doi:10.1016/j.jmb.2009.07.050 J. Mol. Biol. (2009) 392, 1133–1144

JMB

Available online at www.sciencedirect.com

ScienceDirect

ELSEVIER

One Is Not Enough

Robert Daber, Kim Sharp and Mitchell Lewis*

Department of Biochemistry and Biophysics, University of Pennsylvania School of Medicine, 37th and Hamilton Walk, Philadelphia, PA 19104-6059, USA

Received 3 March 2009; received in revised form 15 July 2009; accepted 16 July 2009
Available online 22 July 2009

Edited by J. Karn

In both prokaryotic and eukaryotic organisms, repressors and activators are responsible for regulating gene expression. The *lac* operon is a paradigm for understanding how metabolites function as signaling molecules and modulate transcription. These metabolites or allosteric effector molecules bind to the repressor and alter the conformational equilibrium between the induced and the repressed states. Here, we describe a set of experiments where we modified a single inducer binding site in a dimeric repressor and examined its effect on induction. Based upon these observations, we have been able to calculate the thermodynamic parameters that are responsible for the allosteric properties that govern repressor function. Understanding how effector molecules alter the thermodynamic properties of the repressor is essential for establishing a detailed understanding of gene regulation.

Keywords: *lac* repressor; allostery; induction; repression; inducer binding

Introduction

Cellular proteins that perform specific metabolic tasks are often regulated to meet the needs of the organism. In many instances, regulating the flux through a pathway is achieved by adjusting the concentration of an enzyme that controls the rate-determining step in a pathway. In both prokaryotic and eukaryotic organisms, enzyme concentration is regulated by transcription, which is frequently controlled by repressors and activators. These molecules either directly or indirectly monitor the accumulation or diminution of a metabolite and respond like a molecular switch—increasing or decreasing the rate of gene expression. Effector molecules are the chemical signals that convey the metabolic state of the cell to the genetic machinery. They can function as inducers, anti-inducers, or co-repressors, and in each instance, they alter the equilibrium between two unique conformational states of the molecular switch.

The switch that regulates the *lac* operon is arguably the most well characterized and therefore serves as the paradigm for understanding gene regulation (recently reviewed by Wilson *et al.*[1]). The switch is a two-component system and consists of a repressor molecule and an operator. The repressor is a 360-amino-acid protein that has a modular structure composed of an NH_2-terminal or "headpiece" domain (~60 residues) and a COOH-terminal "core" domain. The headpiece contains the classic helix–turn–helix motif that recognizes and binds to an operator sequence, while the core domain is responsible for inducer binding and contains the dimerization interface. The second component of the switch is the operator, a short stretch of DNA that is pseudo-palindromic. In the *lac* operon, the primary operator is positioned just upstream of the gene for β-galacotosidase.[2] As a negative regulator, the repressor associates with the operator sequence and physically blocks transcription of the genes that are necessary for lactose metabolism. The inducer molecule relieves repression by altering the repressor–operator equilibrium, stabilizing a conformation that is incompatible with operator binding.[3] The natural inducer of the *lac* operon is an analogue of lactose, allolactose, but a gratuitous inducer, such as 1-isopropyl-β,D-thiogalactopyranoside (IPTG), can also effectively decrease the binding affinity between the repressor and its operator. Since the inducer binds to a site that is distal to the DNA binding domain, the signal is transmitted through the molecule by altering the conformation of the repressor; this structural rearrangement of the repressor results in an allosteric transition.

Monod, Wyman, and Changeux first described how structural changes could alter a protein's ability to perform a given function and coined the term *allostery*.[4] The Monod–Wyman–Changeux MWC model assumes that (a) allosteric proteins are oligomeric with at least one axis of symmetry, (b) these proteins adopt two distinct conformations

*Corresponding author. E-mail address: lewis@mail.med.upenn.edu.

Abbreviations used: MWC model, Monod–Wyman–Changeux model; GFP, green fluorescent protein.

(designated as "R" and "T" by Monod *et al*. and described here as "R" and "R*"), and (c) the conformational states have different affinities for the substrates. In this instance, the R and R* conformations have different affinities for the inducer and the operator; (d) each substrate (i.e., operator or inducer) binds to independent sites on each monomer and, finally, (e) the molecular symmetry is conserved in the transition from one state to another. The *lac* repressor is an allosteric protein; it is oligomeric, and each monomer has two well-separated substrate binding sites, one for the operator and another for the inducer. Biochemical and structural studies have also demonstrated that the repressor adopts two distinct quaternary structures that correspond to the induced and repressed states (see Wilson *et al*.[1]).

Here, we describe a set of experiments that allowed us to elucidate the *in vivo* thermodynamic parameters that characterize the binding of inducer to the two conformational states of the repressor, as well as the equilibrium constant between these two states. By measuring the effect of zero, one, and two inducers binding to the repressor, we were able to extract equilibrium constants and develop a thermodynamic model to describe the allosteric process.

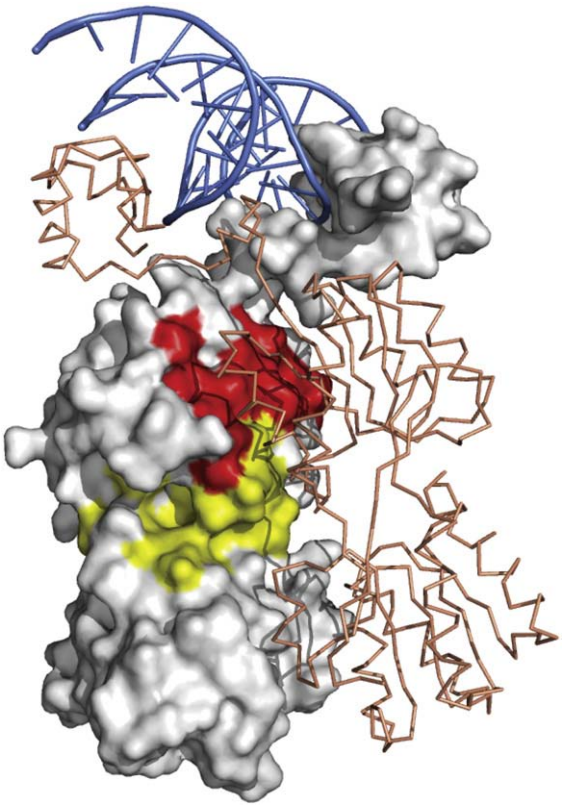

Fig. 1. The surface of the dimeric repressor depicting the locations of mutants with the I^s phenotype. The mutants shown in yellow surround or are within the inducer binding pocket. The surface shown in red represents I^s mutants believed to interfere with transmission of the allosteric signal.

Results

In vivo induction analysis

Several mutant repressor molecules that bind to the operator DNA with wild-type affinity but are incapable of induction have been identified.[5] These substitutions, classified by an I^s phenotype, either are defective in inducer binding or cannot transmit the allosteric signal to the DNA binding domain. The position of the I^s point mutations appear in five general locations with respect to the primary sequence and includes residues 70–80, 90–100, 190–200, 245–250, and 272–277. When the positions of these mutations are mapped onto the protein structure, as is illustrated in Fig. 1, they cluster either at the dimer interface or in close proximity to the inducer binding site. Presumably, mutants within the ligand binding pocket interfere with the ability of the repressor to bind the inducer, while mutants at the dimer interface disrupt the signaling process.[6] The crystal structure of the repressor bound to the inducer illustrates that IPTG forms hydrogen bonds to the amino acid side chains of Asp149, Arg197, Asn246, and Asp274, as well as van der Waals interactions with a hydrophobic surface created by Ile79, Leu148, Phe161, Asn291, Phe293, and Leu296.[7] In fact, mutating any of these residues frequently results in the I^s phenotype.[5] However, as one would anticipate, not all mutations in the inducer binding pocket affect induction equally. One particularly potent mutation in the ligand pocket, R197G, completely disrupts inducer binding but does not alter the repressor's ability to fold or bind to the operator.[8] The side chain of R197 anchors the galactose ring of the inducer by forming two hydrogen bonds with the C2 and C3 hydroxyls. To establish how inducer binding affects the allosteric properties of the repressor, we used the R197G mutation and created heterodimeric repressors with zero, one, or two functional inducer binding sites.

As described previously,[9] heterodimeric repressors were created by altering the C-terminal dimer interface of the wild-type dimeric *lac* repressor. These heterodimeric repressors have a unique monomer–monomer interface that allows us to probe the effect of asymmetric changes in the ligand pocket. The R197G mutation was incorporated into one or both monomers, producing heterodimeric repressors with zero, one, or two nonfunctional inducer binding pockets (Table 1). Subsequently, the heterodimeric repressors were introduced into cells containing the GFPmut3.1 reporter, whose transcript is controlled by the binding of the repressor to a chimeric operator.[9] Cells transformed with plasmids containing the heterodimers were then grown in the presence and in the absence of 2.5 mM IPTG, and the levels of transcription were determined by measuring green fluorescent protein (GFP) fluorescence. As shown in Fig. 2a, the heterodimeric repressors exhibit tight repression, but disruption

Table 1. Sequence of repressors used in *in vivo* assays

Repressor	Monomer 1		Monomer 2								Functional inducer sites
	197	282	17	18	22	251	255	278	281	282	
Wild type[a]	R	Y	Y	Q	R	L	R	D	C	Y	2
R197A	A	Y	Y	Q	R	L	R	D	C	Y	0
R197G	G	Y	Y	Q	R	L	R	D	C	Y	0
Y282S and Het2S	R	S	T	A	N	K	L	T	W	T	2
Y282S/R197A and Het2S	A	S	T	A	N	K	L	T	W	T	1
Y282S/R197G and Het2S	G	S	T	A	N	K	L	T	W	T	1

This table represents the various repressor constructs used in experimentation. The top sequence corresponds to the wild-type sequence found in the *lac* repressor.

[a] Both a full-length tetrameric repressor and a dimeric construct truncated after residue 332 were analyzed. All constructs containing Y282S, R197G/A or Het2S were also truncated after residue 332.

of the ligand binding sites alters the relative induction. The heterodimeric repressor with both sites intact has near-full induction (76% maximal expression), similar to both the dimeric form and the tetrameric form of the wild-type *lac* repressors (Fig. S1). When either of the inducer binding sites is eliminated, induction is far from optimal and repression is only slightly relieved, leading to less than 20% maximal induction. As expected, when both inducer binding pockets are destroyed, repression can only be slightly relieved at the highest concentration of the inducer tested. Since this heterodimeric repressor binds to an asymmetric operator, we also explored if the orientation of the

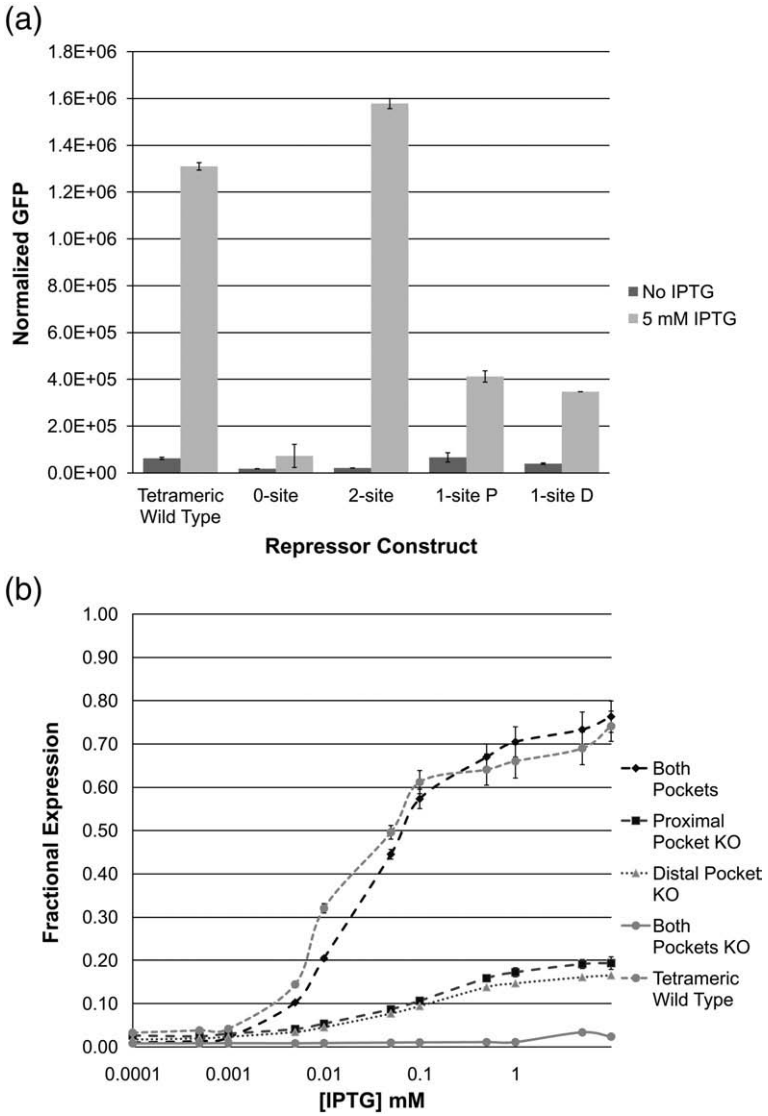

Fig. 2. Repression and induction data. (a) The histogram illustrates the levels of GFP production in the induced and the repressed states. For comparison, GFP production is shown for the wild-type (tetrameric) repressor and for the natural operator. The next four pairs of bar graphs illustrate the level of GFP in the induced and the repressed states when an R197G mutation is placed in (a) both pockets, (b) neither pocket, (c) the pocket proximal to the promoter, and (d) the pocket distal to the promoter in the heterodimeric repressor. Repression by the heterodimers is with respect to the chimeric operator. (b) Response of each repressor mutant to a range of IPTG concentrations. Fractional expression is defined as the signal of the sample divided by the signal when no repressor is present. Higher concentrations of IPTG tested produced mixed results with no increase in signal for the two-site case. Each construct with an R197G mutation began inducing again, consistent with reduced IPTG affinity for that mutation.

repressor containing a single functional ligand binding site with respect to the promoter affected induction. The asymmetric operator was inserted in both orientations into the reporter plasmid. The levels of induction were identical irrespective of whether the inducible monomer was proximal or distal to the promoter (Fig. 2a).

The induction measurements were repeated over a wide range of inducer concentrations (0.1 nM to 50 mM) to ensure complete induction (Fig. 2b). Again, when both inducer sites were functional, we observed 80% of the maximal expression of GFP even at the highest concentration of inducer; the fluorescent signal never reaches the same level as that obtained when no repressor is present. This suggests that even when both inducer sites are fully occupied, there is still residual binding to the operator resulting in incomplete induction. Consistent results were also obtained with both the tetrameric and the dimeric wild-type repressors, and the lack of full induction may represent the repressor–operator–inducer complex recently identified with small-angle scattering.[10] These wild-type repressor constructs behaved similarly since the auxiliary operator sites required for increased repression by the tetrameric repressor were absent.

Induction was further limited by inactivating one of the binding sites. Repressors with one functional ligand binding site never exceed 20% of the maximal induction even at very high concentrations of the inducer. The titration curves suggest that although repression can be partially relieved upon binding a single inducer, two molecules of IPTG bound to the repressor are needed for complete induction, such that one is not enough. To ensure that the observed phenotype was not an artifact of the R197G mutation, we also tested additional constructs containing the R197A mutation. The mutations demonstrated identical phenotypes (Fig. S2), supporting the working hypothesis that mutating residue R197 affects inducer binding alone. By fitting the induction data to a "two-state model," we can establish the conformational equilibrium constant in the MWC model, as well as the binding affinities of the inducer to the repressed and induced conformations.

Extracting the thermodynamic parameters

A fundamental tenet of the MWC model is that allosteric proteins exit in two conformations with different relative activities. In the case of the *lac* repressor, the two conformations are differentiated by their preferential binding to either the operator, O, or the inducer ligand, I. Here, we designate the conformation that prefers to bind DNA as the active form R and the conformation that prefers to bind inducer as the inactive form R*. The two conformations are in equilibrium (Eq. (1)), such that $K_{R^*R} = [R^*]/[R]$:

$$R \overset{K_{R^*R}}{\Leftrightarrow} R^* \tag{1}$$

Both of these repressor conformations can bind to the operator in an equilibrium manner, but equilibrium values are not identical:

$$R + O \overset{K_{RO}}{\Rightarrow} RO \tag{1a}$$

$$R^* + O \overset{K_{R^*O}}{\Rightarrow} R^*O \tag{1b}$$

where $K_{RO} = [RO]/[R][O]$ and $K_{R^*O} = [R^*O]/[R^*][O]$ are the respective association constants. Although both conformations can bind to the operator, the ratio of equilibrium constants, $s = K_{R^*O}/K_{RO}$, must be less than 1 for induction to occur.

In the classic MWC model, the inducer also binds to both conformations of the repressor with different affinities; it binds to the active (R) conformation with an affinity given by the binding constant $K_{IR} = [RI]/[R][I]$ and to the inactive (R*) conformation with an affinity given by the binding constant $K_{IR^*} = [R^*I]/[R^*][I]$. Since the functional unit of the repressor is dimeric, the repressor can bind two inducer molecules. The two binding sites are distant from one another, and there is no *direct* interaction between inducers such that the inducer molecules bind independently to each of the binding pockets with identical affinities (i.e., there is no contribution to cooperativity from direct ligand–ligand interaction). This assumption is borne out by the observation that the dependence of the apparent rate of dissociation of the repressor–operator complex with respect to inducer concentration is noncooperative.[11] In accordance with the MWC model, inducer binding to the repressor drives the repressor equilibrium toward the inactive R* conformation and, therefore, induction only requires that the inducer has a greater affinity for the inactive conformation R* than for the active conformation R (i.e., $K_{IR^*} > K_{IR}$).

The level of transcript produced in our GFP assay can be modeled using a standard ligand binding isotherm. The signal (E) is a function of the concentration of unbound repressor with respect to its operator dissociation constants (the repressor concentrations at which there would be 50% site occupancy).

$$\frac{E}{E_{max}} = \frac{1}{\left(1 + \frac{[R_{tot}]}{[R_{50}]}\right)} \tag{2}$$

In the absence of repressor, the fluorescent signal (E_{max}) is constitutively produced in this system. Since the repressor can adopt two distinct conformations, R and R*, that have different operator binding dissociation constants, then the binding isotherm takes the following form:

$$\frac{E}{E_{max}} = \frac{1}{\left(1 + \frac{[R_a]}{[R_{50}]} + \frac{[R^*_a]}{[R^*_{50}]}\right)} \tag{3}$$

The relative binding affinity is a function of the concentration of unbound repressor species in the two

forms $[R_a]$ and $[R^*_a]$ and their respective dissociation constants for operator binding ($[R_{50}] = 1/K_{RO}$ and $[R^*_{50}] = 1/K_{R^*O}$). As the concentration of the repressor in the active form increases, the fractional expression decreases. Similarly, increasing the affinity of the repressor for its operator also decreases the fractional expression.

The fractional expression described in Eq. (3) can be expressed as the product of two dimensionless ratios: the product of the ratio of the total concentration of the repressor to its operator binding dissociation constant in the active form, $[R_{tot}]/[R_{50}]$, and the fractional amount of unbound repressor in the active form relative to the total concentration of repressor, $[R_a]/[R_{tot}]$. Similarly, $[R^*_a]/[R_{tot}]$ is the fractional amount of unbound repressor in the induced form relative to the total concentration of repressor, such that Eq. (3) takes the following form:

$$\frac{E}{E_{max}} = \frac{1}{\left(1 + \frac{[R_{tot}]}{[R_{50}]}\frac{[R_a]}{[R_{tot}]} + \frac{[R_{tot}]}{[R^*_{50}]}\frac{[R^*_a]}{[R_{tot}]}\right)} \tag{4}$$

We defined a variable f as $[R_a]/[R_{tot}]$, which depends upon both the inducer concentration and the conformational equilibrium constant, and another variable r as the $[R_{tot}]/[R_{50}]$ ratio. Since the amount of repressor species bound to operator is a small fraction of the unbound species, we can assume that $f^* = 1 - f$ and that the ratio of the operator affinities for the two repressor forms for operator is simply $K_{R^*O}/K_{RO} = s$. This allows us to express the relative binding affinity in terms of three variables, r, f, and s, where $e = E/E_{max}$.

$$e = \frac{1}{(1 + r(f + s(1 - f)))} \tag{5}$$

The linked equilibria that describe the binding of the repressor to its operator and to the inducer are illustrated in Fig. 3. There are 12 repressor species that contribute to the total repressor concentration.

$$\begin{aligned}[R_{tot}] = {} & [R] + 2[RI] + [R^*] + 2[R^*I] + \{[RI_2] + [R^*I_2]\} \\ & + [RO] + 2[RIO] + \{[RI_2O]\} + [R^*O] \\ & + 2[R^*IO] + \{[R^*I_2O]\}\end{aligned} \tag{6a}$$

When there are many more copies of the repressor than operators, the species bound to the operator (RO, RIO, RI_2O, R*O, R*IO, and R^*I_2O in Eq. (6)) represent a small fraction of the total repressor concentration (an approximation of repressor and operator copy numbers was derived from plasmid construction; see Materials and Methods) and can be neglected. This simplifies the subsequent equations and permits a closed-form solution. The concentration of unbound repressor that exists in the R and R* conformations is the total of all unbound repressor species $[R_a]$ and $[R^*_a]$, respectively:

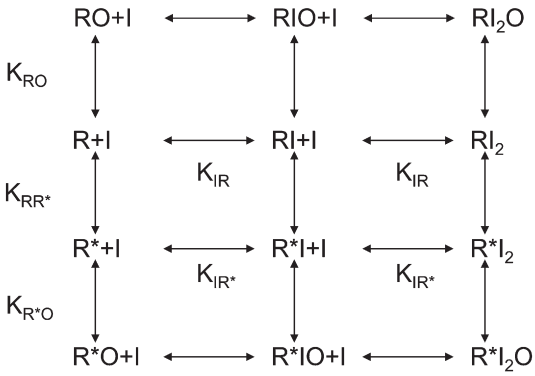

Fig. 3. A diagram illustrating the linked equilibria that result from inducer binding and operator binding. K_{RR^*} is the equilibrium constant that describes the induced and the repressed states. K_{IR} and K_{IR^*} refer to the binding of the inducer to these two conformational states. K_{RO} is the equilibrium constant that describes the operator binding.

$$[R_a] = [R] + \{2\}[RI] + \{[RI_2]\} \tag{6b}$$

$$[R^*_a] = [R^*] + \{2\}[R^*I] + \{[R^*I_2]\} \tag{6c}$$

For the heterodimeric repressor with two inducer binding sites, there are six unbound species in equilibrium, namely R, RI, RI_2, R*, R*I, and R^*I_2 (the two-site case). When one inducer binding site is removed, the terms contained in curly brackets in Eq. (6a) to Eq. (6c) disappear, and there are only four unbound species: R, RI, R*, and R*I (the one-site case). In both instances, the amount of repressor [R] that can bind to the operator depends upon the value of the conformational equilibrium constant K_{RR^*}, the inducer concentration [I], and relative inducer binding affinities to the R and R* conformations. These three parameters are responsible for establishing the relative amounts of active and inactive repressors, R and R*, as described below. When the heterodimeric repressor has a single inducer binding pocket, the amount of the various repressor/inducer species is governed by the linked equilibria on the left half of Fig. 3. The concentrations of unbound repressor in the active and inactive forms are the sum of the first two terms in Eq. (6a) and Eq. (6b), respectively, while the total repressor concentration is the sum of the first four terms in Eq. (6c). When both ligand pockets are functional, the linked equilibria include the right half of Fig. 3 and the total repressor concentrations include the curly bracket terms in Eq. (6a) to Eq. (6c).

The concentrations of all repressor species can be evaluated in terms of the concentration of one species, namely the free active repressor $[R_a]$, using the appropriate equilibrium constants and the fraction of repressor in the unbound R form, $f = [R_a]/[R_{tot}]$, which is given by the ratio of two binding polynomials. For the heterodimeric repressor that

has two functional inducer binding pockets, the ratio of two binding polynomials is given by:

$$f_2 = \frac{[R_a]}{[R_{tot}]} = \frac{(1+[I]K_{IR})^2}{(1+[I]K_{IR})^2 + K_{RR^*}(1+[I]K_{IR^*})^2} \quad (7a)$$

We note that, as in the original MWC model, the presence of factors having the form $(1+[L]K)^n$ results from independent ligand binding at n identical sites, and it properly accounts for the statistical factor arising from the macroscopic indistinguishability of distinct microscopic species IR, RI, etc. Similarly, the fraction of repressor in the R conformation when there is a single inducer binding site can be described by the ratio of corresponding polynomials:

$$f_1 = \frac{[R_a]}{[R_{tot}]} = \frac{(1+[I]K_{IR})}{(1+[I]K_{IR}) + K_{RR^*}(1+[I]K_{IR^*})} \quad (7b)$$

When both pockets are eliminated or in the absence of inducer, the fraction of the repressor in the R conformation reduces to:

$$f_0 = \frac{1}{1+K_{RR^*}} \quad (7c)$$

At saturating concentrations of inducer, the fraction of free repressor depends only upon the conformational equilibrium, K_{RR^*}, and the relative binding of the inducer to the repressor in the R and R* conformations, as described by Eq. (8a) and Eq. (8b):

$$f_1' = \lim_{[I]\to\infty} f_1(I) = \frac{1}{1+K_{RR^*}\frac{K_{IR^*}}{K_{IR}}} \quad (8a)$$

$$f_2' = \lim_{[I]\to\infty} f_2(I) = \frac{1}{1+K_{RR^*}\left\{\frac{K_{IR^*}}{K_{IR}}\right\}^2} \quad (8b)$$

By substituting Eq. (7) and Eq. (8) for each of the 0, 1 and 2 site cases into Eq. (5), we can relate the relative amount of transcript produced at limiting conditions to four parameters: (a) the ratio of total repressor concentration to the active form operator binding dissociation constant, r; (b) the ratio of inducer binding to the inactive and active repressor conformations, $x = \frac{K_{IR^*}}{K_{IR}}$; (c) the equilibrium between these two different conformations, K_{RR^*}; and (d) the ratio of operator affinities of the two repressor forms, s:

$$e_n = 1/\left(1 + r\left(1/1 + K_{RR^*}x^n\right) + rs\left(1 - 1/\left(1 + K_{RR^*}x^n\right)\right)\right) \quad (9)$$

The three values of e_i, for $n=0$, 1, and 2, are the experimentally observed levels of GFP expression with zero inducer (or with zero sites) at a high inducer concentration for the one-site repressor and a high inducer concentration for the two-site repressor, respectively. Rearranging Eq. (9), we define the variable y_i as:

$$y_n = 1/e_n - 1 = r(1/(1+K_{RR^*}x^n) + s(1 - 1/(1+K_{RR^*}x^n))) \quad (10)$$

By evaluating the ratio of the one-site and two-site cases to the zero inducer case, we can eliminate the variable r and obtain two equations in the three variables s, x, and K_{RR^*}.

$$y_n/y_o = \left(1/\left(1+K_{RR^*}x^n\right) + s\left(1 - 1/\left(1+K_{RR^*}x^n\right)\right)\right) / \left(1/\left(1+K_{RR^*}\right) + s\left(1 - 1/\left(1+K_{RR^*}\right)\right)\right) \quad (11)$$

where $n=1$ or 2.

The two parameters x and K_{RR*} were first fit to the experimentally observed values y_1 and y_2 on the left-hand side of Eq. (11) by a straightforward grid search assuming there is no binding of the inactive repressor (R*) to the operator (O) (i.e., $s=0$). We then solved for r using Eq. (10). Subsequently, the absolute value of the inducer binding constant, K_{IR^*}, was determined by fitting the expression level curves *versus* inducer concentration for both the one-site and two-site dimers using the parameter values x, r, and K_{RR^*}. The value of K_{IR^*} $[1/K_{d(IR^*)}]$ was refined by bisection to minimize the mean absolute error in fractional expression. The other inducer binding constant was then determined as $K_{IR} = xK_{IR^*}$.

Analysis of the extracted parameters

The value of K_{RR^*} was determined to be 2 ± 0.5, which is in good agreement with previous *in vitro* measurements[12] and consistent with our observation that the repressor prefers to crystallize in the induced R* conformation in the absence of the ligand, regardless of the crystallization conditions.[7] The ratio of the inducer binding constants, $x = K_{IR^*}/K_{IR} = 15\pm3$, illustrates that the inducer preferentially binds to the R* conformation. The *in vivo* ratio of equilibrium constants is however somewhat less than that observed by O'Gorman *et al*. *in vitro*.[12] The calculated dissociation constant for active repressor–inducer binding, $K_{d(IR)}$, is 4 ± 2 μM. Since the ratio of the two binding constants, K_{IR^*}/K_{IR}, is ~15, we can infer that the corresponding dissociation constant for the inducer to the inactive repressor, $K_{d(IR^*)}$, is 15×4 μM or 60 μM. Overall, the values of K_{RR^*} and x are consistent with our structural observation that the conformations of the apo repressor and the repressor–IPTG complex are essentially isomorphous.[7]

The ratio of the total repressor concentration to its operator dissociation constant, $[R_{tot}]/[R_{50}]$, is a complex quantity that reflects the binding affinity of the repressor for the operator and the concentrations of the repressor and operator in the cell. We found that this ratio is 150 ± 50. This implies that there is a vast excess of repressor compared with that required for 50% operator binding, which is

consistent with our *in vivo* assay system (see Materials and Methods). The model also predicts that the switch is leaky. In the absence of inducer, only a fraction, $f_0 = 1/(1 + K_{RR^*})$, of the repressor molecules adopt the R conformation, leading to a net active repressor concentration of $f_0 = 1/(1+2)$, which is "only" ~50-fold above the repressor–operator dissociation constant. Although the repressor is operating under saturating conditions and the vast majority of the operators are bound by the repressor, a small amount of transcript, roughly 2%–3%, would be produced since $E/E_{max} = 1/(1+50)$. In the absence of inducer, a detectable level of background expression is observed, consistent with the values calculated from the model.

According to the MWC model, the contribution the inactive conformation makes to the binding of the repressor to the operator is negligible. The previously discussed parameters were determined under this assumption that $s = 0$; however, to test its validity, we repeated the above fitting procedure, allowing the induced conformation R* to bind O with a range of affinities ($0 \le s \le 1$). The resulting parameters are tabulated in Table 2. When the R* conformation binds the operator with an affinity of 1/10,000 or less than the active form R, the model produces quality fits very similar to the limiting expression and IPTG titration curves, with almost identical output parameters. Assuming a somewhat higher relative affinity of 1/1000 produces a slightly poorer fit to the limiting expression data but an equally good fit to the whole induction curve. At a higher relative affinity of $s = 1/100$, the fit is poor. Essentially, when the relative affinity of R* for O binding is large, the inducer cannot pull the repressor off the operator and, as a consequence, the system is not inducible. Overall, the best-fit parameters are slightly different from those when we assume $s = 0$; notably, the conformational equilibrium is somewhat closer to unity, but the differences fall within the estimated uncertainty ranges and the same qualitative picture emerges: there is a small but significant preference for the inactive form of repressor in the absence of inducer, a large excess of repressor relative to its K_d for operator binding, and binding of inducer to the inactive repressor form that is tighter by 1 order of magnitude. We therefore conclude that the s value must be on the order of 1/1000 or less to explain the *in vivo* data. This allows us to effectively ignore R* binding to operator.

The quality of the model and the accuracy of the parameters can be assessed by comparing the calculated and observed induction curves (Fig. 4). The experimental data fit the calculated model very well, yielding a mean unsigned error of 2%. The best-fit model slightly overestimates the slope of expression *versus* I at the midpoint. This could result from systematic overestimation of the effective IPTG activity *in vivo*—that is, the actual activity of IPTG is somewhat lower than the nominal concentration values used in plotting the data. Another measure of the validity of the derived constants was assessed by determining the ratio of the $[I]_{50}$ values for the one-site to two-site cases. A calculated value of 2.5 agrees well with the experimentally observed value of ≈3. The agreement between the observed and calculated ratios is a good measure of self-consistency since this ratio is not an adjustable parameter. In the fitting to the model (although not in the model itself), we assumed that the concentration of repressor was much greater than that of the operator such that the amount of repressor bound to operator is small compared with the total and free amounts of repressor. Estimates of the repressor copy number are imprecise, but the ratio of repressor to operator in a similar system was estimated to be 100:1.[13] This ratio is sufficiently large such that the fractional error that results from neglecting the concentration of the bound repressor species in Eq. (6a) is 1/100, less than a few percentages. This permits us to obtain closed-form expressions for limiting expression levels and a straightforward physical interpretation in terms of the basic model parameters with little numerical error. A more exact model may be applied for low repressor copy number by keeping all the terms in Eq. (6a) and fitting iteratively, assuming that the total repressor remains constant at each step.

The expression of GFP in the reporter system depends upon the inducer concentration and its relative affinity for the R* and R conformations of the repressor, which we estimate to be 4 and 60 μM, respectively. This difference in binding affinities

Table 2. Fitted model parameters

s	x	K_{RR^*}	r	$K_{d(IR^*)}$ (μM)[a]	Fit error for (limiting expression values (%)[b]	Fit error for entire induction curves (%)[b]	$K_{d(RI)}$ (μM)[c]
0	15±3	2.3±0.5	150±50	4±2	4	1.9	60
1/100,000	15±3	2.3±0.5	170±50	4±2	4	1.9	60
1/10,000	15±3	2.4±0.5	170±50	4±2	4	1.8	60
1/1000	19±5	1.3±0.5	106±30	4±2	6	2.1	79
1/100	1160±600	0.01±0.005	41±25	0.3–1.1	13	2.7	812

Uncertainty limits were determined from parameter ranges that double the best-fit error.

[a] Defined as $K_{d(IR^*)} = 1/K_{IR^*}$.

[b] Percentage of error fitting limiting expression values e_0, e_1, and e_2 with the parameters x, K_{RR^*}, and r and percentage of error fitting entire induction curves with the same three parameters plus K_{R^*I}.

[c] Dependent parameter fixed by $K_{d(IR^*)}/x$.

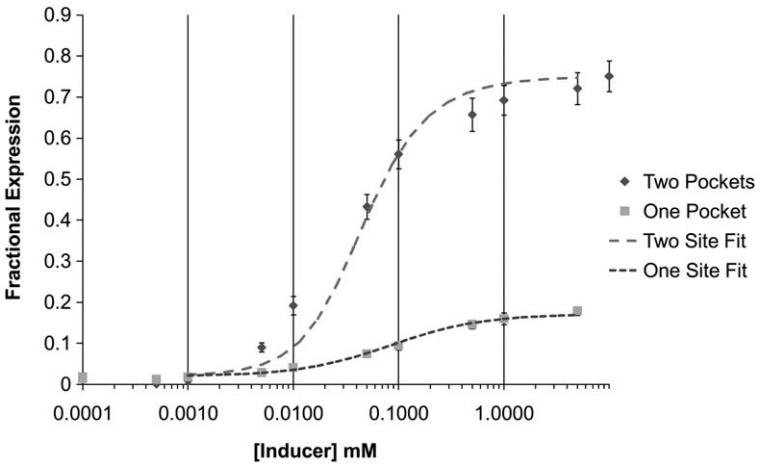

Fig. 4. The graph is a plot of the fractional induction as a function of the log of the inducer concentration for the heterodimeric repressors with a single binding pocket and two intact pockets. The curves are the expected values calculated from the model.

results in curves that are slightly sigmoidal when expression/repression values are plotted against inducer concentration. The midpoint, somewhere between the two individual dissociation constant values, is ~25 μM, and the plot has a Hill coefficient of ~1.25 (Fig. S3), consistent with numerous studies.[14–16] This mild cooperativity arises from the inducer-promoted shift of repressor conformational equilibrium toward the higher-affinity form. However, since these curves are all characterized by a single inflection point, without further independent data, only a single "apparent" inducer affinity constant is extractable.

In the above model of heterodimeric repressor action, we have assumed that the binding pocket mutation R197G/A in one of the monomers acts by blocking binding of inducer to one site and that this is responsible for the drop in maximal expression from 75% to 18%. Inducibility of the repressor would also be reduced if the mutation caused the conformational equilibrium to shift toward the R form (i.e., K_{RR^*} decreases). While it is unlikely that a binding pocket mutation would exert its effect through K_{RR^*} rather than K_{IR} and K_{IR^*}, we can consider this possibility in our model. Using the two-site Eq. (5a) rather than Eq. (5b) to fit the "one-site mutation" induction curve and adjusting only K_{RR^*}, we can model a reduced maximal expression level of 18% by setting $K_{RR^*} = 0.15$. In this model, the heterodimeric repressor with one-site R197G/A mutation still binds two inducers with unchanged affinity, but the mutation has pushed the repressor further into the active conformation. However, this model predicts that at zero inducer, there is no detectable expression (i.e., the switch is not leaky), which is inconsistent with the experimental data.

Discussion

A thermodynamic perspective

The *lac* repressor is a model system for exploring gene regulation and the mechanism by which effector molecules regulate transcription. Measuring induction using the heterodimeric repressors with zero, one, and two inducer pockets allowed us to unambiguously extract three important thermodynamic quantities: the equilibrium constant between the induced and the repressed conformations of the repressor K_{RR^*}, the relative affinities of each conformation for inducer (K_{IR^*}/K_{IR}), and the amount of repressor in the cell relative to its operator binding dissociation constant. In addition, determining these parameters allowed us to begin exploring how each aspect of the repressor contributes to its functionality by modeling theoretical changes in each parameter.

From modeling the *in vivo* data, it appears that effective repression in the absence of inducer (<4% expression) requires a large excess of active repressor with respect to its operator binding constant (50-fold). While a background expression of 4% may be less than ideal, given the other thermodynamic parameters of the repressor, an r value of 150 appears to be optimal. To better understand the relationship between the ratio of active repressor concentration and operator affinity, we generated a series of simulated induction curves by varying the $[R_{tot}]/[R_{50}]$ ratio (Fig. 5a). By increasing the ratio to 500 or 1000, which can be achieved by increasing either the active repressor concentration or the repressor–operator affinity, constitutive expression can be minimized to less than 1%. However, increases in the ratio also affect the dynamic range of the switch. Increased repression comes at the expense of induction, and the inducer cannot relieve repression to more than 50% maximal expression. Conversely, if the ratio is too small, then the switch can be induced but background expression is very high. When the ratio of r is 150, the background expression appears to be minimized while balancing maximal induction near 75%. Experimentally, both the one-site and the two-site inducer binding cases show a plateau in expression level at transcriptional levels less than 100%, suggesting that there is always a small fraction of the repressor that can adopt the active (R) conformation and block transcription.

The equilibrium distribution of repressors in both the active conformation and the inactive conformation is an additional parameter that controls the repression and induction profile of the repressor. From modeling

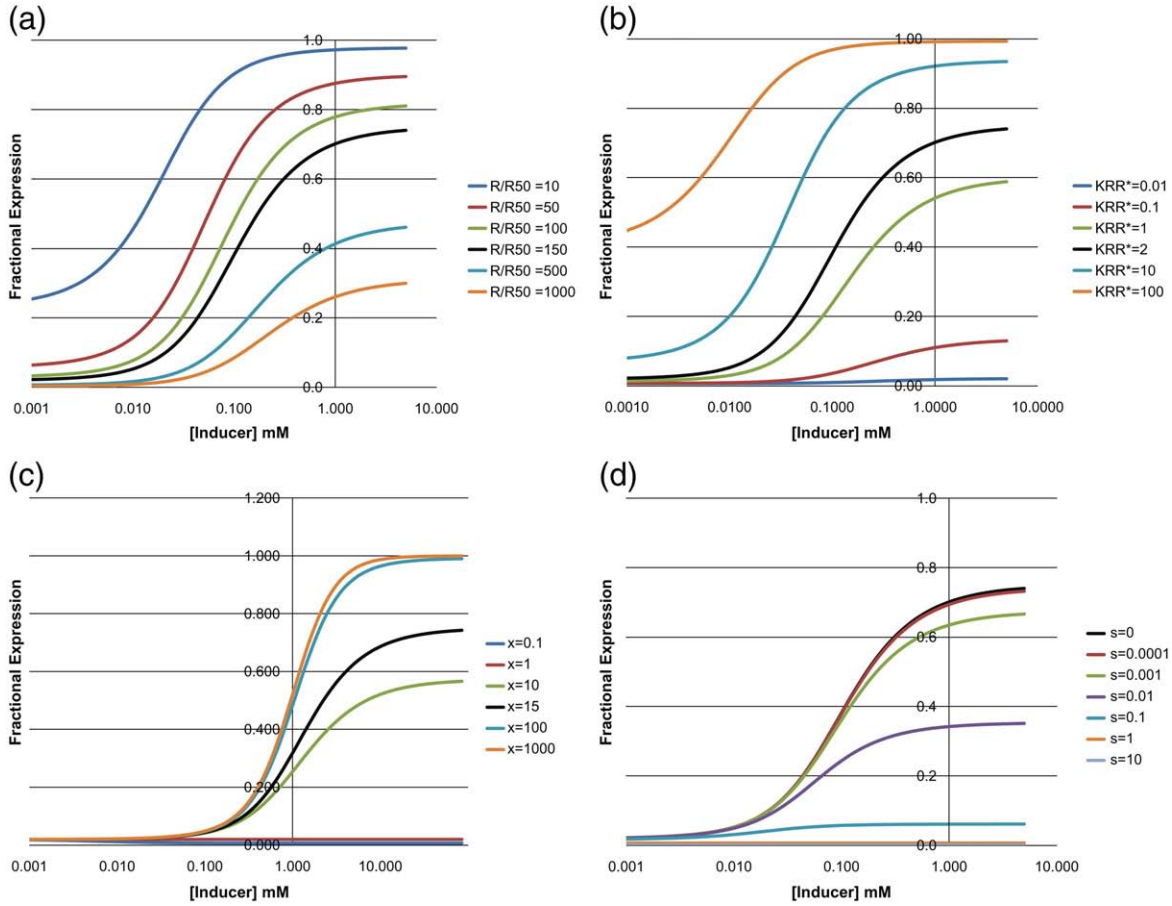

Fig. 5. Simulated plots of fractional induction with respect to inducer concentration (plotted on a log axis). With the use of the determined thermodynamic parameters for the repressor, theoretical plots were simulated for stepwise changes in (a) r, the ratio of $[R_{TOT}]/[R_{50}]$; (b) K_{RR^*}, the conformational equilibrium (L); (c) x, the ratio of inducer binding affinities for both the active and the inactive repressor conformations; and (d) s, the ratio of operator binding affinities for each of the repressor conformations. The plots shown in black correspond to the value derived from the experimental data.

the experimental data, there appears to be a rather small energy difference between the two conformations of the repressor (R and R*), $\Delta G = kT\ln(K_{RR^*}) = 0.4$ kcal/mol. This modest difference in energy allows the repressor to easily switch between active and inactive states. Again, we simulated induction curves by altering the equilibrium constant (designated as "L" according to the MWC nomenclature) to explore the impact of this parameter on repressor activity (Fig. 5b). Since the active conformation of the repressor has preferential binding to the operator, when the apparent equilibrium constant is less than 1 and the active conformation of the repressor is dominant, repression is strong and there is minimal background expression. However, altering K_{RR^*} to increase the concentration of active repressor also increases the ratio r discussed previously. This limits the dynamic range of expression and permits only a fraction of the possible induction. In contrast, when the equilibrium constant is large, the majority of the repressor is in the R* conformation, resulting in a switch that is leaky but fully inducible. At the fixed values for the other thermodynamic properties, the dynamic range of the switch is optimal when L (K_{RR^*}) is between 1 and 10. There is a low level of basal expression in the absence of inducer and a nearly full induction at high concentrations of inducer.

Since repressor activity is a product of linked equilibria, the distribution of active and inactive repressors can be altered by adding either operator or inducer ligands. The system is repressed when there is an excess of active repressor with respect to its operator dissociation constant. Induction, on the other hand, requires an excess of inactive repressor. The inducer relieves repression by shifting the apparent conformational equilibrium, decreasing the concentration of active repressor. A single inducer binds 15 times more tightly to the R* conformation than the R conformation, which in energetic terms is about 1.6 kcal/mol. In the presence of a single bound inducer, the apparent equilibrium constant increases, $K'_{RR^*} = \exp^{(0.4+1.6)/kY} \approx 30$, stabilizing the R* conformation; the $[R_a]/[R_{50}]$ ratio drops by ~5-fold (150/30). The additional energy shifts the apparent equilibrium further toward the R* conformation, thereby reducing the effective concentration of the repressor in the R conformation. Since there is less active repressor, polymerase has a

greater opportunity to transcribe its message. Therefore, in the presence of high concentrations of inducer, the single-site repressor can achieve a fractional induction level of $1/(1+5)=16\%$, which is consistent with our experimental observation. A second inducer molecule shifts the apparent equilibrium more dramatically. An additional 1.6 kcal/mol of binding energy increases K_{RR^*} to ~450, driving the repressor further toward the R* conformation and lowering the $[R_a]/[R_{50}]$ ratio to less than unity, ~0.4. As a consequence, in the presence of high levels of inducer, we expect $1/(1+0.33)=75\%$ induction, which is very close to the experimental observation. Since the calculated energetic effects appear to be purely additive, we would suspect cooperativity to be of little importance when it comes to induction. In fact, the Hill coefficient (at the inducer midpoint) is a rather modest 1.25 and concludes that cooperativity is not an essential requirement of the switch. Cooperativity is important for shifting the window in which an effector can elicit a response to some relatively narrow range of nonzero concentrations. For example, the cooperativity seen with hemoglobin is probably due to the fact that the partial pressures of O_2 are never zero and may only range from 40 to 80 mmHg. In contrast, in noncooperative (Langmuir isotherm type) binding, the rate of increase is greatest starting from a zero effector, which works better as a switch in the *lac* system given the actual inducer concentration range.

In the previous examples, the experimentally determined value of $x=15$ was used to model induction as a consequence of altering the other parameters. Since the repressor's response to inducer is dependent upon the ratio of relative affinities of the inducer for each conformation of the repressor, simulated induction curves were also produced with variations in x (Fig. 5c). For $x=1$, the inducer binds to both the active and the inactive conformations with the same affinity and expression levels become insensitive to inducer. When the ratio is less than unity, the inducer binds more tightly to the active conformation and actually decreases induction. The inducer ligand would then function more like a corepressor. Only when the ratio is significantly greater than 1 does the system work as a true switch—having full repression at low inducer concentrations and complete induction at high concentrations of inducer. Unlike the other parameters that alter both repression and induction, no such balance exists for the value of x and the switch functions even better as x increases to greater values. The measurements and analysis described were made using the gratuitous inducer IPTG; it would be interesting to see if the value for x, and thus the maximal expression level, is higher when the natural inducer, allolactose, is utilized. Regardless, this analysis suggests that induction of the *lac* repressor with IPTG can be improved by finding mutations that increase the ratio of binding affinities beyond 15.

As the MWC model suggests, the repressor adopts two distinct conformations, but only one of the conformations is functional. Therefore, the switch also depends on the relative operator binding of the two repressor conformations. To explore the importance of this parameter on repressor activity, we examined how variations in s alter the induction curves (Fig. 5d). When the value of s is greater than unity, the inactive conformation of the repressor binds to the operator and the inducer again behaves as a corepressor by increasing repression. At unity, both conformations have the same affinity for the operator and inducer has no effect. The switch will function only when the value of s is $\ll 1$. As s decreases, there is greater induction that quickly approaches the limit as $s=0$. Based upon these curves, when s is less than 0.0005, the system functions as if $s=0$.

A kinetic perspective

The thermodynamic description of the switch is extremely useful for understanding the molecular mechanism of gene regulation. However, to fully appreciate how this switch functions, we must consider its activities in the context of its kinetic properties. The binding of the *lac* repressor to its operator has been well studied, and all of the data are consistent with a bimolecular reaction (Eq. (12)) having an association constant, K_{RO}, of 1.5×10^{10} M^{-1} under standard conditions.[11,15] The equilibrium constant, of course, is the ratio of the association and disassociation rate constants:

$$K_{RO} = [RO]/[R][O] = K_a/K_d \qquad (12)$$

The dissociation constant has been measured experimentally for the wild-type repressor and its natural operator, as $K_d=6\times10^{-4}$ s^{-1}; using the above value of K_{RO}, this gives $K_a=7.5\times10^6$ M^{-1} s^{-1}.[15,17,18] A considerably greater association rate has been measured directly by Riggs *et al.*, who found a value of 7×10^9 M^{-1} s^{-1}.[15] In general, the decrease by 3 orders of magnitude in affinity that results from inducer binding can arise from a decrease in the on-rate, an increase in the off-rate, or any combination of the two. Changes in the off-rate have been estimated to range from 6- to 18-fold at inducer concentrations of 10 μM, well above the inducer midpoint.[15,16] This increase, while significant, implies that a large portion of the affinity loss is due to a decrease in the on-rate. We therefore deduce that the inducer stabilizes the R* conformation of the repressor by decreasing the on-rate conservatively by 10-fold and, arguably, by 100-fold or more. It is instructive to speculate how altering these kinetic parameters can affect gene regulation and whether modulation of affinity through the association rate rather than the dissociation rate has any functional significance.

On average, the fraction of the operator that is free of repressor is $f_0=1/(1+[R]K_{RO})$ for a given repressor concentration [R]. Since operator binding is actually a stochastic process, at any instance, the operator is either occupied or not and there will be a distribution of occupation times. From the theory of stochastic processes,[19] the time that the operator is

occupied by repressor is exponentially distributed with a mean of $\tau_{on}=1/K_d$. Similarly, the mean time the operator is free of repressor also varies exponentially, with a mean of $\tau_{off}=1/K_a[R]$. Based upon the data of Riggs *et al.*,[15] we can estimate that in the absence of inducer, the mean occupancy time is 1600 s. With an estimated repressor concentration of 20 nM (O'Gorman *et al.*[12]), the mean time the operator is free of repressor in the absence of inducer is 6 s using the lower estimate of K_a and is 7 ms using the upper estimate.

Transcription initiation takes a finite amount of time. Consequently, the fraction of time the operator is free and the duration of time it is free are both relevant for effective transcription. We hypothesize that the promoter must be free of repressor long enough that the transcriptional machinery can progress to the point where it is not affected by repressor binding. The kinetics of transcriptional initiation at the *lac* promoter has been well studied.[20] RNA polymerase binds to the promoter region, forming a "closed" complex, which then undergoes an "isomerization" to form the open (strand separated) initiating complex. Assuming the isomerization step, which has a first-order rate constant of 1.6×10^{-2} s^{-1}, approximates the time to initiate transcription, and then the promoter must be free from the repressor for roughly 60 s. This is considerably longer than the mean time it is free under high repression conditions. Now if the effect of the inducer were solely on the dissociation rate of the repressor, the mean occupancy time would be significantly reduced, but the mean time the operator is free of repressor would be unchanged; therefore, on average, the time frame is still too short for transcription to initiate. If the inducer affected repressor affinity by depressing the on-rate, then the average time the promoter is free of repressor would increase, allowing transcription to initiate.

Conclusions

Elucidating the thermodynamic properties of the molecular switch is essential for developing a complete understanding of gene regulation. The MWC model for the allosteric transition applies beautifully to the *lac* molecular switch and accounts for the conformations of the repressor observed in the crystal structures. Here, we have experimentally determined the equilibrium constants that account for the observed properties of the molecular switch. Our analysis suggests that the transition from the induced conformation to the repressed conformation requires a relatively small amount of energy (0.4 kcal/mol) and only subtly favors the induced conformation. This soft equilibrium allows the activity of the repressor to be modified by adjusting the conformational equilibrium through the mechanisms of linked equilibria. Since both the inducer and the operator preferentially bind to different conformations, binding of inducer provides energy to compete with operator binding. By increasing the effective concentration of the inactive repressor, inducer binding reduces the number of repressors in the active conformation and therefore reduces repression. Our analysis demonstrates that a single inducer provides enough energy to increase expression levels to 20% and that two inducers are required for maximal induction. In addition, from altering various parameters of the repressor and monitoring the effect on repressor activity, we have gained valuable insight into designing improved switches. Tuning the allosteric properties of the repressor will allow us to create better and novel molecular switches.

Materials and Methods

Repressor plasmid construction

Constructions of the dimeric wild-type repressor plasmid (pBD21004), the heterodimeric plasmid (pBD22010), and the plasmids containing each of the constituent monomers (pBD21008 and pBD21903) have been described previously.[9] To introduce the R197A and R197G mutations into the pBD21004 and pBD21008 plasmids, we used full-circle PCR mutagenesis with primers BD030 (5′ GCT CTG GCT GGC TGG CAT AAA TAT C 3′) and BD031 (5′ CAG ACG CGC CGA GAC AGA AC 3′) and primers BD032 (5′ GGT CTG GCT GGC TGG CAT AAA TAT C 3′) and BD031, respectively. After introducing the point mutations into each of these repressor genes, we reintroduced the two monomeric repressor genes into the heterodimeric plasmid. In short, the pBD21903 plasmid was digested with DrdI and the repressor gene was isolated via gel purification. Subsequently, the vector containing the recently mutated R197A or R197G gene (pBD21012, pBD21016) was digested with DrdI and purified. These vectors (also containing the Y282S monomeric mutation) were then ligated with the insert containing the complementary repressor gene, creating plasmids pBD22011 and pBD22012. The repressor plasmid is derived from the pACYC vector with a copy number around 15. Expression of *lacI* is under the control of the constitutive *lac* promoter. Estimates of repressor copy number are comparable with those of Oehler *et al.*[13]

Reporter plasmid construction

Construction of the reporter plasmid with GFPmut3.1 under the control of the *lac* promoter/operator has been described previously.[9] In short, GFPmut3.1 was introduced into the pBR322 vector and placed under control of the *lac* promoter/operator. To introduce operator changes producing chimeric operators, we used full-circle PCR mutagenesis (operator 212:411, primers BD024 and BD021; operator 411:212, primers BD020 and BD025). To explore the role of directionality on induction with repressor mutants containing asymmetric ligand pocket knockouts, we created and utilized two reporters. Each of these reporters contained a chimeric operator sequence with the orientation of the half sites inverted with respect to each other. In assays using the wild-type (tetrameric or dimeric) repressor, the reporter containing the natural operator was used.

In vivo repression/induction assay

To analyze the phenotypes of various repressor–operator combinations, we used an *in vivo* fluorescent assay. To quantify the level of fluorescence and therefore indirectly measure the degree of transcription, we grew then analyzed cells in a Perkin Elmer Victor3 plate reader. In short, combinations of repressors and operators were transformed and colonies were selected in triplicate for overnight culture growth. In addition, cells controlling the reporter only were also chosen to establish the level of maximal expression under nonrepressing conditions. For induction analysis, samples were grown in the absence of IPTG, as well as in the presence of various amounts of inducer. Once samples reached saturation, 200-μl aliquots were taken and introduced into flat-bottom 96-well plates. A dilution plate was also prepared so that the optical density of the cultures could more accurately be determined. Each of these plates was then measured for GFPmut3.1 fluorescence (495-nm excitation wavelength and 510-nm emission wavelength) and optical density (A_{590}) on a Perkin Elmer Victor3 Plate reader. Fluorescence data were then processed to remove background fluorescence and normalize the GFP signal to the cell count. The normalized signals for biological replicates were then averaged to provide a relative fluorescent signal for each sample. Errors for each sample were determined from the standard deviation of the biological replicates. Fractional expression levels were calculated by dividing the average GFP signal for the sample with a given repressor by the signal of the sample containing no functional repressor.

Supplementary Data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jmb.2009.07.050

References

1. Wilson, C. J., Zhan, H., Swint-Kruse, L. & Matthews, K. S. (2007). The lactose repressor system: paradigms for regulation, allosteric behavior and protein folding. *Cell. Mol. Life Sci.* **64**, 3–16.
2. Jacob, F. & Monod, J. (1961). Genetic regulatory mechanisms in the synthesis of proteins. *J. Mol. Biol.* **3**, 318–356.
3. Barkley, M. D. & Bourgeois, S. (1980). Repressor recognition of operator and effectors. In (Miller, J. H. & Reznikoff, W. S., eds), pp. 177–220, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
4. Monod, J., Wyman, J. & Changeux, J.-P. (1965). On the nature of allosteric transitions: a plausible model. *J. Mol. Biol.* **12**, 88–118.
5. Markiewicz, P., Kleina, L. G., Cruz, C., Ehret, S. & Miller, J. H. (1994). Genetic studies of the *lac* repressor: XIV. Analysis of 4000 altered *Escherichia coli lac* repressors reveals essential and non-essential residues, as well as "spacers" which do not require a specific sequence. *J. Mol. Biol.* **240**, 421–433.
6. Pace, H. C., Kercher, M. A., Lu, P., Markiewicz, P., Miller, J. H., Chang, G. & Lewis, M. (1997). *Lac* repressor genetic map in real space. *Trends Biochem. Sci.* **22**, 334–339.
7. Daber, R., Stayrook, S., Rosenberg, A. & Lewis, M. (2007). Structural analysis of *lac* repressor bound to allosteric effectors. *J. Mol. Biol.* **370**, 609–619.
8. Spotts, R. O., Chakerian, A. E. & Matthews, K. S. (1991). Arginine 197 of *lac* repressor contributes significant energy to inducer binding. Confirmation of homology to periplasmic sugar binding proteins. *J. Biol. Chem.* **266**, 22998–23002.
9. Daber, R. & Lewis, M. (2009). A novel molecular switch. *J. Mol. Biol.* In press. doi:10.1016/j/jmb.2009.06.039.
10. Taraban, M., Zhan, H., Whitten, A. E., Langley, D. B., Matthews, K. S., Swint-Kruse, L. & Trewhella, J. (2008). Ligand-induced conformational changes and conformational dynamics in the solution structure of the lactose repressor protein. *J. Mol. Biol.* **376**, 466–481.
11. Barkley, M. D., Riggs, A. D., Jobe, A. & Burgeois, S. (1975). Interaction of effecting ligands with *lac* repressor and repressor–operator complex. *Biochemistry*, **14**, 1700–1712.
12. O'Gorman, R. B., Rosenberg, J. M., Kallai, O. B., Dickerson, R. E., Itakura, K., Riggs, A. D. & Matthews, K. S. (1980). Equilibrium binding of inducer to *lac* repressor.operator DNA complex. *J. Biol. Chem.* **255**, 10107–10114.
13. Oehler, S., Amouyal, M., Kolkhof, P., von Wilcken-Bergmann, B. & Muller-Hill, B. (1994). Quality and position of the three *lac* operators of *E. coli* define efficiency of repression. *EMBO J.* **13**, 3348–3355.
14. Daly, T. J. & Matthews, K. S. (1986). Allosteric regulation of inducer and operator binding to the lactose repressor. *Biochemistry*, **25**, 5479–5484.
15. Riggs, A. D., Bourgeois, S. & Cohn, M. (1970). The *lac* repressor–operator interaction: 3. Kinetic studies. *J. Mol. Biol.* **53**, 401–417.
16. Swint-Kruse, L., Zhan, H. & Matthews, K. S. (2005). Integrated insights from simulation, experiment, and mutational analysis yield new details of LacI function. *Biochemistry*, **44**, 11201–11213.
17. Riggs, A. D., Newby, R. F. & Bourgeois, S. (1970). *lac* repressor–operator interaction: II. Effect of galactosides and other ligands. *J. Mol. Biol.* **51**, 303–314.
18. Riggs, A. D., Suzuki, H. & Bourgeois, S. (1970). *lac* repressor–operator interaction: I. Equilibrium studies. *J. Mol. Biol.* **48**, 67–83.
19. van Kampen, N. G. (1992). *Stochastic Processes in Physics and Chemistry. Revised and Enlarged Edition.* North-Holland Personal Library, North-Holland, Amsterdam, The Netherlands.
20. Schlax, P. J., Capp, M. W. & Record, M. T., Jr (1995). Inhibition of transcription initiation by *lac* repressor. *J. Mol. Biol.* **245**, 331–350.