doi:10.1016/j.jmb.2011.03.057 J. Mol. Biol. (2011) 409, 76–87

Contents lists available at www.sciencedirect.com

Journal of Molecular Biology

journal homepage: http://ees.elsevier.com.jmb

Thermodynamic Analysis of Mutant lac Repressors

Robert Daber, Matthew A. Sochor and Mitchell Lewis*

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

Received 25 March 2011; accepted 25 March 2011
Available online
1 April 2011

Edited by M. Gottesman

Keywords:
allostery;
gene regulation;
lac repressor;
induction

The lactose (lac) repressor is an allosteric protein that can respond to environmental changes. Mutations introduced into the DNA binding domain and the effector binding pocket affect the repressor's ability to respond to its environment. We have demonstrated how the observed phenotype is a consequence of altering the thermodynamic equilibrium constants. We discuss mutant repressors, which (1) show tighter repression; (2) induce with a previously noninducing species, orthonitrophenyl-β-D-galactoside; and (3) transform an inducible switch to one that is corepressed. The ability of point mutations to change multiple thermodynamic constants, and hence drastically alter the repressor's phenotype, shows how allosteric proteins can perform a wide array of similar yet distinct functions such as that exhibited in the Lac/Gal family of bacterial repressors.

Introduction

Monod *et al.* suggested that proteins can enhance or diminish their activity as a consequence of environmental changes.[1] These proteins respond to changes in metabolite concentrations by altering their structure and are therefore allosterically regulated. In general, allosteric proteins adopt two or more distinct conformational states that have differing activities. Most allosteric proteins are oligomeric and have regulatory and active sites that are physically distinct. The regulatory site recognizes changes in the environment (the accumulation or diminution of a metabolite) and responds by transmitting an intramolecular or intermolecular signal to the active site, modifying a particular function of the protein. The molecular basis for how metabolites or effector molecules enhance or diminish a protein's activity was mathematically formalized by Monod *et al.* and is referred to as the MWC (*M*onod–*W*yman–*C*hangeux) model for concerted transitions.[2]

The lactose (lac) repressor of *Escherichia coli* is an allosteric protein that regulates the transcription of the lac operon by responding to changes in the concentration of a metabolite, lactose.[3] In the absence of the metabolite, the repressor binds to a specific site within the promoter (an operator), blocking transcription of the genes required for metabolizing lactose. The effector molecule that regulates the repressor molecule is allolactose, an analog of lactose created by a side reaction of β-galactosidase. Although the wild-type repressor is tetrameric, it is functionally a dimer of dimers where each monomer is composed of a DNA binding domain and a regulatory domain.[4] The repressor binds to a specific pseudo-symmetric operator sequence using a DNA binding domain that contains a classical helix–turn–helix motif.[5] The regulatory or 'core' domain is responsible for effector binding. In the absence of the effector, the repressor associates tightly with the operator and blocks the transcription of downstream genes. The switch that regulates the lac operon has been well characterized (for a recent review, see Wilson *et al.*[6]). Other transcriptional regulators in the Lac/Gal family function in an analogous fashion, where effector molecules function as inducers or

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

Abbreviations used: GFP, green fluorescent protein; ONPF, orthonitrophenyl-β-D-fucoside; ONPG, orthonitrophenyl-β-D-galactoside.

corepressors and alter the affinity of the repressor for its operator.[7]

The lac repressor exists in at least two distinct conformational states: one corresponding to the active form of the molecule and the other(s) corresponding to the inactive (or less active) form of the molecule:

$$R \overset{K_{RR^*}}{\Longleftrightarrow} R^* \tag{1}$$

These two conformations, R and R* (designated as *R* and *T* by MWC), are in equilibrium and are related by a constant K_{RR^*} (designated as *L* by MWC), which is an inherent characteristic of the protein.[2] The apparent K_{RR^*} is changed upon ligand binding and is modeled using a linked equilibrium (Fig. 1). The wild-type repressor binds to its operator more tightly when it assumes the R conformation than when it assumes the R* conformation, such that the operator binding affinities are related by $K_{RO} \gg K_{R^*O}$. Effector ligands also bind to both the R conformation and the R* conformation of the repressor with different intrinsic affinities. The role of the effector molecule is to shift the apparent conformational equilibrium by either increasing or decreasing the population of repressor molecules in the active or inactive conformation. The system is described as inducible when the effector ligand increases the concentration of the inactive repressor (R*), and as corepressible when the effector increases the concentration of the active repressor (R). The lac repressor is an inducible protein that binds the effector more tightly in the R* conformation ($K_{ER^*} > K_{ER}$), thereby decreasing the concentration of active species, which allows the regulated genes to be transcribed.

The ability of the lac repressor to modulate transcription can be described using a standard binding isotherm:

$$\frac{e_{[E]}}{e_{max}} = \frac{1}{\left(1 + [R_a]K_{RO} + \left[R_a^*\right]K_{R^*O}\right)} \tag{2}$$

where $[R_a]$ and $[R_a^*]$ are the concentrations of unbound repressor species in the active and inactive conformations. In the absence of repressor, RNA polymerase has access to the promoter, and the DNA is constitutively transcribed. The amount of transcript produced depends on the promoter efficiency and is defined as e_{max}. When the lac repressor binds to the operator, the quantity of transcript produced is attenuated, and the expression $e_{[E]}$ decreases. The fractional expression $\frac{e_{[E]}}{e_{max}}$ is a normalized ratio that depends on the effector concentration [E]. Increasing the inducer concentration results in a decrease in the concentration of the active repressor ($[R_a]$) and a corresponding increase in the concentration of the inactive repressor expression ($[R_a^*]$). Since the repressor binds more tightly to the operator in the R conformation such that $K_{RO} \gg K_{R^*O}$, the fractional expression increases with increasing concentration of the inducer.

As described previously,[8] the products $[R_a]K_{RO}$ and $[R_a^*]K_{R^*O}$ are complex quantities that ultimately depend on the thermodynamic equilibrium constants:

$$\frac{e_{[E]}}{e_{max}} = \frac{1}{(1 + r(f_2 + s(1 - f_2)))} \tag{3}$$

where

$$s \equiv K_{R^*O} / K_{RO} \tag{3a}$$

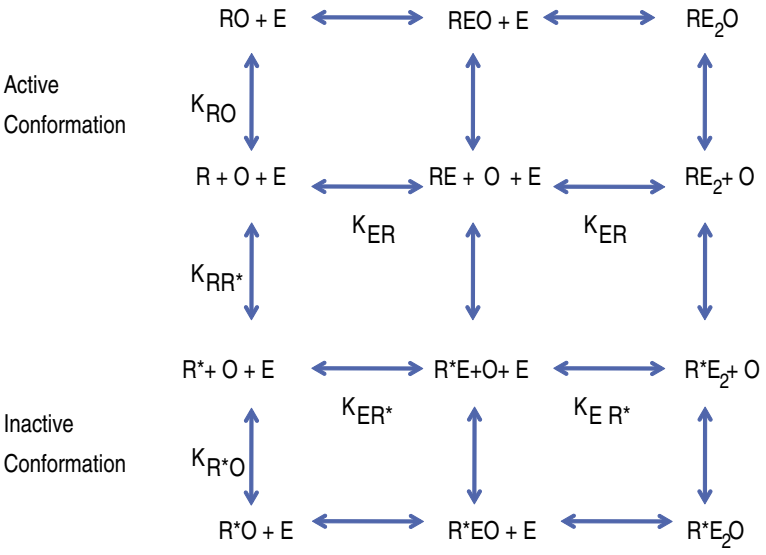

Fig. 1. The diagram illustrates the linked equilibria that result from effector (E) and operator (O) binding. R and R* correspond to the active and inactive conformations of the repressor and are in equilibrium (K_{RR^*}). Each conformation has a distinct binding affinity for both the effector ligand and the operator ligand.

$$r \equiv [R_{TOT}]K_{RO} \tag{3b}$$

$$f_2 \equiv \frac{(1 + [E]K_{ER})^2}{(1 + [E]K_{ER})^2 + K_{RR*}(1 + [E]K_{ER*})^2} \tag{3c}$$

These equations were derived assuming that the total concentration of the repressor R_{TOT} is constant and that the repressor concentration is much larger than the concentration of the operator: $[O] \ll [R_{TOT}]$. Only a fraction of the total repressor concentration can adopt the active conformation f_2, which depends on the effector concentration [E], the binding affinity of the effector molecules K_{ER} and K_{ER*}, and the conformational equilibrium K_{RR*}. The ratio of the operator binding affinities s is on the order of 1:1000 for the wild-type repressor, such that $[R_a^*]K_{R*O}$ contributes nominally to the fractional expression,[8] and Eq. (3) can be simplified to:

$$\frac{e_{[E]}}{e_{max}} = \frac{1}{(1 + rf_2)} \tag{4}$$

An effector dose–response profile can be determined by measuring the ability of the repressor to block transcription at different effector concentrations. The shape of the response curve can be described by three parameters: the basal level of the transcript in the absence of effector (leakiness), the difference between the basal level of transcript and the amount of transcript produced at saturating effector concentrations (dynamic range), and the concentration of effector that is required for the half-maximal expression E_{50}. Leakiness, dynamic range, and E_{50} are three unique phenotypic properties that are governed by the genotypic composition of the repressor. As described in Appendix A, these three features of the response curve depend on one or more of the thermodynamic equilibrium constants.

Many mutants of the lac repressor have been previously identified using a phenotypic β-galactosidase assay.[9] Unfortunately, quantitative data for various effectors at varied concentrations for these mutants are not available. Furthermore, underlying changes to the thermodynamic parameters that define phenotype were not determined. In this article, we performed a series of experiments that allowed us to explore the thermodynamic properties at the foundation of allostery. Specific mutations were introduced into the lac repressor to partially recreate and expand the library of lac repressor mutants already defined. The changes in the thermodynamic equilibrium constants were determined from the experimentally measured expression profiles. Some of the mutations produced repressors with phenotypes that were dramatically different from the wild-type repressor. These observations are discussed in the context of how the altered thermodynamic parameters affect phenotype and, consequently, gene regulation.

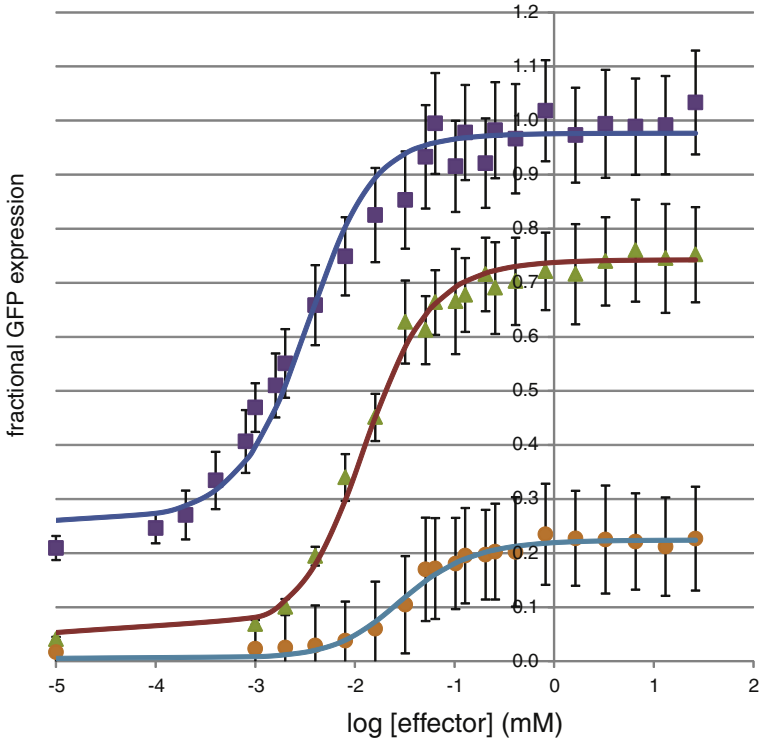

Fig. 2. Fractional expression was measured for the wild-type repressor (green triangles) and for the two mutants Q18M (orange circles) and Q18A (purple squares). The experimental data and the fit for each mutant are included. Error bars are included from replicate measurements. Q18M (the tight-binding mutant) is less leaky but induces poorly. Q18A induces well but is leakier than the wild-type repressor and the Q18M mutant.

Results

Mutations in the DNA binding domain affect the basal level of expression and dynamic range

As described previously,[8] a library of $\sim 20^3$ or 8000 mutant repressors was created by introducing all 20 amino acids into the three key residues involved in operator binding (Y17, Q18, and R22). The ability of specific mutants in this library to repress transcription was evaluated using an *in vivo* assay. By placing a modified green fluorescent protein (GFP), GFPmut3.1, under the control of the natural lac operator, we can accurately measure GFP expression in the absence and in the presence of a gratuitous inducer, IPTG. Altering the three residues that make a direct contact with the operator produces repressor molecules that are very leaky and are characterized with an I^- phenotype, as previously observed.[9,10] Less than 1% of the mutant repressor molecules in the library retain the ability to repress transcription (Table S1), although one mutant (Q18M) repressed GFP production better than the wild-type repressor and was also shown to bind more tightly to the operator *in vitro*.[8,11,12]

Although many of these mutants had been identified previously and a functional phenotype had been described in a qualitative manner,[9] the thermodynamic changes responsible for altered phenotypes were not identified. To investigate the thermodynamic properties that account for the change in phenotype, we measured IPTG response curves for the wild-type repressor, the tight-binding mutant (Q18M), and the leaky mutant (Q18A) (Fig. 2). The wild-type repressor allows for a small basal level of transcript to be produced (slightly leaky) and demonstrates a significant induction reaching 70–80% of the constitutive level (large dynamic range). The tight-binding mutant is significantly less leaky than the wild-type repressor; however, the dynamic range is greatly reduced, and expression plateaus at 20% of the constitutive level. In contrast, the Q18A mutant is very leaky (25% expression), and expression reaches constitutive levels upon induction (Table 1). Even though the dynamic range of this mutant is similar to that of wild type, it is not a potent repressor due to excessive leakiness.

Fitting the wild-type and headpiece mutant data to Eq. (4) using a genetic algorithm, we were able to determine the change in the thermodynamic equilibrium constants that results from a mutation in the DNA binding domain. The genetic algorithm computed the change in the equilibrium constant using a fitness function that minimized the differences between experimental and calculated expression levels weighted by experimental error. Since the mutations are in the helix–turn–helix domain, we anticipate that the change in the midpoint of induction E_{50} for these mutants is largely due to altered binding affinities, while the conformational equilibrium is essentially unaffected. The value of $[R_{TOT}]K_{RO}$ for the tight-binding mutant Q18M is about 10-fold greater than the value for the wild type, while the weaker-binding mutant Q18A demonstrated a 10-fold decrease (Table 1). If we assume that the concentration of the repressor is unaffected by the point mutation, then the differences in the IPTG response curves reflect a change in the repressor–operator affinity K_{RO}, consistent with previous *in vitro* binding data. As is evident from the fitted curves, there is a correlation between leakiness and dynamic range when only K_{RO} is altered. Tight-binding mutants have a restricted dynamic range; the repressor binds too tightly to the DNA. Very-weak-binding mutants also have a restricted dynamic range, illustrating that there is a tradeoff between repression and induction.

Table 1. Thermodynamic parameters altered by effector molecules

Construct	Effector	R	K_{RE}	K_{RR^*}	K_{ER^*}	X	Leakiness	Dynamic range	$\text{Log}_{10}\ E_{50}$
Wild type	IPTG	150 ± 50	60 ± 0.2	5.8 ± 0.07	500 ± 5	8.28	0.044	0.7	−1.94
Wild type	ONPG	150 ± 50	13 ± 1.5	5.8 ± 0.07	1.3 ± 0.9	0.13	0.044	0.0	−1.45
Y17I	IPTG	8.6 ± 0.07	60 ± 0.2	5.8 ± 0.07	500 ± 5	8.28	0.44	0.5	−2.66
Q18A	IPTG	18.0 ± 0.2	60 ± 0.2	5.8 ± 0.07	500 ± 5	8.28	0.28	0.7	−2.49
Q18M	IPTG	1404 ± 1	60 ± 0.2	5.8 ± 0.07	500 ± 5	8.28	0.005	0.2	−1.56
L148D	ONPG	150 ± 50	2.4 ± 0.2	525 ± 7	0.46 ± 0.05	0.19	0.78	−0.7	−0.12
L148W	ONPG	150 ± 50	18 ± 0.7	325 ± 6	3.2 ± 0.2	0.18	0.69	−0.6	−1.13
F161N	IPTG	150 ± 50	23 ± 0.8	0.98 ± 0.34	670 ± 127	28.54	0.013	0.8	−1.64
F161S	ONPG	150 ± 50	0.25 ± 0.01	26.5 ± 0.01	0.54 ± 0.02	2.14	0.16	0.3	0.57
F161T	IPTG	150 ± 50	82 ± 6.5	2.7 ± 1.4	1415 ± 293	17.16	0.024	0.8	−2.21
F161W	IPTG	150 ± 50	18 ± 9	0.47 ± 0.23	615 ± 307	33.89	0.01	0.8	−1.42
Q291I	IPTG	150 ± 50	26 ± 2.8	0.07 ± 0.02	1889 ± 294	71.99	0.007	0.7	−1.46
Q291K	IPTG	150 ± 50	11 ± 0.3	5.9 ± 0.4	86 ± 5	8.02	0.044	0.7	−1.18
Q291M	IPTG	150 ± 50	44 ± 3.1	8.9 ± 0.6	515 ± 51	11.71	0.062	0.8	−2.07
Q291R	IPTG	150 ± 50	251 ± 19	1.33 ± 0.03	296 ± 25	1.18	0.015	0.0	−2.36
L296W	IPTG	150 ± 50	3.8 ± 0.1	0.04 ± 0.01	409 ± 38	107.28	0.007	0.73	−0.66
L296W	ONPG	150 ± 50	0.4 ± 0.01	0.04 ± 0.01	17.5 ± 1.7	48.81	0.007	0.36	0.65

Mutations in the effector binding pocket lead to phenotypic diversity

Effector molecules bind to the repressor and alter the apparent conformational equilibrium. There is a broad range of effector molecules that are recognized by the lac repressor; virtually all of these molecules are galactosides.[13] These sugars bind to the repressor by forming a conserved set of interactions[14] (Fig. 3). Residues R197, N246, and D274 in the regulatory domain of the repressor form hydrogen bonds to the O2 and O3 hydroxyls of the sugar. These interactions are essential for effector binding, and removing any one of the sugar hydroxyls from the effector or epimerizing them to alternate stereochemical configurations prevents the effectors from binding to the repressor. As a consequence, we focused on the portion of the effector binding pocket that is responsible for specificity. The differences between inducing, non-inducing, and anti-inducing galactosides appear to be the substituent groups on the first and sixth carbons of the sugar ring.[13] As a consequence, we mutated the six residues in the effector binding pocket that, based on the structure, appear to interact with the substituent groups at the C1 position (Fig. 3b). One hundred fifteen unique repressor molecules were created by introducing all 20 amino acids at positions 79, 148, 161, 291, 293, and 296. These mutant repressors were analyzed; their phenotypes were evaluated *in vivo* (Table S2) and, for the most part, are consistent with the data of Markiewicz *et al.*[9] and Kleina and Miller.[10]

Mutant lac repressors traditionally have been classified based on their phenotypic response. Mutants that failed to repress transcription were classified as I^-, and mutants that did not respond to the effector were labeled as superrepressors, denoted as I^s. A plot of the leakiness *versus* dynamic range for the 115 effector pocket mutants illustrates that these two phenotypes are not mutually exclusive (Fig. 4). The majority of the single-site mutants exhibit a phenotype similar to that of the wild-type repressor, displaying modest changes in either dynamic range and/or leakiness. A few mutants repress better than the wild type, but their level of induction is less; these mutants would be classified as I^s. Another fraction of the mutants (~10%) are unable to repress transcription (leakiness >50%) and display a traditional I^- phenotype. The remaining point mutations (20%) weakly repress transcription (15%<leakiness<50%) but are still inducible and would be described as weakly I^- in the original classification. In some instances, the level of

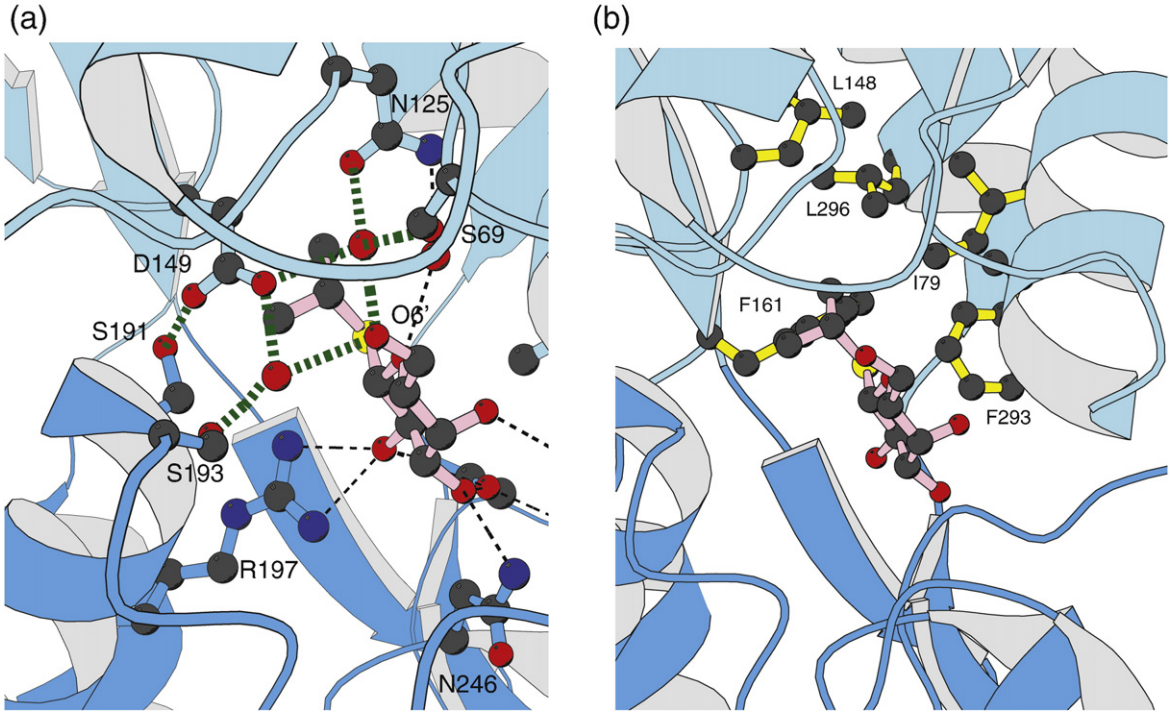

Fig. 3. Structural view of the ligand binding pocket. (a) Structure of IPTG bound to the repressor, illustrating the hydrogen-bonding network. Light blue corresponds to the N-terminal domain, and dark blue corresponds to the C-terminal domain. (b) The residues in yellow interact with the constituent group off the C1 carbon of the galactoside effectors and were mutated to produce 115 mutant repressors. Note that the aqua residues in (a) interact with the inducer ligand but were not altered because they anchor the galactose ring or are implicated in the allosteric signaling pathway.

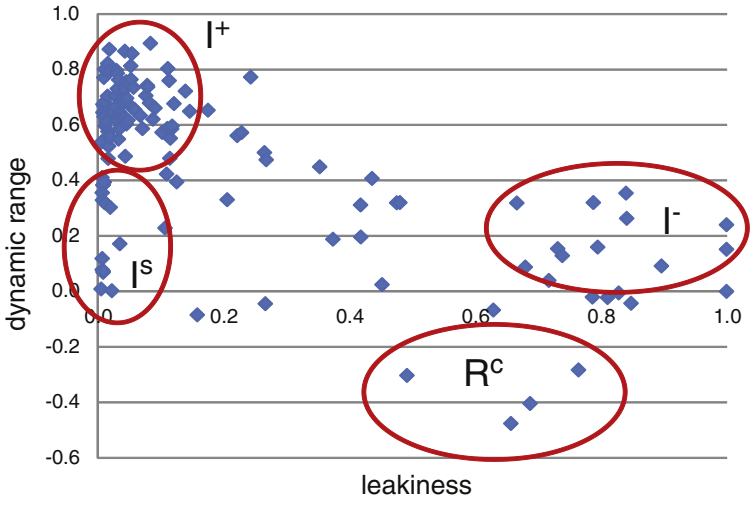

Fig. 4. The phenotype of the mutant repressors is plotted as a function of their leakiness (*x*-axis) and dynamic range (*y*-axis). The mutant repressors broadly fit into four classical phenotypes.

expression is effector independent, suggesting that IPTG could no longer bind to the mutant repressor or was binding to the active and inactive forms with equal affinity (dynamic range ~0). We observed four mutants that showed an inverted response to IPTG (dynamic range <0). Since all of these mutants are also leakier than the wild type, they exhibit a reversible constitutive (R^c) phenotype;[10] IPTG functions like a corepressor since there is a decrease in GFP expression upon addition of effector.

Mutating amino acids in the effector binding pocket affects leakiness and dynamic range by altering the fraction of repressor molecules that can adopt the active conformation f_2. The value of f_2 depends on the thermodynamic equilibrium constants K_{ER}, K_{ER*}, and K_{RR*}. IPTG response curves were measured for a subset of mutants, and the data were fitted using a genetic algorithm assuming that the total repressor concentration R_{TOT} and the repressor–operator affinity K_{RO} were unaffected. The interplay of these three thermodynamic constants alters the dose–response curves in predictable ways. When a mutation alters the relative binding affinities of the effector (i.e., changing the K_{ER*}/K_{ER} ratio), the induction curves are altered only at high effector concentrations, and the response becomes more pronounced as the magnitude of the ratio increases (Fig. 5a). For example, the point mutant F161T appears similar to the wild type in the absence of inducer, but at high levels of IPTG, it is more fully induced and therefore has a greater dynamic range. The equilibrium parameters that best fit the data have a K_{ER*}/K_{ER} ratio that is twice that of the wild type (Table 1).

We observed mutations that alter the effector affinity for both R and R* conformations, leaving the K_{ER*}/K_{ER} ratio unaffected. These mutants have induction profiles that display a left shift or a right shift with no appreciable change in the dynamic range of the response; they change the E_{50} (Fig. 5b). A mutant repressor that has a greater affinity can be induced at lower concentrations of the effector, which is reflected by a left shift in the binding curve (smaller E_{50}). Mutant repressors that bind weakly exhibit a right shift in the binding curve (larger E_{50}). E_{50} is also a function of the conformational equilibrium, but changes to K_{RR*} are analogous to changes in K_{RO} such that leakiness and dynamic range will be altered. Interestingly, one point mutant (Q291K) was identified to have increased E_{50}, but with the dynamic range and leakiness unaffected. The measured equilibrium parameters showed that the ratio of effector binding K_{ER*}/K_{ER} and K_{RR*} are quite similar to the wild-type repressor, but the magnitudes of both K_{ER*} and K_{ER} are six times smaller (Fig. 5b).

Although we initially suspected that mutations in the effector binding pocket would only disturb the equilibrium constants K_{ER*} and K_{ER}, we observed that altering the effector binding site can also produce changes in the conformational equilibrium constant K_{RR*}. Modifying the conformational equilibrium will manifest itself by altering the entire induction profile, affecting leakiness, dynamic range, and E_{50} (Fig. 5c). The Q291I mutant represses more tightly than the wild-type repressor but does not induce as well a phenotype similar to that observed in the DNA binding domain. However, this mutant has an induction profile that is consistent with an altered conformational equilibrium constant. By fitting the data, we observed that the thermodynamic parameter K_{RR*} that best fits the observed induction profile has a value 50 times lower than that of the wild-type repressor.

We found many point mutations that simultaneously alter multiple thermodynamic constants. The mutation F161T increases the concentration of the active form of the repressor by decreasing the conformational equilibrium to half that of the wild

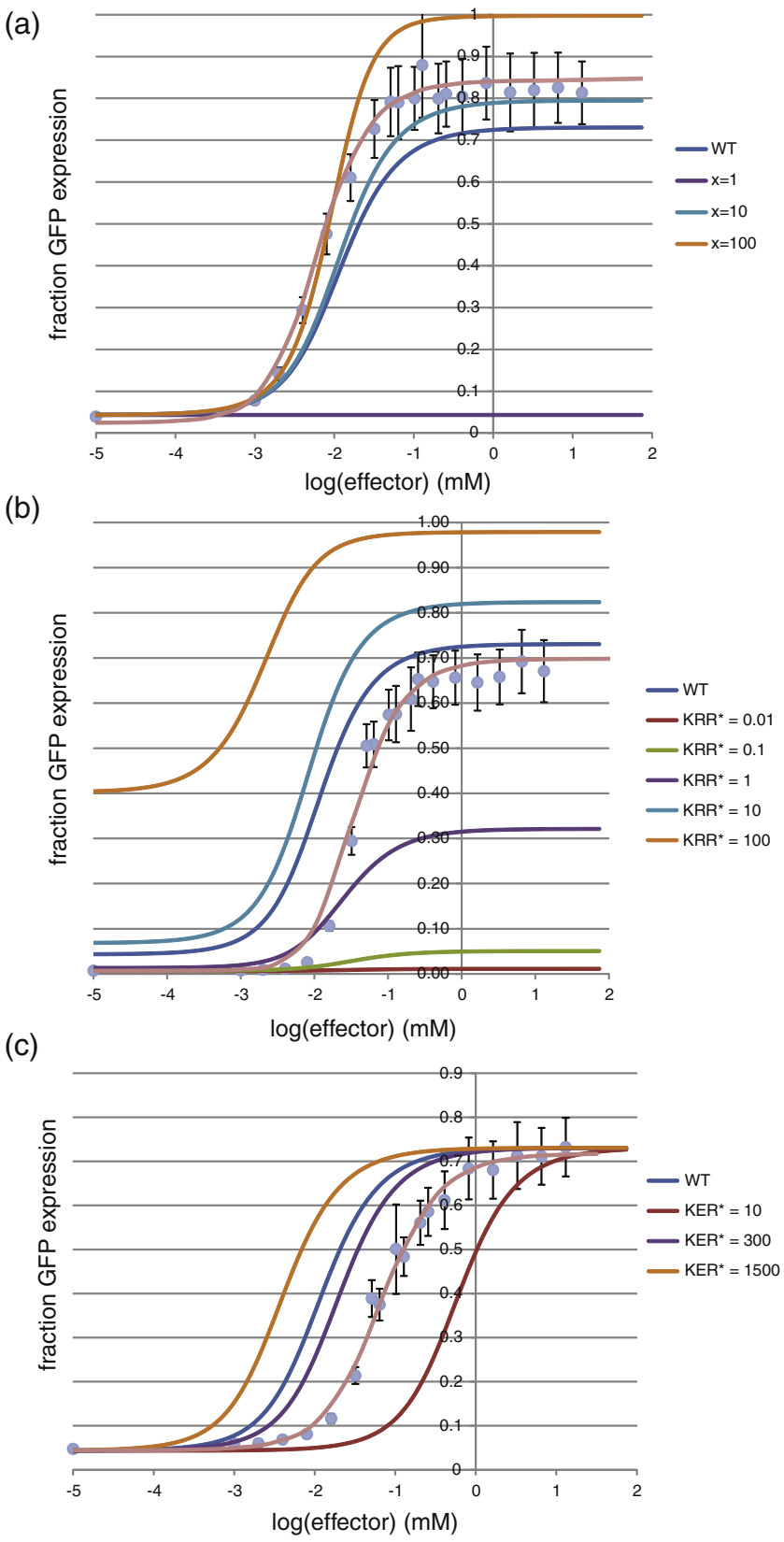

Fig. 5. Changes in thermodynamic parameters affect GFP expression. Simulated plots of fractional induction and experimental data are shown as a function of effector concentration (plotted on a log axis). Using the determined thermodynamic parameters for the wild-type repressor, we simulated theoretical plots for stepwise changes in the following: (a) the ratio of inducer binding affinity for active repressor conformation to inducer binding affinity for inactive repressor conformation ($X = K_{ER^*}/K_{ER}$). The dark blue curve describes the wild-type repressor (see Fig. 2). The experimental values for the F161T mutant show an increase in the ratio compared to the wild type: (b) the conformational equilibrium K_{RR^*}. The mutant Q291I illustrates that decreasing K_{RR^*} produces a less leaky repressor; the curve is right-shifted compared to the wild type and does not induce completely. (c) The absolute values of the effector binding affinities K_{ER} and K_{ER^*} at a fixed ratio X. The mutant Q291K is an example of a mutant that results in a right shift of the binding curve but does not alter the leakiness or the dynamic range.

type; leakiness is thereby decreased. The affinity of the active form for IPTG is comparable to that for wild type; however, the inactive conformation's affinity is increased threefold. This increases the K_{ER*}/K_{ER} ratio, which in turn increases the dynamic range, and the increase in magnitude lowers E_{50}. The net effect of this single-point mutation is a repressor that exhibits less basal expression than wild type but also induces to higher levels with less effector, essentially behaving as a more ideal on–off switch.

Changing effector molecules leads to phenotypic diversity

Effector molecules that shift the apparent K_{RR*} equilibrium towards the inactive conformation are classified as inducers and have values of $K_{ER*}/K_{ER}>1$. In contrast, effector molecules that increase the population of active repressors function as anti-inducers or corepressors, and these compounds have a K_{ER*}/K_{ER} ratio of <1. The most potent anti-inducer, orthonitrophenyl-β-D-fucoside (ONPF), is a galactoside that increases the affinity of the repressor for the operator by roughly a factor of 2–3.[3] There are other galactosides that are referred to as neutral effector molecules, such as orthonitrophenyl-β-D-galactoside (ONPG), that bind to the repressor but do not alter the equilibrium. These compounds bind to the active and inactive conformations of the repressor with similar affinity[13] and therefore have a value of $K_{ER*}/K_{ER}=1$. Although ONPG has traditionally been classified as a neutral effector, we observed that ONPG functions as an anti-inducer for the wild-type repressor (Fig. 6a). In fact, of the 115 mutants analyzed, we observed that ONPG actually functions as an anti-inducer and increases repression for a large number of the mutants. There are, however, a few mutants that are insensitive to ONPG, showing no change in expression levels.

Quite surprisingly, roughly a dozen mutants were inducible with the addition of 10 mM ONPG. (Fig. 6b). Although induction levels varied, one specific mutant, L296W, exhibits a reasonable dynamic range. In the absence of effector molecule, the L296W mutant has a background expression of 1% and reaches a maximal induction of 41% in the presence of 10 mM ONPG. This dynamic range is not as large as that of the wild type, but the L296W mutant demonstrates a clear shift in response to ONPG and marks the first mutant repressor to be induced with a previously noninducing ligand. Interestingly, this mutant shows a clear preference for induction by ONPG and could not be induced by other anti-inducers, suggesting that the 6′ hydroxyl and nitro groups on the phenyl ring are both important for stabilizing the inactive repressor conformation (data not shown). The experimental effect or response profiles were measured for the mutants L296W and F161S. The conformational equilibrium K_{RR*} and the effector binding constants K_{ER} and K_{ER*} were fitted to the data (Table 1). Consistent with ONPG now acting as an inducer, the K_{ER*}/K_{ER} ratio is greater than 1 (F161S: $K_{ER*}/K_{ER}\sim 2$; L296W: $K_{ER*}/K_{ER}\sim 48$) compared to the wild-type value of $K_{ER*}/K_{ER}=0.1$. The L296W point mutant binds ONPG more tightly than the F161S mutant (larger magnitude of K_{ER} and K_{ER*}); however, the F161S mutant exhibits a higher induction, which is linked to a larger conformational equilibrium and, hence, greater leakiness. These are just two instances where a single amino acid change in the inducer binding pocket produces a change in phenotype that can be traced back to distinct changes in the equilibrium constants. Although these two mutants do not have a dynamic range as large as the wild-type repressor in the absence and in the presence of the gratuitous inducer IPTG, they are nonetheless inducible.

Altered thermodynamic constants lead to an inverted switch

Repressors in bacterial operons are classified as either inducible or corepressible based on the intrinsic degree of leakiness and the relative change in expression upon addition of a given ligand. A corepressible system has an effector dose–response curve that is flipped relative to the induction curves. In the absence of a metabolite, the level of gene expression is high. The metabolite alters the conformation of the repressor, decreasing the leakiness. We identified a small number of mutants where a single amino acid change converts an inducible repressor into one that is corepressible. Starting from a pool of mutant repressor molecules having an I^- phenotype (Fig. 4), we looked for molecules that would block transcription in the presence of an effector, going from on to off.

Of the 11 I^- mutants analyzed, three displayed significant corepression with the anti-inducer ONPF, ONPG, or phenyl-β-D-galactoside (Fig. 6c). The L148W substitution demonstrated the greatest dynamic range. In the absence of effector, the expression level was 64% of e_{max}, and the addition of 10 mM ONPF reduced the expression to 4% of e_{max}. The dynamic range of this inverted switch is comparable to the absolute expression levels seen for the wild-type switch in the off–on states, suggesting that the L148W mutant is truly an inverted switch. The thermodynamic parameters for this mutant that best fitted the effector dose–response curves were determined (Fig. 2). As would be expected, the conformational equilibrium is shifted considerably in favor of the R* conformation ($K_{RR*}=325$), such that the concentration of the repressor that adopts the DNA binding

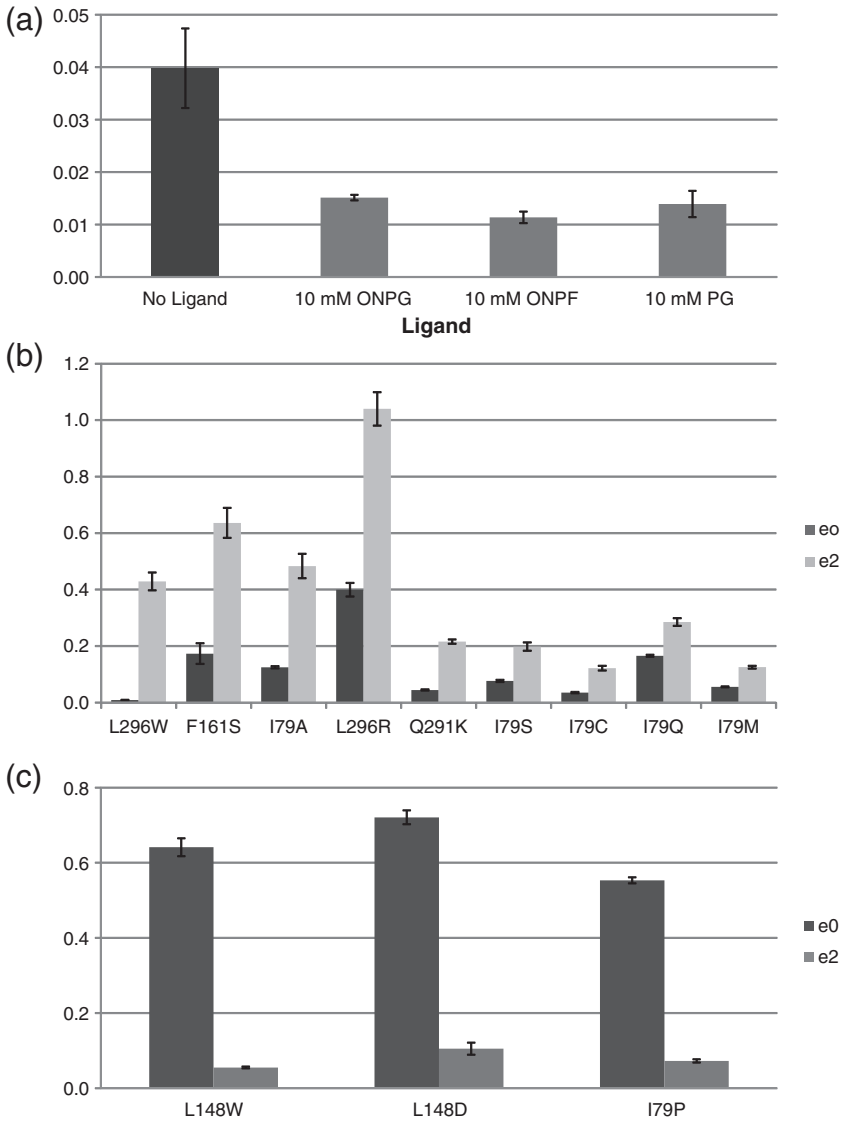

Fig. 6. Different effectors paired with mutant repressors have unique phenotypes. (a) Addition of anti-inducers to the wild-type repressor results in a decrease in the fractional expression level by several percentages. (b) Nine mutants demonstrated a change in effector specificity when incubated in the presence of 10 mM ONPG (e_2). While induction of many mutants is slightly greater than twofold, one mutant (L296W) that was capable of a much larger induction was identified. (c) When 10 mM ONPF is added to several mutants displaying the I^- phenotype (e_2), repression of the natural OR1 operator is restored. While ONPF acts as a corepressor with both wild type and I^- mutants, a more dramatic change in expression occurs with the I^- mutants.

conformation is minimal. Similar to the wild-type repressor, the anti-inducer binds preferentially to the R conformation, but here it does so with a K_{ER*}/K_{ER} ratio of about 1:5. This single-site substitution dramatically changes the functionality of the repressor from an inducible system to a corepressible system. This would suggest that repressors in the Lac/Gal superfamily that are structurally very similar[15] can respond to a diverse set of effectors in both positive and negative directions with only modest changes in the amino acid sequence.

Conclusions

Regulating the flux of metabolic pathways is necessary for cellular homeostasis and, in many instances, is controlled by altering gene expression. For nearly 50 years, the lac operon has served as a model system for understanding how the rate of transcription can be modulated in response to cellular conditions. We have illustrated how mutations in the repressor alter the thermodynamic equilibrium constants and, consequently, the

phenotype. Mutations were engineered into both the DNA binding domain and the effector binding site of the lac repressor, and their effects on the repressor function were quantitatively assessed with an *in vivo* assay. By experimentally measuring expression profiles, we were able to expand upon previous phenotypic analyses of many lac mutants by modeling changes in the equilibrium constants that are responsible for their phenotypes. Moreover, we have illustrated how mutant phenotypes can be traced back to thermodynamic components.

Introducing single amino acid changes can radically alter the phenotype. By measuring the expression data for a large number of mutants, we were able to identify mutants that perform better than wild-type switch. These mutants are less leaky and induce with a larger dynamic range. The improved functionality of many of these mutants was previously undetectable due to their qualitative phenotypic assessment.[9] We were also able to expand on previous analyses of these mutants by measuring the effect of other effectors on their function. This resulted in the discovery of a set of mutant repressors that induce with a previously noninducing species, ONPG. This marks the first successful attempt at reengineering the ligand binding site of the lac repressor to induce with a new ligand.[16,17] Due to the quantitative nature of the *in vivo* assay, we were also able to demonstrate that ONPG acts as an anti-inducer and is not a neutral effector, as was described in previous *in vitro* experiments.[13]

The final sets of mutants described are those that act in a nearly ideal inverted fashion with high leakiness and large negative dynamic range upon addition of an anti-inducer. These mutants resemble other lac family members that function in a corepressible manner. Interestingly, all of these phenotypes were created by introducing single changes to the repressor. While identifying these unique phenotypes is exciting, the strength of this work is in elucidating the thermodynamic properties that are at the foundation of the phenotypic changes. Elucidating the thermodynamic properties of the molecular switch is essential for developing a complete understanding of gene regulation.

Methods

Reporter construction

The reporter plasmid used in these experiments was derived from the pBD1200 plasmid described previously.[8,18] After sequence validation, the reporter plasmid was transformed into DH5α *E. coli* for selection experiments and mutant characterization.

Generation of repressor mutant plasmids

A complete description of the construction of the wild-type repressor plasmid and the DNA binding mutant library can be found elsewhere.[8,18] Starting with the wild-type repressor plasmid, we introduced mutations into each of the six residues of the effector binding site using independent QuickChange (Stratagene) PCRs. Clones were identified, and plasmid DNA was prepared (Clontech) for each of the mutants. After verifying the sequence in each mutant repressor plasmid, we transformed the plasmids into reporter cells and grew them for phenotypic screening.

In vivo repression/induction assay

To analyze the phenotypes of the mutant repressors, we used an *in vivo* fluorescent assay. Cells were grown and analyzed in a Perkin-Elmer Victor3 plate reader to quantitate the level of fluorescence and therefore to indirectly measure the degree of transcription. In short, each of the mutant repressors was transformed, and colonies were selected in triplicate for overnight culture growth. In addition, cells containing 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 effector and in the presence of various amounts of IPTG, ONPG, PG, or ONPF. Once samples had 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 (A590) in a Perkin-Elmer Victor3 Plate reader. The signal from the blank sample was subtracted and the resulting fluorescent signals were normalized to the cell optical density to normalize the signal for each sample. The normalized signals from each of the replicates were averaged, and the standard deviations were calculated (error bars on plots). For dose–response curves, a given mutant was grown in triplicate on the same plate, with wells containing different amounts of a given effector. The data were then processed in the same manner mentioned above.

Global fitting

Induction profiles for mutants were globally fitted to measure the following parameters from Eq. (1): r, K_{RR^*}, K_{R^*E}, K_{RE}, and s. The parameter r was estimated previously and set to 150 for wild type.[19] The parameter s was assumed to be much smaller than r and thus was set to 0. For non-headpiece mutants (any mutation not in residues 1–62), the affinity of the repressor for DNA is assumed to be unchanged (K_{RO} and K_{R^*O}); therefore, r and s were assumed to be the same as wild type ($r = 150$; $s = 0$). For headpiece mutants, only the affinity of the repressor for DNA was assumed to change. Therefore, every parameter was assumed to be the same as wild type, with the exception of r. For mutants measured with multiple effectors, the K_{RR^*} parameter, which is only

dependent on the thermodynamics of the repressor folding alone, is assumed to be unchanged. Using these rules, we determined a minimum set of thermodynamic constants to fully describe every experimental curve. A genetic algorithm was used to simultaneously fit every curve within the data set. The fitness function minimized the absolute value of the difference between experimental expression levels and predicted expression levels. The percentage error of the measurement was used to weight the individual terms. Global fitness was calculated by summing the individual fitness terms:

$$\text{fitness}_j = \log_{10}\left(\sum_i |\exp_i - \text{pred}_i| \frac{\text{err}_i}{\exp_i}\right) \quad (5)$$

$$\text{global fitness} = \sum_j \text{fitness}_j \quad (6)$$

where j denotes the experiment number, and i is the individual measurement.

The genetic algorithm used the following options—rank fitness scaling, stochastic uniform selection function, scattered cross-over function, population of 200, elite count of 2, and cross-over fraction of 0.6745—and adapted feasible mutation function. The parameter search space was over 8 orders of magnitude (10^{-4}–10^{4}) for each fit parameter, and 10 rounds of genetic algorithms were used to estimate error. All codes were implemented in MATLAB (Mathworks).

Appendix A

The basal level of transcript produced at zero effector concentration ([E]=0) defines the leakiness of the switch. It depends on the affinity of the repressor for the operator K_{RO}, the conformational equilibrium K_{RR*}, and the amount of repressor that this presents:

$$\text{Leakiness } (L) \equiv \frac{e}{e_{max}}([E]=0) = \frac{1}{\left(1+\frac{r}{1+K_{RR*}}\right)} \quad (A1)$$

The dynamic range is defined as the difference between the basal level of expression and the maximal level of expression found at a saturating effector concentration ([E]→∞):

$$\text{Dynamic range } (D) \equiv \frac{e_{[E]}}{e_{max}}([E]\to\infty) - \frac{e_{[E]}}{e_{max}}([E]=0) \quad (A2a)$$

$$D = \frac{1}{\left(1+\frac{r}{1+K_{RR*}X^2}\right)} - \frac{1}{\left(1+\frac{r}{1+K_{RR*}}\right)} \quad (A2b)$$

where X is the ratio of the effector binding constants ($X=K_{ER*}/K_{ER}$). The final parameter derived from the effector dose–response curves is the concentration of the effector for half-maximal expression, which is defined as one-half the dynamic range plus the leakiness:

$$E_{50} = L + D/2 \quad (A3a)$$

It is a function of all of the thermodynamic constants:

$$E_{50} = \frac{1}{\left(1+\frac{r(1+K_{ER}[E])^2}{(1+[E]K_{ER})^2 + K_{RR*}(1+[E]K_{ER*})^2}\right)} \quad (A3b)$$

Based on previous models, we assume that the affinity between the inactive repressor and the operator is negligible ($s\sim 0$). This simplifies Eq. (3) and allows us to solve for the effector concentration corresponding to $e_{\frac{1}{2}max}$:

$$\frac{1}{\left(1+\frac{r(1+[E]K_{ER})^2}{(1+[E]K_{ER})^2 + K_{RR*}(1+[E]K_{ER*})^2}\right)} = L + D/2 \quad (A3c)$$

Inverting and subtracting give:

$$\frac{r(1+[E]K_{ER})^2}{(1+[E]K_{ER})^2 + K_{RR*}(1+[E]K_{ER*})^2} = \frac{1}{L+D/2} - 1 \quad (A3d)$$

We now define the right-hand side of Eq. (A3d) as:

$$\frac{1}{\varphi} \equiv \frac{1}{L+D/2} - 1 \quad (A3e)$$

After further algebraic rearrangement:

$$[E]^2\left(K_{ER}^2(\varphi r - 1) - K_{RR*}K_{ER}^2\right) + [E] \times \left(2K_{ER}(\varphi r - 1) - 2K_{RR*}K_{ER*}\right) + \varphi r - K_{RR*} - 1 = 0 \quad (A3f)$$

We then make the three following simplifying definitions:

$$A \equiv K_{ER}^2(\varphi r - 1) - K_{RR*}K_{ER*}^2 \quad (A4)$$

$$B \equiv 2K_{ER}(\varphi r - 1) - 2K_{RR*}K_{ER*} \quad (A5)$$

$$C \equiv \varphi r - K_{RR*} - 1 \quad (A6)$$

The effector needed for half-maximal expression is then the positive root:

$$E_{50} \equiv -B \pm \sqrt{\frac{B^2 - 4AC}{2A}} \quad (A7)$$

Supplementary Data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jmb.2011.03.057

References

1. Monod, J., Changeux, J. P. & Jacob, F. (1963). Allosteric proteins and cellular control systems. *J. Mol. Biol.* **6**, 306–329.
2. Monod, J., Wyman, J. & Changeux, J. P. (1965). On the nature of allosteric transitions: a plausible model. *J. Mol. Biol.* **12**, 88–118.
3. Jobe, A. & Bourgeois, S. (1973). Lac repressor–operator interaction: VIII. Lactose is an anti-inducer of the lac operon. *J. Mol. Biol.* **75**, 303–313.
4. Lewis, M., Chang, G., Horton, N. C., Kercher, M. A., Pace, H. C., Schumacher, M. A. *et al.* (1996). Crystal structure of the lactose operon repressor and its complexes with DNA and inducer [see comment]. *Science*, **271**, 1247–1254.
5. Boelens, R., Scheek, R. M., van Boom, J. H. & Kaptein, R. (1987). Complex of lac repressor headpiece with a 14 base-pair lac operator fragment studied by two-dimensional nuclear magnetic resonance. *J. Mol. Biol.* **193**, 213–216.
6. 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.
7. Swint-Kruse, L. & Matthews, K. S. (2009). Allostery in the LacI/GalR family: variations on a theme. *Curr. Opin. Microbiol.* **12**, 129–137.
8. Daber, R. & Lewis, M. (2009). Towards evolving a better repressor. *Protein Eng. Des. Sel.* **22**, 673–683.
9. 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.
10. Kleina, L. G. & Miller, J. H. (1990). Genetic studies of the lac repressor: XIII. Extensive amino acid replacements generated by the use of natural and synthetic nonsense suppressors. *J. Mol. Biol.* **212**, 295–318.
11. Lehming, N., Sartorius, J., Kisters-Woike, B., von Wilcken-Bergmann, B. & Muller-Hill, B. (1990). Mutant lac repressors with new specificities hint at rules for protein–DNA recognition.[erratum appears in EMBO J. 1990 May;9(5):1674]. *EMBO J.* **9**, 615–621.
12. Lehming, N., Sartorius, J., Niemoller, M., Genenger, G., v Wilcken-Bergmann, B. & Muller-Hill, B. (1987). The interaction of the recognition helix of lac repressor with lac operator. *EMBO J.* **6**, 3145–3153.
13. Barkley, M. D., Riggs, A. D., Jobe, A. & Bourgeois, S. (1975). Interaction of effecting ligands with lac repressor and repressor–operator complex. *Biochemistry*, **14**, 1700.
14. Daber, R., Stayrook, S., Rosenberg, A. & Lewis, M. (2007). Structural analysis of lac repressor bound to allosteric effectors. *J. Mol. Biol.* **370**, 609–619.
15. Weickert, M. J. & Adhya, S. (1992). A family of bacterial regulators homologous to Gal and Lac repressors. *J. Biol. Chem.* **267**, 15869–15874.
16. Chang, W. I. & Matthews, K. S. (1995). Role of Asp274 in lac repressor: diminished sugar binding and altered conformational effects in mutants. *Biochemistry*, **34**, 9227–9234.
17. Satya Lakshmi, O. & Rao, N. M. (2009). Evolving Lac repressor for enhanced inducibility. *Protein Eng. Des. Sel.* **22**, 53–58.
18. Daber, R. & Lewis, M. (2009). A novel molecular switch. *J. Mol. Biol.* **391**, 661–670.
19. Daber, R., Sharp, K. & Lewis, M. (2009). One is not enough. *J. Mol. Biol.* **392**, 1133–1144.