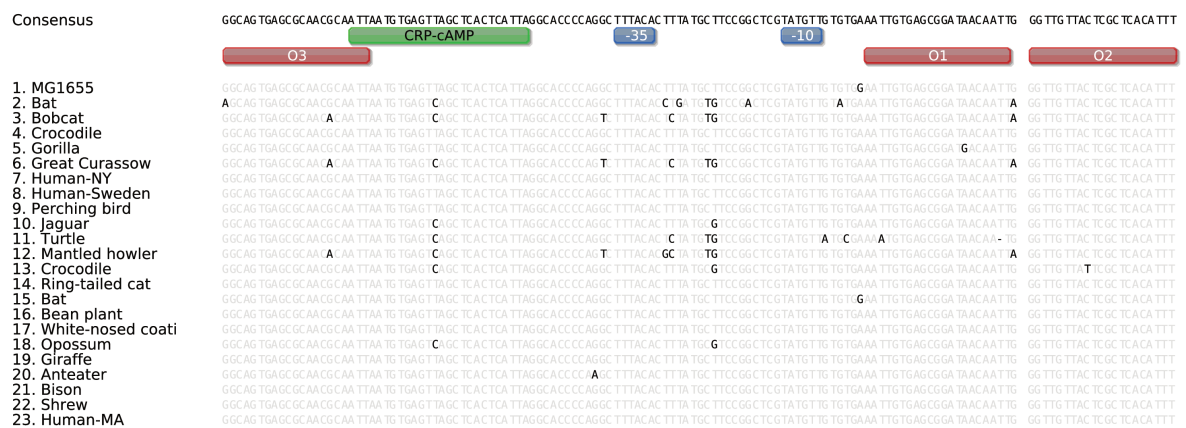


# Supplemental Material

## S1. Alignment of promoter sequences

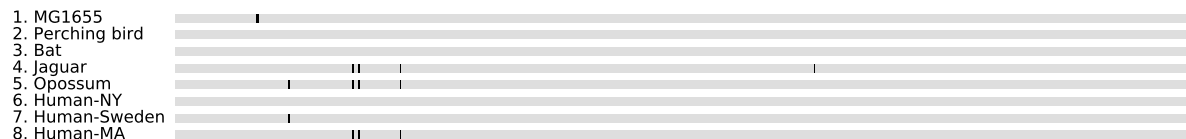
Figure S1 shows the alignment of the promoter regions of the *E. coli* wild isolates sequenced.



**Figure S1.** Promoter alignment of the sequenced strains. Highlighted bases differ from the consensus sequence on top. Colored boxes indicate the relevant binding sites for the Lac repressor (red), CRP (green) and RNAP (blue)

## S2. 16S rRNA sequences

To confirm the identity of the strains we analyzed 490 bp of the 16S rRNA. Figure S2 shows a schematic representation of the sequences. Colored basepairs represent mutations with respect to the consensus sequence. All sequences were found to be  $\geq 99\%$  similar to the reference *E. coli* MG1655 sequence.



**Figure S2.** 16S sequence alignment. Black lines represent mutations with respect to the consensus sequence.

### S3. Model parameters

Table S1 shows the values of the reference parameters for MG1655 obtained from different sources.

**Table S1.** Reference parameters for the strain MG1655.

Parameter	Symbol	Value	Units	Reference
$O_1$ repressor operator binding energy	$\Delta\epsilon_r^{O_1}$	-15.3	$k_B T$	[1]
$O_2$ repressor operator binding energy	$\Delta\epsilon_r^{O_2}$	-13.9	$k_B T$	[1]
$O_3$ repressor operator binding energy	$\Delta\epsilon_r^{O_3}$	-9.7	$k_B T$	[1]
Repressor copy number	$R$	20	tetramer/cell	Measured
Activator binding energy	$\Delta\epsilon_a$	-13	$k_B T$	[2, 3]
Number of active activators	$A$	55	active molecules/cell	[2]
RNAP binding energy for the <i>lac</i> promoter	$\Delta\epsilon_p$	-5.35	$k_B T$	[4]
RNAP copy number	$P$	5500	active molecules/cell	[5]
Number of nonspecific binding sites	$N_{NS}$	$4.6 \times 10^6$	-	GenBank: U00096.2
Looping free energy between $O_1 - O_2$	$\Delta F_{loop(l_{12})}$	4.7	$k_B T$	Fit to data from [6, 7]
Looping free energy between $O_1 - O_3$	$\Delta F_{loop(l_{13})}$	9	$k_B T$	[8]
Looping free energy between $O_2 - O_3$	$\Delta F_{loop(l_{23})}$	5.2	$k_B T$	Fit to data from [6, 7]
RNAP-CRP interaction energy	$\Delta\epsilon_{ap}$	-5.3	$k_B T$	[9, 2]
Lac repressor - CRP interaction energy	$\Delta\epsilon_{ar}$	-5.5	$k_B T$	Fit to data from [6, 7]

### S4. Derivation of the repression level equation

Thermodynamic models of gene regulation consider that the gene expression level is proportional to the probability of finding the RNAP bound to the promoter region [3, 10, 11, 12]. This biologically simplistic but powerful predictive tool allows us to study the effect of different transcription factors in different promoter architectures. In the case of the wild-type (WT) *lac* operon promoter architecture, where we have two different transcription factors involved in the regulation - the activator CRP and the Lac repressor.

The Lac repressor molecule, when bound to the main operator  $O_1$ , blocks the polymerase from binding to the promoter region, stopping the transcription of the operon. CRP plays a double role in the regulation of the operon, activating transcription by recruiting RNAP to the promoter region, and as several experiments have shown, enhancing repression by facilitating the formation of the upstream loop between the  $O_1 - O_3$  operators [13, 14, 15]. Enhanced repression by CRP is due to pre-bending the DNA between  $90^\circ$  and  $120^\circ$  [16], thereby increasing the probability of looping by bringing the *lac* operators closer together. The model captures this effect by adding an interaction term  $\Delta\epsilon_{ar}$  in the states where CRP is bound and the Lac repressor forms a loop between operators  $O_1$  and  $O_3$ .

Assuming quasi-equilibrium conditions for the relevant processes involved in transcription, we can use the Boltzmann distribution to compute the probability of finding the RNAP bound to the promoter region, obtaining

$$\text{GE} \propto \frac{\frac{P}{N_{NS}} e^{-\beta \Delta \varepsilon_p} \left\{ 1 + \frac{2R}{N_{NS}} \left[ e^{-\beta \Delta \varepsilon_r^{O2}} + e^{-\beta \Delta \varepsilon_r^{O3}} \left( 1 + \frac{A}{N_{NS}} e^{-\beta \Delta \varepsilon_a} \right) \right] + \frac{4R(R-1)}{N_{NS}^2} e^{-\beta (\Delta \varepsilon_r^{O2} + \Delta \varepsilon_r^{O3})} \left( 1 + \frac{A}{N_{NS}} e^{-\beta \Delta \varepsilon_a} \right) + \frac{A}{N_{NS}} e^{-\beta (\Delta \varepsilon_a + \Delta \varepsilon_{ap})} \left( 1 + \frac{2R}{N_{NS}} e^{-\beta \Delta \varepsilon_r^{O2}} \right) \right\}}{Z_{tot}}, \quad (1)$$

where GE stands for gene expression,  $Z_{tot}$  represents the partition function for the states shown in Figure 2 in the main text. The presence of CRP in the promoter region is not assumed to influence the kinetics of promoter escape, only the probability of RNAP binding. Tagami and Aiba [17] found that the role of CRP in the activation of the *lac* operon is restricted to the steps up to the formation of the open complex, in other words, the interaction between CRP and the RNAP are not essential for transcription after the formation of the open complex. In our model we capture this effect by including an interaction energy between CRP and the RNAP,  $\Delta \varepsilon_{ap}$ , that has been measured experimentally [2, 9].

In the activation mechanism proposed by Tagami and Aiba [17] CRP bends the DNA and RNAP recognizes the CRP-DNA bent complex. This model would imply that RNAP makes additional contacts with the upstream region of the promoter. Based on this model we assume that the presence of the Lac repressor bound on the  $O_3$  operator and CRP bound on its binding site (without forming a DNA loop between  $O_1 - O_3$ ) allows transcription to occur. Since the RNAP cannot contact the upstream region of the promoter because of the presence of the repressor, the interaction energy between CRP and RNAP is not taken into account in these states.

In order to quantify the influence of Lac repressor on expression levels, we measure repression, which is the fold change in gene expression as a result of the presence of the repressor. This metric has the benefit of normalizing to a strain with an identical genetic background, thus isolating the role of the repressor in regulation. This relative measurement is defined as

$$\text{repression} \equiv \frac{\text{gene expression } (R = 0)}{\text{gene expression } (R \neq 0)}, \quad (2)$$

where  $R$  is the Lac repressor copy number. Computing this we obtain

$$\text{repression} = \frac{\frac{\frac{P}{N_{NS}} e^{-\beta \Delta \varepsilon_p} \left[ 1 + \frac{A}{N_{NS}} e^{-\beta (\Delta \varepsilon_a + \Delta \varepsilon_{ap})} \right]}{1 + \frac{A}{N_{NS}} e^{-\beta \Delta \varepsilon_a} + \frac{P}{N_{NS}} e^{-\beta \Delta \varepsilon_p} \left( 1 + \frac{A}{N_{NS}} e^{-\beta (\Delta \varepsilon_a + \Delta \varepsilon_{ap})} \right)}}{\frac{\frac{P}{N_{NS}} e^{-\beta \Delta \varepsilon_p} \left\{ 1 + \frac{2R}{N_{NS}} \left[ e^{-\beta \Delta \varepsilon_r^{O2}} + e^{-\beta \Delta \varepsilon_r^{O3}} \left( 1 + \frac{A}{N_{NS}} e^{-\beta \Delta \varepsilon_a} \right) \right] + \frac{4R(R-1)}{N_{NS}^2} e^{-\beta (\Delta \varepsilon_r^{O2} + \Delta \varepsilon_r^{O3})} \left( 1 + \frac{A}{N_{NS}} e^{-\beta \Delta \varepsilon_a} \right) + \frac{A}{N_{NS}} e^{-\beta (\Delta \varepsilon_a + \Delta \varepsilon_{ap})} \left( 1 + \frac{2R}{N_{NS}} e^{-\beta \Delta \varepsilon_r^{O2}} \right) \right\}}{Z_{tot}}}. \quad (3)$$

This can be further simplified, resulting in

$$\text{repression} = \frac{\frac{1 + \frac{A}{N_{NS}} e^{-\beta(\Delta\epsilon_a + \Delta\epsilon_{ap})}}{1 + \frac{A}{N_{NS}} e^{-\beta\Delta\epsilon_a} + \frac{P}{N_{NS}} e^{-\beta\Delta\epsilon_p} \left(1 + \frac{A}{N_{NS}} e^{-\beta(\Delta\epsilon_a + \Delta\epsilon_{ap})}\right)}}{1 + \frac{2R}{N_{NS}} \left[ e^{-\beta\Delta\epsilon_r^{O2}} + e^{-\beta\Delta\epsilon_r^{O3}} \left(1 + \frac{A}{N_{NS}} e^{-\beta\Delta\epsilon_a}\right) \right] + \frac{4R(R-1)}{N_{NS}^2} e^{-\beta(\Delta\epsilon_r^{O2} + \Delta\epsilon_r^{O3})} \left(1 + \frac{A}{N_{NS}} e^{-\beta\Delta\epsilon_a}\right) + \frac{A}{N_{NS}} e^{-\beta(\Delta\epsilon_a + \Delta\epsilon_{ap})} \left(1 + \frac{2R}{N_{NS}} e^{-\beta\Delta\epsilon_r^{O2}}\right)}, \quad (4)$$

$Z_{tot}$

the expression we use to predict the repression level of the natural isolates.

#### S4.1. Estimating the number of active CRP molecules

The Catabolite Activator Protein, also known as cAMP-receptor protein (CRP) is a global transcriptional regulator in *E. coli* [18]. As it exists in two forms, the cAMP-CRP complex which is considered as the active state and the inactive state without cAMP bound, the number of active molecules is a function of the cAMP cellular concentration. From a thermodynamic perspective we can estimate this number as

$$[CRP - cAMP] = [CRP] \frac{[cAMP]}{K_{cAMP} + [cAMP]}, \quad (5)$$

where  $[CRP - cAMP]$  is the concentration of active proteins,  $[CRP]$  is the total concentration of this transcription factor,  $[cAMP]$  is the cellular concentration of cAMP and  $K_{cAMP}$  is the in vivo dissociation constant of the cAMP-CRP complex.

Kuhlman et al. [2] reported the values for the CRP concentration ( $[CRP] \approx 1500$  nM) and the dissociation constant ( $K_{cAMP} = 10 \mu\text{M}$ ). Epstein et al. [19] measured the intracellular cAMP concentration in different media, including minimal media with glucose and casamino acids ( $[cAMP] \approx 0.38 \mu\text{M}$ ). Using these values we calculate the number of active CRP molecules as

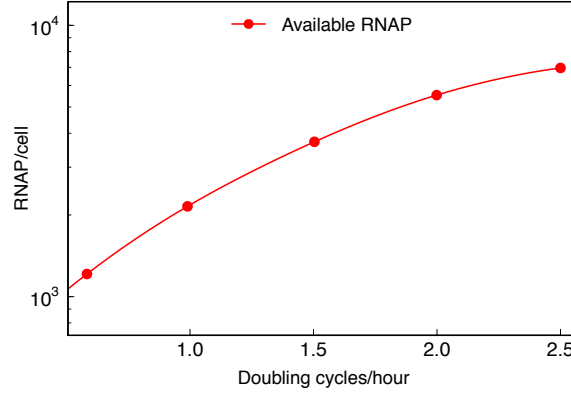
$$A = 1500 \left( \frac{0.38 \mu\text{M}}{10 \mu\text{M} + 0.38 \mu\text{M}} \right) \approx 55 \frac{\text{molecules}}{\text{cell}}, \quad (6)$$

where we used the rule of thumb that  $1 \text{ nM} \approx 1 \frac{\text{molecule}}{E. coli}$ . This rule of thumb is enough for our predictions since the repression level is predicted to be largely insensitive to the activator copy number as shown in Figure 3 in the main text.

#### S4.2. Estimating the number of available RNAP

In order to estimate the available number of RNAP molecules, we appeal to the work of Klumpp and Hwa [5] where they calculated the total number of RNAP molecules as well as the fraction of these molecules available for transcription as a function of the growth rate. Figure S3 shows the number of available RNAP as a function of the doubling cycles per hour.

Using these results, we estimate  $5500 \frac{RNAP}{cell}$  for cells grown in 0.6% glucose + 0.2% casamino acids (with a doubling time of  $\approx 30$  min.). We interpolate between these data to obtain the RNAP copy number for each of the natural isolates.



**Figure S3.** Adapted from Klumpp and Hwa [5]. RNAP available for transcription as a function of the number of doubling cycles per hour.

#### S4.3. Estimating CRP's Binding energy

The activator binding energy was estimated as reported by Bintu et al. [3]. Using the reported dissociation constants from the specific binding site,  $K_{CRP}^{NS}$ , and nonspecific sequences,  $K_{CRP}^S$ , we can compute the binding energy as

$$\frac{\Delta\epsilon_a}{k_B T} = \ln \left( \frac{K_{CRP}^{NS}}{K_{CRP}^S} \right). \quad (7)$$

Bintu et al. also reported the following values for both dissociation constants ( $K_{CRP}^{NS} = 10^4$  nM and  $K_{CRP}^S = 0.02$  nM), which gives us  $\Delta\epsilon_a \approx -13 k_B T$ .

### S5. Fitting parameters and testing the model

The three unknown parameters, the looping energies for the  $O_1 - O_2$  and  $O_3 - O_2$  loops and the decrease in the looping free energy when CRP and Lac repressor are bound at the same time, were inferred from the classic work of Oehler *et al.* [7, 6]. In these papers Oehler and collaborators measured the repression level of different *lac* operon constructs with either mutagenized or swapped Lac repressor binding sites while changing the repressor copy number. Because they reported the mutagenized sequences for the repressor binding sites we used the *sort-seq* derived energy matrix to calculate the residual energies of these modified binding sites. The three unknown parameters were fitted by minimizing the mean square error of the measurements,

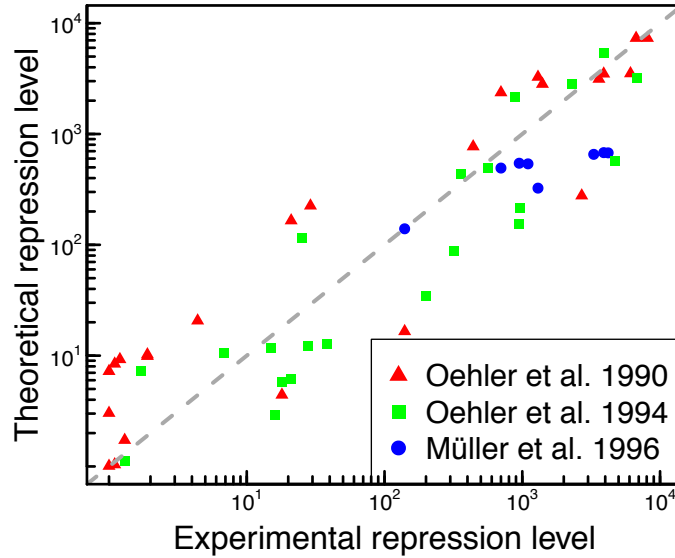
$$f(\mathbf{x}^*) \leq f(\mathbf{x}) \quad \forall \mathbf{x} \in \mathbb{R} \quad (8)$$

$$f(\mathbf{x}^*) = \left\{ \min \sum_{i=1}^N \frac{(Y_i(\mathbf{x}) - \bar{Y}_i)^2}{N} : \mathbf{x} \in \mathbb{R} \right\} \quad (9)$$

where  $Y_i$  is the predicted value,  $\bar{Y}_i$  is the experimental repression level for each of the constructs measured by Oehler *et al.* and  $\mathbf{x}$  are the fitting parameters. Using this method we fit for the values of  $\Delta F_{loop(l_{13})}$ ,  $\Delta F_{loop(l_{23})}$ , and  $\Delta \varepsilon_{ar}$  using the data from references [7, 6]. The three parameter values are listed in Table S1.

## S6. Testing the model with different data

We used the model to predict the repression level of constructs reported by Oehler *et al.* [7, 6] and Müller *et al.* [20]. Figure S4 shows the comparison of the model predictions and the experimental results. The calculations were done using the model whose states are depicted in Figure 2, assuming a wild type repressor copy number of 10 repressors per cell, and calculating all the residual binding energies with the Lac repressor *sort-seq* derived energy matrix.



**Figure S4.** Comparing the experimental data from Oehler *et al.* [7, 6] and Müller *et al.* [20] with the model prediction.

## S7. Error propagation

To calculate a confidence interval of the model, we used the *law of error propagation* [21] where we compute the contribution of the uncertainty in parameters to the uncertainty of the repression level as

$$\sigma_{\text{repression}} = \sqrt{\sum_i \left( \frac{\partial \text{repression}}{\partial x_i} \right)^2 \sigma_i^2}, \quad (10)$$

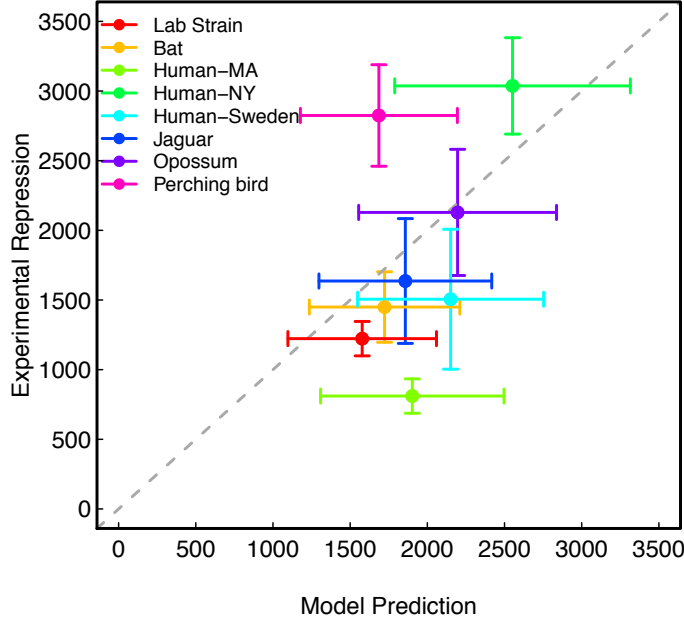
where  $x_i$  represents each of the parameters of the model (binding energies, transcription factors copy number, looping energies, etc.) and  $\sigma_i$  represents the standard deviation of each of these parameters.

Paradoxically, calculating the contribution of each parameter to the uncertainty of the model requires “certainty” about the variability of these parameters. This means that we can only include the uncertainty of the parameters whose uncertainty measurements represents the natural variability in its value and not mostly error due to experimental methods. Table S2 lists the uncertainty of the parameters considered in this analysis given that the *in vivo* error was reported in the listed bibliography.

**Table S2.** Standard deviation of the parameters considered for the calculation of the confidence interval.

Parameter	Deviation	Units	Reference
$R$	Measured for each strain	LacI/cell	-
$\Delta\epsilon_r^{O_1}$	$\pm 0.2$	$k_B T$	[1]
$\Delta\epsilon_r^{O_2}$	$\pm 0.2$	$k_B T$	[1]
$\Delta\epsilon_r^{O_3}$	$\pm 0.1$	$k_B T$	[1]
$\Delta\epsilon_a$	$\pm 1.1$	$k_B T$	[2]

We used a customized *Mathematica* script (Wolfram Research, Champaign, IL) to calculate the partial derivatives. Figure S5 reproduces Figure 7 from the main text, including the predicted standard deviation.



**Figure S5.** Comparison of the model prediction with the experimental measurement. Vertical error bars represent the standard deviation of at least three independent measurements each with three replicates. Horizontal error bars represent the 68% confidence interval of the model calculated by using the *law of error propagation* with the parameter uncertainties listed in Table S2.

## S8. Measuring repression level decouples growth rate effects in translation from effects in transcription

From previous work it was determined that one key regulatory parameter that is influenced by growth rate is the RNAP copy number [22]. However other cellular parameters such as ribosomal copy number and the dilution of mRNA concentration due to growth are also impacted. These parameters will influence protein copy number by influencing the efficiency of mRNA translation. In a very simple dynamical model of transcription, we can imagine that the change in the number of messenger RNA (mRNA) is proportional to the transcription rate and the degradation rate of the mRNA,

$$\frac{dmRNA}{dt} = k_t \cdot p_{bound} - \beta_{mRNA} \cdot mRNA, \quad (11)$$

where  $k_t$  is the maximum transcription rate when the operon is fully induced and  $p_{bound}$  is the probability of finding the RNAP bound to the relevant promoter, as derived using statistical mechanics,  $\beta_{mRNA}$  is the mRNA degradation rate and  $mRNA$  is the number of transcripts of the gene per cell. This equation assumes that the most relevant effect for mRNA depletion is the degradation of the transcripts, compared with the dilution effect due to the growth rate. It is known that this degradation term is not strongly affected by the growth rate [22], so we assume that this term remains constant. In steady

state, when cells are in the exponential growth phase, the concentration of mRNA is

$$mRNA = \frac{k_t \cdot p_{bound}}{\beta_{mRNA}}. \quad (12)$$

The Miller assay (LacZ assay) quantifies the level of LacZ expression, and we assume that the number of proteins is directly proportional to the mRNA copy number. Due to the relatively fast doubling time we assume that dilution is the relevant effect diminishing protein copy number, leading us to

$$\frac{dLacZ}{dt} = \gamma \cdot mRNA - \mu \cdot LacZ, \quad (13)$$

where  $\gamma$  is the proportionality constant of how many proteins per mRNA are produced,  $\mu$  is the growth rate, and  $LacZ$  is the  $\beta$ -galactosidase enzyme copy number.  $\gamma$  can be a function of the growth rate due to the changes in the number of available ribosomes, but still we argue that measuring the repression level should reduce the importance of these effects. If we substitute Equation 12 into 13 and assume steady state we obtain

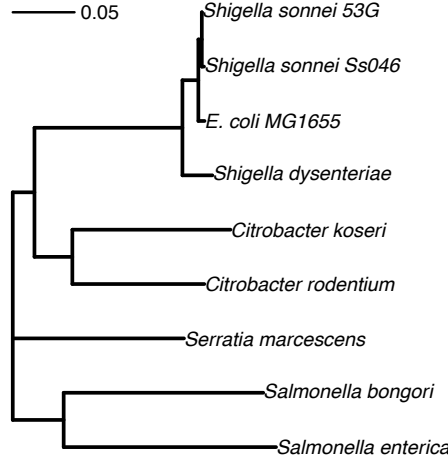
$$LacZ = \frac{\gamma \cdot k_t \cdot p_{bound}}{\mu \cdot \beta_{mRNA}}. \quad (14)$$

By computing the repression level as measured in the LacZ assay we obtain

$$\text{repression} = \frac{LacZ(R=0)}{LacZ(R \neq 0)} = \frac{p_{bound}(R=0, P)}{p_{bound}(R \neq 0, P)}. \quad (15)$$

In this ratio  $\gamma$ ,  $k_t$ ,  $\mu$ , and  $\beta_{mRNA}$  cancel each other leaving only a ratio of  $p_{bound}$ 's.

## S9. Related microbial species *lac* operon phylogenetic tree



**Figure S6.** *lac* operon phylogenetic tree of diverse species with a similar *lac* promoter architecture done with the Neighbor-Joining algorithm. The scale bar represents the relative number of substitutions per sequence.

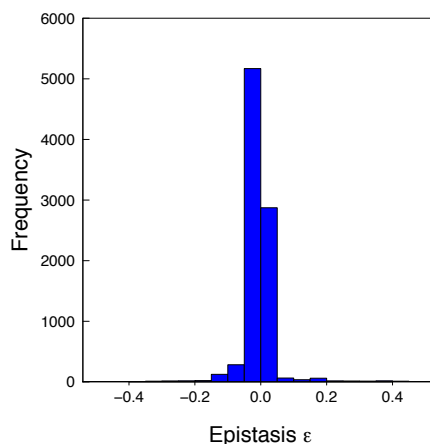
## S10. Epistasis Analysis

Epistasis can be defined as the effect of mutations on the phenotypes caused by other mutations. Our theoretical model explicitly ignores possible interactions between mutations when calculating the transcription factor binding energies with the *sort-seq* energy matrices; but the same cannot be directly assumed for the phenotypic output. As shown in Figure 3 in the main text, the phenotypic response depends on the model parameters in a highly non-linear way. Given this non-linear relation we decided to perform an epistasis analysis on the data, where we defined epistasis as [23, 24]

$$\varepsilon = W_{xy} - W_x \cdot W_y \quad (16)$$

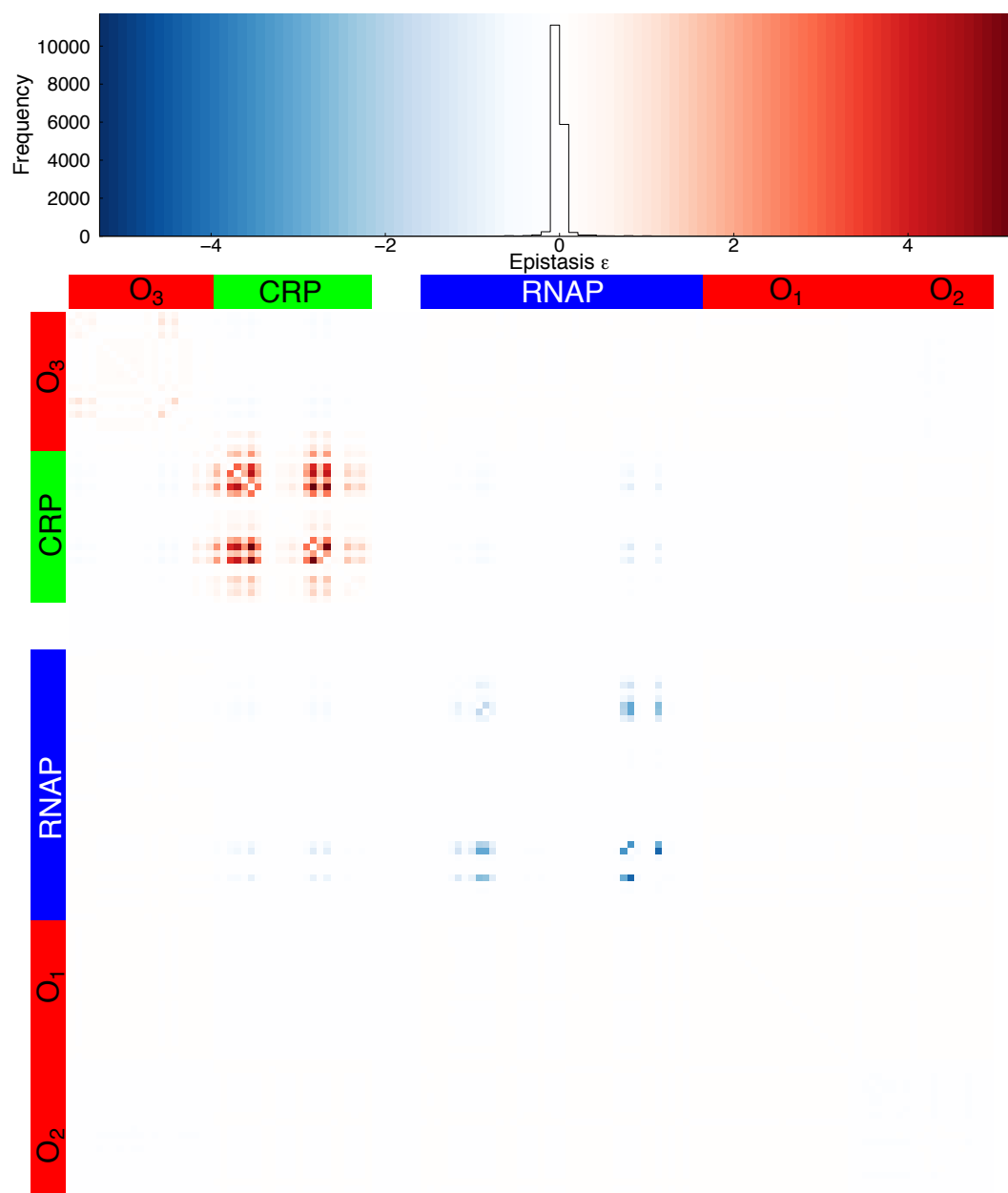
where  $\varepsilon$  is the epistasis,  $W_{xy}$  is the repression value for the double mutant at positions  $x$  and  $y$  normalized to the reference MG1655 repression level, and  $W_x$  and  $W_y$  are the repression values for the single mutants in their respective positions also normalized to the same reference value. This multiplicative epistasis model indicates the type of interaction between mutations;  $\varepsilon = 0$  indicates no epistasis,  $\varepsilon < 0$  indicates antagonistic epistasis and  $\varepsilon > 0$  indicates synergistic epistasis [23].

We calculated this epistasis metric for all the double mutants of the 134 base-pairs considered in the regulatory region of the *lac* operon including the  $O_2$  downstream repressor binding site. For each pair of bases we calculated the epistasis for the two nucleotides with the biggest change with respect to our reference strain MG1655. Figure S7 shows the distribution of the epistasis values for the 8911 possible double mutants. As we initially assumed, most of the base-pairs do not interact with each other. Only 0.5% of the double mutants have an  $\varepsilon < -0.5$ , and 1% have an  $\varepsilon > 0.5$ .



**Figure S7.** Epistasis level (Equation 16) distribution of all the possible double mutants of the *lac* operon regulatory region.

In order to find the base-pairs in the regulatory region predicted to have the biggest interactions Figure S8 shows the heat-map of the  $\epsilon$  values. It is interesting to note that the few regions predicted to have significant epistasis fall mostly within a single binding site, i.e., basically no interaction is predicted between mutations located in different binding sites. The RNAP binding site is predicted to have antagonistic epistasis ( $\epsilon < 0$ ), while the CRP binding site is predicted to have strong synergistic epistasis ( $\epsilon > 0$ ). The  $O_3$  binding site also presents synergistic interactions. This predicted epistasis can be attributed to the highly non-linear dependence of the repression level on these binding energies. Since, for example, the linear regime of the  $O_1$  binding energy extends over a larger range of values (Figure 3 on the main text) two mutations are unable to move this parameter to the non-linear region and no epistasis would be expected at this binding site. Interestingly the only interactions between different binding sites are predicted to be between CRP and RNAP.



**Figure S8.** Epistasis level heat-map for all the possible double mutants. The binding sites positions are indicated with the lateral color bars.

## References

- [1] Garcia H. G and Phillips R. 2011. Quantitative dissection of the simple repression input-output function. *Proceedings of the National Academy of Sciences of the United States of America*, 108(29):12173–8.
- [2] Kuhlman T, Zhang Z, Saier M. H, and Hwa T. 2007. Combinatorial transcriptional control of the lactose operon of *Escherichia coli*. *Proceedings of the National Academy of Sciences of the United States of America*, 104(14):6043–8.
- [3] Bintu L, Buchler N. E, Garcia H. G, Gerland U, Hwa T, Kondev J, and Phillips R. 2005.

- Transcriptional regulation by the numbers: models. *Current opinion in genetics & development*, 15(2):116–24.
- [4] Brewster R. C, Jones D. L, and Phillips R. 2012. Tuning Promoter Strength through RNA Polymerase Binding Site Design in *Escherichia coli*. *PLoS Computational Biology*, 8(12):e1002811.
  - [5] Klumpp S and Hwa T. 2008. Growth-rate-dependent partitioning of RNA polymerases in bacteria. *Proceedings of the National Academy of Sciences of the United States of America*, 105(51):20245–50.
  - [6] Oehler S, Eismann E. R, Krämer H, and Müller-Hill B. 1990. The three operators of the *lac* operon cooperate in repression. *EMBO Journal*, 9(4):973–979.
  - [7] Oehler S, Amouyal M, Kolkhof P, Wilcken-Bergmann B. V, and Müller-Hill B. 1994. Quality and position of the three *lac* operators of *E. coli* define efficiency of repression. *EMBO Journal*, 13(14):3348–3355.
  - [8] Boedicker J. Q, Garcia H. G, and Phillips R. 2013. Theoretical and Experimental Dissection of DNA Loop-Mediated Repression. *Physical Review Letters*, 110(1):018101.
  - [9] Kinney J. B, Murugan A, Callan C. G, and Cox E. C. 2010. Using deep sequencing to characterize the biophysical mechanism of a transcriptional regulatory sequence. *PNAS*, 107(20):9158–63.
  - [10] Ackers G. K, Johnson A. D, and Shea A. M. 1982. Quantitative model for gene regulation by lambda phage repressor. *Proceedings of the National Academy of Sciences of the United States of America*, 79(4):1129–33.
  - [11] Buchler N. E, Gerland U, and Hwa T. 2003. On schemes of combinatorial transcription logic. *Proceedings of the National Academy of Sciences of the United States of America*, 100(9):5136–41.
  - [12] Sherman M. S and Cohen B. A. 2012. Thermodynamic state ensemble models of cis-regulation. *PLoS computational biology*, 8(3):e1002407.
  - [13] Hudson J. M and Fried M. G. 1990. Co-operative interactions between the catabolite gene activator protein and the lac repressor at the lactose promoter. *Journal of molecular biology*, 214(2):381–96.
  - [14] Hudson J. M and Fried M. G. 1996. DNA looping and lac repressor-CAP interaction. *Science*, 82:2–3.
  - [15] Vossen K. M, Stickle D. F, and Fried M. G. 1996. The mechanism of CAP-lac repressor binding cooperativity at the *E. coli* lactose promoter. *Journal of molecular biology*, 255(1):44–54.
  - [16] Schultz S. C, Shields G. C, Steitz T. A, and Steitz T. A. 1991. Crystal Structure of a CAP-DNA Complex: The DNA of Is Bent by 90. *Science (New York, N.Y.)*, 253(5023):1001–1007.
  - [17] Tagami H and Aiba H. 1995. Role of CRP in transcription activation at *Escherichia coli lac* promoter: CRP is dispensable after the formation of open complex. *Nucleic acids research*, 23(4):599–605.
  - [18] Martínez-Antonio A and Collado-Vides J. 2003. Identifying global regulators in transcriptional regulatory networks in bacteria. *Current Opinion in Microbiology*, 6(5):482–489.
  - [19] Epstein W, Rothman-Denes L. B, and Hesse J. 1975. Adenosine 3':5'-cyclic monophosphate as mediator of catabolite repression in *Escherichia coli*. *Proceedings of the National Academy of Sciences of the United States of America*, 72(6):2300–4.
  - [20] Müller J, Oehler S, and Müller-Hill B. 1996. Repression of lac promoter as a function of distance, phase and quality of an auxiliary lac operator. *Journal of molecular biology*, 257(1):21–9.
  - [21] Ku H. H. 1966. Notes on the Use of Propagation of Error Formulas. *Journal of Research of the National Bureau of Standards*, 70C(4):75–79.
  - [22] Klumpp S, Zhang Z, and Hwa T. 2009. Growth rate-dependent global effects on gene expression in bacteria. *Cell*, 139(7):1366–75.
  - [23] Segrè D, Deluna A, Church G. M, and Kishony R. 2005. Modular epistasis in yeast metabolism. *Nature genetics*, 37(1):77–83.
  - [24] Mani R, St Onge R. P, Hartman J. L, Giaever G, and Roth F. P. 2008. Defining genetic interaction.

*Proceedings of the National Academy of Sciences of the United States of America*, 105(9):3461–6.