Molecular Cell, *Volume 46*

Supplemental Information

Mode of Regulation and the Insulation of Bacterial Gene Expression

Vered Sasson, Irit Shachrai, Anat Bren, Erez Dekel, Uri Alon

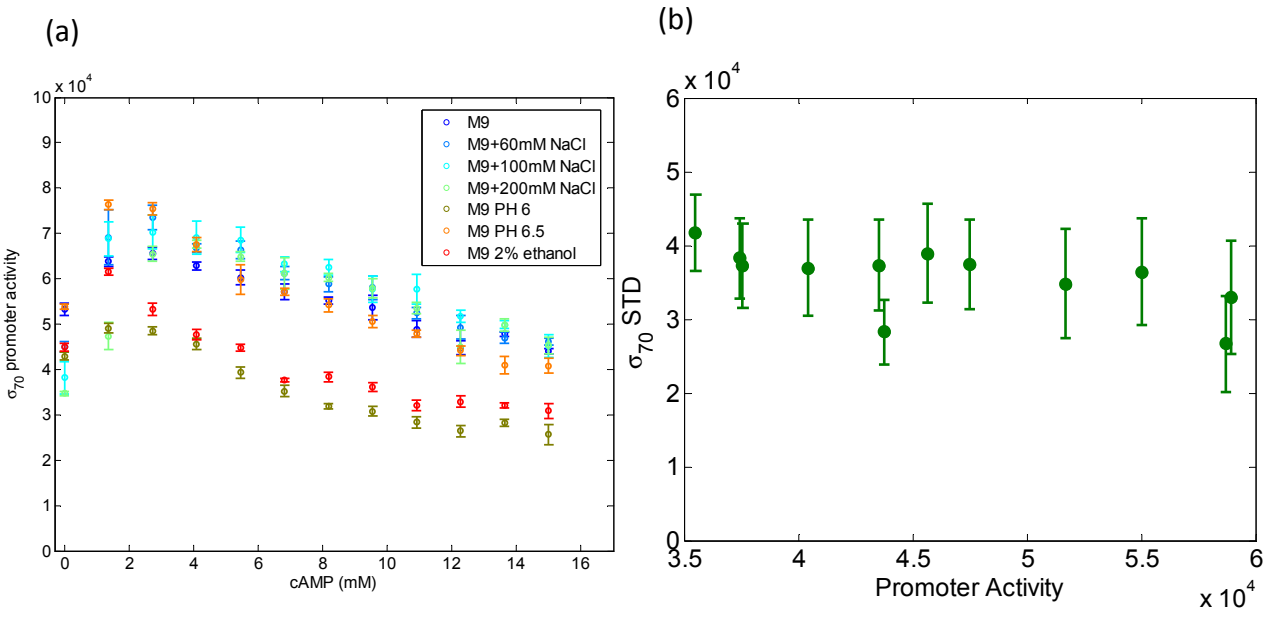


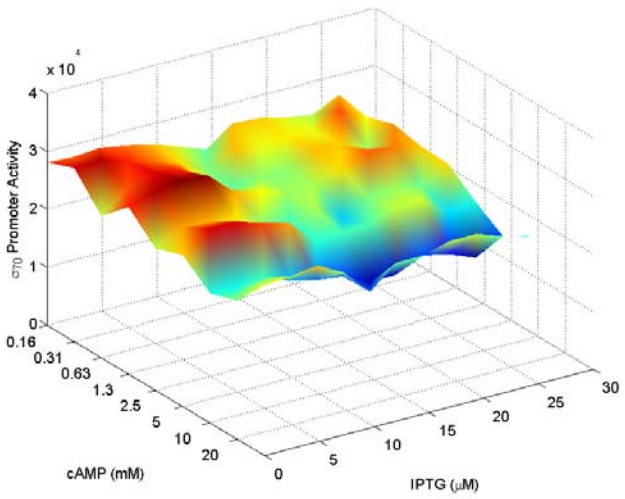
Figure S1. σ_{70} Reporter Activity Shows Weak Dependence on cAMP Levels, Related to Figures 1 and 2

(a) Promoter activity of σ_{70} reporter (U449) used as a normalization control in the present study as a function of cAMP. Error bars are STE between repeats.

(b) STD of σ_{70} reporter activity across different conditions as a function of mean promoter activity is constant within error bar. Each point corresponds to one cAMP level. Error bar are 95% confidence using bootstrapping.

This data shows that the control reporter varies only weakly with cAMP. The large and systematic variation in the CRP constructs in the main text (Fig 1,2) is therefore due to the promoter of interest and not to the U449 normalizing control.

(a)



(b)

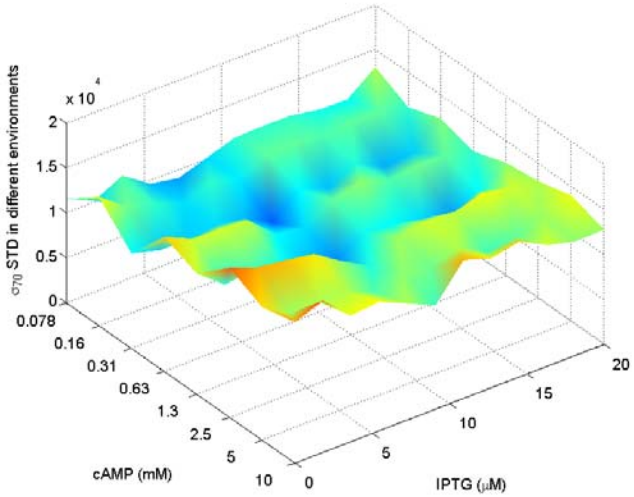


Figure S2. σ_{70} Reporter Activity Used as a Normalizing Control Shows Weak Dependence on cAMP an IPTG Levels, Related to Figure 4

(a) Promoter activity of σ_{70} reporter (U371) in M9 Glucose is almost constant as a function of cAMP and IPTG. (b) The variation across conditions of the σ_{70} reporter construct does not show large systematic changes with the signals IPTG and cAMP. This data shows that the systematic variation in the lac promoter in the main text (Fig 4,5) is due to the promoter of interest and not to the U371 normalizing control.

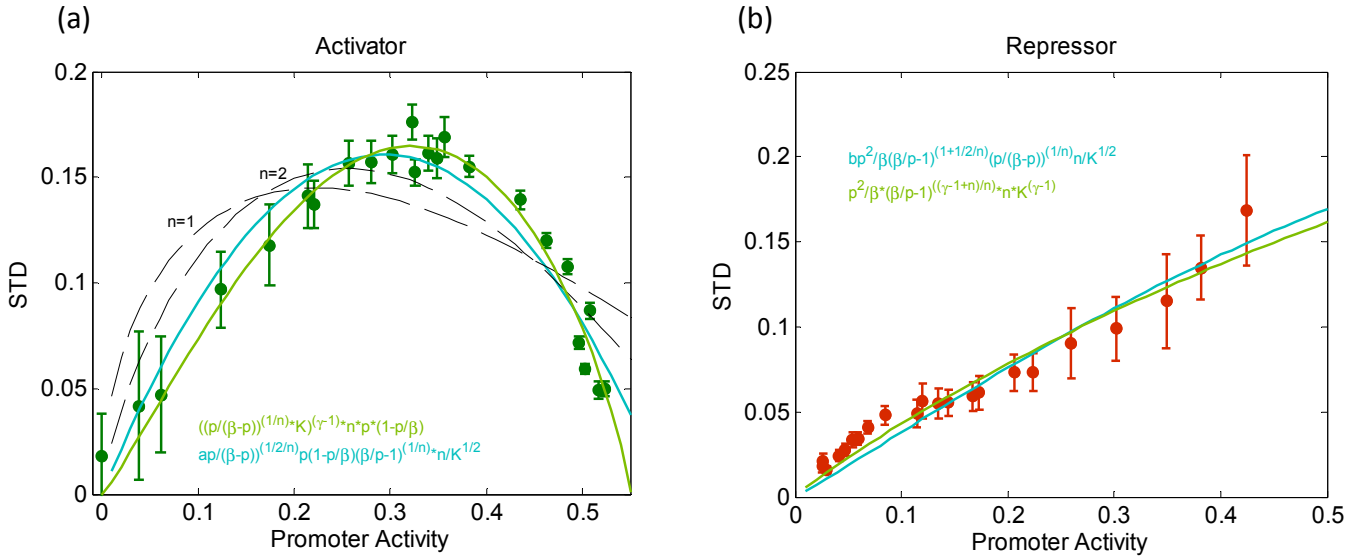


Figure S3. Models in which Cognate Transcription Factor Activity Varies between Conditions Explain Data More Poorly than a Model with Non-specific Binding, Related to Figures 3 and 6

(a) Standard deviation (STD) of activator construct promoter activity (U435). Full line (blue): Fit to model of variability in transcription factor (X) assuming that its scales with activity as $bX^{1/2}$, best fit parameters are $b=3*10^{-4}$, $\beta=0.58$, $k=1.1*10^{6}$, $n=3.5*10^{6}$. Full line (green) assuming a more general dependence of variability on activity, namely aX^{γ}. This general dependence can include, for example, cases where variability does not change with mean activity ($\gamma=0$). Best fit parameters are $a=0.14$, $\beta=0.55$, $k=1.54$, $n=5.4$ and $\gamma=1.9$. Also shown are best fits, fixing the Hill coefficient at $n=1$ and $n=2$. A good fit requires unreasonably high values for the Hill coefficient- as opposed to the model in the main text based on nonspecific binding which fits the data to higher accuracy and does so with reasonable parameters. The model in the main text (nonspecific binding) also has the virtue that its solution is independent on the Hill-coefficient of the induction function.

(b) Standard deviation (STD) of repressor promoter activity (U436). Full line: Fit to model of variability in transcription factor activity (X) assuming that the variability scales with mean activity as $bX^{1/2}$ with $b=0.002$, $\beta=1.9$, $k=7*10^{-4}$, $n=6.5$. Full line (green) assuming a more general scaling of variability with activity, namely aX^{γ}, whose best fit parameters are $a=0.09$, $\beta=3.2$, $k=0.3$, $n=8.4$ and $\gamma=1.7$. The best fit does not capture the plateau in the data at intermediate cAMP levels, which is well described by the model in the main text based on nonspecific binding.

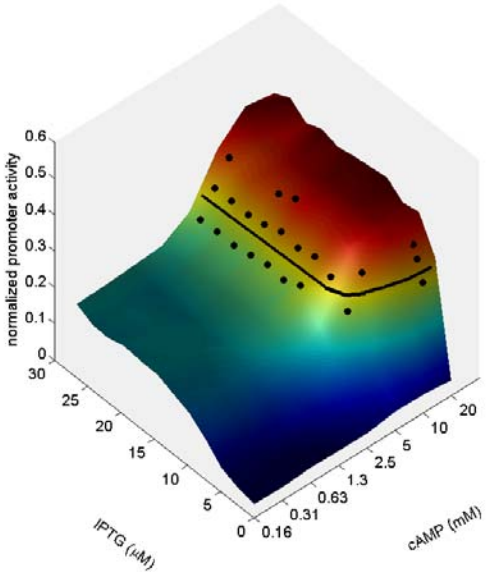


Figure S4. Promoter Activity of Data Points at High Slope Regions of the Input Function, Defined as Points Whose Slope Exceeds 75% of the Maximal Slope of the Input Function, Related to Figure 5

The variation in the measured promoter activity comes from two sources: the condition effect on the input function, and a measurement error due to experimental noise. Experimental noise is expected to be higher in high-slope regions of the input function. For example, errors in the concentration of IPTG or cAMP in the medium will affect the input function in high slope areas more than in low slope area. For stringency, we therefore removed from the analysis points that were in high slope areas: points with gradients higher than 75% of the maximal gradient were removed from the analysis (white point in figure 5). The STD of these points between different environments is significantly higher than other points.

Figure S5. Detailed Derivation of Model Equations, Related to Figures 3 and 6

In this section, we detail the calculations for the estimated variation between conditions of an induction curve, based on various sources of variation in the regulatory mechanism. We begin with the model in the main text, which describes for effect of nonspecific binding of factors that of activator and repressor regulated promoters.

Consider an activator with a Michaelis-Menten induction curve, Eq 1 below, where promoter activity P_A depends on activator activity X. Halfway induction is at K, and maximal expression is β. We add two sources of nonspecific binding, as random variables that change between conditions. N describes nonspecific binding that interferes with RNA polymerase binding (or more generally with transcription initiation), and N' interferes with binding of the cognate transcription factor (Eq 2). In the case of an activator, both N and N' do not appear in the numerator because they interfere with expression. We next write N and N' as <N>+dN and <N'>+dN' ; the terms in brackets are the means and dN and dN' are random variables with mean zero. Assuming for simplicity that dN and dN' are independent, the squared standard deviation of the induction curve is the sum of the squares of the standard deviations of N and N' , Eq (3). [if N and N' are not independent, one adds a covariance term which does not materially affect the qualitative conclusions]. To evaluate the standard deviations of N and N', we assume that the variations dN and dN' are much smaller than K+<N>+<N'>. This assumption is plausible because non-specific binding is expected to be weaker than cognate binding. As a result, one can use a Taylor expansion to find that standard deviations in promoter activity are the derivative of the induction curve with respect to the varying quantity times its standard deviation (Eq 4,5). Plugging this into Eq 3 results in the final expression for the STD of the promoter activity as a function of promoter activity (Eq 6,7).

Equivalent derivation for a repressor is given by Eqs (8-14). Note that here, N' appears in the numerator because interfering with repressor binding contributes to expression.

Repeating the calculations with Hill-curves instead of Michaelis-Menten curves yields exactly the same formulae Eq 7,14. In other words, the predicted variation between conditions, in the case of non-specific binding, is independent on the Hill coefficient of the induction curve.

$$(1)\ P_A = \frac{\beta \cdot X}{K + X} \qquad (8)\ P_R = \frac{\beta \cdot K}{K + X}$$

$$(2)\ P_A = \frac{\beta \cdot X}{K + X + N + N'} \qquad (9)\ P_R = \frac{\beta \cdot (K + N')}{K + X + N + N'}$$

$$(3)\ \sigma_{P_A} = \sqrt{(\sigma_{P_A}^{N})^2 + (\sigma_{P_A}^{N'})^2} \qquad (10)\ \sigma_{P_R} = \sqrt{(\sigma_{P_R}^{N})^2 + (\sigma_{P_R}^{N'})^2}$$

$$(4)\ \sigma_{P_A}^{N} = \frac{dP_A}{dN} \cdot \sigma_N = \frac{-P_A}{K + X + N + N'} \cdot \sigma_N \qquad (11)\ \sigma_{P_R}^{N} = \frac{dP_R}{dN} \cdot \sigma_N = \frac{-P_R}{K + X + N + N'} \cdot \sigma_N$$

$$(5)\ \sigma_{P_A}^{N'} = \frac{dP_A}{dN'} \cdot \sigma_{N'} = \frac{-P_A}{K + X + N + N'} \cdot \sigma_{N'} \qquad (12)\ \sigma_{P_R}^{N'} = \frac{dP_R}{dN'} \cdot \sigma_{N'} = \frac{\beta \cdot (X + N)}{(K + X + N + N')^2} \cdot \sigma_{N'}$$

$$(6)\ \sigma_{P_A} = P_A \cdot (1 - \frac{P_A}{\beta}) \cdot \underbrace{\frac{\sqrt{(\sigma_{P_A}^{N})^2 + (\sigma_{P_A}^{N'})^2}}{K + N + N'}}_{c}$$

$$(13)\ \sigma_{P_R} = \sqrt{[\frac{P_R^2}{\beta} \cdot \underbrace{\frac{\sigma_{P_R}^{N}}{K + N'}}_{a}]^2 + [P_R \cdot (1 - \frac{P_R}{\beta}) \cdot \underbrace{\frac{\sigma_{P_R}^{N'}}{K + N'}}_{b}]^2}$$

$$(7)\qquad \sigma_{P_A} = P_A \cdot (1 - \frac{P_A}{\beta}) \cdot c \qquad (14)\qquad \sigma_{P_R} = \sqrt{[\frac{P_R^2}{\beta} \cdot a]^2 + [P_R \cdot (1 - \frac{P_R}{\beta}) \cdot b]^2}$$

We finally describe the models in which the activity of the cognate regulator X varies between conditions (whose best fit results are plotted in Fig S3). The variation in X can be due to condition-specific factors that affect X expression, localization, stability or impinge in other ways on its activity. We use the random variable dX to describe these effects, so that activity is X+dX, and again assume that dX is much smaller than X. The induction curves are assumed to be Hill-functions with Hill coefficient n. We also allowed the variation in X activity to scale with mean activity, such that the STD of X is $\sigma_x = bX^{\gamma}$. The case $\gamma=0$ describes no dependence of variation on mean activity, $\gamma=1/2$ describes variation that scales as the square root of activity, etc. The derivation proceeds along the same lines as described above. The resulting equations are given below. Note that unlike the case of non-specific binding, the results of this model show explicit dependence on the Hill coefficient.

$$\sigma_{P_{ax}} = bP\left(1 - \frac{P}{\beta}\right) nK^{\gamma-1} \left(\frac{P}{\beta - P}\right)^{\frac{\gamma-1}{n}} \qquad \sigma_{P_{rx}} = b\frac{P^2}{\beta} nK^{\gamma-1} \left(\frac{\beta}{P} - 1\right)^{\frac{\gamma+n-1}{n}}$$