LETTERS

A bottom-up approach to gene regulation

Nicholas J. Guido[1]*, Xiao Wang[2]*, David Adalsteinsson[3], David McMillen[5], Jeff Hasty[6], Charles R. Cantor[1], Timothy C. Elston[4] & J. J. Collins[1]

The ability to construct synthetic gene networks enables experimental investigations of deliberately simplified systems that can be compared to qualitative and quantitative models[1–23]. If simple, well-characterized modules can be coupled together into more complex networks with behaviour that can be predicted from that of the individual components, we may begin to build an understanding of cellular regulatory processes from the 'bottom up'. Here we have engineered a promoter to allow simultaneous repression and activation of gene expression in *Escherichia coli*. We studied its behaviour in synthetic gene networks under increasingly complex conditions: unregulated, repressed, activated, and simultaneously repressed and activated. We develop a stochastic model that quantitatively captures the means and distributions of the expression from the engineered promoter of this modular system, and show that the model can be extended and used to accurately predict the *in vivo* behaviour of the network when it is expanded to include positive feedback. The model also reveals the counterintuitive prediction that noise in protein expression levels can increase upon arrest of cell growth and division, which we confirm experimentally. This work shows that the properties of regulatory subsystems can be used to predict the behaviour of larger, more complex regulatory networks, and that this bottom-up approach can provide insights into gene regulation.

We engineered the O_RO_{lac} promoter in *E. coli* as a tool for investigating the interactive effects of positive and negative transcriptional regulation. The basis for the engineered promoter is the P_{RM} promoter of bacteriophage λ. In the native promoter, the operator sites O_R1 and O_R2 cooperatively bind the λ repressor protein, CI, which in turn acts as a transcriptional activator by recruiting RNA polymerase[24]. In the engineered O_RO_{lac} promoter, the third operator site of the native P_{RM} promoter, O_R3, was mutated to significantly reduce CI binding, thereby removing the repressive function of this site. A *lac* operator, which can bind the repressor protein LacI, was placed upstream of O_R1. The result is a promoter that can be repressed by LacI and activated by CI; it thus enables the combined effects of repression and activation to be investigated.

We inserted the O_RO_{lac} promoter in a high-copy plasmid that lacked the *cI* activator and *lacI* repressor genes to create an unregulated system (Fig. 1a), to establish the promoter's basal properties. The green fluorescent protein gene (*gfp*) was placed under the control of the O_RO_{lac} promoter as a readout for transcriptional induction. A repressor-only system was constructed by adding a strongly constitutively expressed *lacI* gene, under the control of the P_LtetO1 promoter, to the unregulated system (Fig. 1b). The LacI protein binds to the *lac* operator as a tetramer, which reduces the binding of the RNA polymerase, and may affect the binding of CI to its operator site. When isopropyl-β-D-thiogalactopyranoside (IPTG) is added to the media, it binds to the LacI tetramer and weakens binding of the LacI protein to the O_{lac} operator site. This allowed us to tune the degree of O_RO_{lac} repression. An activator-only system was constructed by adding the *cI* gene under the control of the *pBAD* promoter, which is activated by arabinose (Fig. 1c). The CI protein forms a dimer that binds sequentially and cooperatively to the O_R1 and O_R2 sites of the O_RO_{lac} promoter. Finally, both repression and activation, as implemented in the individual modules above, were placed together on one plasmid, to create a repressor–activator system (Fig. 1d).

A mathematical model was developed to account for the *in vivo* behaviour of this modular system, initially using equilibrium modelling of the underlying biochemical reaction scheme to capture the mean transcriptional response of the three regulated systems: the repressor-only, activator-only and repressor–activator systems. These considerations allow us to derive explicit expressions for the probabilities of each of the six binding states as a function of the inducer levels (see Supplementary Information). These expressions were globally fitted to the transcriptional induction data using the unregulated system as a baseline, leading to the mean fluorescence results (blue lines) shown in Fig. 1f–h.

The deterministic model was then extended to include stochastic effects[25]. Preliminary investigations with the stochastic model revealed that fluctuations in the concentrations of transcription factors have very minor effects on the variability in the expression levels of the GFP reporter protein. Therefore, synthesis, degradation, and multimerization of CI and LacI were not included in our baseline model. All stochastic modelling was carried out using the BioNetS software[26], which uses a highly optimized version of the Gillespie Monte Carlo algorithm. To create the distributions of the GFP reporter protein, synthesis and degradation of GFP messenger RNA and protein were tracked explicitly. To include cell growth and division in the model, the cell volume was treated as a random variable that undergoes exponential growth[26] with a mean doubling time of 20 minutes, as observed in the experimental system. At cell division, the volume is halved and the mRNA and GFP molecules are divided between daughter cells on the basis of a binomial distribution[21].

Because the plasmid used in our experimental system has a high-copy ColE1 origin of replication, intercellular copy-number variability can be large[27]. Rather than model the complex copy-number control system explicitly, we assumed a gamma distribution at the time of cell division (see Supplementary Information). The mean was assumed to be 50 plasmid copies per cell[28]. The variance was adjusted so that the model accurately captures the fluctuations in GFP levels observed in the activator-only system, and then kept at this value

[1]Department of Biomedical Engineering, Bioinformatics Program, Center for BioDynamics and Center for Advanced Biotechnology, Boston University, 44 Cummington Street, Boston, Massachusetts 02215, USA. [2]Department of Statistics and Operations Research, [3]Department of Mathematics, [4]Department of Pharmacology, University of North Carolina, Chapel Hill, North Carolina 27599, USA. [5]Institute for Optical Sciences and Department of Chemical and Physical Sciences, University of Toronto at Mississauga, 3359 Mississauga Rd, Mississauga, Ontario L5L 1C6 Canada. [6]Department of Bioengineering, University of California, San Diego, 9500 Gillman Drive, La Jolla, California 92093, USA.
*These authors contributed equally to this work.

when the repressor-only and repressor–activator systems were considered. Comparison of the experimental results and the stochastic simulations in Fig. 1e and Fig. 1i–k shows that the model can accurately capture the *in vivo* behaviour of the unregulated system and all three regulated systems.

The coefficients of variation (CVs) of the three regulated systems are shown in the insets of Fig. 1f–h. As can be seen, the stochastic model slightly underestimates the experimentally observed variability. This discrepancy of approximately 0.1 in CV for each system is largely due to the model's inability to capture the long tails that extend into high fluorescence levels in the experimental flow cytometry histograms, and indicates that our relatively simple model does not account for all sources of variability. For example, while the model includes extrinsic noise[5] that arises from cell growth and division and fluctuations in plasmid copy number, it does not account for cell-to-cell variability in the transcriptional and translational machinery.

An important test for any mathematical model is its ability to predict the outcome of novel experiments. We designed and conducted several new experiments on the basis of the model's predictions. First, the stochastic model results revealed that a large amount of the variability in GFP levels is attributable to fluctuations in the plasmid copy number. To test this prediction, we placed the repressor–activator system on a low-copy plasmid that is tightly regulated, maintaining 3–4 copies of the plasmid per cell[28]. Experiments were then conducted using both arabinose and IPTG as an inducer, and as expected, the cells with the low copy number showed reduced expression levels. In the stochastic model, the plasmid copy number was fixed at three, and the concentration of LacI was proportionally reduced to reflect this change. The functional relationship between arabinose level and CI concentration was taken to be the same in the low-copy case as in the high-copy case, to match our experimental observation that reducing the plasmid copy number did not affect the shape of the system's arabinose response curve (see Supplementary Information). No additional changes were made to the model. As seen in Fig. 2, the experimental distributions for GFP expression levels closely match those predicted by the model, indicating that the model accurately captures fluctuations in expression levels that are attributable to varying plasmid copy number.

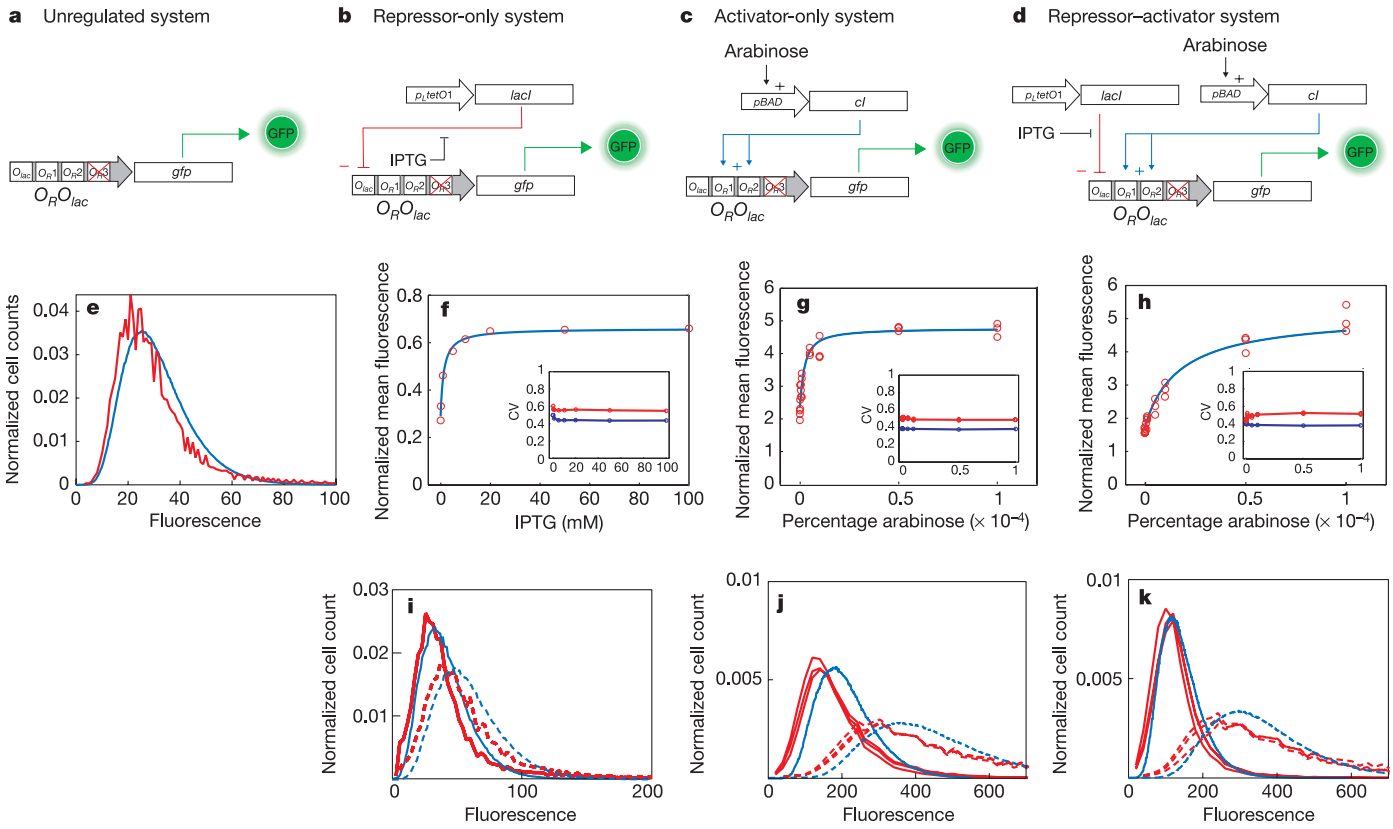

Figure 1 | Unregulated, repressor-only, activator-only and repressor–activator systems on high-copy plasmids. Schematic designs, experimental data and model data are shown. **a**, The unregulated system with the engineered O_RO_{lac} promoter controlling the reporter *gfp*. The white boxes in the O_RO_{lac} promoter represent the operator sites at which CI and LacI proteins bind. The red cross at O_R3 represents the point mutation that reduces CI binding at that operator site. **b**, The repressor-only system consists of the P_LtetO1 constitutive promoter controlling *lacI* and the O_RO_{lac} promoter controlling *gfp*. **c**, The activator-only system consists of the *pBAD* promoter controlling *cI* and the O_RO_{lac} promoter controlling *gfp*. **d**, The repressor–activator system represents a combination of the repressor-only and activator-only systems. **e**, Histograms of the unregulated system: experimental data (red) and stochastic model data (blue). The *x* axis represents arbitrary fluorescence units from flow cytometry, and the *y* axis represents the frequency of cells producing the corresponding fluorescence level. **f**, Repressor-only system results: GFP expression represented as normalized fluorescence versus IPTG level; red circles are experimental data and the blue lines are the results of the deterministic model. The inset shows CV versus IPTG level: experimental data (red) and stochastic model data (blue). **g**, Activator-only system results: normalized fluorescence versus arabinose level, with an inset showing CV versus arabinose level. **h**, Repressor–activator system results: normalized fluorescence versus arabinose level. The inset shows CV versus arabinose level, with 10 mM IPTG in each case. **i**, Histograms of normalized cell counts versus arbitrary fluorescence units, of experimental data (red) and stochastic model data (blue) for the repressor-only system. The solid lines in the histograms are the results for no inducer (IPTG in this case) and the dashed lines are the results for the highest level of IPTG. **j**, Histograms (as above), showing results for no inducer (solid lines) and the highest level of arabinose (dashed lines) for the activator-only system. **k**, Histograms (as above), showing results for no arabinose (solid lines) and the highest level of arabinose (dashed lines), with 10 mM IPTG in each case for the repressor–activator system.

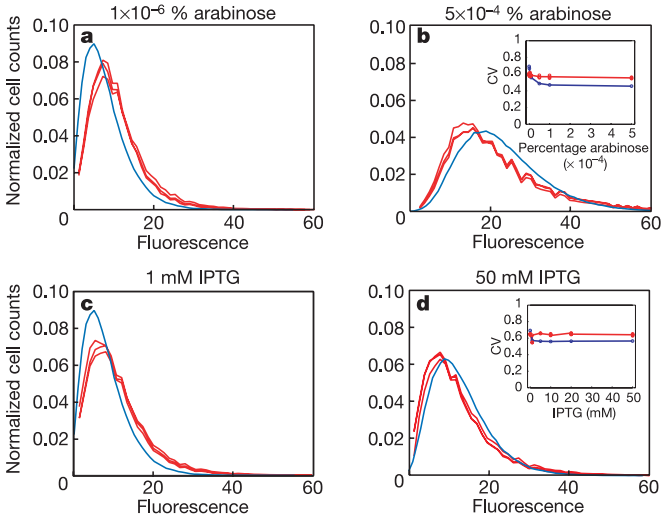

Figure 2 | Histograms of model data (blue lines) and experimental data (red lines) for the repressor–activator system on a low-copy plasmid. The *x* axis represents arbitrary fluorescence units from flow cytometry, and the *y* axis represents the frequency of cells producing the corresponding fluorescence level for: **a**, 1×10^{-6}% arabinose, no IPTG; **b**, 5×10^{-4}% arabinose, no IPTG; **c**, 1 mM IPTG, no arabinose; **d**, 50 mM IPTG, no arabinose. The inset in **b** shows CV versus arabinose level, and the inset in **d** shows CV versus IPTG level; model data (blue lines) and experimental data (red lines).

The insets of Fig. 2b and d show a comparison of the CVs for the experimental and model results. The model accurately predicts the behaviour of the CV for the low-copy case, taking into account the 0.1 discrepancy discussed above. At very low inducer levels, the model predicts a larger CV than is observed experimentally. We attribute this discrepancy to autofluorescence in the experimental system. This has the effect of artificially increasing the mean fluorescence, thereby decreasing the CV. The effect is only important at low fluorescence levels, where autofluorescence makes up a significant portion of the measured mean value.

To test the model's predictive power in a more complex system, we added positive feedback to the repressor–activator system. To accomplish this, we expanded the system so that the *cI* gene was transcribed polycistronically with *gfp* under the control of the O_RO_{lac} promoter (Fig. 3a). Adding feedback to the stochastic model requires the explicit inclusion of the CI mRNA and protein molecular abundances and the biochemical processes that affect their levels (synthesis, degradation, multimerization). The parameters that control these processes were chosen so that in the absence of feedback the expanded model produced CI protein levels identical to that of the simpler model (see Supplementary Information). As can be seen in Fig. 3b–g, the agreement between the behaviour predicted by the model and the experimental results is very good, validating the bottom-up approach to understanding gene regulatory networks.

We also used the model to conduct a systematic analysis of how the different sources of noise contribute to the overall variability (Supplementary Fig. 7). Surprisingly, we found that if we run simulations with no cell growth or division using an ensemble of cells with varying plasmid copy number, the CV increases above the level at which cell growth and division proceed normally. This result was unexpected, because cell growth and division are generally thought to add variability to expression levels. A theoretical explanation of this counterintuitive prediction is provided in the Supplementary Information. In very general terms, this phenomenon can be understood as follows: in a population of cells with varying plasmid copy number, the intercellular variability increases as expression levels increase from new protein synthesis. Cell division not only limits the mean protein level, it also causes the distribution to be more tightly centred about the mean (Supplementary Fig. 8). This leads to an overall reduction of the variability in protein expression for the high-copy system where plasmid copy-number fluctuations account for a significant portion of the noise. In

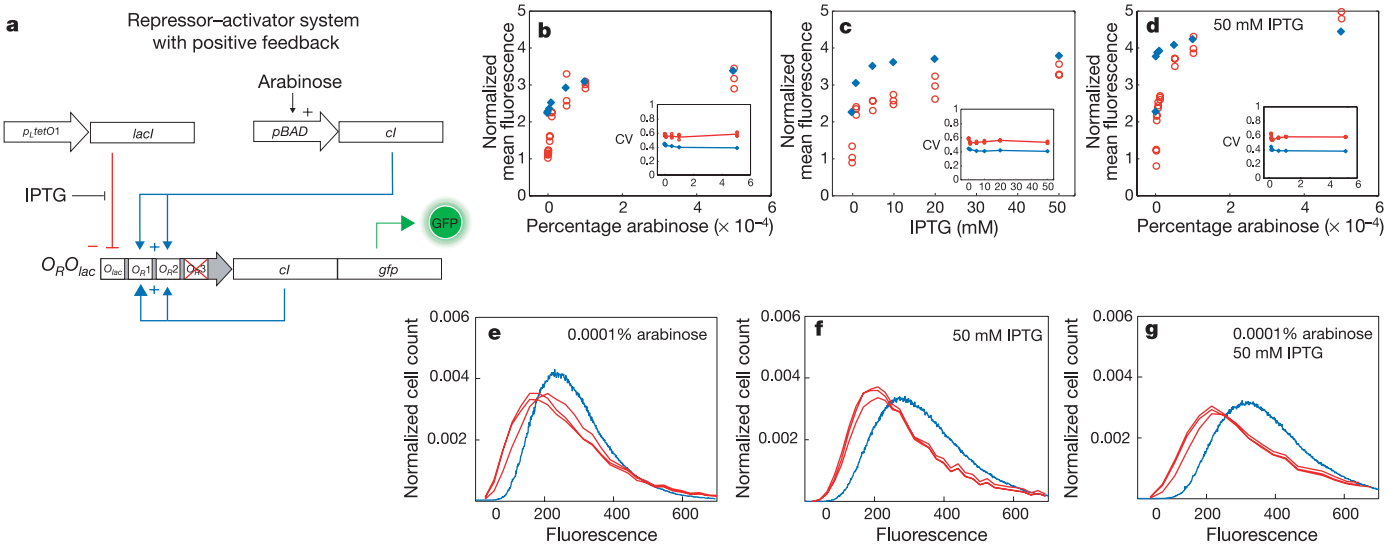

Figure 3 | Repressor–activator system with positive feedback. **a**, Schematic of the positive feedback construct, where the *cI* gene is added to the repressor–activator system on a high-copy plasmid. The *cI* gene is incorporated, along with *gfp*, in a polycistronic region controlled by the engineered O_RO_{lac} promoter. Model predictions and experimental results for the repressor-activator system with positive feedback: **b**, Normalized mean (arbitrary fluorescence units) versus arabinose level, where the red circles are experimental data and the blue diamonds are model predictions. The inset shows CV versus arabinose level. **c**, Normalized mean versus IPTG level, with an inset showing CV versus IPTG level. **d**, Normalized mean versus arabinose level, with 50 mM IPTG in each case, with an inset showing CV versus arabinose level for 50 mM IPTG. **e**, Histograms, normalized cell counts versus arbitrary fluorescence units, of experimental data (red) and model predictions (blue) for 0.0001% arabinose. **f**, Histograms of experimental data (blue) and model predictions (red) for 50 mM IPTG. **g**, Histograms of experimental data (blue) and model predictions (red) for 0.0001% arabinose and 50 mM IPTG.

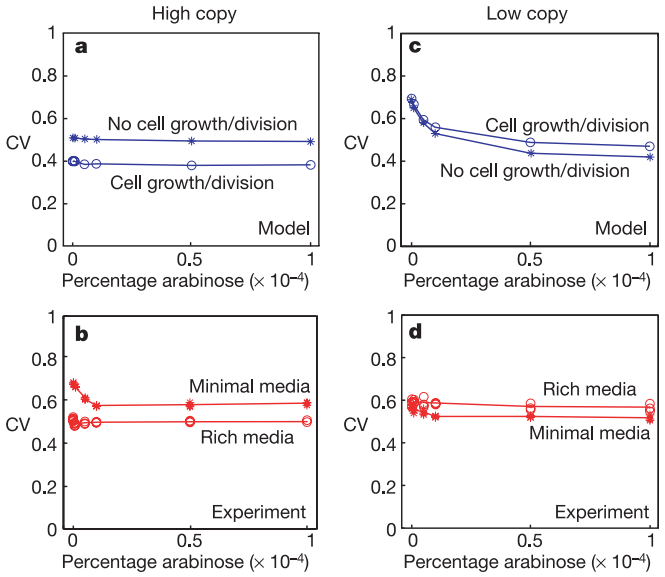

Figure 4 | CV versus arabinose levels for the repressor–activator system. **a**, Model results for the high-copy case with cell growth and division (circles) and with no cell growth or division (stars). **b**, Experimental results for the high-copy case for cells in rich media (circles) and minimal media (stars). **c**, Modelling results for the low-copy case with cell growth and division (circles) and with no cell growth or division (stars). **d**, Experimental results for the low copy case for cells in rich media (circles) and minimal media (stars).

contrast, for the low-copy case where plasmid copy number is tightly controlled, the model predicts that cell growth and division should not reduce noise levels and instead, halting cell growth and division is expected to decrease the variability.

We considered ways to experimentally test these model predictions. Most methods of achieving arrest of cell growth and division (such as addition of the drug cephalexin[29]) cause drastic changes in cell morphology or protein production, both of which would be detrimental to our experimental design. We found cell behaviour in minimal media to be the best approximation of our model conditions. In minimal media, cell growth and division slow considerably, while protein overproduction from plasmids can continue.

Cells were placed in M9 minimal media and 2% glucose with incubator conditions identical to those in the rich media experiments. Our model predicts that the variability in protein expression levels can increase if cells with the repressor–activator system on a high-copy plasmid do not grow or divide (Fig. 4a). The results in Fig. 4b show that the cells in minimal media with high-copy plasmids do in fact have a higher CV than those that are growing and dividing regularly in rich Lauria broth (LB) media. Additionally, our model predicts that the variability in expression levels for the low-copy case decreases when cell growth and division are stopped (Fig. 4c). This effect is validated experimentally in Fig. 4d where the results show that cells with low-copy plasmids in minimal media have lower CVs than similar cells growing and dividing in rich media. (See Supplementary Information for similar results with stationary phase cells.)

Previous work in synthetic biology[1–23] has shown that one can develop a mathematical model for a given synthetic gene network and use the model to describe the behaviour of that network. These efforts set the stage for the possibility of using quantitative models of regulatory subsystems to predict the behaviour of larger, more complex regulatory networks, one of the main goals of synthetic biology. In this study, we show that such a bottom-up approach is indeed achievable and can lead to insights and testable predictions that are biologically meaningful. Specifically, we show that we can fit a stochastic model to experimental data from a modular, synthetic gene network, and then extend the model and use it to predict quantitatively the behaviour of a more complex network involving regulatory feedback. The model also accurately predicted the effects of changing plasmid copy numbers on the experimentally observed levels of variability in GFP expression and suggested that halting cell growth and division in cells with high-copy plasmids should increase fluctuations in protein expression, which was confirmed experimentally. This synthesis of theory and experiment has much to offer in investigations of cellular behaviour, and such approaches may eventually allow us to assemble a genuine bottom-up picture of the intricate processes of gene regulation.

METHODS

Plasmid construction. Plasmid construction began with the pZE21-MCS1 modular construct of Lutz and Bujard[28]. This plasmid backbone contains the $P_L tetO1$ promoter, *T1T2* terminator, kanamycin resistance gene, *ColE1* origin of replication and multiple cloning sites. The low copy origin of replication was from the pZS*24-MCS1 plasmid[28]. To create the $O_R O_{lac}$ promoter, the O_R region of bacteriophage λ was altered with the G-to-T or3-r2 point mutation in the $O_R 3$ operator site[24] (greatly reducing its affinity for the CI protein), and an O_{lac} operator site was added upstream of the O_R region. The promoter sequence is shown in Supplementary Fig. 12. The point mutation and operator insertion were carried out by polymerase chain reaction (PCR; using PfuTurbo DNA polymerase from Stratagene and an MJ Research PTC-100 thermal cycler) with primers designed such that the point mutation and operator site were placed in the primers.

The *lacI* gene was PCR-amplified, with PfuTurbo and an MJ Research thermal cycler, from pTrc99a, the *clts* and O_R region were from pGW7 (ATCC), the *gfpmut3* was from pJBA111 (J. B. Andersen, Technical University of Denmark), and the pBAD promoter was from pBADHisA (Clontech).

Cell growth and expression experiments. All plasmids were inserted into the *E. coli* strain JM2.300 (−λ, *lacI22 rpsL135* (StrR), *thi-1*) via the transformation and storage solution heat-shock transformation protocol[30]. Cells for experimentation were grown overnight and cultures were inoculated 1:300 in LB media and with 0.03 mg ml^{-1} kanamycin antibiotic. Varying levels of IPTG and arabinose were added separately and in combination to affect the behaviour of the inducible promoters. Cultures were grown in a 37 °C incubator shaking at 300 r.p.m. for approximately three hours, at which point the cells reached a steady state (see Supplementary Information), and the absorbance at 600 nm ($A_{600\,nm}$) was between 0.3 and 0.4. The minimal media cultures used the same conditions as LB cultures except that the minimal media cultures were grown in M9 minimal media with 2% glucose.

Data acquisition and analysis. Culture samples were spun down at 8,000 r.p.m. to pellet cells, and supernatant was removed. The samples were then resuspended in 0.75 ml of 1 × PBS, and fluorescence measurements were taken via flow cytometry. The Becton Dickinson FACScalibur flow cytometer was used to measure 50,000 cells of each sample from a culture representing one level of inducer. Excitation of GFP was achieved via a 488-nm argon excitation laser and fluorescence measured with the 515–545 nm emission filter. The flow cytometer generates logscale values using a 10-bit analogue-to-digital converter, yielding integers in the range 0 to 1,023 for each of three measurements: fluorescence intensity, forward-scattering, and side-scattering. Cells were collected within a small forward-scatter and side-scatter gate, in the Cellquest software designed by BD Biosciences, to minimize fluorescence variation due to cell size.

Modelling and stochastic simulations. The model parameters were estimated using Matlab's least-squares fitting routine. The mean expression levels were fitted to the repressor-only, activator-only and repressor–activator system results. To initialize the least-squares routine, we used 2,000 random initial guesses for the model parameters, generated from a uniform distribution within a biologically reasonable regime of parameter space. A discussion of the sensitivity of the model's output to the values of the parameters is given in the Supplementary Information. The best parameter set was chosen from those 2,000 runs. All the fits were based on data normalized to the mean GFP expression level of the unregulated system. All stochastic simulations were done in BioNetS[26] using the Gillespie algorithm. See Supplementary Information for details.

Received 7 October; accepted 18 November 2005.

1. Elowitz, M. B. & Leibler, S. A synthetic oscillatory network of transcriptional regulators. *Nature* **403**, 335–338 (2000).

2. Gardner, T. S., Cantor, C. R. & Collins, J. J. Construction of a genetic toggle switch in *Escherichia coli*. *Nature* **403,** 339–342 (2000).
3. Becskei, A. & Serrano, L. Engineering stability in gene networks by autoregulation. *Nature* **405,** 590–593 (2000).
4. Becskei, A., Séraphin, B. & Serrano, L. Positive feedback in eukaryotic gene networks: cell differentiation by graded to binary response conversion. *EMBO J.* **20,** 2528–2535 (2001).
5. Elowitz, M., Levine, A., Siggia, E. & Swain, P. Stochastic gene expression in a single cell. *Science* **297,** 1183–1186 (2002).
6. Ozbudak, E., Thattai, M., Kurtser, I., Grossman, A. & van Oudenaarden, A. Regulation of noise in the expression of a single gene. *Nature Genet.* **31,** 69–73 (2002).
7. Rosenfeld, N. Y., Elowitz, M. B. & Alon, U. Negative autoregulation speeds the response times of transcription networks. *J. Mol. Biol.* **323,** 785–793 (2002).
8. Guet, C. C., Elowitz, M. B., Hsing, W. & Leibler, S. Combinatorial synthesis of genetic networks. *Science* **296,** 1407–1470 (2002).
9. 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).
10. Blake, W. J., Kaern, M., Cantor, C. R. & Collins, J. J. Noise in eukaryotic gene expression. *Nature* **422,** 633–637 (2003).
11. Atkinson, M. R., Savageau, M. A., Myers, J. T. & Ninfa, A. J. Developement of genetic circuitry exhibiting toggle switch or oscillatory behaviour in *Escherichia coli*. *Cell* **113,** 597–607 (2003).
12. Weiss, R. *et al.* Genetic circuit building blocks for cellular computation, communications, and signal processing. *Natural Comput.* **2,** 47–84 (2003).
13. Basu, S., Mahreja, R., Thiberge, S., Chen, M. T. & Weiss, R. Spatiotemporal control of gene expression with pulse generating networks. *Proc. Natl. Acad. Sci.* **17,** 6355–6360 (2004).
14. You, L., Cox, R. S. III, Weiss, R. & Arnold, F. H. Programmable population control cell-cell communication and regulated killing. *Nature* **428,** 868–871 (2004).
15. Kramer, B. P. *et al.* An engineered epigenetic transgene switch in mammalian cells. *Nature Biotechnol.* **22,** 867–870 (2004).
16. Kobayashi, H. *et al.* Programmable cells: interfacing natural and engineered gene networks. *Proc. Natl. Acad. Sci.* **101,** 8414–8419 (2004).
17. Isaacs, F. J. *et al.* Engineered riboregulators enable post-transcriptional control of gene expression. *Nature Biotechnol.* **22,** 841–847 (2004).
18. Fung, E. *et al.* A synthetic gene-metabolic oscillator. *Nature* **435,** 118–122 (2005).
19. Hooshangi, S., Thiberge, S. & Weiss, R. Ultrasensitivity and noise propagation in a synthetic transcriptional cascade. *Proc. Natl. Acad. Sci. USA* **102,** 3581–3586 (2005).
20. Pedraza, J. M. & van Oudenaarden, A. Noise propagation in gene networks. *Science* **307,** 1965–1969 (2005).
21. Rosenfeld, N. Y., Young, J. W., Alon, U., Swain, P. S. & Elowitz, M. B. Gene regulation at the single cell level. *Science* **307,** 1962–1965 (2005).
22. Isalan, M., Lemerle, C. & Serrano, L. Engineering gene networks to emulate *Drosophila* embryonic pattern formation. *PLoS Biol.* **3**(3), e64 (2005).
23. Kramer, B. P. & Fussenegger, M. Hysteresis in a synthetic mammalian gene network. *Proc. Natl. Acad. Sci. USA* **102,** 9517–9522 (2005).
24. Ptashne, M. *A Genetic Switch: Phage λ and Higher Organisms* (Cell Press & Blackwell Scientific, Cambridge, Massachusetts, 1992).
25. Kepler, T. & Elston, T. C. Stochasticity in transcriptional regulation: origins, consequences, and mathematical representations. *Biophys. J.* **81,** 3116–3136 (2001).
26. Adalsteinsson, D., McMillen, D. & Elston, T. C. Biochemical Network Stochastic Simulator (BioNetS): software for stochastic modeling of biochemical networks. *BMC Bioinform.* **5,** 24 (2004).
27. Paulsson, J. & Ehrenberg, M. Noise in a minimal regulatory network: plasmid copy number control. *Q. Rev. Biophys.* **34,** 1–59 (2001).
28. 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).
29. Greenwood, D. & O'Grady, F. Comparison of the response of *Escherichia coli* and *Proteus mirabilis* to seven β-lactam antibiotics. *J. Infect. Dis.* **128,** 211–222 (1973).
30. Sambrook, J., Fritsch, E. & Maniatis, T. *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory Press, Plainview, New York, 1989).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements This work was supported by the NIH, NSF and DARPA.

Author Contributions T.C.E. and J.J.C. are co-senior authors.

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to J.J.C. (jcollins@bu.edu).