27. M. C. Saraf, A. R. Horswill, S. J. Benkovic, C. D. Maranas, *Proc. Natl. Acad. Sci. U.S.A.* **101**, 4142 (2004).
28. J. M. Christie, M. Salomon, K. Nozue, M. Wada, W. R. Briggs, *Proc. Natl. Acad. Sci. U.S.A.* **96**, 8779 (1999).
29. M. Salomon, J. M. Christie, E. Knieb, U. Lempert, W. R. Briggs, *Biochemistry* **39**, 9401 (2000).
30. J. M. Christie, T. E. Swartz, R. A. Bogomolni, W. R. Briggs, *Plant J.* **32**, 205 (2002).
31. S. M. Harper, J. M. Christie, K. H. Gardner, *Biochemistry* **43**, 16184 (2004).
32. X. Yao, M. K. Rosen, K. H. Gardner, *Nat. Chem. Biol.* **4**, 491 (2008).
33. C. A. Fierke, K. A. Johnson, S. J. Benkovic, *Biochemistry* **26**, 4085 (1987).
34. J. T. Chen, K. Taira, C. P. Tu, S. J. Benkovic, *Biochemistry* **26**, 4093 (1987).
35. D. M. Epstein, S. J. Benkovic, P. E. Wright, *Biochemistry* **34**, 11037 (1995).
36. J. Kuriyan, D. Eisenberg, *Nature* **450**, 983 (2007).
37. We thank members of the Ranganathan and Benkovic labs for discussions, M. Rosen and K. H. Gardner for discussions and sharing of unpublished data and materials, and J. Wang for solvent accessibility calculations. Supported by the Robert A. Welch foundation (R.R.) and a grant from the Defense Advanced Research Projects Agency (R.R. and S.J.B.).

Supporting Online Material
www.sciencemag.org/cgi/content/full/322/5900/438/DC1
Materials and Methods
Figs. S1 to S8
Tables S1 to S4
References

14 April 2008; accepted 9 September 2008
10.1126/science.1159052

A Stochastic Single-Molecule Event Triggers Phenotype Switching of a Bacterial Cell

Paul J. Choi,* Long Cai,*† Kirsten Frieda,‡ X. Sunney Xie§

By monitoring fluorescently labeled lactose permease with single-molecule sensitivity, we investigated the molecular mechanism of how an *Escherichia coli* cell with the *lac* operon switches from one phenotype to another. At intermediate inducer concentrations, a population of genetically identical cells exhibits two phenotypes: induced cells with highly fluorescent membranes and uninduced cells with a small number of membrane-bound permeases. We found that this basal-level expression results from partial dissociation of the tetrameric lactose repressor from one of its operators on looped DNA. In contrast, infrequent events of complete dissociation of the repressor from DNA result in large bursts of permease expression that trigger induction of the *lac* operon. Hence, a stochastic single-molecule event determines a cell's phenotype.

Genetically identical cells in the same environment can exhibit different phenotypes, and a single cell can switch between distinct phenotypes in a stochastic manner (*1*–*4*). In the classic example of lactose metabolism in *Escherichia coli*, the *lac* genes are fully expressed for every cell in a population under high extracellular concentrations of inducers, such as the lactose analog methyl-β-D-thiogalactoside (TMG). However, at moderate inducer concentrations, the *lac* genes are highly expressed in only a fraction of a population, which may confer a fitness advantage for the entire population (*5*). We studied the molecular mechanism that controls the stochastic phenotype switching of a single cell.

Lactose metabolism is controlled by the *lac* operon (*6*, *7*), which consists of the *lacZ*, *lacY*, and *lacA* genes encoding β-galactosidase, lactose permease, and transacetylase, respectively. Expression of the operon is regulated by a transcription factor, the *lac* repressor (*8*), which dissociates from its specific binding sequences of DNA, the *lac* operators, in the presence of an inducer to allow transcription (Fig. 1A). The production of the permease increases inducer influx (*9*), which results in positive feedback on the permease expression level. Above a certain threshold of permease numbers, a cell will remain in a phenotype that is capable of lactose metabolism, and below this threshold, a cell will remain in a phenotype that is incapable of lactose metabolism (*10*, *11*). The former has high fluorescence from the cell membrane when the permease is labeled with a yellow fluorescent protein (YFP), whereas the latter has low fluorescence. The image in Fig. 1B shows the coexistence of both phenotypes in a cell population at intermediate inducer concentrations, characterized by the bimodal distributions of fluorescence intensity in Fig. 1C.

Although it is known that the bistability in the *lac* operon arises from positive feedback (*12*, *13*), the molecular mechanism underlying the initiation of switching between two phenotypes remains unclear. Novick and Weiner deduced that switching from the uninduced state to the induced state occurs through a single rate-limiting molecular process (*12*) rather than a multistep process. They further hypothesized that the random expression of one molecule of permease was enough to trigger induction. However, this has never been observed experimentally because of insufficient sensitivity.

Our group has previously shown that a single fluorescent protein molecule can be visualized in a living bacterial cell by using the method of detection by localization (*14*–*16*). Because a permease molecule is a membrane protein with slow diffusion, its fluorescence label is highly localized as compared with a fluorescent protein in the cytoplasm, which allows for its detection above the background of cellular autofluorescence. Thus, we generated strain SX700, which possesses intact *lac* promoter elements and expresses a functional LacY-YFP fusion protein (fig. S1) from *lacY*'s native chromosomal position (fig. S2). Figure 1B shows a fluorescence image of SX700 cells that allows us to directly count the number of single LacY-YFP molecules present in a cell.

Figure 1D shows the histogram of permease copy numbers in the uninduced fraction of cells, free from the complication of autofluorescence background in the lower peak of Fig. 1C. Evidently, the uninduced cells have 0 to 10 LacY molecules with significant probability, independent of the inducer concentration. Had one permease been enough to trigger induction, all cells in the uninduced subpopulation would have contained zero LacY molecules. Thus, we conclude that a single copy of LacY is not sufficient to induce switching of the phenotype, and the threshold for induction must be much higher than several molecules per cell.

We then set out to determine the threshold of permease molecules for induction. It is experimentally difficult to capture the rare events of phenotype switching in real time. To overcome this difficulty, we prepared cells covering a broad range of permease copy numbers by first fully inducing the cells and subsequently washing out the inducer. We then allowed the cells to divide for one to six generations, during which the initial permeases were partitioned into daughter cells (*17*). Figure 2A shows fluorescence time traces of the cells, normalized by cell size, upon the reintroduction of 40 μM TMG. Although the fluorescence in cells with low permease numbers continued to decay because of cell division and photobleaching, cells with more than 300 initial permease molecules induced again within 3 hours and showed increased fluorescence. Figure 2B shows the probability of induction as a function of initial permease number, which is well fit by a Hill equation with a Hill coefficient of 4.5 and a threshold of ~375 molecules. The large value of the threshold indicates that hundreds of permease molecules are necessary to switch the phenotype.

If the induction is due to a single rate-limiting event as Novick and Weiner argued, there must be a single large burst of permease expression to reach this high threshold. However, only small

Department of Chemistry and Chemical Biology, Harvard University, Cambridge, MA 02138, USA.
*These authors contributed equally to this work.
†Present address: Division of Biology, California Institute of Technology, Pasadena, CA 91125, USA.
‡Present address: Biophysics Program, Stanford University, Stanford, CA 94305, USA.
§To whom correspondence should be addressed. E-mail: xie@chemistry.harvard.edu

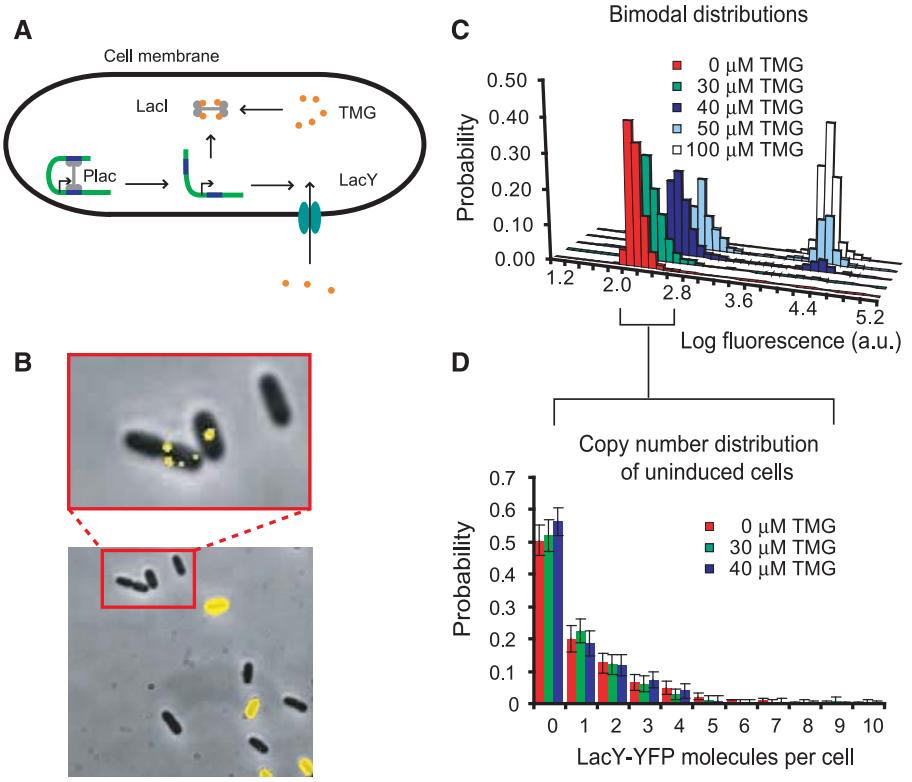

Fig. 1. The expression of lactose permease in *E. coli*. (**A**) The repressor LacI and permease LacY form a positive feedback loop. Expression of permease increases the intracellular concentration of the inducer TMG, which causes dissociation of LacI from the promoter, leading to even more expression of permeases. Cells with a sufficient number of permeases will quickly reach a state of full induction, whereas cells with too few permeases will stay uninduced. (**B**) After 24 hours in M9 medium containing 30 μM TMG, strain SX700 expressing a LacY-YFP fusion exhibits all-or-none fluorescence in a fluorescence-phase contrast overlay (bottom, image dimensions 31 μm × 31 μm). Fluorescence imaging with high sensitivity reveals single molecules of permease in the uninduced cells (top, image dimensions 8 μm × 13 μm). (**C**) After 1 day of continuous growth in medium containing 0 to 50 μM TMG, the resulting bimodal fluorescence distributions show that a fraction of the population exists either in an uninduced or induced state, with the relative fractions depending on the TMG concentration. (**D**) The distributions of LacY-YFP molecules in the uninduced fraction of the bimodal population at different TMG concentrations, measured with single-molecule sensitivity, indicate that one permease molecule is not enough to induce the *lac* operon, as previously hypothesized (*12*). More than 100 cells were analyzed at each concentration. Error bars are SE determined by bootstrapping.

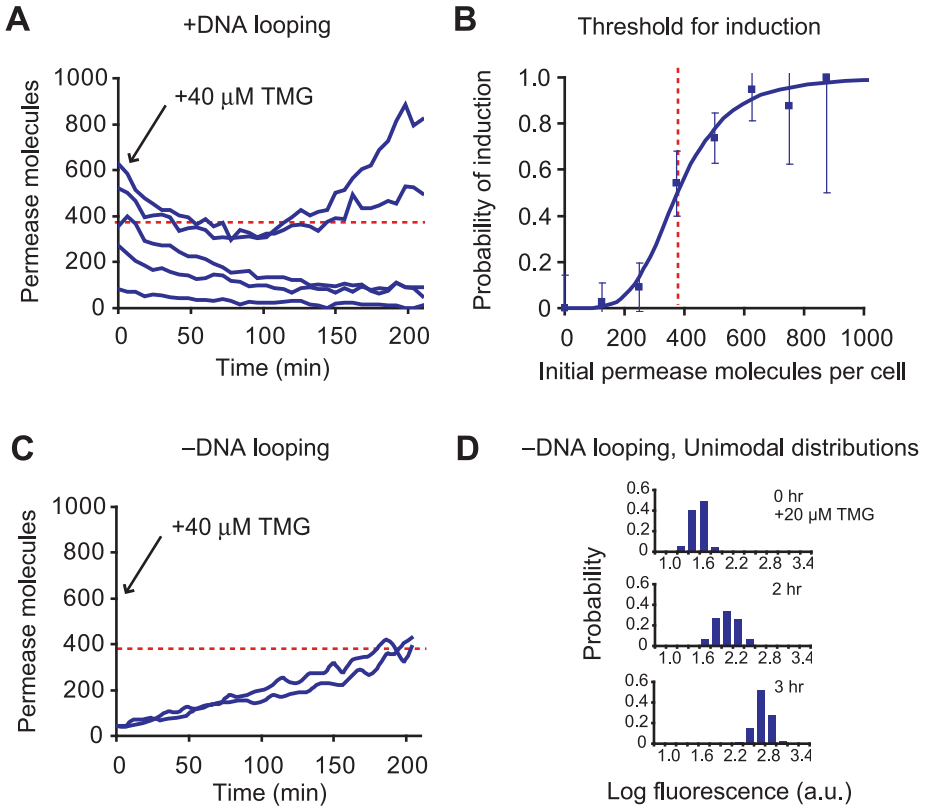

Fig. 2. Measurement of the threshold of permease molecules for induction. (**A**) Single-cell time traces of fluorescence intensity, normalized by cell size, starting from different initial permease numbers. The initial LacY-YFP numbers were prepared through dilution by cell division of fully induced cells after removal of the inducer. Upon adding 40 μM TMG at time 0, those cells with low initial permease numbers lost fluorescence with time as a result of dilution by cell division and photobleaching, whereas those cells with high initial permease numbers exhibited an increase in fluorescence as a result of reinduction. Permease molecule numbers were estimated from cell fluorescence (*28*). The dashed red line indicates the determined threshold. (**B**) The probability of induction of a cell within 3 hours as a function of the initial permease number was determined using traces from 90 cells. The probability of induction (p) was fit with a Hill equation $p = y^{4.5}/(y^{4.5} + 375^{4.5})$ for the initial permease number y. The threshold of permease numbers for induction was thus determined to be 375 molecules. Error bars are the inverse square root of the sample size at each point. (**C**) To prove that the complete dissociation of tetrameric repressor from two operators triggers induction, we constructed strain SX702 with auxiliary operators removed (no DNA looping). The figure shows single-cell traces of permease numbers in single cells grown in 40 μM TMG as a function of time. Unlike the looping strain SX700, the rapid induction of SX702 is no longer dependent on the initial number of permease molecules. This proves that phenotype switching is the result of a complete dissociation of the tetrameric repressor, as shown in (B). (**D**) In the absence of DNA looping, the entire population of strain SX702 rapidly induces in a coordinated manner from far below the threshold for a concentration as low as 20 μM TMG. DNA looping is necessary for bistability of the *lac* operon under these conditions.

bursts have been observed in previous studies of the repressed *lac* promoter (*14*, *18*). Large bursts, if any, would be infrequent and insufficiently sampled in the previous work. To explain why both small and large bursts may occur from the *lac* promoter, we considered the tetrameric repressor that is doubly bound to two operators, forming a DNA loop (*19*, *20*). We hypothesized that more frequent partial dissociations of the tetrameric repressor from one operator lead to single transcripts and small bursts of protein expression, whereas rare complete dissociations from all operators lead to multiple transcripts and large bursts of expression. In this model, the complete dissociation of the tetrameric repressor would be the molecular event that causes a change in phenotype.

To test this model experimentally, we first removed the auxiliary operators from SX700 to generate SX702 (fig. S2), which cannot form DNA loops. As a result, every dissociation event of LacI from the remaining O1 operator results in a complete dissociation event that would generate a large burst of expression. In this case, we predicted frequent and large bursts of expression with

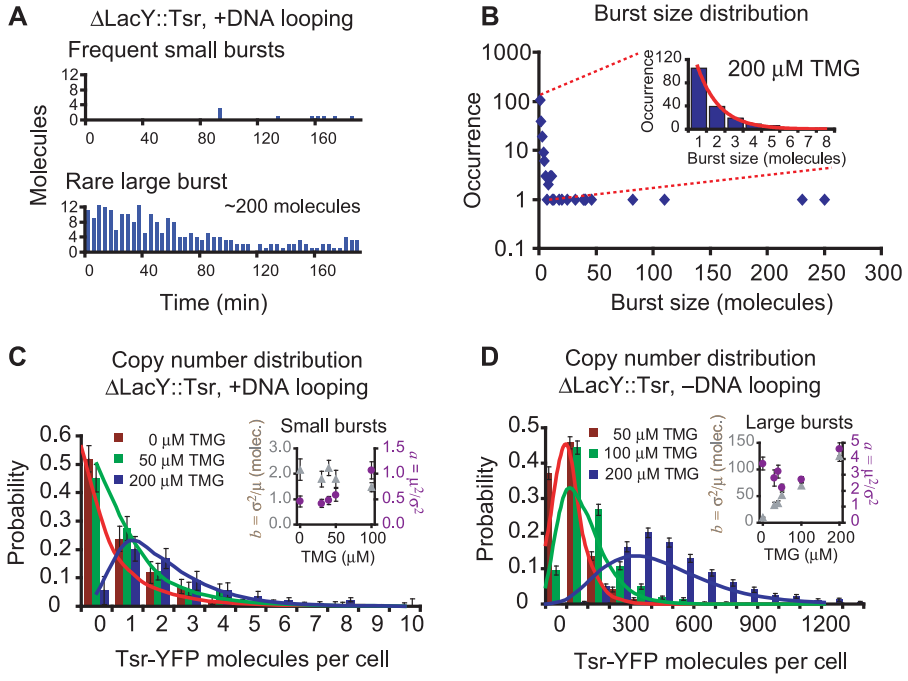

Fig. 3. Small and large bursts in the absence of positive feedback. In order to eliminate positive feedback from permease transport, we constructed strain SX701 by replacing the lactose permease with the membrane protein fusion Tsr-YFP. (**A**) Real-time traces of protein production in SX701 in 200 μM TMG. The fluorescent proteins were photobleached immediately after detection to ensure that only newly produced proteins were measured after each 4-min interval. Representative traces of single cells show frequent small bursts, associated with partial dissociation events, and a rare large burst, associated with a complete dissociation event, of protein production. (**B**) Distribution of burst sizes determined for 208 bursts from the real-time experiment depicted in (A). Although the majority of bursts are small, a number of unusually large bursts are observed. The former are attributed to partial dissociations and the latter to complete dissociations of the tetrameric repressor. (Inset) The occurrence of burst sizes smaller than 10 molecules is well-fit by an exponential distribution (red line). (**C**) Analysis of small bursts from partial dissociations of the tetrameric repressor shows that the distribution of YFP molecules in the range of 0 to 10 molecules does not change in the range of 0 to 200 μM TMG. A small percentage of cells have much more than 10 molecules and do not appear on the axis of this plot (fig. S7C). Over 100 cells were analyzed for each concentration. (Inset) The number of bursts per cell cycle (purple circles) increases slightly whereas the average number of proteins per burst (gray triangles) remains approximately constant. These kinetic parameters are determined from the steady-state distribution of YFP in those cells with fewer than 10 molecules (see SOM text). Error bars are standard errors determined by bootstrapping. Gamma distributions using the determined parameters are overlayed as solid lines. The gamma distribution for 200 μM is normalized for the subpopulation of cells with fewer than 10 molecules. (**D**) In addition to having the permease replaced with Tsr-YFP, strain SX703 has the auxiliary operators removed, eliminating DNA looping so that every dissociation event is a complete dissociation. The protein number distributions should reflect the bursts of protein expression from complete dissociations alone. (Inset) the noise parameters μ^2/σ^2 (purple circles) and σ^2/μ (gray triangles) suggest that the frequency of large bursts is independent of the TMG concentration but that the burst size is not.

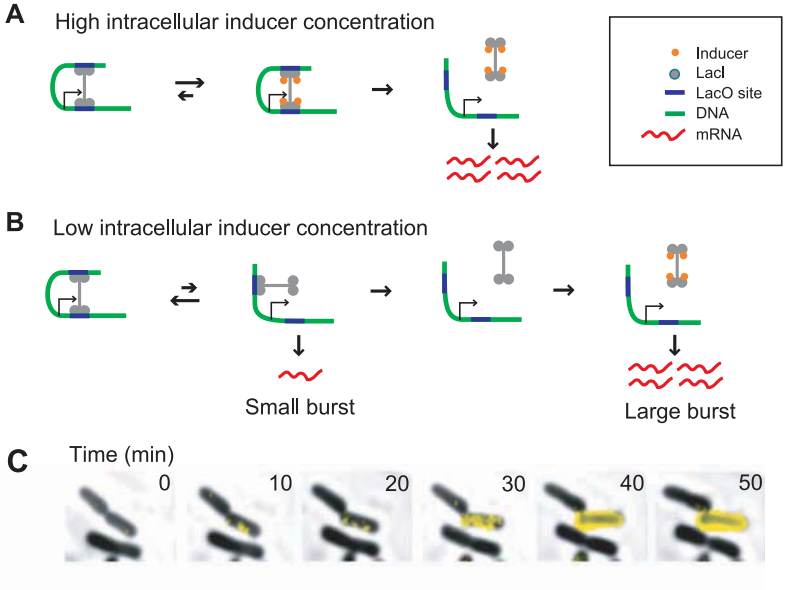

Fig. 4. Complete dissociation of the tetrameric repressor triggers induction. (**A**) A high concentration of intracellular inducer can force dissociation of the repressor from its operators, as described by Monod and Jacob (*6*). (**B**) At low or intermediate concentrations of intracellular inducer, partial dissociation from one operator by the tetrameric LacI repressor is followed by a fast rebinding. Consequently, no more than one transcript is generated during such a brief dissociation event. However, the tetrameric repressor can dissociate from both operators stochastically and then be sequestered by the inducer so that it cannot rebind, leading to a large burst of expression. (**C**) A time-lapse sequence captures a phenotype-switching event. In the presence of 50 μM TMG, one such cell switches phenotype to express many LacY-YFP molecules (yellow fluorescence overlay) whereas the other daughter cell does not (movie S1).

a faster rate of phenotype switching. Indeed, Fig. 2C shows that without DNA looping, SX702 induces rapidly upon the addition of inducer, even though the initial permease numbers are much smaller than the threshold in Fig. 2, A and B [see supporting online material (SOM) text]. The removal of DNA looping eliminated bistability. This was manifested by the unimodal distributions of an inducing population in Fig. 2D at 20 μM TMG.

To directly observe large bursts, we replaced the *lacY* gene of the *E. coli* chromosome with a membrane-localized protein, Tsr, fused to YFP, which generated strain SX701 (fig. S2). Eliminating LacY's positive feedback served two purposes: to allow the resolution of distinct large bursts and to prove that large bursts do not require permease activity. The Tsr-YFP fusion functioned as a surrogate reporter for protein production with single-molecule sensitivity.

When we acquired time-lapse fluorescence microscopy movies of SX701 cells with 200 μM TMG, we observed that Tsr-YFP was produced mostly in small bursts, but occasionally in a large burst (Fig. 3A). The distribution of burst sizes (Fig. 3B) taken from these real-time traces shows both the frequent exponentially distributed small bursts and the rare unusually large bursts.

We next analyzed the inducer concentration dependence of burst frequency and size by using population distributions. Figure 3C shows that SX701 has a protein copy number distribution in a cell population very similar to the distributions of uninduced cells in Fig. 1D within the range of 0 to 10 molecules. We have shown that these copy-number distributions manifest the stochasticity in gene expression characterized by two parameters, transcription rate (a) and the burst size of proteins per mRNA (b), and can be well fit with a gamma distribution $p(n) = b^{-a}n^{a-1}e^{-n/b}/\Gamma(a)$ (*18*, *21*). The inset in Fig. 3C shows a and b determined in this fashion for cells with fewer than 10 molecules for different TMG concentrations. The small burst size is independent of the inducer concentration, whereas the small burst frequency has only a weak concentration dependence. In fact, within the range of 0 to 50 μM TMG, the small burst frequency does not change appreciably, suggesting that the partial repressor dissociation is predominantly spontaneous at low inducer concentrations.

Because characterization of the rare, large bursts is difficult for the wild-type operators, we generated strain SX703, in which the permease gene was replaced with Tsr-YFP and the O2 and O3 operators were removed to eliminate DNA looping. Every dissociation event in the SX703 should be a complete dissociation, and hence should lead to a large burst. Because there is a 20- to 100-fold difference in the inducer-binding affinity to the repressor in DNA-bound form (~1 mM) relative to the free form (~10 μM) (*22*, *23*), we expected to observe the stochastic dissociation of the repressor from the sole operator site and sequestration of the free repressor by the inducers at the concentrations we used (0 to 50 μM) rather than inducer-driven dissociation events.

Consequently, we expected the frequency of dissociation to remain constant as the inducer concentration increased but expected the length of the dissociation time to increase. When we analyzed the steady-state protein distributions at different inducer concentrations (Fig. 4C) by applying a generalized interpretation of the *a-b* model, which estimates the number of proteins per expression bursts (SOM text), we found that inducer concentration affects the size of large bursts but not the frequency. This observation supports our model that the inducer sequesters the repressor after it stochastically dissociates from the operator and prolongs its lifetime in the non–DNA-bound state, leading to larger burst sizes (SOM Text).

Figure 4 summarizes the model of induction in the *lac* operon. In the case of high inducer concentration (Fig. 4A), the repressor is actively pulled off both operator sites by the inducer, as described in Monod and Jacob's model (*6*). Under low or intermediate TMG concentrations, however, the repressor can stochastically dissociate from one operator independently of the inducer, as shown in Fig. 4B. When the repressor partially dissociates from one operator, a small protein burst from a single copy of mRNA is generated (*14*) before the repressor rapidly rebinds to the vacant operator. When the repressor completely dissociates from both operators, multiple mRNAs are transcribed, which leads to a large protein burst that surpasses the LacY threshold, initiates positive feedback, and maintains a switch in phenotype.

Why do complete dissociation events give rise to large bursts? Our group has recently shown that if a repressor dissociates from DNA, it takes a time scale of minutes for the repressor to rebind to the operator because the repressor spends most of its time binding to nonspecific sequences and searching through the chromosomal DNA (*15*). In addition, there are only a few copies of the tetrameric repressors (*8*). Such a slow repressor rebinding time, relative to transcript-initiation frequencies (*24*), would allow multiple copies of *lacY* mRNA to be made following a complete repressor dissociation event. Furthermore, in the presence of inducer, the nonspecific binding constant remains unchanged (*25*), but the affinity of the inducer-bound repressor to the operator is substantially reduced, rendering specific rebinding unlikely. The large burst that results from slow repressor rebinding is an example of how a single-molecule fluctuation under out-of-equilibrium conditions can have considerable biological consequences, which has been discussed theoretically in the context of cell signaling (*26*) and gene expression (*27*) but has not previously been experimentally observed.

Because the binding affinity of inducers for operator-bound repressors is weaker than for free repressors by a factor of 20 to 100 (*22*, *23*), the inducer's role under low concentrations is not to force a dissociation event but to simply sequester repressors that stochastically dissociated from their operators in order to aid in creating a large burst (SOM text) (Fig. 3D). When the inducer concentration increases, the size of the large bursts increases because the duration of complete dissociations increases, improving the probability that a large burst can cross the positive-feedback threshold. Consequently, it is this higher probability of successful switch-on events that shifts the bimodal population toward the fully induced state as inducer concentration increases.

The biological importance of DNA looping has been discussed in the literature in terms of facilitating interactions between distance sequences and enhancing the local concentration (*19*, *20*). Here, we show that DNA looping allows the control of gene regulation on multiple time scales through different kinds of dissociation events. The presence of DNA looping allows the use of rare complete dissociation events to control a bistable genetic switch.

We have shown that a stochastic single-molecule event can cause a change in phenotype. It is not difficult to imagine that similar molecular events might determine more complicated phenotypes of other cells or organisms. The ability to observe and probe the properties of genetic switches at the molecular level is crucial for understanding how cells make decisions.

Referenes and Notes

1. A. Arkin, J. Ross, H. H. McAdams, *Genetics* **149**, 1633 (1998).
2. M. Ptashne, *A Genetic Switch: Phage Lambda Revisited* (Cold Spring Harbor Laboratory Press, Woodbury, NY, 2004).
3. M. Kaern, T. C. Elston, W. J. Blake, J. J. Collins, *Nat. Rev. Genet.* **6**, 451 (2005).
4. D. Dubnau, R. Losick, *Mol. Microbiol.* **61**, 564 (2006).
5. E. Dekel, U. Alon, *Nature* **436**, 588 (2005).
6. J. Monod, F. Jacob, *J. Mol. Biol.* **3**, 318 (1961).
7. B. Müller-Hill, *The Lac Operon: A Short History of a Genetic Paradigm* (de Gruyter, New York, 1996).
8. W. Gilbert, B. Müller-Hill, *Proc. Natl. Acad. Sci. U.S.A.* **56**, 1891 (1966).
9. G. N. Cohen, J. Monod, *Bacteriol. Rev.* **21**, 169 (1957).
10. J. M. G. Vilar, C. C. Guet, S. Leibler, *J. Cell Biol.* **161**, 471 (2003).
11. T. Mettetal, D. Muzzey, J. M. Pedraza, E. M. Ozbudak, A. van Oudenaarden, *Proc. Natl. Acad. Sci. U.S.A.* **103**, 7304 (2006).
12. A. Novick, M. Weiner, *Proc. Natl. Acad. Sci. U.S.A.* **43**, 553 (1957).
13. E. M. Ozbudak, M. Thattai, H. N. Lim, B. I. Shraiman, A. van Oudenaarden, *Nature* **427**, 737 (2004).
14. J. Yu, J. Xiao, X. Ren, K. Lao, X. S. Xie, *Science* **311**, 1600 (2006).
15. J. Elf, G. W. Li, X. S. Xie, *Science* **316**, 1191 (2007).
16. X. S. Xie, P. J. Choi, G. W. Li, N. K. Lee, G. Lia, *Annu. Rev. Biophys.* **37**, 417 (2008).
17. N. Rosenfeld, J. W. Young, U. Alon, P. S. Swain, M. B. Elowitz, *Science* **307**, 1962 (2005).
18. L. Cai, N. Friedman, X. S. Xie, *Nature* **440**, 358 (2006).
19. H. Krämer *et al.*, *EMBO J.* **6**, 1481 (1987).
20. S. Oehler, E. R. Eismann, H. Krämer, B. Müller-Hill, *EMBO J.* **9**, 973 (1990).
21. N. Friedman, L. Cai, X. S. Xie, *Phys. Rev. Lett.* **97**, 168302 (2006).
22. M. D. Barkley, A. D. Riggs, A. Jobe, S. Bourgeois, *Biochemistry* **14**, 1700 (1975).
23. M. Dunaway *et al.*, *J. Biol. Chem.* **255**, 10115 (1980).
24. D. Kennell, H. Riezman, *J. Mol. Biol.* **114**, 1 (1977).
25. A. Revzin, P. H. von Hippel, *Biochemistry* **16**, 4769 (1977).
26. M. N. Artyomov, J. Das, M. Kardar, A. K. Chakraborty, A. K. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 18958 (2007).

27. J. E. M. Hornos *et al.*, *Phys. Rev. E* **72**, 051907 (2005).
28. Materials and methods are available as supporting material on *Science* Online.
29. We would like to thank G. Church, A. Miyawaki, and B. Wanner for bacterial strains and plasmids; J. Hearn for technical assistance; and J. Elf, N. Friedman, and G. W. Li for helpful discussions. This work was supported by the NIH Director's Pioneer Award. P.J.C. acknowledges the John and Fannie Hertz Foundation.

Supporting Online Material
www.sciencemag.org/cgi/content/full/322/5900/442/DC1
Materials and Methods
SOM Text
Figs. S1 to S8
Movie S1
References

5 June 2008; accepted 12 September 2008
10.1126/science.1161427

Remeasuring the Double Helix

Rebecca S. Mathew-Fenn,[1,2]* Rhiju Das,[2,3]*† Pehr A. B. Harbury[1,2]‡

DNA is thought to behave as a stiff elastic rod with respect to the ubiquitous mechanical deformations inherent to its biology. To test this model at short DNA lengths, we measured the mean and variance of end-to-end length for a series of DNA double helices in solution, using small-angle x-ray scattering interference between gold nanocrystal labels. In the absence of applied tension, DNA is at least one order of magnitude softer than measured by single-molecule stretching experiments. Further, the data rule out the conventional elastic rod model. The variance in end-to-end length follows a quadratic dependence on the number of base pairs rather than the expected linear dependence, indicating that DNA stretching is cooperative over more than two turns of the DNA double helix. Our observations support the idea of long-range allosteric communication through DNA structure.

Since the double helical structure of DNA was discovered 50 years ago (*1*), its average structure and internal fluctuations have been objects of intense study. Near its equilibrium structure, the DNA duplex is generally viewed as an ideal elastic rod. Current estimates put the bending rigidity B at ~230 pN·nm^2, the torsional rigidity C at 200 to 500 pN·nm^2, and the stretching modulus S (the extrapolated force required to double the length of the DNA) at ~1000 pN (*2*–*7*). Recent experimental observations, however, have called into question the accuracy of this simple mechanical picture. For example, single-molecule measurements show that overtwisting of DNA induces helix stretching (*8*). This twist-stretch coupling leads to a revised picture of DNA in which the helix core is modeled as an elastic rod while the phosphodiester backbone is modeled as a rigid wire. Analysis of DNA bending on short length scales has also yielded surprises. Specifically, ~100–base pair (bp) DNA helices circularize two to four orders of magnitude faster than would be predicted by the elastic rod model, leading to the idea that discrete kinks contribute to DNA bending (*9*, *10*).

The most straightforward way to characterize DNA structural fluctuations would be to directly visualize them under nonperturbing solution conditions. Kilobase-sized DNA structures have been imaged in real time, but it has not been possible to resolve bending, twisting, and stretching fluctuations at the microscopic level. Alternatively, analyzing the motions of very short DNA fragments simplifies the problem by limiting the contributions from bending. In practice, this has proved technically challenging. The experimental tools suited to the job, molecular rulers, provide an indirect readout of distance that is difficult to relate quantitatively to variation in end-to-end length. Indeed, short DNA duplexes are often assumed to be completely rigid and are used as length standard controls for new molecular rulers (*11*–*14*).

Our investigations apply a recently developed technique for measuring distance distribu-

[1]Biophysics Program, Stanford University, Stanford, CA 94305, USA. [2]Department of Biochemistry, Stanford University, Stanford, CA 94305, USA. [3]Department of Physics, Stanford University, Stanford, CA 94305, USA.
*These authors contributed equally to this work.
†Present address: Department of Biochemistry, University of Washington, Seattle, WA 98195, USA.
‡To whom correspondence should be addressed. E-mail: harbury@cmgm.stanford.edu

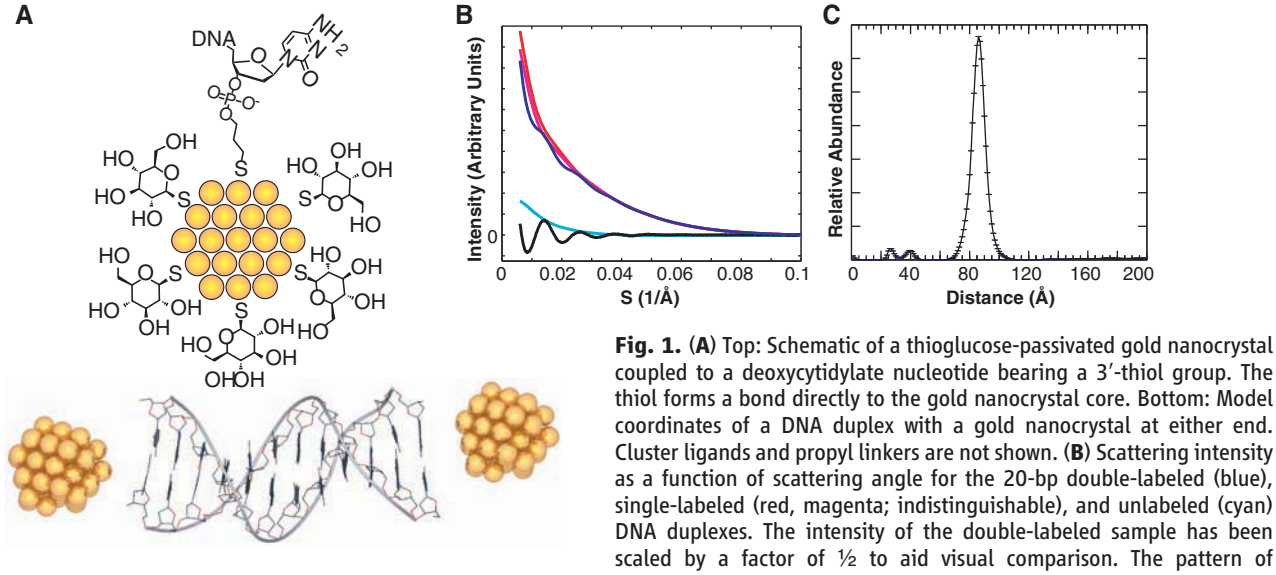

Fig. 1. (**A**) Top: Schematic of a thioglucose-passivated gold nanocrystal coupled to a deoxycytidylate nucleotide bearing a 3′-thiol group. The thiol forms a bond directly to the gold nanocrystal core. Bottom: Model coordinates of a DNA duplex with a gold nanocrystal at either end. Cluster ligands and propyl linkers are not shown. (**B**) Scattering intensity as a function of scattering angle for the 20-bp double-labeled (blue), single-labeled (red, magenta; indistinguishable), and unlabeled (cyan) DNA duplexes. The intensity of the double-labeled sample has been scaled by a factor of ½ to aid visual comparison. The pattern of scattering interference between the two nanocrystal labels (black) is obtained by summing the intensities of the double-labeled and unlabeled samples, then subtracting the intensities of the two single-labeled samples (*15*). The data were obtained at 200 μM DNA and are averages of 10 exposures of 1 s each. Measurements were made at 25°C in the presence of 70 mM Tris-HCl (pH 8.0), 100 mM NaCl, and 10 mM ascorbic acid. The scattering parameter S is defined as $(2 \sin \theta)/\lambda$, where 2θ is the scattering angle and λ is the x-ray wavelength. (**C**) Transformation of the nanocrystal scattering interference pattern into a weighted sum of sinusoidal basis functions (corresponding to different interprobe distances) yields the probability distribution for nanocrystal center-of-mass separation (*15*).