molecular
systems
biology

REPORT

Accurate prediction of gene feedback circuit behavior from component properties

Nitzan Rosenfeld[1,2,6], Jonathan W Young[3,4], Uri Alon[1,2], Peter S Swain[5] and Michael B Elowitz[3,4,*]

[1] Department of Molecular Cell Biology, Weizmann Institute of Science, Rehovot, Israel, [2] Department of Physics of Complex Systems, Weizmann Institute of Science, Rehovot, Israel, [3] Division of Biology, California Institute of Technology, Pasadena, CA, USA, [4] Department of Applied Physics, California Institute of Technology, Pasadena, CA, USA and [5] Department of Physiology, Centre for Non-linear Dynamics, McGill University, Montreal, Quebec, Canada
[6] Present address: Rosetta Genomics Ltd, 10 Plaut Street, Rehovot 76706, Israel
* Corresponding author. Division of Biology and Department of Applied Physics, California Institute of Technology, M/C 114-96, Pasadena, CA 91125, USA.
Tel.: +1 6263958871; Fax: +1 6263955972;
E-mail: melowitz@caltech.edu

A basic assumption underlying synthetic biology is that analysis of genetic circuit elements, such as regulatory proteins and promoters, can be used to understand and predict the behavior of circuits containing those elements. To test this assumption, we used time-lapse fluorescence microscopy to quantitatively analyze two autoregulatory negative feedback circuits. By measuring the gene regulation functions of the corresponding repressor–promoter interactions, we accurately predicted the expression level of the autoregulatory feedback loops, in molecular units. This demonstration that quantitative characterization of regulatory elements can predict the behavior of genetic circuits supports a fundamental requirement of synthetic biology.
Molecular Systems Biology 13 November 2007; doi:10.1038/msb4100185
Subject Categories: synthetic biology
Keywords: auto-regulation; feedback; GRF; quantification; regulatory elements

Introduction

A major challenge in systems biology is to create quantitative, predictive models of gene-circuit dynamics (Alon, 2007). It has recently become possible to construct and analyze synthetic gene regulatory circuits (Becskei and Serrano, 2000; Elowitz and Leibler, 2000; Gardner *et al*, 2000; Sprinzak and Elowitz, 2005; Andrianantoandro *et al*, 2006; Guido *et al*, 2006). A basic assumption underlying such 'synthetic' biology is that the properties of individual genetic components can be used to understand and quantitatively predict circuit-level behaviors. Nevertheless, even seemingly simple genetic circuits have produced unexpected results (Becskei and Serrano, 2000; Guet *et al*, 2002; Paulsson, 2004; Hooshangi and Weiss, 2006). Most characterizations of genetic circuit elements involve uncalibrated, relative units (Becskei and Serrano, 2000; Guet *et al*, 2002; Rosenfeld *et al*, 2002; Austin *et al*, 2006; Guido *et al*, 2006), which makes comparisons between experiments difficult and limits the design of larger circuits. Models are often restricted to qualitative phase space analysis or require numerous free parameters to fit observed behavior (Paulsson, 2004; Sprinzak and Elowitz, 2005; Andrianantoandro *et al*, 2006; Hooshangi

and Weiss, 2006). Finally, stochastic noise in the operation of genetic circuit elements can further complicate analysis. Therefore, it is critical to test the assumption that genetic circuit elements can be reconnected in a predictable fashion.

Previously, we described quantitative measurement of gene regulation functions (GRFs), defined as the rate of protein production as a function of transcription factor concentration in single cells (Rosenfeld *et al*, 2005). The simplest circuit whose behavior depends on this GRF is the negative autoregulatory circuit in which a repressor regulates its own expression (Figure 1A). The behavior of this circuit can be predicted from the GRF with no fitting parameters. Here, we show that *in vivo* biochemical characterization of a promoter–repressor interaction enables accurate prediction of the output of a negative feedback circuit built from these elements.

To measure the GRF, we constructed a chimeric gene, *cI-yfp*, consisting of the phage lambda repressor gene (*cI*) fused to the yellow fluorescent protein gene (*yfp*). We inserted a cyan fluorescent protein gene (*cfp*) in the chromosome under the control of each of two variants of the lambda P_R promoter (Rosenfeld *et al*, 2005). Because the promoter is inserted at single-copy on the *Escherichia coli* chromosome, effects of plasmid copy
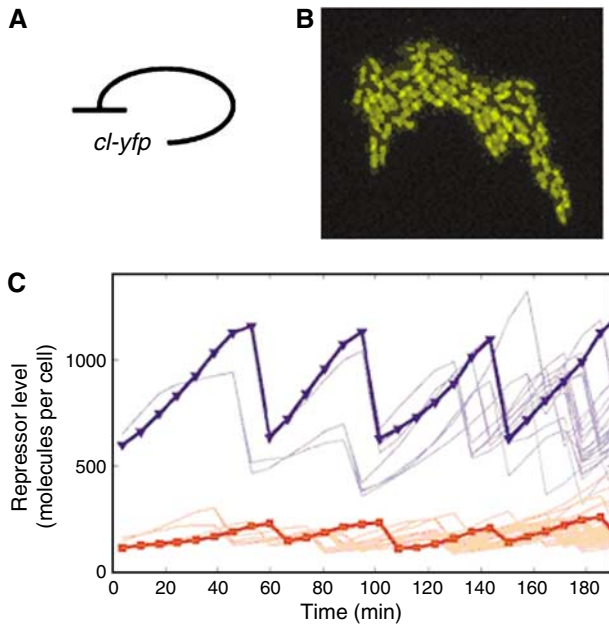
A

cI-yfp

B

C

Figure 1 Measurements of synthetic negative autoregulation circuits. (**A**) Schematic diagram of a negative autoregulatory circuit. *cI-yfp* encodes a fusion protein of the lambda repressor and *yfp* genes. This chimeric gene is placed under the control of one of the two variants of the lamba P_R promoter (see Supplementary Figure S1). (**B**) Snapshot of a growing bacterial colony carrying the negative autoregulatory circuit (O_R2^* variant). Localization of CI-YFP to cell nucleoids can be observed (Bakk and Metzler, 2004). (**C**) CI-YFP fluorescence plotted over time for cell lineages containing the two negative autoregulatory P_R-*cI-yfp* circuit variants. Circuit with wild-type P_R promoter is shown in red, while its O_R2^* variant, which has weaker repressor binding is shown in blue. One cell lineage is highlighted for each bacterial strain.

number fluctuations are avoided. We expressed *cI-yfp* from a tetracycline-inducible promoter and monitored the rate of expression of *cfp* as a function of the level of CI-YFP in individual cells. These experiments used dilution of CI-YFP during growth to systematically vary the levels of repressor in individual cell lineages. We characterized two variants of the lambda P_R promoter which differ in affinity for repressor because of a single point mutation in one operator region (Rosenfeld *et al*, 2005). This study produced a quantitative GRF, including measurements of effective affinity, cooperativity, and promoter strength, and determined the levels and timescales of fluctuations in gene expression controlled by the target promoter.

These GRFs contain sufficient information to predict the behavior of corresponding autoregulatory circuits, in which CI-YFP represses its own expression. To test this prediction, we inserted the *cI-yfp* gene under each of the two characterized P_R promoters, so that CI-YFP represses its own production (Supplementary Figure S1). We then acquired time-lapse fluorescence microscopy movies of growing microcolonies of *E. coli* carrying these synthetic genetic circuits (Figure 1B and Supplementary movies SM1 and SM2). We used quantitative image analysis (Rosenfeld *et al*, 2005) and a fluorescence calibration method based on fluctuations in protein partitioning (Rosenfeld *et al*, 2006) to determine the protein levels and production rates in individual cells over time in absolute numbers of repressor molecules (Figure 1C). We compared the steady-state protein levels in the autoregulatory circuits with predictions based on the GRFs.

Results and discussion

We repeated our measurement of the GRFs using the strains and procedures described previously (Supplementary Figure S2; Rosenfeld *et al*, 2005). We converted fluorescence units to protein numbers and quantified fluorescent protein expression in units of proteins per cell (Rosenfeld *et al*, 2006). This calibration has an approximately twofold systematic uncertainty in concentration values (Rosenfeld *et al*, 2005). The mean GRF is well described by a Hill function: the rate of protein production is $\beta/(1 + (R/k)^n)$, where β is the maximal production rate, n indicates the degree of effective cooperativity in repression, and k is the concentration of repressor yielding half-maximal expression. The values we obtained for β, k, and n from the present data (Supplementary Figure S2) lie within the confidence limits of our previous measurement (Rosenfeld *et al*, 2005). The effective binding affinities of CI-YFP to the lambda P_R promoters are in the range of affinities estimated for the native lambda cI repressor (Meyer *et al*, 1980; Ptashne, 2004; Darling *et al*, 2000; Dodd *et al*, 2004).

We used these GRFs (Supplementary Figure S2) to derive quantitative predictions for the corresponding autoregulatory circuits. At steady state, the protein production rate equals the protein dilution rate, as there is no detectable protein degradation in these conditions (Rosenfeld *et al*, 2005). In the continuous limit, both rates are functions of the total cellular repressor level, R_T: to compare dilution and production rates in the same units, we converted the GRF from $f(R)$, a function of the repressor concentration (Supplementary Figure S2) to $f_T(R_T)$, a function of total repressor levels (Figure 2). We set $R=R_T/V$, where V is the volume of the cell, estimated from its size in phase contrast images (see Materials and methods).

We plotted the repressor production rate $f_T(R_T)$ (Figure 2, green and black for the two variants of P_R) and the repressor dilution rate $D(R_T)=\alpha \cdot R_T$ (Figure 2, magenta), where α denotes the bacterial growth rate (Rosenfeld *et al*, 2002). Its average value of $\alpha=0.019\,\text{min}^{-1}$ was obtained directly from the image data, by measuring the mean cell-cycle period, $\tau=36\,\text{min}$ (Figure 1C). We predict that the repressor expression level in the autoregulatory circuit should be centered on the intersection of the two functions: $f_T(R_T)=D(R_T)$.

To test this prediction, we acquired and quantified time-lapse fluorescence microscopy movies of growing colonies of *E. coli* strains containing the autoregulatory feedback circuits (Supplementary movies SM1 and SM2). In these circuits, the CI-YFP repressor–reporter fusion protein represses its own production by binding to the P_R promoter. The negative autoregulatory synthetic circuits were inserted at the same chromosomal locus (*galK*) that we used for the P_R measurement constructs (see Materials and methods). We found that repressor levels increased steadily over each cell cycle and dropped by approximately half at each cell division, maintaining concentrations at an approximately constant value (Figure 1C). The wild-type P_R promoter is more tightly repressed than the O_R2^* variant (Meyer *et al*, 1980; Rosenfeld *et al*, 2005), and the repressor levels are therefore lower in the wild-type negative autoregulatory circuit (Figure 1C).

As predicted, repressor levels and production rates for the negative autoregulatory circuits were centered at the intersection of the dilution rate and the GRF curves (Figure 2). The
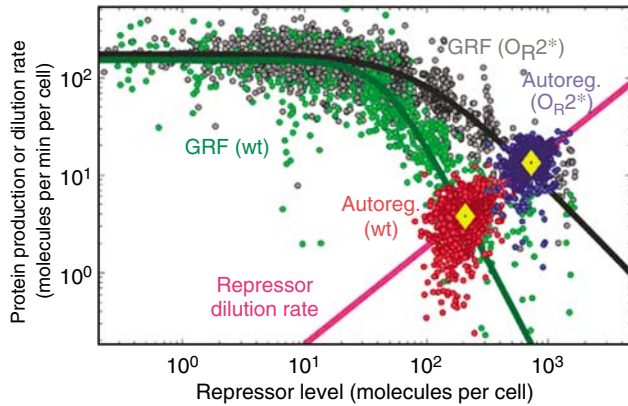
Figure 2 Predicted and actual behavior of the synthetic negative auto-regulatory circuits. GRFs are plotted as a function of repressor level R_T for wild-type P_R promoter (green) and the O_R2^* variant (black). GRF is the protein production rate f_T (R_T), as a function of repressor level and was measured using the 'λ-cascade' strains (Supplementary Figure S2) (Rosenfeld et al, 2005). The average repressor dilution rate D(R_T) is linear in R_T, and is plotted in magenta. Its slope, α, was obtained from direct measurements of cell growth rates in these movies. Data obtained from the negative autoregulatory circuits are super-imposed. Red and blue dots represent wild-type and O_R2^* circuit variants, respectively. Yellow diamonds indicate the mean protein concentration and production rates of the two autoregulatory circuits. As predicted, these mean values occur where the production rate, given by the GRFs (green and black), intersects the repressor dilution rate (magenta line).

steady-state level and production rate of CI-YFP (mean ± s.d.) were, respectively, 210 ± 70 molecules per cell and 4 ± 2 molecules per cell per min for the wild-type negative feedback circuit, and 720 ± 190 molecules per cell and 14 ± 5 molecules per cell per min, respectively, for the O_R2^* variant. The single nucleotide change from the wild-type P_R promoter to O_R2^* thus increased the steady-state level of the negative auto-regulatory circuit approximately fourfold.

Previously, we found that intracellular noise causes the GRF to fluctuate significantly about its mean value. Thus, the GRF cannot be described by a deterministic function alone (Rosenfeld et al, 2005). Although negative autoregulatory circuits have been shown to produce faster response times (Savageau, 1974; Rosenfeld et al, 2002), they are often considered to act to reduce gene-expression noise (Becskei and Serrano, 2000; Thattai and van Oudenaarden, 2001). Theoretical (Paulsson, 2004; Swain, 2004; Austin et al, 2006) and simulation (Hooshangi and Weiss, 2006) studies show that this reduction does not always occur but depends strongly on parameter values and external effects. Using the circuits described here, we could not detect a measurable decrease in gene expression variability. In the O_R2^* circuit, protein concentration variation was equal to that observed in the corresponding GRF (coefficient of variation of 15% in the mean cell fluorescence and 37% in the protein production rate, for both strains). The more strongly repressed autoregulation circuit, however, exhibited approximately fourfold lower protein levels and approximately 1.5-fold greater variability (Supplementary Figure S3). Theoretically, intrinsic noise would be expected to increase at most by the square root of the change in expression level, or approximately twofold in this case. The effects of feedback can reduce the magnitude of this noise increase (Swain, 2004). Thus, the observed noise

increase is consistent with the theoretical expectations (Becskei and Serrano, 2000; Thattai and van Oudenaarden, 2002; Elowitz et al, 2002; Ozbudak et al, 2002; Swain et al, 2002; Paulsson, 2004; Rosenfeld et al, 2005; Hooshangi and Weiss, 2006).

These results show that accurate quantitative measurements of GRFs and other component properties can allow prediction of the behaviors of autoregulatory genetic feedback circuits. It will now be important to investigate the degree of complexity and the range of cellular environments over which such quantitative descriptions can be extended. An advantage of the lambda phage system is that it has very specific interactions with its target promoters in E. coli. It can thus be treated as a relatively independent circuit module. It will be interesting to see how many other similarly modular systems can be constructed. As the complexity of components increases, GRFs must be more complex than Hill functions to effectively describe combinatorial interactions (Bintu et al, 2005; Mayo et al, 2006; Guido et al, 2006; Libby et al, 2007). For sufficiently complex systems, completely characterizing the GRF may become impractical and other simplifications may be required. Nevertheless, these results support the view that quantitatively predictive system models are feasible, and that such models can be used for bottom-up design of synthetic circuits (Guido et al, 2006).

Materials and methods

Bacterial strains and constructs

λ-cascade strains were described previously (Rosenfeld et al, 2005). E. coli strains encoding chromosomally integrated negative autoregulatory circuits (Figure 1A; Supplementary Figure S1) were constructed as follows. First, complementary oligonucleotides encoding the lambda P_R promoter were designed with appropriate cohesive ends to replace the $P_L$lacO1 promoter driving cI-yfp expression in previously described plasmid pZE21-clyfp (Rosenfeld et al, 2005). The P_R promoter contains O_R2 and O_R1 repressor-binding sites (Meyer et al, 1980), conserves the spacing between them (Liu and Little, 1998), and also contains the following sequence: …ataaatatc[taacaccgtgcGtgttg]actattt[tacctctggcg gtgata]atggttgcatgtactagaattcattaaagaggagaaaggtaccATG… (capital ATG marks translation start; brackets indicate O_R2 and O_R1; capital G marks the base that is changed in the O_R2^* variant). Site-directed mutagenesis of this promoter (G is changed to T; see Supplementary Figure S1) yielded the O_R2^* promoter variant, previously designated as 'vN' (Meyer et al, 1980). These steps resulted in the plasmid-based negative autoregulatory circuits. Homologous recombination of the construct into the chromosome of strain MC4100Z1 was performed using a 'recombineering' technique. Briefly, PCR products of both of the complete transcriptional units, along with an adjacent amp^R marker, were generated with ends homologous to the galK locus, as described previously (Elowitz et al, 2002). Electroporation of these products into heat-shocked MC4100Z1 containing pSIM5 (Yu et al, 2000) and subsequent recovery at 37°C on Amp plates resulted in the desired strains. Plasmid pSIM5, a generous gift of DL Court, contains the genes necessary for recombination under heat-shock control on a plasmid with a tsSC101 origin that is lost at 37°C.

Bacterial growth and medium

Cultures were grown overnight in LB + 15 µg/ml kanamycin at 37°C from single colonies, and diluted 1:100 in MSC medium (M9 minimal medium + 0.6% succinate + 0.01% casamino acids + 0.15 µg/ml bio-tin + 1.5 µM thiamine). Cultures were grown to OD$_{600}$ ∼0.1 at 32°C, and then diluted to give ∼1 cell per visual field when placed between a coverslip and 1.5% low-melt MSC agarose.

Quantitative fluorescence microscopy

Growth of bacterial microcolonies was observed by fluorescence microscopy at 32°C using an automated Leica DMIRB/E microscope with ×100 phase contrast objective, an Orca ERG-cooled CCD camera (Hamamatsu), and custom acquisition software. Typical intervals between subsequent exposures were 8–9 min. Custom software was developed using MATLAB (The Mathworks Inc) to analyze time-lapse movie data (Rosenfeld *et al*, 2005). The cell length (l, typical values are 3–4 μm) and width (w, narrowly distributed around 0.75 μm) were recorded, and cell volume was calculated by modeling the cell as a cigar-shape cylinder of length ($l-w$) and radius ($w/2$), capped by two hemispheres of radius ($w/2$). Cell volumes typically varied from 1–2 to 2–3 μm^3. Fluorescence levels were translated into units of fluorescent proteins using a fluctuation method, which compares the distribution of sister-cell fluorescence values after cell-division to a hypothesis of binomial protein segregation (Rosenfeld *et al*, 2005, 2006). Protein production rate was averaged between subsequent movie frames. Segmentation errors can contribute a relative error of a few percent, and calibration errors can contribute a systematic additive error in the order of 10 molecules per cell.

Supplementary information

Supplementary information is available at the *Molecular Systems Biology* website (www.nature.com/msb).

References

Alon U (2007) *An Introduction to Systems Biology: Design Principles of Biological Circuits*. Chapman & Hall/CRC: Boca Raton, FL

Andrianantoandro E, Basu S, Karig DK, Weiss R (2006) Synthetic biology: new engineering rules for an emerging discipline. *Mol Syst Biol* **2**: 2006 0028

Austin DW, Allen MS, McCollum JM, Dar RD, Wilgus JR, Sayler GS, Samatova NF, Cox CD, Simpson ML (2006) Gene network shaping of inherent noise spectra. *Nature* **439**: 608–611

Bakk A, Metzler R (2004) *In vivo* non-specific binding of lambda CI and Cro repressors is significant. *FEBS Lett* **563**: 66–68

Becskei A, Serrano L (2000) Engineering stability in gene networks by autoregulation. *Nature* **405**: 590–593

Bintu L, Buchler NE, Garcia HG, Gerland U, Hwa T, Kondev J, Phillips R (2005) Transcriptional regulation by the numbers: models. *Curr Opin Genet Dev* **15**: 116–124

Darling PJ, Holt JM, Ackers GK (2000) Coupled energetics of lambda cro repressor self-assembly and site-specific DNA operator binding II: cooperative interactions of cro dimers. *J Mol Biol* **302**: 625–638

Dodd IB, Shearwin KE, Perkins AJ, Burr T, Hochschild A, Egan JB (2004) Cooperativity in long-range gene regulation by the lambda CI repressor. *Genes Dev* **18**: 344–354

Elowitz MB, Leibler S (2000) A synthetic oscillatory network of transcriptional regulators. *Nature* **403**: 335–338

Elowitz MB, Levine AJ, Siggia ED, Swain PS (2002) Stochastic gene expression in a single cell. *Science* **297**: 1183–1186

Gardner TS, Cantor CR, Collins JJ (2000) Construction of a genetic toggle switch in *Escherichia coli*. *Nature* **403**: 339–342

Guet CC, Elowitz MB, Hsing W, Leibler S (2002) Combinatorial synthesis of genetic networks. *Science* **296**: 1466–1470

Guido NJ, Wang X, Adalsteinsson D, McMillen D, Hasty J, Cantor CR, Elston TC, Collins JJ (2006) A bottom-up approach to gene regulation. *Nature* **439**: 856–860

Hooshangi S, Weiss R (2006) The effect of negative feedback on noise propagation in transcriptional gene networks. *Chaos* **16**: 026108

Libby E, Perkins TJ, Swain PS (2007) Noisy information processing through transcriptional regulation. *Proc Natl Acad Sci USA* **104**: 7151–7156

Liu Z, Little JW (1998) The spacing between binding sites controls the mode of cooperative DNA–protein interactions: implications for evolution of regulatory circuitry. *J Mol Biol* **278**: 331–338

Mayo AE, Setty Y, Shavit S, Zaslaver A, Alon U (2006) Plasticity of the *cis*-regulatory input function of a gene. *PLoS Biol* **4**: e45

Meyer BJ, Maurer R, Ptashne M (1980) Gene regulation at the right operator (OR) of bacteriophage lambda. II. OR1, OR2, and OR3: their roles in mediating the effects of repressor and cro. *J Mol Biol* **139**: 163–194

Ozbudak EM, Thattai M, Kurtser I, Grossman AD, van Oudenaarden A (2002) Regulation of noise in the expression of a single gene. *Nat Genet* **31**: 69–73

Paulsson J (2004) Summing up the noise in gene networks. *Nature* **427**: 415–418

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

Rosenfeld N, Young JW, Alon U, Swain PS, Elowitz MB (2005) Gene regulation at the single-cell level. *Science* **307**: 1962–1965

Rosenfeld N, Elowitz MB, Alon U (2002) Negative autoregulation speeds the response times of transcription networks. *J Mol Biol* **323**: 785–793

Rosenfeld N, Perkins TJ, Alon U, Elowitz MB, Swain PS (2006) A fluctuation method to quantify *in vivo* fluorescence data. *Biophys J* **91**: 759–766

Savageau MA (1974) Comparison of classical and autogenous systems of regulation in inducible operons. *Nature* **252**: 546–549

Sprinzak D, Elowitz MB (2005) Reconstruction of genetic circuits. *Nature* **438**: 443–448

Swain PS (2004) Efficient attenuation of stochasticity in gene expression through post-transcriptional control. *J Mol Biol* **344**: 965–976

Swain PS, Elowitz MB, Siggia ED (2002) Intrinsic and extrinsic contributions to stochasticity in gene expression. *Proc Natl Acad Sci USA* **99**: 12795–12800

Thattai M, van Oudenaarden A (2001) Intrinsic noise in gene regulatory networks. *Proc Natl Acad Sci USA* **98**: 8614–8619

Thattai M, van Oudenaarden A (2002) Attenuation of noise in ultrasensitive signaling cascades. *Biophys J* **82**: 2943–2950

Yu D, Ellis HM, Lee EC, Jenkins NA, Copeland NG, Court DL (2000) An efficient recombination system for chromosome engineering in *Escherichia coli*. *Proc Natl Acad Sci USA* **97**: 5978–5983