Detailed map of a cis-regulatory input function

Y. Setty*†, A. E. Mayo*†, M. G. Surette‡, and U. Alon*†§

Departments of *Molecular Cell Biology and †Physics of Complex Systems, The Weizmann Institute of Science, Rehovot 76100, Israel; and ‡Department of Microbiology and Infectious Diseases, University of Calgary, Calgary, AB, Canada T2N 4N1

Most genes are regulated by multiple transcription factors that bind specific sites in DNA regulatory regions. These cis-regulatory regions perform a computation: the rate of transcription is a function of the active concentrations of each of the input transcription factors. Here, we used accurate gene expression measurements from living cell cultures, bearing GFP reporters, to map in detail the input function of the classic *lacZYA* operon of *Escherichia coli*, as a function of about a hundred combinations of its two inducers, cAMP and isopropyl β-D-thiogalactoside (IPTG). We found an unexpectedly intricate function with four plateau levels and four thresholds. This result compares well with a mathematical model of the binding of the regulatory proteins cAMP receptor protein (CRP) and LacI to the *lac* regulatory region. The model is also used to demonstrate that with few mutations, the same region could encode much purer AND-like or even OR-like functions. This possibility means that the wild-type region is selected to perform an elaborate computation in setting the transcription rate. The present approach can be generally used to map the input functions of other genes.

Transcription factor proteins regulate genes by binding to specific sites in their DNA cis-regulatory regions (1–9). Often, multiple transcription factors regulate the same gene. The regulatory regions of a gene must therefore perform a computation: the rate of transcription is a function of the active concentrations of each of the input transcription factors (6–9, 47, 48). This point has been demonstrated, for example, for the *endo-16* gene during sea-urchin development, where multiple transcription factors combine to perform an intricate logical computation (6). Cis-regulatory regions are usually studied by genetic methods, by deleting the various transcription factors or mutating sites in the regulatory region. The picture that emerges from such studies is often stated in terms of logic gates such as AND and OR gates (4, 10). However, the precise computations performed by regulatory regions may be more complex than logic gates. In general they are multidimensional functions of the active transcription factor concentrations. Knowing these functions is essential to understand the dynamics of the transcription networks that control cell responses.

Here, we use accurate gene expression measurements to map in detail the input function of one of the best characterized regulatory regions, that of the *lacZYA* operon (*lac* operon) of *Escherichia coli* (1–3). The *lac* operon encodes for three genes responsible for lactose catabolism and transport. It is transcriptionally regulated by the repressor LacI and the activator cAMP receptor protein (CRP). LacI binds to three operator sites, O1, O2, and O3, located at positions +9, +411, and −84 with respect to the transcriptional start site of the *lacZ* gene (11). The strongest site, O1, overlaps the −10 region of the promoter [the RNA polymerase (RNAP) binding site] (11–13). Repression is enhanced by a DNA loop that forms when LacI tetramers bind two operators, such as O1 and O3 or O1 and O2. When LacI binds artificial inducers such as isopropyl β-D-thiogalactoside (IPTG) or natural inducers such as allolactose, its DNA binding is much reduced and repression is relieved. The activator CRP binds its site only when bound to the inducer cyclic AMP (14). cAMP is a signaling molecule whose level depends on the starvation state of the cell. Glucose intake into the cells suppresses cAMP production and thus inactivates CRP. Therefore, textbook descriptions of this system often employ a logic-gate analogy for *lac* expression: lactose AND NOT glucose.

Beginning with the work of Monod and Jacob (1), there have been many measurements of *lac* expression under various conditions (11–18), usually using substrates for the *lacZ* gene product that have optically detectable readouts (19). The results of different measurements are somewhat varied. It was found that the ratio of fully induced to fully repressed expression is in the range of 300–1,300, and that growth on glucose can repress expression by 3- to 10-fold. A more detailed mapping at four cAMP and four IPTG concentrations was performed by Zubay *et al.* (14) on semipermeabilized cells. The drawback of these methods is that cells need to be lysed or permeabilized to allow the substrate to interact with the LacZ enzyme. This requirement limits the accuracy and amount of data that can be collected. Here, we used green fluorescent protein (GFP) as a reporter to obtain measurements of *lac* promoter activity from populations of living cells, without need of lysis or treatment (20–22). We find that the *lac* cis-regulatory region performs a rather intricate computational function.

Methods

Plasmids and Strains. The *lac* cis-regulatory region was amplified from MG1655 genomic DNA by using PCR at start coordinate 365438 and end coordinate 365669 from the sequenced genome (23). It was subcloned into *Xho*I and *Bam*HI sites upstream of a promoterless gfpmut2 (24) gene in a low-copy pSC101 origin plasmid (20, 21) and was used to transform *E. coli* K12 strain MG1655 (wild type for the *lac* system).

Culture and Measurements. Cultures (1 ml) inoculated from single colonies were grown for 16 h in M9C defined medium [M9, 2 mg/ml glucose, 1 mM MgSO₄, 0.1 mM MgCl₂, 25 μg/ml kanamycin] at 37°C with shaking at 250 rpm. The cultures were diluted to OD₆₀₀ = 0.003 into M9C with different concentrations of cAMP (0–20 mM, Sigma) and IPTG (0–200 μM), at a final volume of 150 μl per well in a flat-bottom 96-well plate (Sarstedt). The cultures were covered with 100 μl of mineral oil (Sigma) to prevent evaporation, and grown in a Wallac Victor2 multiwell fluorimeter at 37°C, set with an automatically repeating protocol of shaking and OD₆₀₀ and fluorescence readings (20–22). Time between repeated measurements was 6 min. Background fluorescence at a given OD was determined from the fluorescence of cells bearing a promoterless GFP vector at the same OD (20–22). Cells growing on glucose with saturating external cAMP and cells growing on glycerol (high endogenous cAMP) without exogenous cAMP show similar *lac* promoter activity and growth rates.

Colorimetric Assay of β-Galactosidase Activity. MG1655 bearing a promoterless vector was grown in M9C for 16 h and diluted (1:200, yielding initial OD₆₀₀ = 0.006) into 1 ml of M9C medium with various concentrations of cAMP (0–20 mM) and IPTG (0–200 μM) in a 2-ml 96-well plate. The cultures were grown at 37°C in a shaker (shaking at 250 rpm), until OD₆₀₀ = 0.12. Each sample was lysed (1 ml of cells, 20 μl of 0.1% SDS, 40 μl of
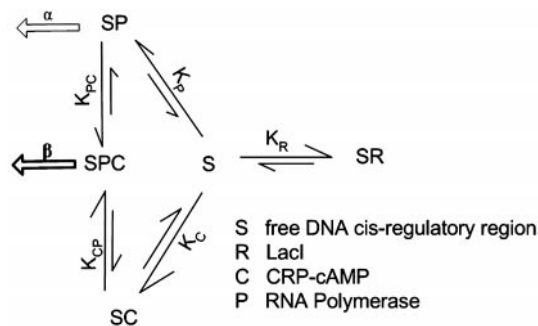
Fig. 1. Simple mathematical model of the *lac* region, protein–DNA level reactions. Transcription occurs from states where RNAP is bound (large open arrows). The activity of the promoter in these states is α and β. An additional parameter γ describes transcription from other states (leakiness). The K parameters are the equilibrium dissociation constants of the various reactions.

chloroform). Twenty microliters of each sample that was grown at high concentration of IPTG (>25 μM) was assayed in a 96-well plate with 80 μl of Z-buffer (19) to a final volume of 100 μl per well. Eighty microliters of each sample that was grown at low concentration of IPTG (<25 μM) was mixed with 20 μl of Z-buffer to a final volume of 100 μl per well. Twenty microliters of *o*-nitrophenyl β-D-galactoside (ONPG; 4 mg/ml) was rapidly added to the wells by using a multichannel pipetter. Absorbance was measured over time in Wallac Victor2 multiwell fluorimeter at 30°C, set with an automatic repeating protocol of shaking (1 mm orbital, normal speed, 0.5 sec shaking) and absorbance measurements. Time between repeated measurements of A_{405} (absorbance at 405 nm) and A_{490} was 2 min. Background reading of A_{405} and A_{490} for Z-buffer with ONPG was subtracted. The product content was calculated as $(A_{405} - r \cdot A_{490})/\text{OD}_{600} \cdot v$, where OD_{600} is the cell density before the assay and v is the volume of culture used in the assay (ml). The factor $r = 0.8$ is the ratio of cell absorbance at 405 nm to that at 490 nm. We find that the product reaches a maximum and then declines. The β-galactosidase activity was determined by the initial slope of product formation.

Simple Model of *lac* Regulatory Region. We analyze a simple model for the activity of the *lac* promoter as function of the inducers, cAMP and IPTG. RNAP, CRP, and LacI are designated by P, C, and R, respectively. A free *lac* regulatory region is designated by S. S represents effective occupation of the *lac* promoter region by RNAP, CRP, and LacI. S concatenated with another letter indicates a site occupied by a protein. For example, SC indicates binding of CRP, SPC indicates binding of both RNAP and CRP, etc. Assumptions: (*i*) Inducer concentrations are constant during growth (25). (*ii*) Regulatory protein concentration are constant across growth conditions. (*iii*) The effect of binding to S on the free concentration of the proteins is negligible. (*iv*) All reactions are in equilibrium. CRP binds cAMP and becomes active according to $\mathcal{A} = [C]/[C_T] = X^n/(1 + X^n)$, in which $X = [\text{cAMP}]/K_{\text{cAMP}}$ and the subscript T indicates total. X designates cAMP concentration in units of its dissociation constant for cAMP. When $X = 1$, half of the total CRP is bound with cAMP. Cooperativity in cAMP binding to a CRP dimer is described by the Hill coefficient n (26). Similarly, for LacI–IPTG binding, $\mathcal{R} = [R]/[R_T] = 1/(1 + Y^m)$, $Y = [\text{IPTG}]/K_{\text{IPTG}}$, where m is the Hill coefficient describing the cooperativity of inducer binding to the LacI tetramer (27). Fig. 1 describes the reactions at the protein-DNA level. Note that R binding excludes binding of C or P. Transcription occurs at rate α from the SP state, β from SPC state, and γ from other states (leakiness). The input function (mRNA production rate) is $f = \alpha[SP] + \beta[SPC] +$

$\gamma([S] + [SC] + [SR])$. Solving the equilibrium reactions, using detailed balance conditions, yields

$$f = \frac{\alpha a + 2\beta b d \mathcal{A} + \gamma(c\mathcal{R} + d\mathcal{A} + 1)}{1 + a + (2b + 1)d\mathcal{A} + c\mathcal{R}},$$

where the dimensionless parameters are as follows:

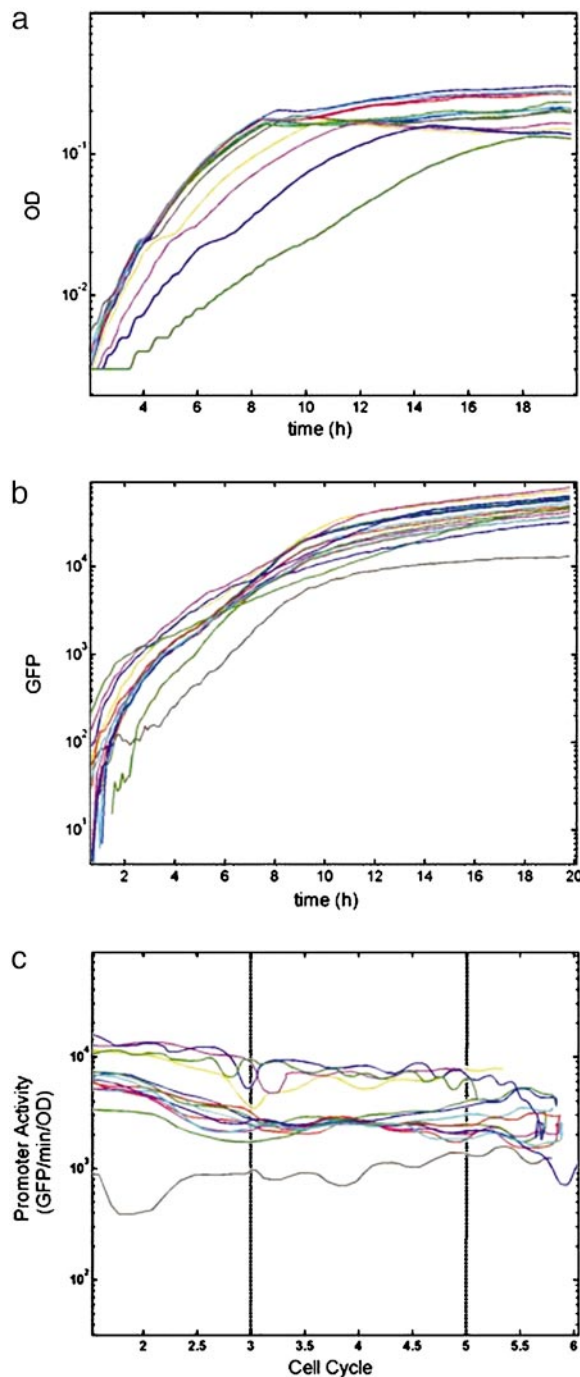$a = [P]/K_P$, RNAP in units of dissociation constant for binding to a free site.

Fig. 2. Growth and GFP fluorescence of the *lac*-reporter strain for 20 different IPTG and cAMP concentrations. (*a*) OD at 600 nm. (*b*) GFP fluorescence (535 nm). (*c*) Promoter activity, $d[\text{GFP}]/dt/\text{OD}_{600}$ (proportional to mRNA production rate). Cell cycles = $\log_2(\text{OD}/\text{OD}_0)$.

GENETICS

$b = [P]/K_{CP}$, RNAP in units of dissociation constant for binding to its site with bound CRP–cAMP.

$c = [R_T]/K_R$, LacI in units of dissociation constants for binding to its site.

$d = [C_T]/K_C$, CRP in units of dissociation constants for binding to its site.

The promoter activity can also be expressed as:

$$f = V_1 \frac{1 + V_2 \mathcal{A} + V_3 \mathcal{R}}{1 + V_4 \mathcal{A} + V_5 \mathcal{R}},$$

where $V_1, \ldots, V_5$ are combinations of the biochemical parameters: $V_1 = (\alpha a + \gamma)/(1 + a)$, $V_2 = d(2\beta b + \gamma)/(\alpha a + \gamma)$, $V_3 = \gamma c/(\alpha a + \gamma)$, $V_4 = d(2b + 1)/(1 + a)$, and $V_5 = c/(1 + a)$.

The present model is not meant to be a detailed account of the full biochemistry of the *lac* system, rather a toy model that captures the essential behavior. For example, multioperator DNA loops in the system (13) may introduce additional effects.

Best-Fit Parameters. We used nonlinear minimization (Matlab 6.1) to fit the model to the data. Diverse initial conditions yielded the same best fit result. The rms relative fit error is about 15%. Some of the best-fit parameters can be compared with biochemical measurements. K_{IPTG} is close to the measured dissociation constant of IPTG ($K_{IPTG} = 1.3 \ \mu M$) (28). The *in vitro* measured affinity of CRP to cAMP is in the micromolar range (26). Thus our results of $K_{cAMP} \approx 2$ mM suggest that at millimolar external cAMP levels, the internal cAMP level is in the micromolar range, which may be accounted for by the action of cAMP degradation and export systems. The parameter c, LacI in units of the dissociation constant for its site, is large ($c \approx 500$). This result qualitatively agrees with

biochemical measurements where $K_R \approx 0.1$–1 pM (29) and the concentration of LacI not nonspecifically bound to DNA is thought to be $R_T \approx 1$ nM (3), giving $c \approx 10^3$ to 10^4. Some biochemical parameters appear only in combinations (αa, $\beta b d$). Therefore the present measurements cannot distinguish between the two roles that CRP–cAMP may have in transcription activation: RNAP stabilization ($b \gg a$) and enhancement of transcription initiation ($\beta \gg \alpha$). Interactions between CRP and LacI (30) can be added to the model, but they do not significantly improve the fit quality.

Model Parameters for AND-Like and OR-Like Input Functions. AND-like and OR-like gates were generated in the model with a few changes in the best-fit parameters. The AND-like gate was generated by increasing c (LacI binding) and decreasing αa (RNAP binding). This adjustment leads to changes in V_1, V_3, and V_5. To generate an OR-like gate required increasing αa (RNAP binding) and d (CRP–cAMP binding) and decreasing c (LacI binding). This adjustment leads to changes in V_1 through V_5.

Results

GFP Reporter Measurement of the *lac* Input Function. We measured *lac* promoter activity (rate of mRNA production) by using a low-copy reporter plasmid in which the *lac* regulatory region controls GFP. The plasmid includes a 232-bp regulatory region spanning the entire region between the *lacI* gene and the *lac* operon, including 18 bases into the *lacI* coding region and 130 bases into the *lacZ* coding region. The reporter strain, MG1655, which is wild type for the *lac* system, was grown in minimal medium under all combinations of 8 different IPTG concentrations and 12 different cAMP concentrations in a 96-well plate. cAMP levels ranged from 0 to 10 mM, and IPTG levels ranged from 0 to 200 μM. The
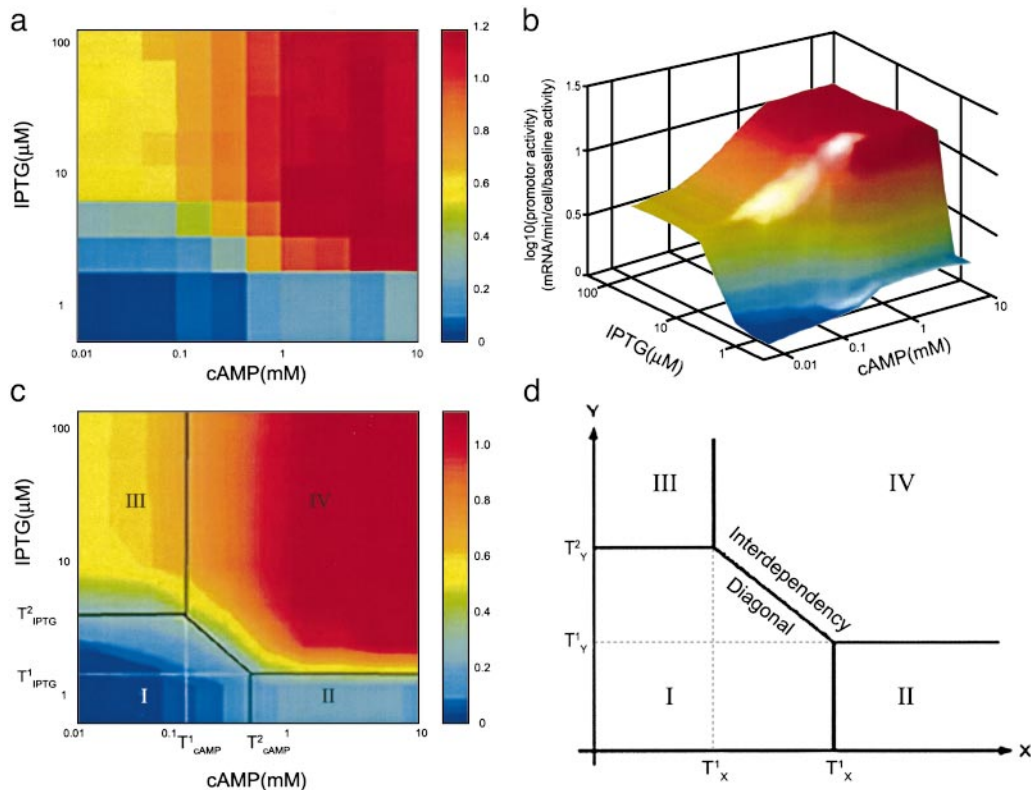

Fig. 3. *lac* cis-regulatory input function (promoter activity) as function of cAMP and IPTG concentrations measured by using the GFP-plasmid system. The promoter activity was taken as the mean over two cell cycles in midexponential phase; to help visualize the lowest plateau there are two rows that represent data at low IPTG. Note the logarithmic scales. (*a*) Promoter activity. (*b*) Smoothed surface plot of the promoter activity. (*c*) Smoothed contour plot of the promoter activity, with the four thresholds and the interdependency diagonal. (*d*) Schematic form of the input function of the *lac* promoter. The four plateaus are designated as follows: I, low cAMP, low IPTG; II, high cAMP, low IPTG; III, low cAMP, high IPTG; IV, high cAMP, high IPTG.

growth medium contained saturating concentrations of glucose as a sole carbon source, to minimize endogenous cAMP levels. For each well, the cell density (OD_{600}) and GFP fluorescence were monitored at a 6-min resolution over growth at 37°C with shaking within an automated multiwell fluorimeter (Fig. 2). The rate of GFP production, divided by the OD, provided a measure of the promoter activity (21, 31). High cAMP levels reduced the growth rate of the cells (Fig. 2a). The promoter activity measurement used OD to normalize the differences in growth. At all conditions, the promoter activity achieved an approximately constant value during about two cell cycles in midexponential growth (Fig. 2c). We computed the promoter activity in each of the 96 growth conditions, by an average of the promoter activity over these two cell cycles.

The *lac* Input Function Has Four Plateau Levels and Four Thresholds. The promoter activity shows four plateau levels (Fig. 3). Plateau IV (relative height 1.0) is at high IPTG and cAMP. Plateau III (relative height 0.29 ± 0.02) is at high IPTG and low cAMP. Plateau II (relative height 0.12 ± 0.008) is at high cAMP and low IPTG. Plateau I, the lowest plateau (relative height 0.07 ± 0.006), is at low IPTG and cAMP.

The *lac* input function appears to have two thresholds for IPTG, one at low cAMP and the other at high cAMP. It also has two thresholds for cAMP, one at low IPTG and the other at high IPTG. At high cAMP levels, the threshold for IPTG is $T_{IPTG}^1 \approx 2 \pm 1 \ \mu M$. At low cAMP levels, the thresholds is about $T_{IPTG}^2 \approx 7 \pm 1 \ \mu M$. Similarly, the cAMP thresholds at high and low IPTG are $T_{cAMP}^1 \approx 0.2 \pm 0.1$ mM and $T_{cAMP}^2 \approx 0.8 \pm 0.2$ mM external cAMP, respectively. There is a diagonal threshold in which the *lac* promoter is induced at intermediate concentrations of cAMP and IPTG, where the higher the concentration of one inducer is, the easier it is for the other inducer to activate expression (interdependency diagonal in Fig. 3 c and d).

The *lac* Chromosomal Input Function Measured with ONPG Assay Is Similar to the GFP Plasmid Results but with Deeper Plateaus. We repeated the experiment with a colorimetric assay for β-galactosidase (the *lacZ* gene product) activity, using the ONPG substrate, from the chromosomal *lac* operon. ONPG is cleaved by β-galactosidase, yielding a yellow product. Cells were grown on 96 IPTG and cAMP combinations in 96-well plates and lysed at a fixed OD corresponding to the middle of the OD region where promoter activity was determined by using the GFP reporters. ONPG was added, and the rate of yellow product appearance was measured by using absorbance at 405 nm at high temporal resolution in a multiwell fluorimeter. The initial rate of product formation allowed a fairly accurate measure of the relative β-glycosidase enzyme concentration at the time of lysis. Throughout, one assumes that because lysis was performed at approximately balanced growth, the enzyme concentration is proportional to the promoter activity at the time of lysis (25, 31). We find that the chromosomal *lac* input function measured with ONPG is qualitatively similar to the results with the GFP reporter plasmid (Fig. 4a). There are four plateau levels, ranked from lowest to highest at low IPTG and cAMP, low IPTG and high cAMP, low cAMP and high IPTG, and high IPTG and cAMP. Furthermore, there appear to be two thresholds for IPTG and two for cAMP at concentrations that were similar to those obtained with the GFP plasmid. However, the height ratios between some of the plateaus were different in the two methods: whereas the ratios of the plateau heights in the GFP reporter assay were about 1:2:4:14, in the ONPG assay we found 1:2:40:300.

We repeated the measurements with lactose instead of IPTG (Fig. 4b). The threshold for activation by lactose in the absence of cAMP is much higher then the threshold for IPTG. This difference is probably due to inducer exclusion, in which the lactose permease is blocked by components of the PTS (phosphotransferase system)
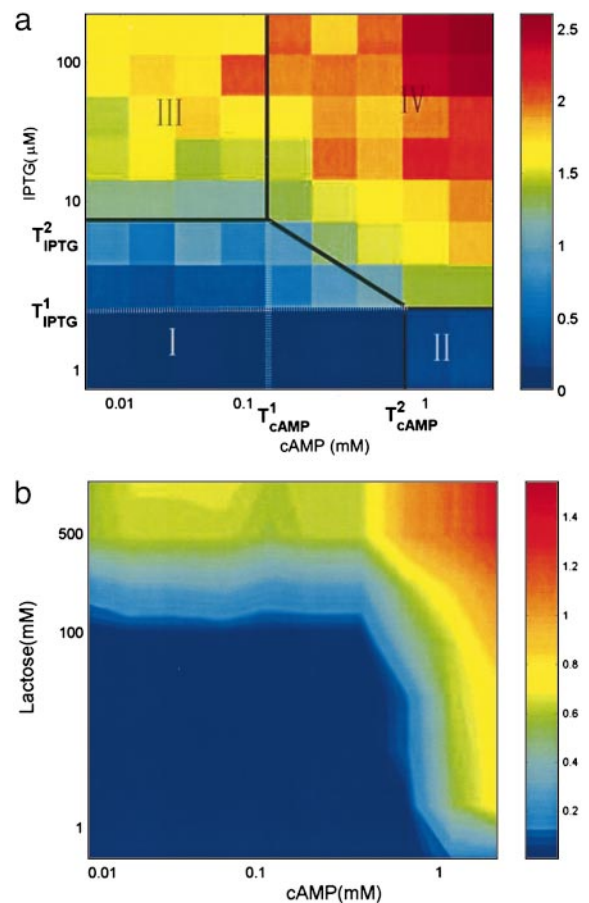

Fig. 4. (a) Assay of chromosomally encoded β-galactosidase activity as a function of cAMP and IPTG concentrations by the ONPG colorimetric method. To help visualize the lowest plateau there are two rows that represent data at low IPTG. (b) Smoothed promoter activity as a function of lactose and cAMP concentration measured by using the GFP-plasmid system. Promoter activity was taken as the mean over two cell cycles in midexponential phase.

(32). IPTG provides control of LacI activity, which is less sensitive to these effects.

Simple Mathematical Model Suggests That Purer AND-Like and OR-Like Functions Can Be Reached with Few Mutations. We constructed a simple model of the *lac* promoter control region including binding of LacI and CRP (Fig. 1) (29, 33–35, 48–50). The model has nine biochemical parameters, including affinities of CRP and LacI to their DNA sites and their inducers, affinity of RNAP to the promoter in the presence and absence of bound CRP–cAMP, and the mRNA production rates in various configurations (Table 1). We obtained best-fit parameters that reproduce the experimental data at a mean relative error of about 15% (Table 1, Fig. 5a). We note that the number of free parameters in the model reasonably matches the content of the data. That is, there is no over-fitting where there are many more free parameters than data. The four plateaus and four thresholds effectively define eight of the parameters; the remaining parameter is determined by the slopes of the input function.

We then attempted to obtain promoter functions that are closer to pure AND and OR gates, by varying as few parameters as possible in the model. A pure AND gate has three equally low plateaus and a distinctively high one. It is also characterized by only one threshold for each inducer. Input functions that are close to this can be obtained in several ways in the model. Perhaps the easiest way is to strengthen LacI binding and to weaken RNAP binding

GENETICS

Table 1. *lac* model parameters that best fit the measurement using the GFP reporter plasmid (wild type) and putative mutants that have purer AND-like and OR-like gates

Parameter	Wild type	AND	OR
m	4 ± 0.6	4	4
n	2 ± 0.4	2	2
K_{IPTG}, μM	1.2 ± 0.2	1.2	1.2
K_{cAMP}, mM	1.8 ± 0.5	1.8	1.8
V_1	3.5 ± 0.7	**1**	**10**
V_2	70 ± 10	70	**1,700**
V_3	170 ± 30	**2,000**	**15**
V_4	17 ± 3	17	**400**
V_5	540 ± 100	**7,000**	**50**

Errors are parameter variations that give 15% deviation from the best-fit results. The values of the parameters that were changed to produce AND- and OR-like gates are in boldface. V_1, V_2, V_3, V_4, and V_5 represent combination of the biochemical parameters a, b, c, d, α, β, and γ (see *Methods*).

(Table 1 and Fig. 5*b*). Stronger LacI binding abolishes plateau II. Weaker RNAP binding increases the dependence on CRP and reduces plateau III. A pure OR gate has three equally high plateaus and one low plateau. Like the AND gate, it is characterized by only two thresholds. An OR-like gate can be obtained in the model by weakening LacI binding and strengthening CRP binding and the RNAP promoter site (Table 1 and Fig. 5*c*). These changes can in principle be obtained with only a small number of mutations in the CRP, LacI, and RNAP binding sites.

Discussion

The *lac* Input Function Is More Elaborate than a Simple AND Gate. We find a rather intricate input function with four plateaus and four thresholds in the *lac* system. Each of the inducers has a low threshold when the other inducer is present, and an $\approx$4-fold higher threshold when the other inducer is absent (Fig. 3). A pure AND gate, in contrast, has only one threshold per inducer and two plateaus level (Fig. 5*b*). Our mathematical analysis suggests that
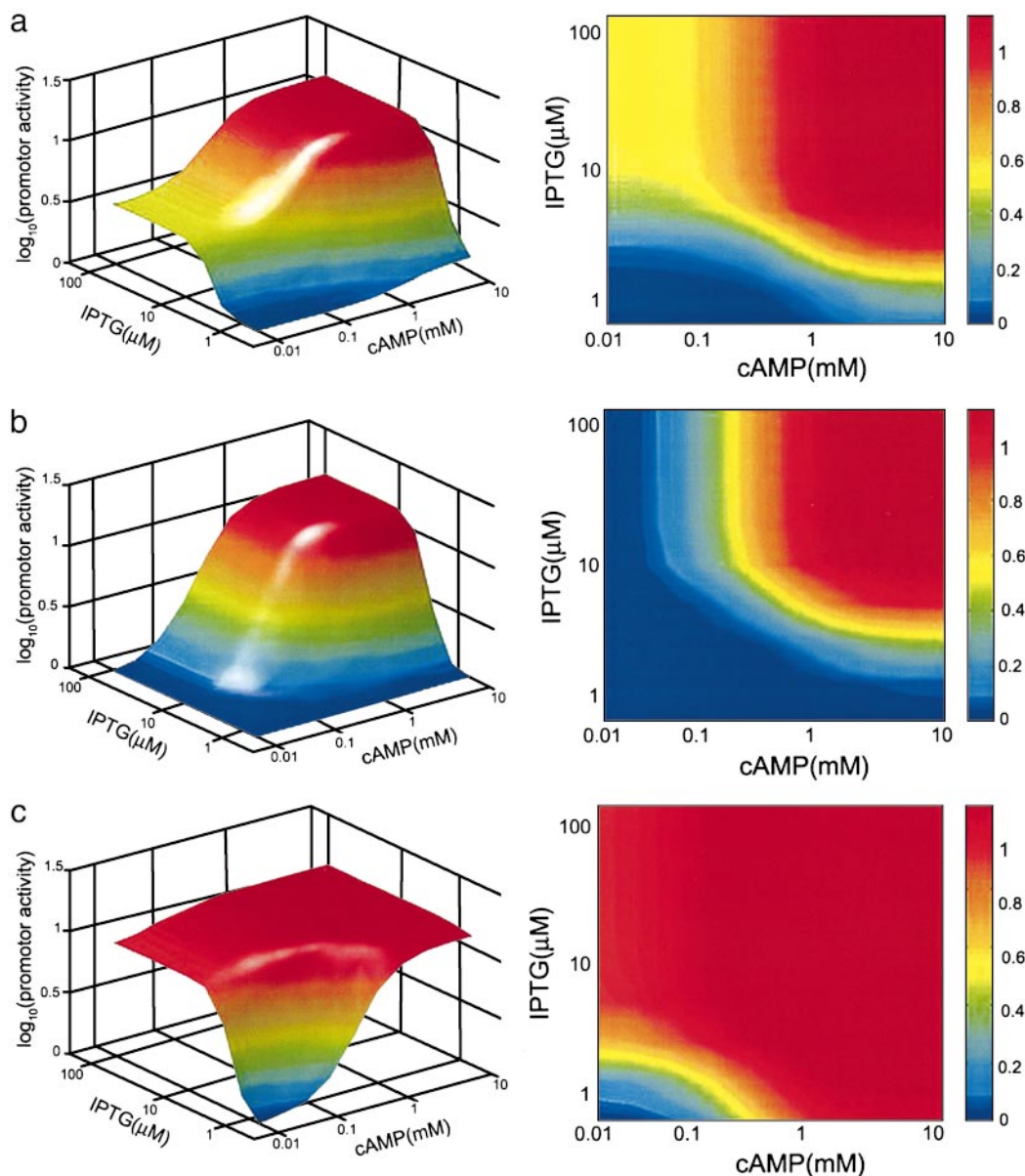

Fig. 5. Input function (promoter activity) in the mathematical model of the *lac* promoter. (*a*) Input function with best-fit parameters to the GFP measurements. (*b*) Input function with parameters for AND-like function. (*c*) Input function with parameters for OR-like function.

more AND-like functions could be reached from the *lac* regulatory region with just a few mutations. This raises the possibility that the intricate function observed in the present study may have been selected over simpler AND-gate-like functions. What might be the function of the four-plateau, four-threshold input function in the *lac* system?

Plateau II represents situations where cells are starved for glucose but no lactose or other inducer of the *lac* system is present. In this regime, cells may benefit from a low level of *lac* system induction: cells that contain some LacY permease proteins will be able to become fully induced upon addition of lactose even at low levels (15). The weak induction of plateau II may be stochastic (15, 36–38), in which only a minority expresses some permease and most cells express none (39).

The rather high plateau III (Fig. 3) represents the effect of induction in the presence of glucose. Lactose is excluded from entering the cells through LacY in the presence of glucose, because the lactose pumps are blocked by components of the glucose phosphotransferase system, an effect known as catabolite exclusion. LacI, however, can be induced by putative naturally occurring nonmetabolizable galactosides that enter the cells by non-LacY routes (40–42). The cell may require the *lac* operon, in particular the galactoside transacetylase LacA, to detoxify these potentially harmful sugar analogues (40–42). In addition, plateau III is useful in the rare cases where in addition to glucose there is a very high concentration of external lactose, which enters the cell by nonspecific transport. We find that very high levels of external lactose (about 100 mM, Fig. 4*b*) can induce the *lac* system even in the presence of glucose. In this case, the cells may benefit from making LacZ enzymes to anticipate the feast ready for them when glucose runs out.

There is an interdependency region at intermediate concentrations where the activation by cAMP is aided by IPTG and vice versa (Fig. 3*d*). The threshold for IPTG activation is effectively lowered by cAMP. Mechanistically this can occur because CRP–cAMP binding strengthens the RNAP binding to the promoter, decreasing the probability of LacI binding to the O1 site, which overlaps the promoter. Therefore, in the presence of cAMP, a lower level of IPTG is required to achieve the same level of promoter activity as in the absence of cAMP. Another possible mechanism is that CRP–cAMP binding disrupts the DNA loop formed by LacI, thus making IPTG more effective in relieving repression. On the interdependency diagonal, the *lac* regulatory region appears to compute curves of equivalent activity for defined combinations of glucose starvation and availability of inducer.

Comparison of GFP Reporter Plasmid and Chromosomal Enzyme Production Assays. The *lac* system has classically been studied by using colorimetric assays for the *lacZ* gene product (12–19). The present GFP reporter plasmid measurement is different in several ways from assays of enzyme activity from the chromosomally encoded operon: (*i*) the low-copy plasmid (pSC101 origin) introduces several extra copies of the promoter region, thus potentially titrating out LacI (43); (*ii*) the promoter region on the plasmid lacks the O2 binding site (at +411), a site whose absence makes shutoff of the promoter about 5-fold weaker (12); and (*iii*) the plasmid DNA may be harder to loop, reducing repression strength. We therefore repeated the experiment with the classical colorimetric assay for *β*-galactosidase, using accurate time-resolved ONPG absorbance measurements (19). The input functions found by the two methods were qualitatively similar with four plateaus and four threshold levels. However, some of the plateaus were deeper in the ONPG assay. This difference presumably reflects the above-mentioned plasmid effects. The four plateau levels are in the range of those previously measured by using ONPG (12–18).

The GFP reporter method has the advantage of not requiring cell lysis and of yielding highly accurate kinetic readouts. It could be improved by integrating the GFP reporter into the chromosome. In the present study the input function was measured on the average of cell populations. It would be important to measure the input function also at the single-cell level.

Using the present approach, one may attempt to map the input functions of operons in bacteria on a genomic scale by using reporter libraries covering most of the organism's genes (44, 45). Similar approaches should apply to eukaryotic cis-regulatory input functions (46). It would be important to see what classes of input functions are found in different gene systems. In particular, it would be fascinating to discover whether there is a limited dictionary of input-function types that recur throughout gene regulation networks, and whether each gene possesses an exquisitely crafted input function that requires several parameters to be fully described.

1. Jacob, F. & Monod, J. (1961) *J. Mol. Biol.* **3**, 318–356.
2. Beckwith, J. & Zipser, D., eds. (1970) *The Lactose Operon* (Cold Spring Harbor Lab. Press, Plainview, NY).
3. Muller-Hill, B. (1996) *The lac Operon* (de Gruyter, Berlin).
4. Savageau, M. A. (1976) *Biochemical Systems Analysis: A Study of Function and Design in Molecular Biology* (Addison–Wesley, Reading, MA).
5. Hartwell, L. H., Hopfield, J. J., Leibler, S. & Murray, A. W. (1999) *Nature* **402**, C47–C52.
6. Yuh, C. H., Bolouri, H. & Davidson, E. H. (1998) *Science* **279**, 1896–1902.
7. Pilpel, Y., Sudarsanam, P. & Church, G. M. (2001) *Nat. Genet.* **29**, 153–159.
8. Thieffry, D., Huerta, A. M., Perez-Rueda, E. & Collado-Vides, J. (1998) *BioEssays* **20**, 433–440.
9. Shen-Orr, S. S., Milo, R., Mangan, S. & Alon, U. (2002) *Nat. Genet.* **31**, 64–68.
10. Thomas, R. (1973) *J. Theor. Biol.* **42**, 563–585.
11. Oehler, S., Amouyal, M., Kolkhof, P., von Wilcken-Bergmann, B. & Muller-Hill, B. (1994) *EMBO J.* **13**, 3348–3355.
12. Eismann, E., von Wilcken-Bergmann, B. & Muller-Hill, B. (1987) *J. Mol. Biol.* **195**, 949–952.
13. Eismann, E. R. & Muller-Hill, B. (1990) *J. Mol. Biol.* **213**, 763–775.
14. Zubay, G., Schwartz, D. & Beckwith, J. (1970) *Proc. Natl. Acad. Sci. USA* **66**, 104–110.
15. Novick, A. & Weiner, M. (1957) *Proc. Natl. Acad. Sci. USA* **43**, 553–566.
16. Pastan, I. & Perlman, R. (1970) *Science* **169**, 339–344.
17. Wei, Y., Lee, J.-M., Richmond, C., Blattner, F. R., Rafalski, J. A. & LaRossa, R. A. (2001) *J. Bacteriol.* **183**, 545–556.
18. Hopkins, J. D. (1974) *J. Mol. Biol.* **87**, 715–724.
19. Miller, J. H. (1972) *Experiments in Molecular Genetics* (Cold Spring Harbor Lab. Press, Plainview, NY).
20. Kalir, S., McClure, J., Pabbaraju, K., Southward, C., Ronen, M., Leibler, S., Surette, M. G. & Alon, U. (2001) *Science* **292**, 2080–2083.
21. Ronen, M., Rosenberg, R., Shraiman, B. I. & Alon, U. (2002) *Proc. Natl. Acad. Sci. USA* **99**, 10555–10560.
22. Rosenfeld, N., Elowitz, M. B. & Alon, U. (2002) *J. Mol. Biol.* **323**, 785–793.
23. Blattner, F. R., Plunkett, G., III, Bloch, C. A., Perna, N. T., Burland, V., Riley, M., Collado-Vides, J., Glasner, J. D., Rode, C. K., Mayhew, G. F., *et al.* (1997) *Science* **277**, 1453–1474.
24. Cormack, B. P., Valdivia, R. H. & Falkow, S. (1996) *Gene* **173**, 33–38.
25. Yagil, G. & Yagil, E. (1971) *Biophys. J.* **11**, 11–27.
26. Herman, J. G. (2001) *Biochim. Biophys. Acta* **1547**, 1–17.
27. Pace, H. C., Kercher, M. A., Lu, P., Markiewicz, P., Miller, J. H., Chang, G. & Lewis, M. (1997) *Trends Biochem. Sci.* **22**, 334–339.
28. Gilbert, W. & Muller-Hill, B. (1970) in *The Lactose Operon*, eds. Beckwith, J. & Zipser, D. (Cold Spring Harbor Lab. Press, Plainview, NY), pp. 93–110.
29. Wong, P., Gladney, S. & Keasling, J. D. (1997) *Biotechnol. Progr.* **13**, 132–143.
30. Perros, M., Steitz, T. A., Fried, M. G., Hudson, J. M. & Lewis, M. (1996) *Science* **274**, 1929–1932.
31. Maaløe, O. & Kjeldgaard, N. (1966) *Control of Macromolecular Synthesis* (Benjamin, New York).
32. Kimata, K., Takahashi, H., Inada, T., Postma, P. & Aiba, H. (1997) *Proc. Natl. Acad. Sci. USA* **94**, 12914–12919.
33. Ackers, G. K., Johnson, A. D. & Shea, M. A. (1982) *Proc. Natl. Acad. Sci. USA* **79**, 1129–1133.
34. Manabe, T. (1981) *J. Theor Biol.* **89**, 271–302.
35. Kremling, A., Bettenbrock, K., Laube, B., Jahreis, K., Lengeler, J. W. & Gilles, E. D. (2001) *Metab. Eng.* **3**, 362–379.
36. Rao, C. V., Wolf, D. M. & Arkin, A. P. (2002) *Nature* **420**, 231–237.
37. McAdams, H. H. & Arkin, A. (1999) *Trends Genet.* **15**, 65–69.
38. Elowitz, M. B., Levine, A. J., Siggia, E. D. & Swain, P. S. (2002) *Science* **297**, 1183–1186.
39. Siegele, D. A. & Hu, J. C. (1997) *Proc. Natl. Acad. Sci. USA* **94**, 8168–8172.
40. Zabin, I., Kepes, A. & Monod, J. (1962) *J. Biol. Chem.* **237**, 253–257.
41. Musso, R. & Zabin, I. (1973) *Biochemistry* **12**, 553–557.
42. Wilson, T. H. & Kashket, E. R. (1969) *Biochim. Biophys. Acta* **173**, 501–508.
43. Lee, S. B. & Bailey, J. E. (1984) *Biotechnol. Bioeng.* **26**, 66–73.
44. Goh, E. B., Yim, G., Tsui, W., McClure, J., Surette, M. G. & Davies, J. (2002) *Proc. Natl. Acad. Sci. USA* **99**, 17025–17030.
45. Van Dyk, T. K., Ayers, B. L., Morgan, R. W. & LaRossa, R. A. (1998) *J. Bacteriol.* **180**, 785–792.
46. Dimster-Denk, D., Rine, J., Phillips, J., Scherer, S., Cundiff, P., DeBord, K., Gilliland, D., Hickman, S., Jarvis, A., Tong, L. & Ashby, M. (1999) *J. Lipid Res.* **40**, 850–860.
47. Bolouri, H. & Davidson, E. H. (2002) *BioEssays* **24**, 1118–1129.
48. Buchler, N. E., Gerland, U. & Hwa, T. (2003) *Proc. Natl. Acad. Sci. USA* **100**, 5136–5141.
49. Kremling, A. & Gilles, E. D. (2001) *Metab. Eng.* **3**, 138–150.
50. Kremling, A., Jahreis, K., Lengeler, J. W. & Gilles, E. D. (2000) *Metab. Eng.* **2**, 190–200.

GENETICS