

Multistability in the lactose utilization network of *Escherichia coli*

Supplementary Information

Positive feedback and bistability. We model the *lac* system using the following equations:

$$\frac{R}{R_T} = \frac{1}{1 + (x/x_0)^n} \quad (S1)$$

$$\tau_y \frac{dy}{dt} = \alpha \frac{1}{1 + R/R_0} - y \quad (S2)$$

$$\tau_x \frac{dx}{dt} = \beta y - x \quad (S3)$$

Here, x is the intracellular TMG concentration, y is the concentration of LacY in green fluorescence units, R_T is the total concentration of LacI tetramers, and R is the concentration of active LacI. The active fraction of LacI is a decreasing sigmoidal function of the TMG concentration x , with half-saturation concentration x_0 , and Hill coefficient n (Eq. S1). This sigmoidal behaviour arises from the fact that the binding of TMG to any one of four possible sites on the LacI tetramer is sufficient to interfere with LacI activity, while higher TMG occupancies cause even further impairment. There is extensive experimental evidence^{S1} showing that $n \approx 2$. The interaction of a single active LacI tetramer with multiple operator sites on the *lac* promoter generates a DNA loop which blocks transcription. The rate of generation of LacY is therefore a decreasing hyperbolic function of R , with maximal value α , half-saturation concentration R_0 , and minimal value α/ρ achieved at $R = R_T$. The repression factor $\rho = 1 + R_T/R_0$ describes how tightly LacI is able to regulate *lac* expression. LacY is depleted in a first-order reaction with time constant τ_y , due to a combination of degradation and dilution (Eq. S2). TMG enters the cell at a rate proportional to y , and is similarly depleted in a first-order reaction with time constant τ_x (Eq. S3). The parameter β measures the TMG uptake rate per LacY molecule. Since we cannot directly measure x , we are free to choose its units so that $x_0 = 1$. Once inside the cell, TMG is able to inactivate LacI, completing the feedback loop. Combining these equations, we obtain the steady state result

$$y = \alpha \frac{1 + (\beta y)^2}{\rho + (\beta y)^2} \quad (S4)$$

Here, ρ , α and β are allowed to be arbitrary functions of our external inputs, the extracellular glucose (G) and TMG (T) levels. As these parameters are varied, the system is capable of generating either one or two stable fixed points, with saddle-node bifurcations¹² separating these two behaviours (Fig. 4d).

Theoretical phase diagram. The boundary between monostability (one stable fixed point) and bistability (two stable fixed points separated by one unstable fixed point) occurs when Eq. S4 admits precisely two solutions: this signifies the onset of a saddle-node bifurcation. Rewriting Eq. S4 as a cubic, we obtain

$$y^3 - \alpha y^2 + (\rho/\beta^2)y - (\alpha/\beta^2) = 0. \quad (S5)$$

On the other hand, a general cubic with two identical roots has the form

$$(y-a)(y-a)(y-\theta a) = y^3 - (2+\theta)ay^2 + (1+2\theta)a^2y - \theta a^3 \quad (S6)$$

where θ is the dimensionless ratio of roots. Comparing coefficients, we find

$$\begin{aligned} \rho &= (1+2\theta)(1+2/\theta), \\ \alpha\beta &= (2+\theta)^{3/2}/\theta^{1/2}. \end{aligned} \quad (S7)$$

These are the parametric equations describing the boundary of the bistable region (Fig. 4d). The critical point occurs where all three roots coincide, so $\theta = 1$.

Calculation of parameters. The two fluorescence values under bistable conditions do not by themselves provide enough information to uniquely specify the underlying parameters. However, by also applying the saddle-node condition at the switching thresholds, we are able to obtain three equations for the three unknowns ρ , α and β . This can be done separately at each threshold (OFF and ON). We can therefore solve for these parameters on the boundaries of the bistable region, thereby determining their complete functional dependence on G and T . However, there are three caveats. First, we find that α is systematically higher (by about 15%) at the OFF threshold than at the ON threshold, though it shows precisely the same linear behaviour at both thresholds, over the entire range (a factor of 5 variation) of red fluorescence levels (Fig. 3a). This small discrepancy could arise from a systematic error in our estimate of the induced fluorescence, since our measurements are performed near but not precisely at the switching threshold. However, it might also arise due to a weak competitive or cooperative interaction between CRP and LacI at the *lac* promoter, which we have neglected in our model. Second, the low fluorescence values at the OFF threshold are very close to background, introducing a large error in the calculation of ρ . We therefore estimate both α and ρ at the ON threshold alone, and use this information to calculate β at both thresholds. Third, we can decompose the net TMG uptake rate as $\beta(T, G) = \beta_T(T) \beta_G(G)$, where T and G are the extracellular glucose and TMG concentrations. Assuming a power-law for $\beta_T(T)$, we use a least-squares fitting routine to extract the functions $\beta_T(T)$ and $\beta_G(G)$. Note that we have normalized units to give $\beta_G(G=0) = 100$, and $\alpha(G=0) = 100$. These calculations produce the following results (Fig. 3):

$$\begin{aligned} \alpha &= \frac{84.4}{1 + (G/8.1)^{1.2}} + 16.1, \quad \rho = 167.1, \\ \beta_T &= (1.23 \times 10^{-3})T^{0.6}, \quad \beta_G(G > 10) \cong 65. \end{aligned} \quad (S8)$$

where G and T are measured in μM . The uncertainties in the estimates of α and β_G are shown in Fig. 3e; the uncertainty in β_T is similar to that shown for β_G . There is a 50% uncertainty in each measurement of ρ , and a 20% uncertainty in its final fitted value.

Value of the repression factor. The most direct way to measure the operon repression factor ρ is to take the ratio of fluorescence levels of fully induced cells (grown in saturating amounts of TMG) and uninduced cells (grown in the absence of TMG). However, because we use a single copy chromosomal reporter, the fluorescence levels of uninduced cells are very close to the measurement background of the camera, and comparable to cell autofluorescence. We therefore chose to determine the repression factor by the fitting technique discussed above. We obtained a repression factor $\rho = 170 \pm 30$ for the wild-type system, whereas previous studies¹⁹ report much higher repression factors, of order 1000. This difference could be due to the following reasons. First: The wild-type *lac* promoter contains three operator sites (O_1 , O_2 , and O_3) to which LacI binds. The promoter used in our reporter construct is missing the O_2 operator site, leading to a decrease in repression efficiency due to a drop in DNA looping activity. Such a promoter has been reported²³ to have a repression factor of only 440. (Note that the native copy of the *lac* promoter still contains all three operator sites, so system response is unaffected.) Second, our fit is performed only at the switching thresholds, far from either the fully induced or uninduced limits. Our fitted value of ρ will probably be different from the value that would be obtained by direct measurement at the two limits, due to small differences between our approximate model and the true system response. However, the fact that we find ρ to be independent of glucose and TMG levels strongly suggests that the true repression factor is similarly independent of these parameters.

Correlated green and red fluorescence levels. We showed in the main text that the mean green fluorescence level of a population of induced cells was directly proportional to the mean red fluorescence level. This carries through to the behaviour of single cells. For example, Fig. S1 shows a scatter plot of green and red fluorescence levels of single cells in a population that is initially uninduced, then grown in 10 μ M glucose and 150 μ M TMG. The population has a bimodal distribution of green

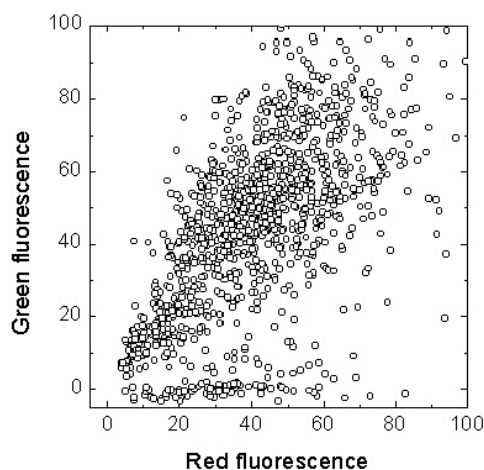


Figure S1 Fluorescence levels of single cells in a bimodal population. Cells in the induced state show strongly correlated green and red fluorescence levels (correlation coefficient = 0.71).

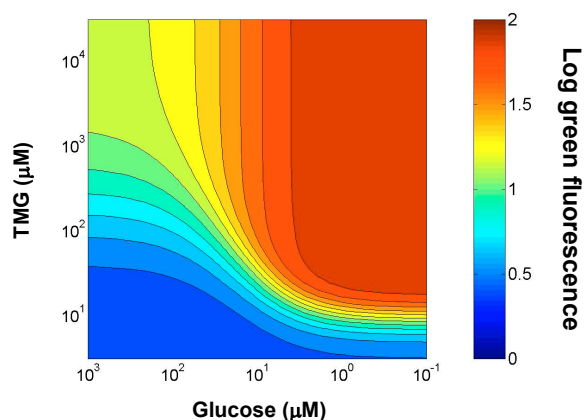


Figure S2 Population averaged *lac* expression levels as a function of glucose and TMG concentrations. These results are obtained by incorporating our fitted parameters into a stochastic version of the positive feedback model. Axes are oriented so that cells are uninduced at the bottom left corner.

fluorescence levels because it is close to a switching threshold. The green and red fluorescence of the induced subpopulation are highly correlated.

Response to IPTG and lactose. For completeness, we conducted a series of experiments using isopropyl-beta-D-thiogalactopyranoside (IPTG) and lactose as inducers in place of TMG.

During induction with IPTG, cells show a persistent bimodal response, but the fluorescence levels of uninduced cells are higher than with TMG. It is known that IPTG is able to enter cells independently of LacY, accounting for the increased fluorescence of the uninduced cells. However, the persistence of bimodality indicates that LacY continues to play a role in the active transport of IPTG^{S2}, preserving positive feedback.

During induction with lactose, initially uninduced cell populations show a transient bimodal distribution of green fluorescence levels at certain glucose concentrations, and a transient unimodal distribution at others. However, the steady state distribution after 4 hours of growth is *always* unimodal, and we never observe hysteresis. By performing extensive measurements, we confirmed this unimodal behaviour to occur for over fifty combinations of glucose and lactose concentrations, upto saturating quantities of each sugar. The difference between the observed responses to TMG and lactose could be due to several causes. First, because lactose is metabolized and therefore affects cell growth rate, it could happen that the induced subpopulation of a bimodal population always grows to dominance. Second, since the metabolism of lactose leads to a drop in CRP-cAMP levels¹⁸, the inducer activity of allolactose might be counteracted to some extent. Third, although an increase in operon expression leads to an increase in lactose uptake and allolactose production, it also leads to an increase in allolactose degradation by β -galactosidase. Intracellular allolactose levels therefore depend very weakly on operon expression levels, reducing the strength of positive feedback and possibly eliminating bistability altogether.

Population averaged fluorescence levels. Many previous studies of *lac* operon expression have focussed on population-averaged measurements¹⁹, whereas we have used single-cell measurements. These two approaches should produce similar results if the cell populations are homogeneous, but not if they are heterogeneous. In order to connect with previous studies, we incorporated our fitted parameters in a stochastic model⁸, thus generating a map of population-averaged *lac* expression levels as glucose and TMG concentrations were varied (Fig. S2). In these calculations, we assumed that the system had reached a steady state distribution between the induced and uninduced sub-populations, corresponding to a timescale much longer than those at which hysteresis would be observed. The population-averaged fluorescence can increase either because the fluorescence level of induced cells increases (due to a decrease in glucose) or because the fraction of cells in the induced state increases (due to an increase in TMG). It is only by performing single-cell measurements that these two effects can be distinguished.

-
- S1. Yagil, G. & Yagil, E. On the relation between effector concentration and the rate of induced enzyme synthesis. *Biophys. J.* **11**, 11-27 (1971).
- S2. Hansen, L. H., Knudsen, S. & Sorensen, S. J. The effect of the *lacY* gene on the induction of IPTG inducible promoters, studied in *Escherichia coli* and *Pseudomonas fluorescens*. *Curr. Microbiol.* **36**, 341-347 (1998).