Published in final edited form as:
*J Mol Biol*. 2009 August 28; 391(4): 671–678. doi:10.1016/j.jmb.2009.06.043.

# Relation of intracellular signal levels and promoter activities in the *gal* regulon of *Escherichia coli*

**Sandeep Krishna**[1], **László Orosz**[2], **Kim Sneppen**[1], **Sankar Adhya**[3], and **Szabolcs Semsey**[2,*]

[1]Center for Models of Life, Niels Bohr Institute, Copenhagen, Denmark

[2]Department of Genetics, Eötvös Lóránd University, H-1117 Budapest, Hungary

[3]Laboratory of Molecular Biology, Center for Cancer Research, National Cancer Institute, National Institutes of Health, Bethesda, MD 20892−4264

## Abstract

Transcription of many genes is regulated by combinations of multiple signals. In *Escherichia coli* combinatorial control is typical in the case of operons related to utilization of different sugars in the absence of glucose. To understand regulation of the transport and metabolic pathways in the galactose system, we measured activities of the six *gal* regulon promoters simultaneously, using an *in vitro* transcription system containing purified components. Input functions were computed based on the experimental measurements. We observed four different shapes of input functions. From the results we can conclude that the structure of the regulatory network is insufficient for the determination of signal integration. It is the actual structure of the promoter and regulatory region, the mechanism of transcription regulation, and the interplay between transcription factors that shape the input function to be suitable for adaptation.

### Keywords

transcription; signal integration; regulatory network; gal regulon; computational biology

## Introduction

Bacteria sense a wide array of signals (minerals, nutrients, stress signals, etc.). A large class of cellular response systems regulates the flux and concentration of small molecules by controlling transport and metabolism pathways via two feedback loops connected by a common transcription factor (TF) that senses the intracellular concentration of the small molecule[1]. The simplest systems consist of two operons, a regulator gene and a regulated operon containing at least two cistrons, one encoding a transporter and the other encoding an enzyme which modifies/degrades the small molecule[2]. In *E. coli* cells expression of several sugar operons are repressed when glucose is abundant in the growth medium. Full expression of these operons occurs when the glucose level is low and the corresponding sugar level is high[3; 4; 5; 6]. Therefore, regulation of these operons occurs by “combinatorial control”, which is common



*For correspondence: Szabolcs Semsey, Department of Genetics, Eötvös Lóránd University, Pázmány P. s. 1/C, H-1117 Budapest, Hungary E-mail semseys@yahoo.com; Tel: +36 1 3722500/8693.

in both prokaryotes and eukaryotes[7; 8]. In case of the above sugar operons the combinatorial control is implemented molecularly by a global (CRP) and a sugar specific transcription factor. Activation by CRP requires the inducer cAMP, which is produced intracellularly as a signal of glucose shortage[9]. In the simplest systems a single operon contains genes encoding a transporter and the enzyme for sugar conversion, therefore transport and metabolism pathways are regulated simultaneously[2]. However, some sugar utilization systems reached higher levels of complexity, e.g. having multiple transporters, regulators, or several enzymes of a metabolic pathway[5; 6]. One of the best studied examples of the latter is the galactose utilization system. The *gal* system contains genes involved in the transport (*galP*, *mglBAC*) and amphibolic utilization (*galETKM*) of the sugar D-galactose. Regulation of the *gal* regulon is governed by two transcription factors, GalR and GalS (Figure 1). GalR and GalS are iso-regulators because of the high similarity in their amino acid sequence (55% identity/ 88% similarity), inducer binding specificity, and DNA binding site recognition[6; 10]. To understand combinatorial control of the transport and metabolic pathways in the galactose system it is essential to determine how the two signals, cAMP and D-galactose concentrations are integrated at each promoter. The correlation between the levels of input signals and the promoter activity is called the input function.

In this study we measured activities of the six *gal* regulon promoters simultaneously, using an *in vitro* transcription system containing purified components. Input functions were computed based on the experimental measurements. Four different shapes of input functions were observed, which are in good agreement with our previous Boolean approximation[10]. We computed probabilities of operator occupancy by GalR and GalS at different levels of input signals at all six promoters. GalR binding dominates at most of the reported physiological intracellular concentrations of input signals.

## Results

### *Effect of TF and input signal concentrations on* gal *regulon promoter activities*

We studied the effect of varying concentrations of GalR, GalS, D-galactose, and cAMP-CRP on the *gal* regulon promoters simultaneously using the pRPGSM plasmid [10] in *in vitro* transcription reactions (Figure 2). The pRPGSM plasmid contains the respective *cis*-regulatory regions and promoters of the *galETKM*, *galP*, *galS*, *galR* and *mglBAC* operons. Reactions contained 80 nM HU protein required for GalR mediated DNA loop formation. RNA products of the *in vitro* transcription reactions were separated on 7% polyacrylamide-urea DNA sequencing gels. Relative promoter activities were determined based on the level of $^{32}$P radioactivity in the separated RNA bands as described in the *Materials and Methods* section.

### Strength of GalR and GalS binding

We constructed a model of the *in-vitro* system (see *Materials and Methods*). By fitting this model to our data, we obtain estimates of all binding strengths. The full results are shown in Table 1, but to summarize:

1. GalR binds to operators in decreasing order of strength: $O_{galP}$>*galE* $O_E$ (loop) > $O_{galR}$ > *galS* $O_E$ > *galE* $O_E$ (no loop) > $O_{mglB}$
2. GalS binds to operators in decreasing order of strength: *galS* $O_E$ > $O_{mglB}$ > $O_{galR}$ > *galE* $O_E$ > $O_{galP}$
3. The Hill coefficients are all around 2 (the range is 1.7 to 2.5).
4. D-galactose binds slightly stronger to GalR than to GalS, and the Hill coefficient is in both cases a little under 1.

5. cAMP-CRP binds to the activator binding sites (*AS*) in the regulatory region of the *gal* regulon promoters in decreasing order of strength: *galP>mglB> galE >galS*, and the Hill coefficients lie in the range of 1 to 2.

## Intracellular promoter input functions

In a cell, the levels of GalR and GalS are not constant. We can use the binding strengths determined above to estimate how promoter activities depend on the intracellular levels of D-galactose and cAMP by describing the dynamics of the GalR and GalS levels using the following differential equations:

$$\frac{dR_{tot}}{dt} = \nu_R A_{galR}(R_{tot}, S_{tot}, g, c) - R_{tot} \quad (1)$$

$$\frac{dS_{tot}}{dt} = \nu_S A_{galS}(R_{tot}, S_{tot}, g, c) - S_{tot} \quad (2)$$

where $A_{galR}$ and $A_{galS}$ are the promoter activities, defined in *Materials and Methods*. $\nu_R$ and $\nu_S$ are maximal production rates of GalR and GalS respectively. From the in vivo observation that in the absence of GalR, $P1_{galE}$ is repressed to 43% of its maximal activity, we previously determined that $\nu_S$=484 nM/cell generation [11]. Similarly, in the absence of GalS, $P1_{galE}$ is repressed to 10% of its maximal level, which leads to $\nu_R$ =125 nM/cell generation. We measure time in units of one cell generation (we assume that GalR and GalS are both stable, so the only way their concentrations can decrease is by dilution due to cell growth.)

Figure 3 shows the steady-state activity of various promoters as a function of D-galactose (*g*) and cAMP (*c*) concentrations. We observe four different shapes of input functions. Four of the input functions have simple shapes resembling Boolean AND ($P_{galS}$, $P_{galP}$, and $P1_{galE}$) or DGAL NIMPLIES CAMP ($P2_{galE}$) logic. The two other input functions have more complex shapes that can be best approximated by the combination of CAMP and AND ($P_{mglB}$) logic or DGAL and TRUE ($P_{galR}$). The input functions are generally monotonic; however, the $P_{mglB}$ transcription shows non-monotonic dependence on cAMP concentration at low galactose levels.

This non-monotonic response is the result of two regulatory features, (i) partial repression of $P_{mglB}$ by GalR[10] and (ii) the indirect repression of $P_{mglB}$ by cAMP-CRP through GalS[12]. The color coding in Figure 3 indicates which of the two regulators dominates (i.e., when the operator is bound by a repressor, what fraction of the time is it GalR) – brown means GalR, blue means GalS. In the case of *mgl* the dominant repressor changes from GalR, at low cAMP, to GalS, at high cAMP. As cAMP is increased from zero, GalS starts competing with GalR for the $O_{mglB}$ operator, but as GalS is a weaker repressor the net repression decreases. As cAMP is increased further, free GalS levels rise much higher (see Figure 4) once again increasing repression of the *mglBAC* operon.

The color coding for dominance in Figure 3 shows that GalS is mainly responsible for self-repression and repression of $P_{mglB}$ at high cAMP. For the rest, blue-regions are present only when activity levels are close to maximum (i.e., neither repressor spends much time bound to the respective operators). In other words, GalS plays little role in the steady-state repression of the *gal* regulon. GalR appears to be the main steady-state regulator[10].

### Functional significance of feed-forward loops in the galactose system

Data presented in Figure 3 allows for the prediction of the functionality of feed-forward loops in the galactose network. In the feed-forward loop network motif a transcription factor regulates a gene both directly and indirectly through regulating another TF. There are two types of feed-forward loops in the galactose network of *E. coli*, incoherent type I1 and incoherent type I2. In the type I1 feed-forward loops cAMP-CRP activates the target gene directly and represses it indirectly through the activation of $P_{galS}$. The significance of type I1 feed-forward loops in the galactose system have been discussed earlier[10; 12]. In the type I2 feed-forward loops one of the repressors (GalR or GalS) represses a target gene directly and also inhibits repression of the target gene indirectly by repressing the transcription of the other repressor gene. However, these type I2 feed-forward loops can function only at medium to high cAMP concentrations that allow for sufficient expression of GalS. Type I2 feed-forward loops where GalR regulates target genes directly and indirectly through repression of $P_{galS}$ are mostly inactive because GalS autoregulation dominates at higher cAMP concentrations. There is a narrow range of cAMP concentration where $P_{galS}$ regulation is dominated by GalR and therefore GalS mediated repression of $P_{mglB}$ is inhibited. This cAMP concentration range overlaps with the non-monotonous region of the input function for $P_{mglB}$. Type I2 feed-forward loops where GalS represses target genes directly and regulate them indirectly through repression of $P_{galR}$ are functional at high cAMP levels, where repression of $P_{galR}$ is strongly affected by GalS. These type I2 feed-forward loops result in slightly increased activity of the $P_{mglB}$, $P_{galP}$, and $P1_{galE}$ promoters at higher cAMP levels.

## Discussion

Regulation of gene expression is a main component in the response to the changing environment[13]. Transcription regulatory networks continuously sense a set of environmental signals and perform computations to adjust the gene expression profile according to the changes of these signals, in order to enhance the cells performance. Transcription of many genes is influenced by multiple signals[14], therefore it is essential to understand the relation of promoter activities and combinations of signal levels. The simplest case of combinatorial control is when the promoter activity depends on only two signals. This is typical in case of operons related to utilization of different sugars in the absence of glucose. In this work we analyzed transcriptional regulation of all six promoters of the galactose regulon in *E. coli*. This study presents the first characterization of signal integration at all promoters in a regulon.

The negative autoregulatory motif at $P_{galR}$, which is affected by a single signal (D-galactose concentration), can itself determine the shape of the input function. However, this is not the case when two or more signals are integrated at a promoter. The network structures around $P_{galP}$, and $P_{mglB}$, which control the two galactose transporters, are identical (Figure 1)[10], however, the input functions have different characteristics. Network motifs in these cases are insufficient for the determination of signal integration. It is the actual structure of the promoter and regulatory region, the mechanism of transcription regulation, and the interplay between transcription factors that shapes the input function to be suitable for adaptation. Our results support the conclusion of theoretical studies that network motif structure does not determine function[15; 16].

The response range of the input function to cAMP and D-galactose concentrations shows good agreement with the potential physiological range of these input signals. The galactose transport system is able to accumulate up to 60 mM D-galactose (nearly 5% of total cell dry weight) when D-galactose is not metabolized intracellularly[17]. However, wild type cells growing in the presence of D-galactose as a single carbon source contain only about 0.38 mM D-galactose as a net effect of the feedback loops involving transport and metabolism[18]. As a comparison, input functions show saturation (reach 90% of maximal activity) at maximum 16mM

concentration of D-galactose. cAMP is synthesized intracellulary by adenylate cyclase as a response to glucose starvation[9]. A sharp rise in intracellular cAMP level occurs when the nutrient concentration in minimal medium drops to approximately 0.3 mM glucose[19]. The level of intracellular cAMP, which is highly unstable[19], was reported to reach a maximum of 40 −240 μM, depending on growth conditions[19; 20; 21]. Input functions of the galactose regulon promoters show saturation (reach 90% of maximal activity) at maximum 28μM concentration of cAMP.

We analyzed the effect of the two galactose specific regulators, GalR and GalS, on the repression of the *gal* regulon promoters at steady state concentrations of D-galactose and cAMP. The strong autoregulation of GalS limits cross-regulation between GalS and GalR by favoring GalS binding to the *galS* $O_E$ operator and by preventing accumulation of GalS in sufficient amount to repress the $P_{galR}$ promoter. Besides autoregulation, GalS also influence transcription of $P_{mglB}$ at low D-galactose and high cAMP concentrations, however, to much less extent.

Input functions obtained in this study show significant differences from the ones observed by controlling extracellular cAMP and D-galactose levels[22; 23]. There are several factors that can cause these differences. The relation of extracellular and intracellular concentrations of small molecules is often non-linear partly because of regulated transport and intracellular conversion[1]. This is particularly true for cAMP[24], which is normally synthesized intracellularly, and for small molecules whose transport depends on the absence or presence of other small molecules in the environment. For example, transport of D-galactose is inhibited in the presence of glucose (inducer exclusion)[25]. Also, transcription of chromosomal transport and metabolic genes is influenced by the number and quality of TF binding sites supplied on plasmids[26] in the above *in vivo* measurements. For instance, supplying a strong GalR binding site (e.g. $O_{galP}$) in multicopy would inhibit GalR mediated repression of all the *gal* regulon promoters, while using a weaker binding site (e.g. $O_{mglB}$) in a similar setup would be less effective in derepression of promoters with a strong GalR binding site. Therefore contradictions in results observed by the *in vivo* and *in vitro* approaches can reveal important aspects of signal transport/ processing.

## Materials and Methods

### Protein purification

Hexahistidine tagged GalR and GalS was purified as described earlier [27]. Briefly, harvested cells were resuspended in 1/40 volume of Lysis I buffer (50 mM sodium phosphate, pH 8.0, 0.5 mg/ml lysozyme) and stored on ice for 30 minutes. Equal volume of Lysis II buffer (50 mM sodium phosphate, pH 8.0, 2 M NaCl, 8 mM imidazole, 20% glycerol) was added and incubated for 30 minutes on ice. The cell debris was removed by centrifugation at 10,000 g for 1 hour. 3% Ni-NTA slurry (Qiagen) was added to the solution, followed by 1 hour incubation at 4°C. A Poly-Prep Chromatography Column (BIO-RAD) was used to collect the protein bound to Ni-NTA agarose from the mixture. Twenty column-volume of washing buffer (50 mM sodium phosphate, pH 8.0, 600 mM NaCl, 60 mM imidazole, and 10% glycerol) was allowed to flow through the column. Hexahistidine tagged proteins were eluted by four column volumes of elution buffer (50 mM sodium phosphate, pH 8.0, 600 mM NaCl, 10% glycerol) containing 250 mM imidazole and fractions containing repressor protein with less than 5% contamination were stored at −80°C in 100 μl aliquots. HU protein was purified according to the method described by Aki et al [28]. CRP was purified as described by Ryu et al [29].

### In vitro transcription

Transcription reactions were performed as described previously [30]. The reaction mixture (50 μl) contained 20 mM Tris acetate, pH 7.8, 10 mM magnesium acetate, 200 mM potassium glutamate, 100 μM cAMP, and 2 nM supercoiled pRPGSM plasmid[10] DNA template. GalR, GalS, and CRP concentrations are as indicated in Figure 2. HU was used at 80 nM concentration. 20 nM RNA polymerase (USB) was added before incubating the reactions at 37°C for 5 minutes. Transcription was started by the addition of 1.0 mM ATP, 0.1 mM GTP, 0.1 mM CTP, 0.01 mM UTP, and 5 μCi of [α-$^{32}$P]UTP (3000 Ci/mmol). Reactions were terminated after 10 minutes by addition of an equal volume of transcription loading buffer (0.025% bromophenol blue, 0.025% xylene cyanol, 0.01 M EDTA and 90% deionized formamide). After heating at 90°C for 3 minutes, the samples were loaded onto 7% polyacrylamide-urea DNA sequencing gels. RNA bands were quantified using the ImageQuant™ PhosphorImager (Molecular Dynamics CA). Band intensities were corrected by the background and normalized to the RNA1 band intensities of the corresponding lanes.

### *Model of the* in-vitro *system*

In each *in-vitro* experiment, there is a fixed amount of GalR ($R_{tot}$), GalS ($S_{tot}$), cAMP ($c$) and galactose ($g$), some of which values could be zero. There are five operators at which GalR and GalS compete to bind: $O_{mglB}$, $O_{galR}$, $O_{galP}$, *galS* $O_E$, and *galE* $O_E$. In addition, GalR can bind to *galE* $O_E$ and $O_I$ simultaneously, forming a DNA loop, which GalS cannot do. We use $R$ and $S$ to denote, respectively, the amount of free GalR and GalS, i.e. the amount not bound to the operators or galactose. Similarly, we denote the amount bound to galactose by *(Rg)* and *(Sg)*. Our model of the promoter activities ($A$) in the *in vitro* system in steady-state then consists of the following equations:

$$\frac{A_i}{A_{i,\max}} = \frac{1}{1+(R/K_{r,i})^{h_{r,i}}+(S/K_{s,i})^{h_{s,i}}} \times \frac{(c/K_{c,i})^{h_{c,i}}}{1+(c/K_{c,i})^{h_{c,i}}}$$

where *i=mglB, galS, galP*. These equations use standard Michaelis-Menten terms to model activation of promoters by cAMP-CRP, as well as repression by GalR and GalS, whose concentrations appear in additive form in the denominator because GalR and GalS compete for binding to the same operator sites in the *mglB, galS* and *galP cis*- regulatory regions.[6]

$$\frac{A_{galR}}{A_{galR,\max}} = \frac{1}{1+\left(R/K_{r,galR}\right)^{h_{r,galR}}+\left((Rg)/K'_{r,galR}\right)^{h_{r,galR}}+\left(S/K_{s,galR}\right)^{h_{s,galR}}}$$

$P^{galR}$ activity ($A^{galr}$) is independent of cAMP-CRP. It is repressed by GalR and GalS which compete for the same binding site ($O^{glaR}$)[10]. Therefore $A^{galR}$ is modeled similarly to $A^i$ above. The only difference is that we also include the possibility for the GalR-galactose complex to bind to the operator site and repress $P^{galR}$. This is included because experimental data shows that even for very large concentrations of galactose, $P^{glaR}$ activity does not reach 100%.

$$\frac{A_{P1}}{A_{P1,\max}} = \frac{1+0.025\left(R/K_{r,OE}\right)^{h_{r,OE}}+0.025\left((Rg)/K'_{r,OE}\right)^{h_{r,OE}}}{1+\left(R/K_{r,OE}\right)^{h_{r,OE}}+\left((Rg)/K'_{r,OE}\right)^{h_{r,OE}}+\left(R/K_{r,loop}\right)^{2h_{r,loop}}+\left(S/K_{s,OE}\right)^{h_{s,OE}}} \times \frac{\left(c/K_{c,OE}\right)^{h_{c,OE}}}{1+\left(c/K_{c,OE}\right)^{h_{c,OE}}}$$

This equation for the activity of $P1^{galE}$ models the following facts: $P1^{galE}$ is activated by cAMP-CRP, repressed by GalR or GalS binding to *galE* $O^E$, and strongly repressed by GalR

forming a DNA loop by binding to both *galE* $O^E$ and *galE* $O^{I30}$; [31; 32]. Similar to $P^{galR}$, $P1^{galE}$ activity does not reach 100% at large galactose concentrations[32], therefore we include the possibility of the GalR-galactose complex binding to *galE* $O^E$ and repressing the promoter[32; 33]. However, it is known that the GalR-galactose complex is unable to form the DNA loop therefore we exclude that possibility[32].

$$\frac{A_{P2}}{A_{P2,basal}}=\frac{1+2\left(R/K_{r,OE}\right)^{h_{r,OE}}+2\left((Rg)/K'_{r,OE}\right)^{h_{r,OE}}+\left(R/K_{r,OE}\right)^{h_{r,OE}}\left(c/K_{c,P2}\right)^{h_{c,P2}}+\left((Rg)/K'_{r,OE}\right)^{h_{r,OE}}\left(c/K_{c,P2}\right)^{h_{c,P2}}+1.2\left(S/K_{s,OE}\right)^{h_{s,OE}}}{\left(1+\left(R/K_{r,OE}\right)^{h_{r,OE}}+\left((Rg)/K'_{r,OE}\right)^{h_{r,OE}}+\left(R/K_{r,loop}\right)^{2h_{r,loop}}+\left(S/K_{s,OE}\right)^{h_{s,OE}}\right)\times\left(1+\left(c/K_{c,P2}\right)^{h_{c,P2}}\right)}$$

$P2^{galE}$ activity has a more complex regulation. Overall, the above equation models the fact that GalR binding to *galE* $O^E$ activates the promoter two fold, whereas both GalR mediated DNA loop formation and cAMP-CRP binding repress it[31; 34].

The levels of GalR and GalS bound to galactose are given by:

$$(Rg)=R\left(g/K_{r,g}\right)^{h_{r,g}}$$

$$(Sg)=S\left(g/K_{s,g}\right)^{h_{s,g}}$$

Finally, the total amounts of GalR and GalS can be split into four pools: free protein, bound to galactose but not to an operator, bound to operator sites, and bound to galactose and to an operator site:

$$R_{tot}=R+(Rg)+R_{bound}+(Rg)_{bound}$$

$$S_{tot}=S+(Sg)+S_{bound}+(Sg)_{bound}$$

From these equations, given values of $R^{tot}$, $S^{tot}$, $g$ and $c$, we can determine all promoter activities, and the amounts of GalR and GalS bound to each operator and to galactose. The equations have 49 parameters. 13 of these have been derived directly from measured data, and the rest we determine by least squares fitting to the *in-vitro* data (see Table 1).

## Acknowledgments

We thank our colleagues in the laboratory for various inputs in the project, in particular Thomas Soares and Mofang Liu for purification of HU and CRP, and Takácsné Botond Judit for technical assistance. This research was supported by the Intramural Research Program of the NIH, National Cancer Institute, Center for Cancer Research, by a Marie Curie International Reintegration Grant within the 6th European Community Framework Program, by OTKA PD75496, and by the Danish National Research Foundation. SS is grateful for the János Bolyai fellowship of the Hungarian Academy of Sciences.

## Abbreviations

GalR, Gal repressor; GalS, Gal isorepressor; cAMP, cyclic AMP; CRP, cyclic AMP receptor protein; HU, heat unstable protein; TF, transcription factor.

## References


1. Krishna S, Semsey S, Sneppen K. Combinatorics of feedback in cellular uptake and metabolism of small molecules. Proc Natl Acad Sci U S A 2007;104:20815–9. [PubMed: 18093927]
2. Monod J, Changeux JP, Jacob F. Allosteric proteins and cellular control systems. J. Mol. Biol 1963;6:306–329. [PubMed: 13936070]
3. Chapon C. Role of the catabolite activator protein in the maltose regulon of Escherichia coli. J Bacteriol 1982;150:722–9. [PubMed: 7040340]
4. Dickson RC, Abelson J, Barnes WM, Reznikoff WS. Genetic regulation: the Lac control region. Science 1975;187:27–35. [PubMed: 1088926]
5. Schleif R. Regulation of the L-arabinose operon of Escherichia coli. Trends Genet 2000;16:559–65. [PubMed: 11102706]
6. Weickert MJ, Adhya S. The galactose regulon of Escherichia coli. Mol Microbiol 1993;10:245–51. [PubMed: 7934815]
7. Buchler NE, Gerland U, Hwa T. On schemes of combinatorial transcription logic. Proc Natl Acad Sci U S A 2003;100:5136–41. [PubMed: 12702751]
8. Davidson EH, Rast JP, Oliveri P, Ransick A, Calestani C, Yuh CH, Minokawa T, Amore G, Hinman V, Arenas-Mena C, Otim O, Brown CT, Livi CB, Lee PY, Revilla R, Rust AG, Pan Z, Schilstra MJ, Clarke PJ, Arnone MI, Rowen L, Cameron RA, McClay DR, Hood L, Bolouri H. A genomic regulatory network for development. Science 2002;295:1669–78. [PubMed: 11872831]
9. Pastan I, Perlman R. Cyclic adenosine monophosphate in bacteria. Science 1970;169:339–44. [PubMed: 4317896]
10. Semsey S, Krishna S, Sneppen K, Adhya S. Signal integration in the galactose network of Escherichia coli. Mol Microbiol 2007;65:465–76. [PubMed: 17630975]
11. Semsey S, Krishna S, Erdossy J, Horvath P, Orosz L, Sneppen K, Adhya S. Dominant Negative Autoregulation Limits Steady-State Repression Levels in Gene Networks. J Bacteriol 2009;191
12. Mangan S, Itzkovitz S, Zaslaver A, Alon U. The incoherent feed-forward loop accelerates the response-time of the gal system of Escherichia coli. J Mol Biol 2006;356:1073–81. [PubMed: 16406067]
13. Covert MW, Schilling CH, Famili I, Edwards JS, Goryanin II, Selkov E, Palsson BO. Metabolic modeling of microbial strains in silico. Trends Biochem Sci 2001;26:179–86. [PubMed: 11246024]
14. Gama-Castro S, Jimenez-Jacinto V, Peralta-Gil M, Santos-Zavaleta A, Penaloza-Spinola MI, Contreras-Moreira B, Segura-Salazar J, Muniz-Rascado L, Martinez-Flores I, Salgado H, Bonavides-Martinez C, Abreu-Goodger C, Rodriguez-Penagos C, Miranda-Rios J, Morett E, Merino E, Huerta AM, Trevino-Quintanilla L, Collado-Vides J. RegulonDB (version 6.0): gene regulation model of Escherichia coli K-12 beyond transcription, active (experimental) annotated promoters and Textpresso navigation. Nucleic Acids Res 2008;36:D120–4. [PubMed: 18158297]
15. Ingram PJ, Stumpf MP, Stark J. Network motifs: structure does not determine function. BMC Genomics 2006;7:108. [PubMed: 16677373]
16. Knabe JF, Nehaniv CL, Schilstra MJ. Do motifs reflect evolved function?--No convergent evolution of genetic regulatory network subgraph topologies. Biosystems 2008;94:68–74. [PubMed: 18611431]
17. Horecker BL, Thomas J, Monod J. Galactose transport in Escherichia coli. I. General properties as studied in a galactokinaseless mutant. J Biol Chem 1960;235:1580–5. [PubMed: 14403098]
18. Death A, Ferenci T. Between feast and famine: endogenous inducer synthesis in the adaptation of Escherichia coli to growth with limiting carbohydrates. J Bacteriol 1994;176:5101–7. [PubMed: 8051023]
19. Notley-McRobb L, Death A, Ferenci T. The relationship between external glucose concentration and cAMP levels inside Escherichia coli: implications for models of phosphotransferase-mediated regulation of adenylate cyclase. Microbiology 1997;143(Pt 6):1909–18. [PubMed: 9202467]
20. Buettner MJ, Spitz E, Rickenberg HV. Cyclic adenosine 3',5'-monophosphate in Escherichia coli. J Bacteriol 1973;114:1068–73. [PubMed: 4351386]
21. Peterkofsky A, Gazdar C. Glucose inhibition of adenylate cyclase in intact cells of Escherichia coli B. Proc Natl Acad Sci U S A 1974;71:2324–8. [PubMed: 4366761]

22. Kaplan S, Bren A, Dekel E, Alon U. The incoherent feed-forward loop can generate non-monotonic input functions for genes. Mol Syst Biol 2008;4:203. [PubMed: 18628744]

23. Kaplan S, Bren A, Zaslaver A, Dekel E, Alon U. Diverse two-dimensional input functions control bacterial sugar genes. Mol Cell 2008;29:786–92. [PubMed: 18374652]

24. Kuhlman T, Zhang Z, Saier MH Jr. Hwa T. Combinatorial transcriptional control of the lactose operon of Escherichia coli. Proc Natl Acad Sci U S A 2007;104:6043–8. [PubMed: 17376875]

25. Adhya S, Echols H. Glucose effect and the galactose enzymes of Escherichia coli: correlation between glucose inhibition of induction and inducer transport. J Bacteriol 1966;92:601–8. [PubMed: 5332079]

26. Irani M, Orosz L, Busby S, Taniguchi T, Adhya S. Cyclic AMP-dependent constitutive expression of gal operon: use of repressor titration to isolate operator mutations. Proc Natl Acad Sci U S A 1983;80:4775–9. [PubMed: 6308647]

27. Semsey S, Geanacopoulos M, Lewis DE, Adhya S. Operator-bound GalR dimers close DNA loops by direct interaction: tetramerization and inducer binding. EMBO J 2002;21:4349–4356. [PubMed: 12169637]

28. Aki T, Choy HE, Adhya S. Histone-like protein HU as a specific transcriptional regulator: co-factor role in repression of gal transcription by GAL repressor. Genes Cells 1996;1:179–188. [PubMed: 9140062]

29. Ryu S, Kim J, Adhya S, Garges S. Pivotal role of amino acid at position 138 in the allosteric hinge reorientation of cAMP receptor protein. Proc Natl Acad Sci U S A 1993;90:75–9. [PubMed: 8380500]

30. Geanacopoulos M, Adhya S. Functional characterization of roles of GalR and GalS as regulators of the gal regulon. J. Bacteriol 1997;179:228–234. [PubMed: 8982002]

31. Choy HE, Adhya S. Control of gal transcription through DNA looping: inhibition of the initial transcribing complex. Proc. Natl. Acad. Sci. U. S. A 1992;89:11264–11268. [PubMed: 1454806]

32. Semsey S, Virnik K, Adhya S. Three-stage regulation of the amphibolic gal operon: from repressosome to GalR-free DNA. J Mol Biol 2006;358:355–63. [PubMed: 16524589]

33. Chatterjee S, Zhou YN, Roy S, Adhya S. Interaction of Gal repressor with inducer and operator: induction of gal transcription from repressor-bound DNA. Proc. Natl. Acad. Sci. U. S. A 1997;94:2957–2962. [PubMed: 9096328]

34. Semsey S, Tolstorukov MY, Virnik K, Zhurkin VB, Adhya S. DNA trajectory in the Gal repressosome. Genes Dev 2004;18:1898–907. [PubMed: 15289461]

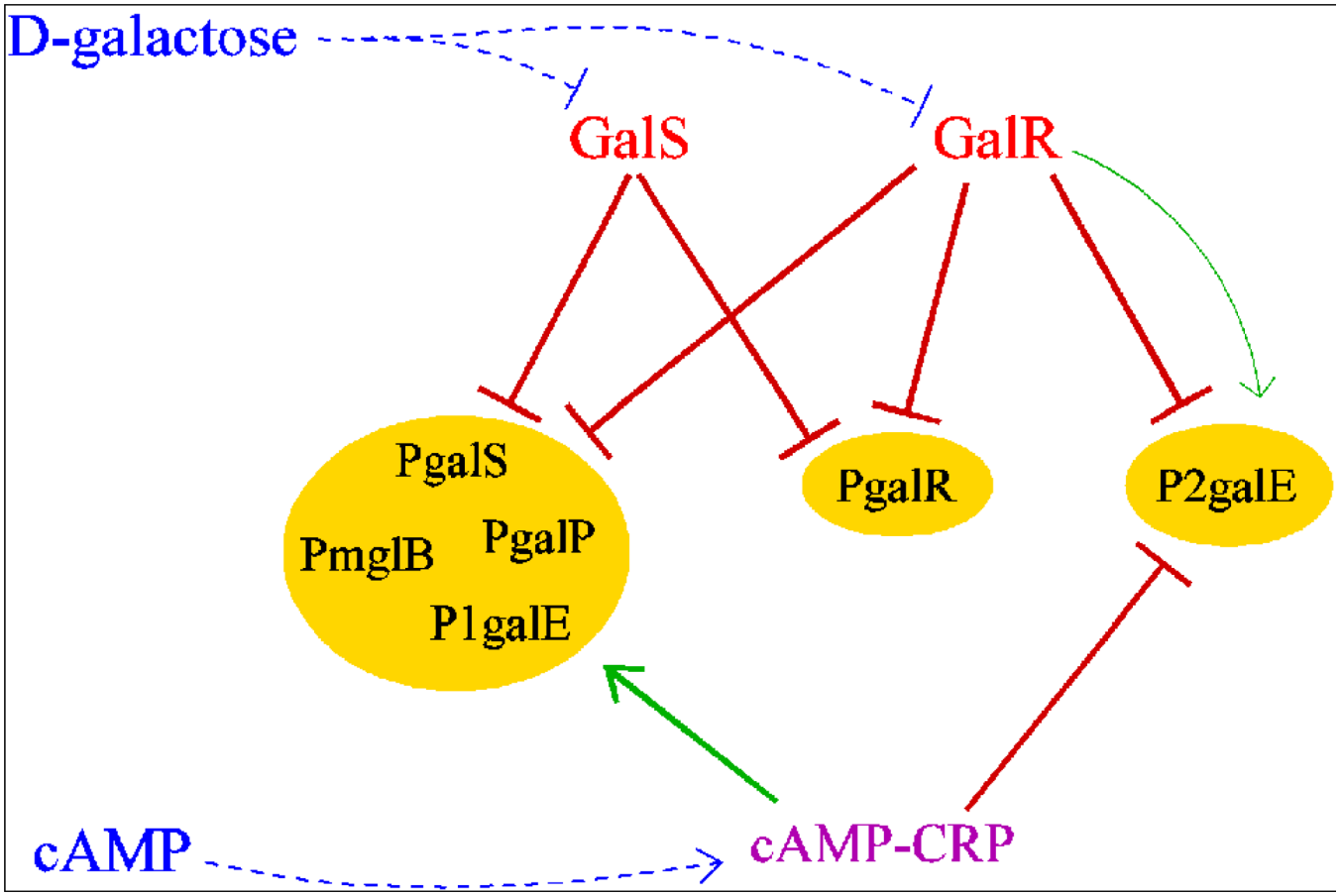


**Figure 1.**
Effect of transcription factors and small molecules on the activity of the *gal* regulon promoters. $P_{galS}$, $P_{galP}$, $P_{mglB}$, and $P1_{galE}$ are repressed by GalR and regulated by an incoherent feed-forward loop[12] involving cAMP-CRP mediated activation and GalS mediated repression. $P_{galR}$ activity is repressed directly by both GalR and GalS, and indirectly by cAMP-CRP through GalS. $P2_{galE}$ transcription can be repressed by either cAMP-CRP or by GalR-mediated DNA looping, and enhanced by GalR in the absence of DNA loop formation (thin green arrow).

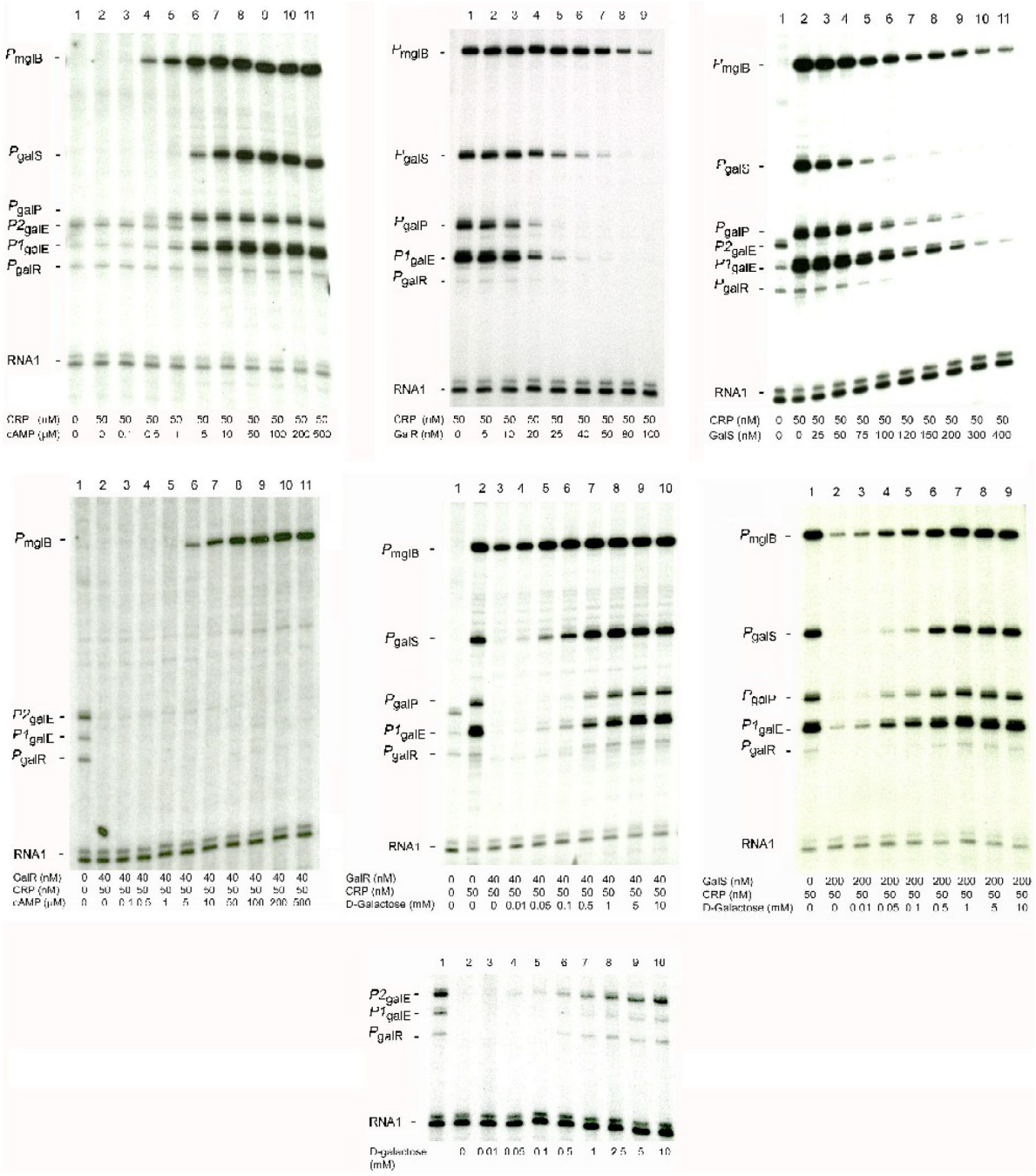


**Figure 2.**

In vitro transcription of the *gal* regulon promoters in the presence or absence of varying amounts of GalR, GalS, CRP, D-galactose and cAMP. cAMP was used at 100 μM concentration when not indicated. HU was present at 80 nM in all reactions to assist GalR mediated DNA loop formation.

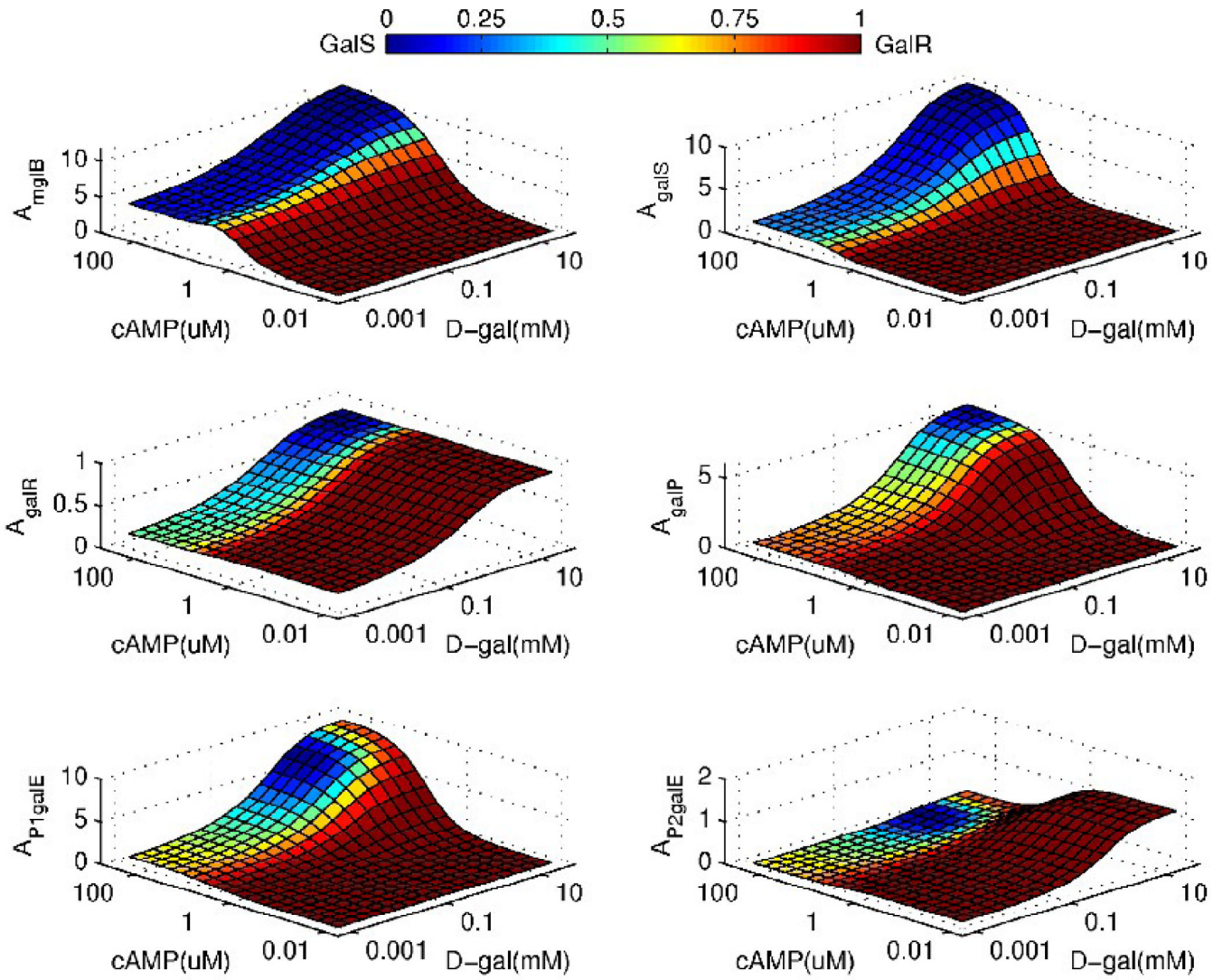


**Figure 3.**

Promoter activities as a function of intracellular galactose and cAMP levels (z-axis values are relative to $P_{galR}$, for which the maximum activity level is chosen to be 1). For each value of cAMP and D-gal, we integrated the differential equations (1) and (2), with parameter values as described in Table 1, until the concentrations reached steady-state. The activity levels were then computed by inserting these steady-state values into the equation for promoter activities given in Materials and Methods. Color coding indicates which repressor dominates the binding, i.e., blue indicates that whenever a repressor is bound to the operator it is always GalS, while green indicates that 50% of the time a repressor is bound to the operator it is GalS, and dark red indicates that GalR dominates repression. Please note that $P2_{galE}$ repression in the yellow and blue areas is the result of cAMP-CRP action. The effect of GalS on $P2_{galE}$ transcription is negligible [30]. We checked that similar behavior was observed when all the fitted parameters were randomly varied by up to 50% (data not shown).

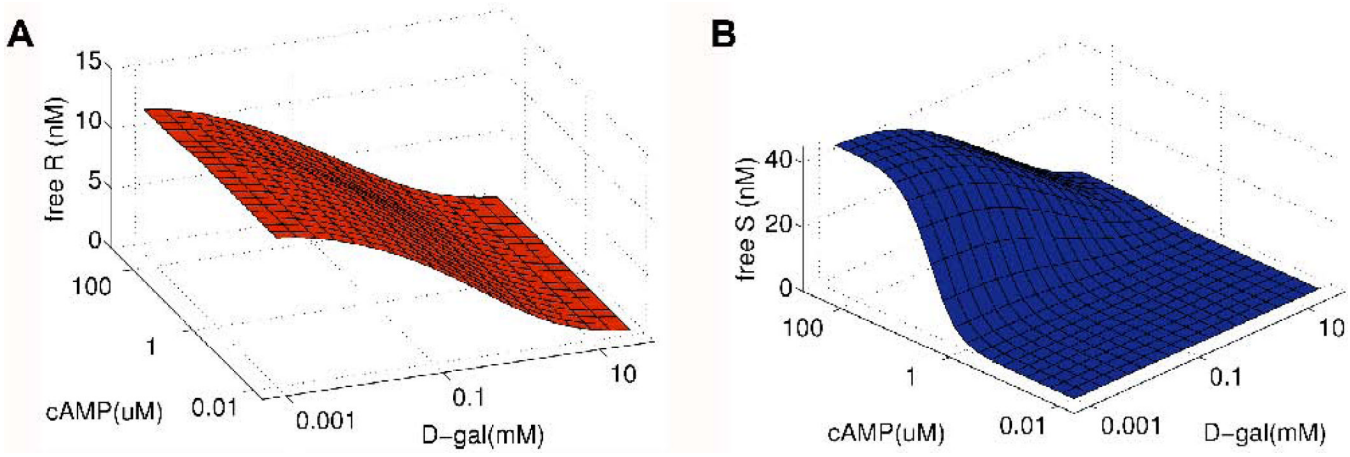


**Figure 4.**
Concentration of (a) free GalR (i.e., not bound to either operators or galactose) and (b) free GalS, derived from the theoretical model, in steady-state conditions, as a function of intracellular galactose and cAMP levels.

## Table 1

Parameter values in the theoretical model of the gal system. The table shows the 36 parameters that were determined by least-square fitting to *in-vitro* data. Uncertainties in these values were determined by repeating the fits 50 times, each time with data points perturbed randomly by up to 10%, which mimics experimental uncertainty. Overall, the errors in the fit parameter values were around 10% for cAMP-CRP binding, and less than 7% for other parameters. Exceptions are the dissociation constants of galactose binding to GalR and GalS, which have errors of 16% and 20%, and the dissociation constant and Hill coefficient of GalR binding to *galE* $O_E$, which have errors of 40% and 14%, respectively (this is not surprising: there is only a small range of GalR concentrations where *galE* $O_E$ binding is significant, without being overshadowed by DNA loop formation, therefore these parameters do not affect the fits much.) Parameters that have been derived from measurements are: $K'_{r,r}=44.77K_{r,r}$; $K'_{r,OE}=10.20K_{r,OE}$; $A_{mgl,max}=11.88$; $A_{galS,max}=10.06$; $A_{galP,max}=5.62$; $A_{P1galE,max}=12.1$; $A_{P2,basal}=1$ (these five are all relative to the maximum promoter activity of *galR*, $A_{max,galR}$, which is normalized to be 1); maximal GalR and GalS protein production rates, $\nu_R=125$ and $\nu_S=484$ nM/cell generation, respectively. Five additional parameters obtained directly from measurements are the relative activity levels of $P1_{galE}$ (when R or (Rg) is bound to *galE* $O_E$ = 0.025 of maximal activity) and of $P2_{galE}$ (when R or (Rg) is bound to *galE* $O_E$ = 2; when S is bound to *galE* $O_E$ = 1.2 of basal activity).

| | **Michaelis constant for GalR binding to operator** | **Michaelis constant for GalS binding to operator** | **Michaelis constant for cAMP activation via cAMP-CRP binding to *AS*** |
|---|---|---|---|
| *mglB* | $K_{r,m}=24.92$nM<br>$h_{r,m}=1.73$ | $K_{s,m}=35.20$nM<br>$h_{s,m}=1.85$nM | $K_{c,m}=2.48$uM<br>$h_{c,m}=1.09$ |
| *galS* | $K_{r,s}=8.50$nM<br>$h_{r,s}=2.41$ | $K_{s,s}=23.74$nM<br>$h_{s,s}=2.50$ | $K_{c,s}=8.36$uM<br>$h_{c,s}=1.83$ |
| *galP* | $K_{r,p}=4.27$nM<br>$h_{r,p}=2.48$ | $K_{s,p}=46.26$nM<br>$h_{s,p}=2.54$ | $K_{c,p}=1.69$uM<br>$h_{c,p}=1.24$ |
| *galR* | $K_{r,r}=6.54$nM<br>$h_{r,r}=1.99$ | $K_{s,r}=42.82$nM<br>$h_{s,r}=2.20$ | |
| *galE* | $K_{r,OE}=20.71$nM<br>$h_{r,OE}=1.84$ | $K_{s,OE}=43.29$nM<br>$h_{s,OE}=1.88$ | $K_{c,1}=4.59$uM<br>$h_{c,1}=1.34$<br>$K_{c,2}=2.75$uM<br>$h_{c,2}=1.37$ |
| *galE loop* | $K_{r,loop}=4.33$<br>$h_{r,loop}=1.17$ | | |
| D-gal | $K_{r,g}=0.02$mM<br>$h_{r,g}=0.74$ | $K_{s,g}=0.03$mM<br>$h_{s,g}=0.71$ | |