[13]. The partial state indicates that expression is not simply switched on and off by OxyR unbinding and binding to the switch region respectively; otherwise full transcription would occur in the absence of OxyR. Furthermore it indicates that DNA methylation is necessary for the on state.

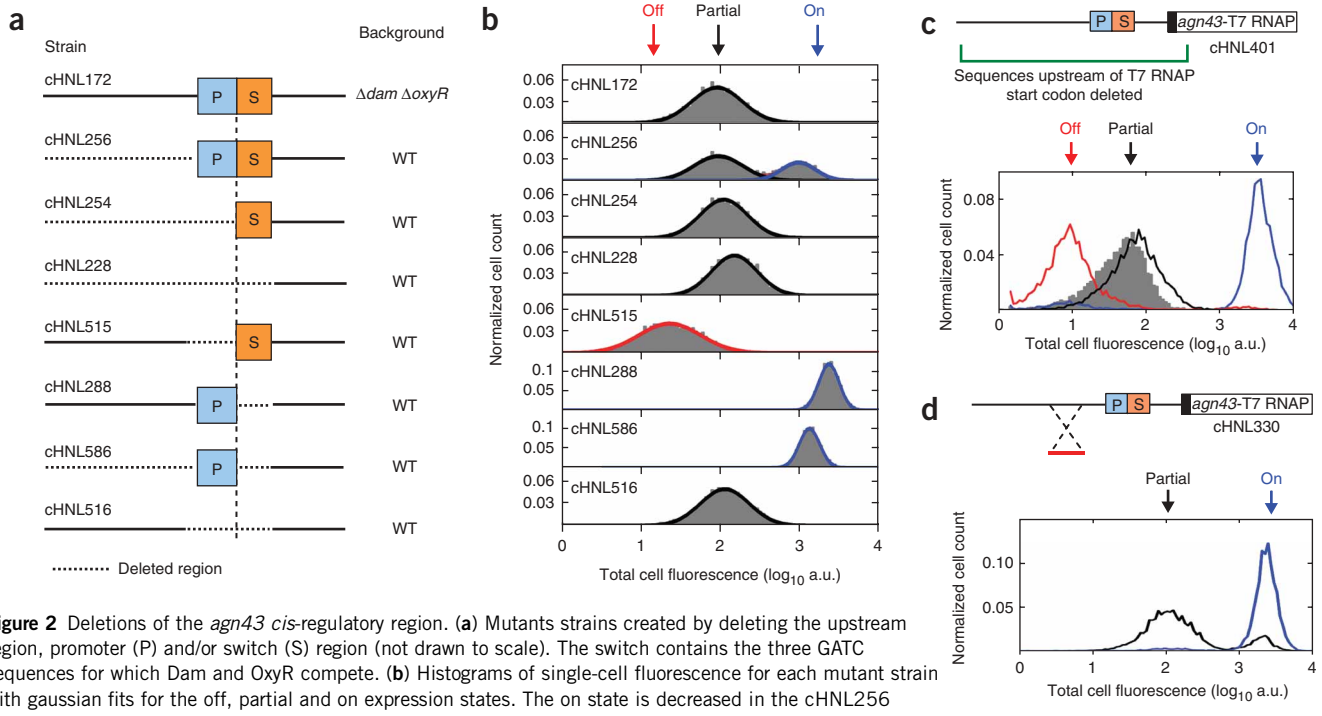

Figure 2 Deletions of the *agn43 cis*-regulatory region. (**a**) Mutants strains created by deleting the upstream region, promoter (P) and/or switch (S) region (not drawn to scale). The switch contains the three GATC sequences for which Dam and OxyR compete. (**b**) Histograms of single-cell fluorescence for each mutant strain with gaussian fits for the off, partial and on expression states. The on state is decreased in the cHNL256 and cHNL586 stains, indicating that sequences upstream of the promoter are required for full transcription. (**c**) Deletion of all *agn43* regulatory sequences (cHNL401) results in a partial state (gray columns) that has slightly lower expression than in the OxyR and Dam double deletion strain, cHNL172 (black line). This lower expression can be explained by loss of the ribosomal binding site (black box). The on and off states for the wild-type strain (cHNL135) are shown for comparison. (**d**) Replacement of a portion of the upstream region with an alternative sequence of identical size (cHNL330). The alternative sequence (red line) is the residual scar that follows excision of the kanamycin resistance cassette by the lambda Red method[30].

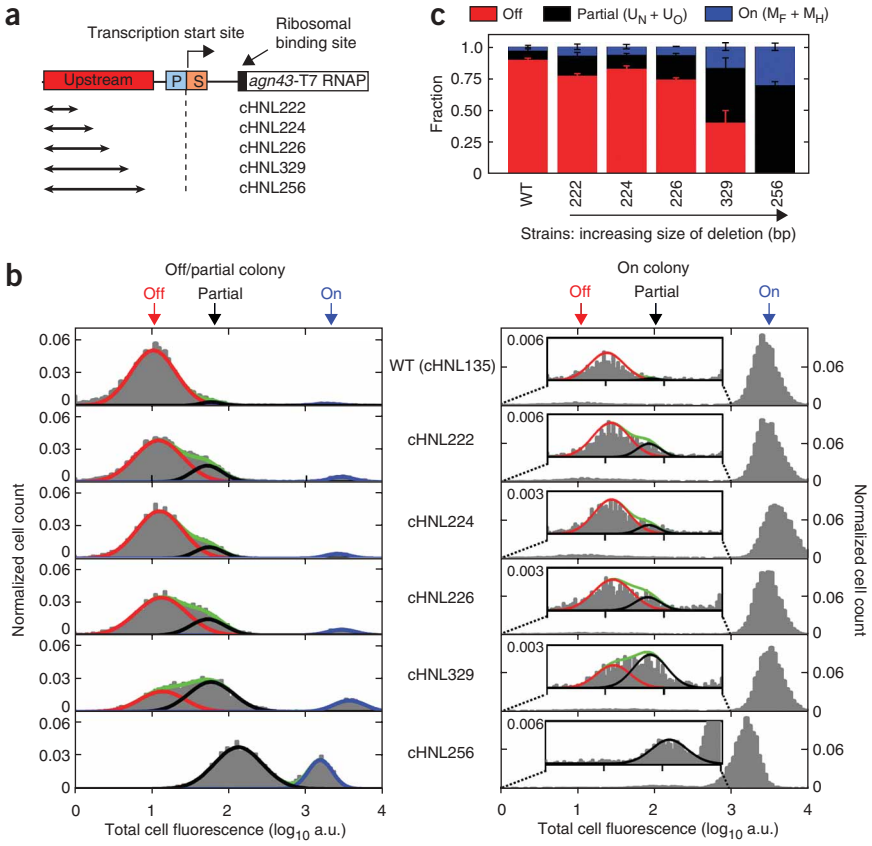

Figure 3 Deletion mapping of sequences upstream of the *agn43* promoter. (**a**) Deletion mapping strategy showing the promoter (P) and switch (S) regions (not drawn to scale). The ribosomal binding site is represented by a black box; the double-headed arrows indicate the region that is deleted. (**b**) Representative histograms of cultures derived from off/partial and on colonies for each strain. Gaussian fits for the off (red), partial (black) and on (blue) expression states were determined for cultures derived from unmethylated colonies. The number of unmethylated cells in the cultures derived from methylated colonies was insufficient for curve fitting; therefore, curves from the off/partial cultures were superimposed to show the similarity in the distributions (inset). Owing to the long half-life of GFP, the partial expression level is influenced by the amount of time that cells spend in the off and partial states. As a result, maximum expression of the partial state should occur in strain cHNL256, where cells cannot switch to the off state, as was observed. (**c**) Relative fraction of cells (±s.e.m.) in the off, partial and on states in cultures derived from unmethylated colonies for each upstream-deletion strain.

The origin of the partial expression state was investigated by deleting the promoter, switch region and the sequence upstream of the promoter in various combinations (**Fig. 2a**). The partial state was observed in mutants without the promoter (cHNL254, cHNL228 and cHNL516; **Fig. 2b**) and when the entire *agn43 cis*-regulatory region was deleted (cHNL401; **Fig. 2c**), showing that partial expression is not specifically due to the *agn43* promoter or to other sequences in this region. In addition, deleting a portion of the upstream sequence and replacing it with an equivalently sized alternative sequence showed that partial expression is not due to the fact that the deletions shift the sequences of upstream genes closer to *agn43* (cHNL330; **Fig. 2d**). Together, these data suggest that partial expression arises from non-specific transcription. The partial expression can be repressed resulting in the off state when OxyR, the upstream sequences and the switch region are present (for example, cHNL135; **Fig. 1c**).

The upstream region is known only to contain the non-essential YeeP pseudogene[21]; there are no known *agn43* regulatory sites. We therefore systematically deleted the upstream region, keeping the distal end of the deletion fixed and extending the proximal end towards the promoter, in successive mutants to identify specific sequences necessary for the off state (**Fig. 3a**). For each strain, single-cell *agn43* expression was measured. We observed that, each time the deletion was extended, the fraction of cells in the off state decreased and the fraction in the partial state increased (**Fig. 3b**,**c**). The decrease in the off fraction was greatest when the most proximal sequences (~69 to 126 bp upstream from the center of the switch region) were deleted (**Fig. 4a**).

The organization of the regulatory region with the promoter positioned between two regions required for repression[22] (proximal and switch), coupled with the observation that deletions far from the transcription start site can decrease the off fraction, suggested that the repression might involve a DNA conformational change. DNase I footprinting studies have not identified an OxyR-binding site in the proximal region[13]; therefore, it is unlikely that DNA looping occurs between OxyR molecules bound at the proximal and switch regions in a manner similar to that described for the Gal and Lac repressors[22,23]. An alternative mechanism is that OxyR, which has DNA-bending properties[24], enables the switch region to function as a hinge that brings the 'proximal' region in contact with a 'downstream' region (**Fig. 4b**). The model predicts that downstream sequences between the switch region and the start codon are crucial for repressing partial expression, and this prediction was confirmed by loss of the off state in a mutant in which these sequences were deleted (cHNL332; **Fig. 4c**).

The above experiments identified two different partial states. One partial state occurs when the OxyR-binding site and OxyR expression are intact but the upstream sequences or downstream are deleted (for example, cHNL330 and cHNL332; **Figs. 2d** and **4c**), which we designate U_O (unmethylated plus OxyR). The other partial state occurs when there is a 'naked' switch region without DNA methylation or bound OxyR (cHNL172), which we designate U_N (unmethylated plus no OxyR; **Fig. 2b**). The essential difference between the U_O and the U_N states is, respectively, the presence or absence of OxyR binding. Both states have the same partial expression, which indicates that OxyR binding by itself does not alter expression. OxyR binding does, however, block DNA methylation. As a result, the U_O partial expression state can stably exist in the presence of Dam (cHNL256, **Fig. 2b**), whereas the naked switch region (U_N state) is unstable because it cannot prevent DNA methylation (cHNL26, **Fig. 1c**). Because OxyR binding can occur in the absence of the repression mechanism, but not vice versa, we propose that the off state is achieved in two consecutive steps: the unmethylated switch region (U_N) is first bound by OxyR (U_O), and then undergoes repression (off; **Fig. 5**).

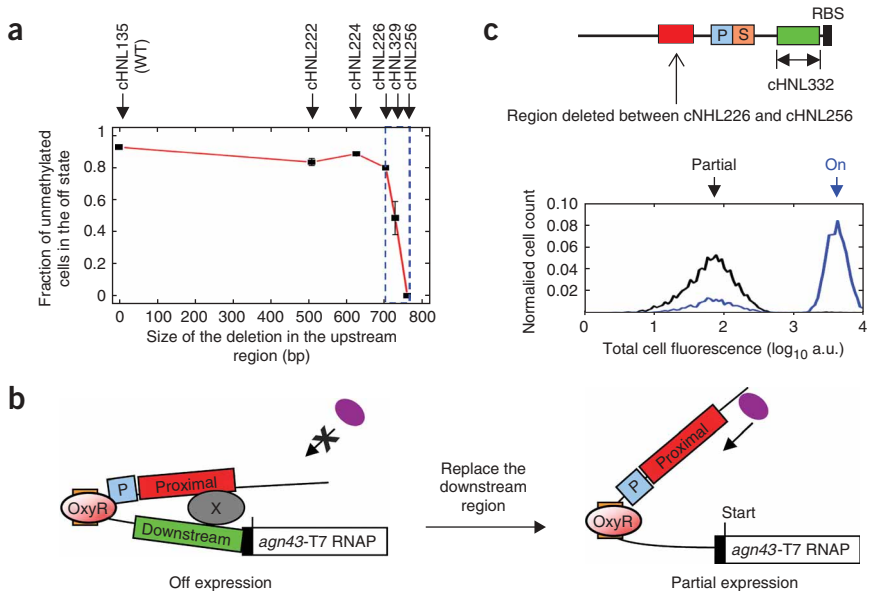

Figure 4 Sequences necessary for maintaining the off state. (**a**) Fraction of unmethylated cells in the off state as a function of the size of the upstream sequence deleted. Data are the mean ± s.e.m.; the connecting red line is a guide for the eye. Deletion of the 'proximal' sequence, ~69 to 126 bp upstream of the center of the switch region (bounded by broken blue lines), caused the greatest reduction in the off fraction. (**b**) Proposed mechanism for generating the off state, involving an interaction between the proximal region, a 'downstream' sequence and OxyR (represented by a single molecule) that blocks partial transcription (represented by the purple ellipse). X is an unknown protein or factor that might mediate this interaction. (**c**) The downstream region (69–186 bp from the center of the switch region) was replaced with the lambda Red scar to produce cHNL332. Red, green and black boxes represent the proximal sequence, downstream sequence and ribosomal binding site (RBS), respectively. Double-headed arrow indicates the region deleted.

We constructed a kinetic model to determine how transitions between the two methylated states (M_F and M_H) and the three unmethylated states (U_N, U_O and off) regulate *agn43* switching (**Fig. 5**). The fraction of hemimethylated cells that become unmethylated per generation (M_H to U_N transition), termed 'methylated loss', is given by $0.5e^{-k_M\tau}$, where k_M is the rate at which the hemimethylated switch region becomes fully methylated and τ is the DNA replication period (**Supplementary Methods** online). The fraction of unmethylated cells that become hemimethylated (U_N to M_H transition), termed 'unmethylated loss', occurs at a rate of $k_H F_{UN}\tau$ cells per generation, where k_H is the rate of U_N to M_H transitions (k_H is estimated to be 30–40 times slower than k_M; ref. 25) and F_{UN} is the fraction of unmethylated cells in the U_N state. From the expressions for methylated and unmethylated loss and reported on-off switching rates[16], we determined the transition rates and the distribution of cells between the states.

The fraction of cells in the off and partial states was the same in cultures derived from methylated and unmethylated colonies, indicating that transitions within the unmethylated states had reached steady state (**Fig. 3b**). The steady-state ratio of the U_O and U_N states provides K_O, the equilibrium constant for OxyR binding (k_O) and unbinding (k_{-O}). Similarly, the ratio of the off and U_O states determines K_R, the equilibrium constant for repression (k_R) and non-repression (k_{-R}) of partial expression. In the wild-type strain, K_O and K_R were 3.7 and 15.8, respectively, corresponding to a cell ratio of about 1:4:58 for the U_N:U_O:off states. The shift from the off state to the partial state caused by the upstream deletions (corresponding to a decreasing K_R) increases the number of cells in the U_O and U_N states (the U_O:U_N ratio is unaltered because the OxyR concentration and its binding site are unchanged). As K_R decreases, therefore, the fraction of unmethylated cells in the U_N state (F_{UN}) increases from 1.6% in

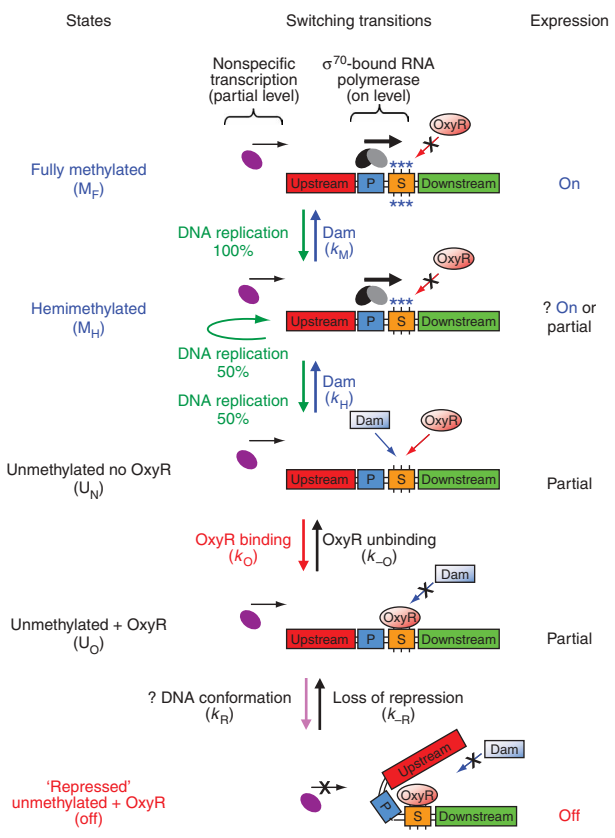

Figure 5 Molecular model of the steps involved in *agn43* on-off switching. Partial transcription arising from non-specific RNA polymerase activity is indicated by the purple ellipse. Transcription in the on state occurs from the promoter site containing the −35 and −10 sites for σ^{70} (black and gray ellipses) when the switch region is fully methylated (blue asterisks on both sides of the switch region). At DNA replication, the fully methylated switch region becomes hemimethylated (blue asterisks on an arbitrary side). The hemimethylated switch can block OxyR binding (less effectively than full DNA methylation); it is unknown whether this transient state has partial or on expression. The hemimethylated switch can be converted to the fully methylated state by Dam; if this does not occur, then half of the hemimethylated switch regions will become unmethylated (U_N state) at the next round of DNA replication. Dam and OxyR compete for binding to the unmethylated switch region. Binding of OxyR to the unmethylated switch region (U_O state) blocks Dam, thereby preventing DNA methylation. The U_O state can undergo a conformational change in DNA, resulting in the off state.

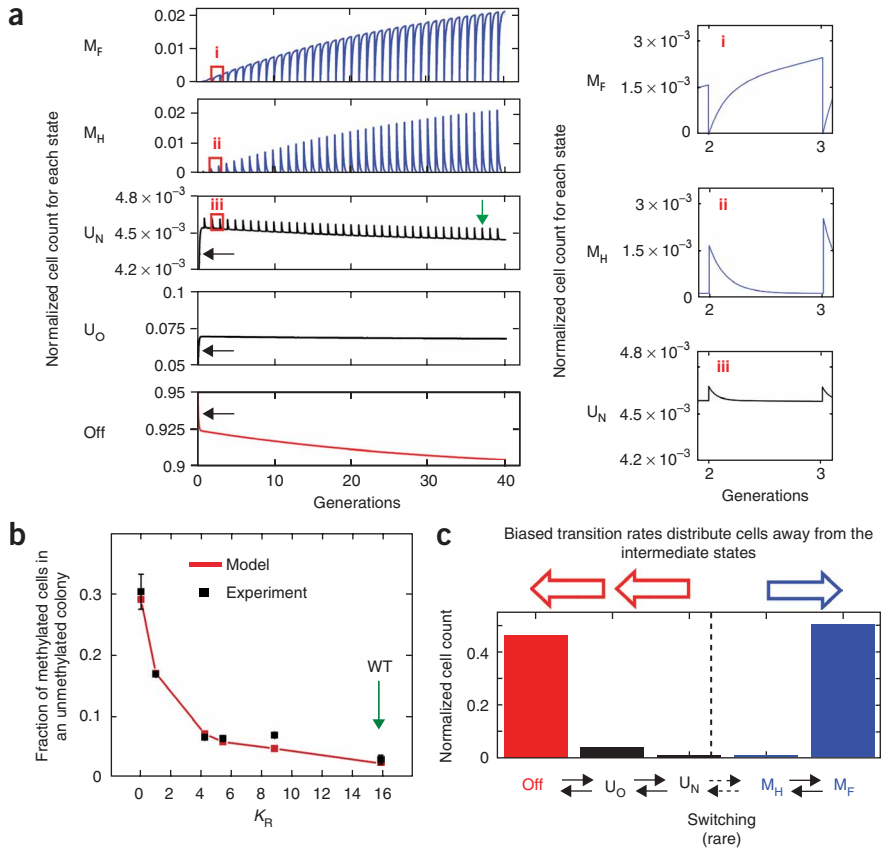

Figure 6 Simulation of *agn43* switching. (**a**) Simulated number of cells (normalized cell count) in each of the five states for a synchronously replicating population initialized with cells in the off state. The M_H and M_F states oscillate, and the U_N state (green arrow) shows spiking owing to transitions that follow DNA replication (shown in greater detail for a single generation on the right). The intermediate states show low occupancy. (**b**) Comparison of the simulated (red squares with connecting line to guide the eye) and observed (black squares with error bars indicating s.e.m.) fraction of methylated cells in an unmethylated culture at the K_R values measured in the wild-type (WT) and mutant strains. (**c**) Distribution of cells (assuming an equal fraction of methylated and unmethylated states) towards the off state and the fully methylated state (M_F) owing to the bias in the transition rate. Broken arrows indicate rare transitions.

the wild-type strain to 26.8% in the strain without the off state (cHNL256).

The increase in F_{UN} in the upstream-deletion strains should increase switching from the unmethylated state to the methylated state in accordance with the expression for unmethylated loss. To compare this prediction quantitatively with the experimental data, the dynamics of the *agn43* system were simulated with the above-derived parameters and cell growth (**Fig. 6a**). The K_R value obtained for each of the upstream-deletion mutant strains was substituted into the simulation to predict the fraction of methylated cells that should arise in an unmethylated colony. The simulations showed that increasing K_R (that is, an increasing fraction of unmethylated cells in the off state) decreases the fraction of methylated cells as predicted (**Fig. 6b**). In addition, the simulated and the experimentally measured fraction of methylated cells were in quantitative agreement (**Fig. 6b**). This observation supports the model's prediction that repressing partial expression enhances the stability of the unmethylated states by reducing the fraction in the U_N state, which consequently reduces the number of cells that can switch to the methylated states (**Fig. 6c**). Switching from the methylated to the unmethylated states is similarly reduced by the rapid conversion of hemimethylated DNA to fully methylated DNA such that few cells are in the M_H state at DNA replication.

The transition bias away from the hemimethylated state and the intermediate unmethylated states towards the fully methylated and off states, respectively, acts as a buffer that 'captures' cells as they leave the on and off states and returns them to their original state. As a consequence, cells that leave the on state owing to DNA replication, or that leave the off state owing to fluctuations in repression of the partial state, rarely switch to the opposite state. The stability of the methylated and unmethylated states is also coupled to the time delay that occurs between one round of DNA replication and the next. Increasing the delay increases the stability of the methylated states because the DNA methylation reactions can proceed for longer (more M_H to M_F transitions). By contrast, decreasing the time delay increases the stability of the unmethylated states because it decreases the time that cells in the U_N state can move to the M_H state. A time delay to regulate the fraction of product formed is also used in kinetic proofreading to increase the fidelity of nucleic acid and protein synthesis[26].

The *agn43* switching mechanism has two key properties that are absent in feedback systems and may prove advantageous in the construction of synthetic circuits. First, cellular memory is determined by the DNA methylation state and not by the concentration of cytoplasmic protein. This property could facilitate the construction of simpler circuits because the protein that is controlled by the switch does not need to regulate its own expression either directly or indirectly. In addition, it is theoretically possible for expression states to be transferred to a cell by genetic material alone, as occurs when viral genomes such as Mu are injected into host cells[27]. Second, unlike feedback regulation, the state of the system is not determined by the expression level. Therefore, the expression states of the *agn43* switch could be modified without affecting switching rates. For example, the system could be engineered such that the promoter is inactive unless a transcriptional activator is also present; in this case, full expression would occur only if both the switch region was methylated and the activator was present.

A similar switching mechanism to *agn43* seems to control expression of the bacteriophage Mu *mom* operon and also involves Dam and OxyR competing for a switch region with three GATC sequences[27]. Other types of repressor or activator protein in *E. coli* and other organisms that block DNA methylation and whose binding is prevented by DNA methylation[1,28] might also form mechanistically similar epigenetic switches. The potential flexibility of the *agn43* switch makes it an attractive starting point for designing artificial switches and indicates that other naturally occurring, multistep epigenetic switches with biased transition rates may exist in various configurations.

METHODS

Genetic amplification system, bacterial strains and plasmids. A plasmid (pHL16) was constructed containing the *agn43* promoter regulating the T7 RNA polymerase gene. The *agn43* regulatory region and the gene encoding T7 RNA polymerase were amplified by PCR from genomic DNA derived from the *E. coli* strains MG1655 (Coli Genetic Stock Center) and BL21-Gold(DE3) (Stratagene), respectively. The chloramphenicol resistance gene (Cam^R) was amplified by PCR from the pZA31 plasmid[29] and inserted adjacent to the T7 RNA polymerase gene. The whole pHL16 insert (*agn43* regulatory sequence, T7 RNA polymerase and Cam^R) was used to replace the *agn43* gene in the MC4100 *E. coli* strain by means of the lambda Red disruption system[30]. The T7 RNA polymerase sequence contained two mutations (leading to the amino acid changes F509S and S686C) that were introduced during cloning and integration. All deletion strains were generated with the lambda Red disruption system and the deletions were confirmed by PCR amplification and DNA sequencing (further details are provided in **Supplementary Tables 1** and **2** and **Supplementary Fig. 1** online). OxyR deletion strains were grown on plates supplemented with catalase[7].

The reporter plasmid (pHL32) was constructed by inserting *GFP* under the control of a T7 RNA polymerase–specific promoter into pZA21 (ref. 29). The T7 RNA polymerase promoter and *GFP* were amplified by PCR from the pET-11a plasmid (Stratagene) and the pGFPmut3.1 plasmid (Clontech), respectively. The stop codon of *GFP* was followed by a T7 RNA polymerase terminator sequence. The OxyR overexpression plasmid (pHL34) was created by PCR amplification of the OxyR coding region and inserted into the multiple cloning site of pZE11 under the control of the pTetO-1 promoter[29].

Data acquisition and analysis. Cells were streaked on LB plates with appropriate antibiotics. Single colonies in the methylated and unmethylated states were selected by fluorescence microscopy and inoculated into M9 media with 0.4% glycerol, 34 μg/ml of chloramphenicol and 50 μg/ml of kanamycin. After 14 h of growth, cells were inoculated into fresh media for a further 6 h of growth, and then single-cell measurements were made with a Becton-Dickinson FACScan flow cytometer (488-nm excitation laser; 525-nm emission filter). Data from 10,000 individual cells were collected for each sample. FACScan data were converted to ASCII format using MFI (E. Martz, University of Massachusetts, Amherst) and analyzed with MATLAB (MathWorks). Histograms were fitted to gaussian distributions with Origin 7.5 (Microcal Software).

The growth rates of on and off cells, as measured by the doubling time, were similar in Luria Bertani media (33.1 ± 0.8 min and 34.5 ± 1.2 min, respectively); however, in M9 media plus 0.4% glycerol, the doubling time of the on cells was longer than that of the off cells (90.1 ± 0.1 min and 82.6 ± 0.1 min, respectively). The difference in growth rate is presumably due to the metabolic burden of high GFP expression in the on state, which becomes apparent in M9 media owing to the lower nutrient availability.

Mathematical model and simulation. A deterministic model was used because the data were collected from a large cell population grown over many generations. The number of cells in each state (M_F, M_H, U_N, U_O and off) is determined by the switching reactions, cell growth and DNA replication. The switching reactions for cells in the M_F, M_H, U_N, U_O and off states are consecutive components of $\frac{d}{dt}\mathbf{x}^j(t)$, which describes the changes in each state in generation j, from the end of one round of DNA replication ($t = 0$) until the start of the next ($t = \tau$).

$$\frac{d\mathbf{x}^j(t)}{dt} = \begin{pmatrix} k_M\mathbf{M_H} + \gamma_M \\ k_H\mathbf{U_N} - k_M\mathbf{M_H} + \gamma_H \\ k_{-O}\mathbf{U_O} - (k_O + k_H)\mathbf{U_N} + \gamma_N \\ k_{-R}\mathbf{Off} + k_O\mathbf{U_N} - (k_{-O} + k_R)\mathbf{U_O} + \gamma_O \\ -k_{-R}\mathbf{Off} + k_R\mathbf{U_O} + \gamma_R \end{pmatrix}$$

k_M, k_H, k_O, k_{-O}, k_R and k_{-R} are rate constants defined as the number of cell transitions per generation. Cell growth is an exponential process occurring at rate γ_M, γ_H, γ_N, γ_O and γ_R for the M_F, M_H, U_N, U_O and off states, respectively, without nutrient depletion. Integrating $\frac{d}{dt}x^j(t)$ at τ specifies the number of cells in each state immediately before DNA replication. Transforming $\mathbf{x}^j(\tau)$ by $\mathbf{D}$, a matrix incorporating the DNA replication–associated transitions, provides the initial conditions for the next generation $j+1$; that is,

$$\mathbf{x}^{j+1}(0) = \mathbf{D}\mathbf{x}^j(\tau),$$

where

$$\mathbf{D} = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0.5 & 0.5 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

Iterative simulations with Matlab determined the population distribution at 36 generations (22 generations on LB agar and 14 generations in M9 medium liquid culture), which was the approximate number in the experiments (further details are given in **Supplementary Methods**).

Note: Supplementary information is available on the Nature Genetics website.

ACKNOWLEDGMENTS

We thank E.A. Oakenfull for logistical assistance; M. Thattai for discussions and assistance with data analysis; M. van der Woude for bacterial strains and plasmids; R. Lutz and H. Bujard for pZ expression plasmids; and A. Becskei, S.V. Godoy and J.M. Pedraza for suggestions and comments. This work was supported by grants from the National Science Foundation (PHY-0548484) and the National Institutes of Health (R01-GM077183).

COMPETING INTERESTS STATEMENT

The authors declare that they have no competing financial interests.

Published online at http://www.nature.com/naturegenetics
Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions

1. Jaenisch, R. & Bird, A. Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. *Nat. Genet.* **33** (Suppl.), 245–254 (2003).
2. Low, D.A., Weyand, N.J. & Mahan, M.J. Roles of DNA adenine methylation in regulating bacterial gene expression and virulence. *Infect. Immun.* **69**, 7197–7204 (2001).
3. Gardner, T.S., Cantor, C.R. & Collins, J.J. Construction of a genetic toggle switch in *Escherichia coli*. *Nature* **403**, 339–342 (2000).
4. Isaacs, F.J., Hasty, J., Cantor, C.R. & Collins, J.J. Prediction and measurement of an autoregulatory genetic module. *Proc. Natl. Acad. Sci. USA* **100**, 7714–7719 (2003).
5. Xiong, W. & Ferrell, J.E., Jr. A positive-feedback-based bistable 'memory module' that governs a cell fate decision. *Nature* **426**, 460–465 (2003).
6. Klose, R.J. & Bird, A.P. Genomic DNA methylation: the mark and its mediators. *Trends Biochem Sci* (2006).
7. Haagmans, W. & van der Woude, M. Phase variation of Ag43 in *Escherichia coli*: Dam-dependent methylation abrogates OxyR binding and OxyR-mediated repression of transcription. *Mol. Microbiol.* **35**, 877–887 (2000).
8. Owen, P., Caffrey, P. & Josefsson, L.G. Identification and partial characterization of a novel bipartite protein antigen associated with the outer membrane of *Escherichia coli*. *J. Bacteriol.* **169**, 3770–3777 (1987).
9. Henderson, I.R., Meehan, M. & Owen, P. Antigen 43, a phase-variable bipartite outer membrane protein, determines colony morphology and autoaggregation in *Escherichia coli* K-12. *FEMS Microbiol. Lett.* **149**, 115–120 (1997).
10. Henderson, I.R. & Owen, P. The major phase-variable outer membrane protein of *Escherichia coli* structurally resembles the immunoglobulin A1 protease class of exported protein and is regulated by a novel mechanism involving Dam and oxyR. *J. Bacteriol.* **181**, 2132–2141 (1999).
11. Waldron, D.E., Owen, P. & Dorman, C.J. Competitive interaction of the OxyR DNA-binding protein and the Dam methylase at the antigen 43 gene regulatory region in *Escherichia coli*. *Mol. Microbiol.* **44**, 509–520 (2002).
12. Henderson, I.R., Meehan, M. & Owen, P. A novel regulatory mechanism for a novel phase-variable outer membrane protein of *Escherichia coli*. *Adv. Exp. Med. Biol.* **412**, 349–355 (1997).
13. Wallecha, A., Munster, V., Correnti, J., Chan, T. & van der Woude, M. Dam- and OxyR-dependent phase variation of agn43: essential elements and evidence for a new role of DNA methylation. *J. Bacteriol.* **184**, 3338–3347 (2002).
14. Owen, P., Meehan, M., de Loughry-Doherty, H. & Henderson, I. Phase-variable outer membrane proteins in *Escherichia coli*. *FEMS Immunol. Med. Microbiol.* **16**, 63–76 (1996).
15. Correnti, J., Munster, V., Chan, T. & Woude, M. Dam-dependent phase variation of Ag43 in *Escherichia coli* is altered in a seqA mutant. *Mol. Microbiol.* **44**, 521–532 (2002).

16. Hasman, H., Schembri, M.A. & Klemm, P. Antigen 43 and type 1 fimbriae determine colony morphology of *Escherichia coli* K-12. *J. Bacteriol.* **182**, 1089–1095 (2000).
17. Kullik, I., Toledano, M.B., Tartaglia, L.A. & Storz, G. Mutational analysis of the redox-sensitive transcriptional regulator OxyR: regions important for oxidation and transcriptional activation. *J. Bacteriol.* **177**, 1275–1284 (1995).
18. Wallecha, A., Correnti, J., Munster, V. & van der Woude, M. Phase variation of Ag43 is independent of the oxidation state of OxyR. *J. Bacteriol.* **185**, 2203–2209 (2003).
19. Stancheva, I., Koller, T. & Sogo, J.M. Asymmetry of Dam remethylation on the leading and lagging arms of plasmid replicative intermediates. *EMBO J.* **18**, 6542–6551 (1999).
20. Campbell, J.L. & Kleckner, N. The rate of Dam-mediated DNA adenine methylation in *Escherichia coli*. *Gene* **74**, 189–190 (1988).
21. Gerdes, S.Y. *et al.* Experimental determination and system level analysis of essential genes in *Escherichia coli* MG1655. *J. Bacteriol.* **185**, 5673–5684 (2003).
22. Geanacopoulos, M., Vasmatzis, G., Zhurkin, V.B. & Adhya, S. Gal repressosome contains an antiparallel DNA loop. *Nat. Struct. Biol.* **8**, 432–436 (2001).
23. Oehler, S., Alberti, S. & Muller-Hill, B. Induction of the *lac* promoter in the absence of DNA loops and the stoichiometry of induction. *Nucleic Acids Res.* **34**, 606–612 (2006).
24. Sun, W. & Hattman, S. *Escherichia coli* OxyR protein represses the unmethylated bacteriophage Mu mom operon without blocking binding of the transcriptional activator C. *Nucleic Acids Res.* **24**, 4042–4049 (1996).
25. Urig, S. *et al.* The *Escherichia coli* dam DNA methyltransferase modifies DNA in a highly processive reaction. *J. Mol. Biol.* **319**, 1085–1096 (2002).
26. Hopfield, J.J. Kinetic proofreading: a new mechanism for reducing errors in biosynthetic processes requiring high specificity. *Proc. Natl. Acad. Sci. USA* **71**, 4135–4139 (1974).
27. Hattman, S. Unusual transcriptional and translational regulation of the bacteriophage Mu mom operon. *Pharmacol. Ther.* **84**, 367–388 (1999).
28. Hernday, A.D., Braaten, B.A. & Low, D.A. The mechanism by which DNA adenine methylase and PapI activate the pap epigenetic switch. *Mol. Cell* **12**, 947–957 (2003).
29. Lutz, R. & Bujard, H. Independent and tight regulation of transcriptional units in *Escherichia coli* via the LacR/O, the TetR/O and AraC/I1–I2 regulatory elements. *Nucleic Acids Res.* **25**, 1203–1210 (1997).
30. Datsenko, K.A. & Wanner, B.L. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc. Natl. Acad. Sci. USA* **97**, 6640–6645 (2000).