Biophysical Chemistry 159 (2011) 66–72

Contents lists available at ScienceDirect

Biophysical Chemistry

journal homepage: http://www.elsevier.com/locate/biophyschem

Allostery in the lac operon: Population selection or induced dissociation?

Kim A. Sharp *

Dept. of Biochemistry and Biophysics, University of Pennsylvania School of Medicine, Philadelphia, PA 19104, United States

ARTICLE INFO

Article history:
Received 23 March 2011
Received in revised form 26 April 2011
Accepted 4 May 2011
Available online 12 May 2011

Keywords:
Allostery
Lac repressor
Induced fit
Population selection
Kinetics of allostery
Tertiary capture

ABSTRACT

Allostery, the modulation of function of a protein at one site by the binding of a ligand at a different site, is a property of many proteins. Two kinetically distinct models have been proposed: i) The induced fit model in which the ligand binds to the protein and then induces the conformational change. ii) The population selection model, in which the protein spontaneously undergoes a conformational change, which is then 'captured' by the ligand. Using measured kinetic constants for the lac repressor the contribution of population selection *vs.* induced dissociation is quantified by simulating the kinetics of allostery. At very low inducer concentration, both mechanisms contribute significantly. Total induction, though, is small under these conditions. At increasing levels of induction the induced dissociation mechanism soon dominates, first due to binding of one inducer, and then from two inducers binding.

1. Introduction

Allostery, the modulation of the function of a protein at one site by the binding of a ligand at a different site, is a property of many proteins. Allostery is a key element of gene regulation, enzyme regulation, switching, signal transduction and many other regulation events. It has become clear that conformational switching and allostery is also an important property of many RNA molecules [1]. The importance of allosteric regulation, and linkage and allostery are illustrated by their frequent recurrence as themes at the Gibbs Conferences on Biothermodynamics. In the now classic Monod–Wyman–Changeux (MWC) model for allostery [2], the protein exists in two states, denoted P and P*. A ligand (L) binds to P, then there is a conformational change or induced fit to form LP*, with a resultant change in protein activity or function. This conceptual framework has been applied to many proteins, including the lac inducer–repressor–operator system [3–8]. Here the ligand is an inducer (I) and the function it modulates is the affinity of the repressor (R) for the operator (O), which in turn controls expression of the genes downstream from the operator. This repressor has been the paradigm for studying prokaryotic transcriptional regulation since the pioneering work of Jacob and Monod in 1961 [9]. Moreover, applications of the lac operon to non-bacterial systems have highlighted its use as a regulator of desired genes *in vivo*. The switch of the lac operon has been used to regulate expression of target genes in mammals [10,11], stem cells [12], and breast cancer lines [13]. The potential use of this molecular switch in gene therapy is promising, illustrated by the integration of this molecular switch into mice. [10]. The lac repressor is also an important system for studying the structural basis for sequence specific recognition of DNA in the context of the classical helix–turn–helix motif [7]. The thermodynamic and kinetic properties of the lac inducer–repressor–operator system have also been extensively studied, [3,4,14–18], which includes work presented at several Gibbs conferences [6]. Analysis of this biothermodynamic data can provide insights into how the lac operon functions as a genetic switch.

Models of allostery have three irreducible features. The first is a minimum of two functionally different states of the protein (or RNA). These may each consist of underlying substates – see below – as long as the 'activity' of the ensemble average forming P differs from that for P*. Second, there must be a more favorable free energy of ligand binding to P* *vs.* P, in order to drive the conformational switching. Finally, the P* state(s) have a higher (average) free energy than the P state(s), *i.e.* the absence of effector, P is the prevalent form. This last is implied otherwise one wouldn't need the effector to shift the equilibrium to P* in the first place. The energetic price of conformational switching to the higher free energy form comes from the ligand binding energy.

The original MWC insight that allosteric effects are transmitted by a conformational change has been broadened in several ways. More nuanced descriptions of allostery incorporate the fact that proteins are dynamic, so each 'state' is typically an ensemble of sub-states [19]. These sub-states will be populated according to the Boltzmann distribution and impart a significant conformational entropy to the protein. Allosteric effectors can act by binding preferentially to more than one sub-state. Re-establishment of the equilibrium Boltzmann distribution involves a shift in the population distribution [20]. The

* Tel.: +1 215 573 3506.
E-mail address: sharpk@mail.med.upenn.edu.

0301-4622/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.bpc.2011.05.007

basic principle is the same as in original MWC model: There must be at least two functionally different states, but there may well be more. The fact that proteins have considerable conformational entropy led to the idea that the effector could work by narrowing the conformational spread of states without changing the mean structure. This is known as allostery without conformational change [21]. In this model the prevalent, effector-free, form P has a lower free energy than P* due to its greater conformational entropy. When the ligand binds to form LP* the protein has a reduced conformational entropy. At a thermodynamic level this model retains the basic MWC principle: The cost of conformational switching from the lower free energy form to higher free energy form comes from the ligand binding energy.

Within this general framework, it is to some extent a matter of perspective as to whether one considers two effective states, P and P*, each with a certain conformational entropy, or subdivides these states into an ensemble of structures. Which viewpoint is most useful depends on whether there is a continuous distribution of state energies and conformations or whether the protein is found to switch between two (or more) well separated sets of states. High resolution structural studies of allosteric proteins, such as hemoglobin [22], lac repressor [7,23], adenylate kinase [24] and aspartamyl trans-carbamylase [25], reveal discrete structural states corresponding to P and P* or in the Hb terminology, to T and R respectively, although sub-ensembles of these states may be quenched or missed in the crystallization process. To date no intermediate structural states have been identified for the lac repressor.

More recently, a picture of allostery that is qualitatively different from the induced fit (IF) model has been introduced: This model has various names including population shift, conformational selection, ensemble/landscape shifting, sequestration and tertiary capture [1,20,26–30]. Here it will be referred to as population selection (PS). The features of the PS model, and evidence for it in a wide variety of cases have recently been reviewed in detail [30]. Of specific interest here, repressor sequestration has been postulated to occur in the lac operon system *in vivo* at low inducer concentrations [28]. Here sequestration refers to spontaneous dissociation of repressor from DNA with no ligand bound, followed by inducer binding to repressor, preventing the latter from rebinding DNA.

In brief, in the PS mechanism, the protein must again exist in at least two forms, one of lower free energy (P) and one of higher free energy (P*). In the absence of effector the latter, according to Boltzmann's principle, is present but with lower probability. The effector, when present, binds P* more tightly, so 'selecting' this minor component out as LP*. The depleted minor component P* is repopulated via thermal re-equilibration, more effector binds to it and so on, until the final equilibration establishes LP* as the prevalent form.

Examination of the IF and PS models shows that the difference between them is purely kinetic. In both models, in the absence of the effector the equilibrium protein is mostly P-form. In the presence of saturating effector, the equilibrium shifts towards LP*. Kinetically, in the IF model one goes from L + P to LP* via LP which then conformationally switches to LP*. In other words, the major flux pathway is:

$$P \rightarrow LP \rightarrow LP^*. \qquad \text{(InducedFit)}$$

In contrast, in the PS models the major flux pathway is:

$$P \rightarrow P^* \rightarrow LP^*. \qquad \text{(PopulationSelection)}$$

It should be noted that the MWC model, which is an equilibrium model, is equally consistent with either IF or PS mechanisms or both. Although there is now abundant evidence that a number of allosteric proteins and RNA's sample, in detectable amounts, the higher energy P* state(s) spontaneously at equilibrium in the absence of effector (for reviews see *e.g.* [1,30]), this alone does not establish the relative contributions of IF and PS mechanisms. This depends, via the relative rate constants, on the non-equilibrium fluxes.

Of course it is possible that significant contributions from both IF and PS mechanisms occur simultaneously. However, distinguishing which mechanism dominates in specific cases has proven difficult. Okazaki and Takada have constructed simplified models of the glutamine-binding protein in which, through suitable selection of parameters, either the population selection or induced fit mechanism dominates [29], but with the lack of suitable kinetic data it is not possible to know which mechanism is actually used by this protein. While equilibrium data is usually available, typically there is not enough kinetic data to distinguish categorically between the IF and PS mechanisms.

The allosteric lac–repressor–operator–inducer genetic switch has been extensively studied by genetics, mutational analysis, structural, thermodynamic and kinetic methods. Thus it is one system for which there is sufficient kinetic data to characterize the kinetics of the repressor–inducer–operator binding and address the question of allosteric mechanism. In this work, the relevant rate constants are used to set up the kinetic equations describing the time dependence of the concentration of all the bound and unbound species of R, I and O. The equations are integrated numerically to high precision in order to obtain the non-equilibrium, time-dependent description of the concentrations. Knowledge of the time dependence of all the species' concentrations allows one to compute the flux through any pathway during induction as a function of inducer concentration. The fluxes are then used to see whether population selection or induced dissociation or both, describes the allosteric mechanism and to gain general insight into the lac system as a genetic switch. Since the downstream events of gene expression have their own kinetics the kinetic differences between these mechanisms would affect the overall kinetic behavior of the system.

2. Methods and theory

2.1. Kinetic framework

Native lac repressor is a tetramer formed of a dimer of dimers. When a *single* operator site is present, the tetramer acts as a functional dimer. Structural and thermodynamic studies now typically use the engineered-dimer form of the repressor, which has nearly identical thermodynamic behavior, and is easier to work with [6,23]. Here the repressor will also be treated as a dimer that binds a single operator and up to two inducer molecules. The repressor (R), inducer (I) and operator (O) form a system of linked equilibria according to the scheme given in Fig. 1.

In this scheme K_{R-O}, K_{RI-O} and K_{RI2-O} are the equilibrium constants for binding of the repressor to the operator with 0, 1 or 2 inducer molecules bound, respectively. K_{I-R} and K_{I-RO} are the equilibrium constants for binding of inducer to free repressor and binding of inducer to operator bound repressor, respectively. Here it is assumed that the binding of the first and second inducer molecules to free R

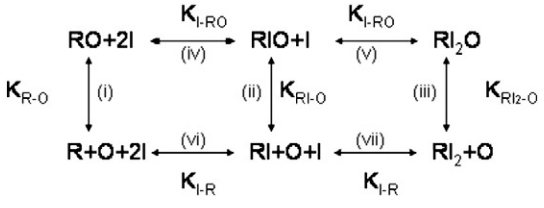

Fig. 1. Scheme for the linked binding of inducer (I) to repressor (R) and repressor to operator DNA (O). K_{R-O}, K_{RI-O} and K_{RI2-O} are the equilibrium constants for binding of the repressor to the operator with 0, 1 or 2 inducer molecules bound, respectively. K_{I-R} and K_{I-RO} are the equilibrium constants for binding of inducer to free repressor and binding of inducer to operator bound repressor, respectively. Each step is composed of a forward (association) and backward (dissociation) step characterized by the relevant second order association rate constant and first order dissociation constant respectively, as given in Table 1.

occurs with the same affinity. Similarly, binding of the two inducer molecules to repressor bound to operator (RO) is assumed to occur with the same affinity. This assumption is reasonable considering the small degree of cooperativity observed (Hill numbers of 1.2 for unbound R, and 1.4 for operator bound R [3,8,18]. Also, the latter slightly higher value is accounted for by coupled inducer binding-DNA dissociation, rather than direct cooperativity between inducer binding sites [8].

2.2. Experimental rate constants

The kinetic scheme has seven distinct steps and so is described by a total of 14 rate constants. There are also several requirements for thermodynamic consistency and consistency among measured equilibrium and kinetic data. These were applied as follows to obtain these 14 constants.

1) Equilibrium constants satisfy $K = k_a/k_d$ where k_a and k_d are the second order association and first order dissociation rate constants, respectively. Experimental values of any two of these three constants can be used to obtain k_a and k_d.
2) The assumption of identical binding constants for the 1st and 2nd inducers to repressor is assumed to apply also to the kinetics, *i.e.* steps (iv) and (v) have identical rate constants, as do steps (vi) and (vii). This reduces the number of independent kinetic constants to 10.
3) Closure of the thermodynamic cycles in Fig. 1 requires that a) $K_{R-O}/K_{RI2-O} = (K_{I-R}/K_{I-RO})^2$ and b) $K_{R-O}/K_{RI-O} = K_{I-R}/K_{I-RO}$. This reduces the number of independent rate constants from 10 to 8. The two closure relations imply a third, dependent relation, namely $K_{R-O}/K_{RI-O} = \sqrt{(K_{R-O}/K_{RI2-O})}$. The ratio of inducer binding constants to the induced form of repressor, *i.e.* R, and the repressing form, *i.e.* RO has been determined *in vitro* [4] and *in vivo* [8] with a range of $K_{I-R}/K_{I-RO} = 15–33$, leading to a K_{R-O}/K_{RI2-O} ratio in the range 225–1089. We choose a round number 'mid-range' value of $K_{R-O}/K_{RI-O} = 500$, and so $K_{I-R}/K_{I-RO} = \sqrt{500} = 22.4$.
4) Dunaway et al. [18] have measured the association rate constants of inducer to free repressor and operator-bound repressor to be 4.6×10^4/M/s and 1×10^4/M/s respectively. Their values for the accompanying dissociation rate constants are 0.2/s and 0.8/s respectively. These yield inducer dissociation equilibrium constants of 4.4 μM and 80 μM ($K_a = 2.3\times10^5$/M and $K_a = 1.25\times10^4$/M respectively) for free and operator bound repressor, respectively. The former dissociation constant has also been determined independently *in vitro* [4] and *in vivo* [8] with a tight range of $K_D^{RI} = 1/K_{I-R} = 3.9–4$ μM. The experimental data for free repressor–inducer binding are very close to self consistent. However, for the four kinetic constants involving inducer to be completely consistent with $K_{I-R}/K_{I-RO} = 22.4$, we used a slightly lower value of 0.812×10^4/M/s for association rate of inducer to operator-bound repressor (Table 1), giving $K_D^{RO-I} = 98.5$ μM ($K_a = 1.01\times10^4$/M).

Association of the lac repressor with the operator has been studied extensively by equilibrium and kinetic measurements, to the extent that there are redundant data for certain binding steps. However, there is some uncertainty in particular values. For example, dissociation constants for free repressor–operator binding vary over two orders of magnitude depending on length and type of DNA used, number of operator sites present, salt conditions, repressor oligomerization state, etc. Early studies give values as low as sub-picomolar [16,17], while more recent studies are in the range of 10 pM to sub-nM [3,4,6,18]. At the same time, there are no direct kinetic measurements for operator binding to repressor dimer with a *single* inducer molecule bound (Step (ii), Fig. 1). Rate constants for repressor–operator binding obtained from experiments in the absence of inducer relate to the R form. Experiments with saturating amounts of inducer then provide the corresponding rate constants for the RI_2 form of repressor. Data at intermediate inducer concentrations would reflect some mixture of R, RI and RI_2 species. To address these issues three different sets of internally consistent rate constants involving operator binding were used.

2.2.1. Set 1

We use the value of $k^a_{R+O\rightarrow RO} = 2\times10^9$/M/s from Dunaway et al. [18] and a value of $k^d_{RO\rightarrow R+O} = 6\times10^{-4}/s$ from Riggs Bourgeois and Cohn [14]. This gives $K_D^{RO} = 0.3$ pM, consistent with earlier equilibrium measurements [16]. From operator–repressor dissociation measurements with saturating inducer [17] we have $k_{RI_2O\rightarrow RI2+O}{}^d = 0.2$/s. The closure relation (3) above ($K_{R-O}/K_{RI2-O} = 500$) then gives $k_{RI_2+O\rightarrow RI2O}{}^a = 1.33\times10^9$/M/s. To obtain the operator association and dissociation rate constants for repressor with one inducer bound, the geometric means of the corresponding values for free repressor and two-inducer bound repressor were used, giving $k^d_{RIO\rightarrow RI+O} = 0.011$/s and $k^a_{RI+O\rightarrow RIO} = 1.63\times10^9$/M/s, which automatically satisfies the third closure requirement $K_{R-O}/K_{RI-O} = \sqrt{(K_{R-O}/K_{RI2-O})}$. Using geometric means for the RI specie values is equivalent to assuming that the effect on the various free energy differences is linear in number of inducers bound.

2.2.2. Set 2

We again use the value of $k^a_{R+O\rightarrow RO} = 2\times10^9$/M/s from Dunaway et al. [18] along with their measured R–O affinity of $K_D^{RO} = 10$ pM, which gives $k^d_{RO\rightarrow R+O} = 0.02$/s. From the closure relation $K_{R-O}/K_{RI2-O} = 500$, the RI_2–O affinity is 5 nM ($K_a = 2\times10^8$). Assuming that the 500-fold drop in affinity comes equally from the association and dissociation, we get $k_{RI_2O\rightarrow RI2+O}{}^d = 0.447$/s and $k_{RI_2+O\rightarrow RI2O}{}^a = 8.94\times10^7$/M/s. The

Table 1
Rate constants[a].

Step	Rate constant	Set 1	Set 2	Set 3
		$K_D^{RO} = 0.3$ pM	$K_D^{RO} = 10$ pM	$K_D^{RO} = 0.6$ pM
(i)	$k^a_{R+O\rightarrow RO}$	2×10^9[b]	2×10^9[b]	1.74×10^9[c]
(i)	$k^d_{RO\rightarrow R+O}$	$6\pm1\times10^{-4}$[d]	0.02[b]	$1.15\pm0.06\times10^{-3}$[c]
(iii)	$k_{RI_2+O\rightarrow RI2O}{}^a$	1.33×10^9[e]	8.94×10^7[e]	4.54×10^8[e]
(iii)	$k_{RI_2O\rightarrow RI2+O}{}^d$	0.2[c]	0.447[e]	0.15[c]
(iv), (v)	$k^a_{RO+I\rightarrow RIO}, k^a_{RIO+I\rightarrow RI_2O}$	8.12×10^3[e]	8.12×10^3[e]	8.12×10^3[e]
(iv), (v)	$k^d_{RIO\rightarrow RO+I}, k^d_{RI_2O\rightarrow RIO+I}$	0.8[b]	0.8[b]	0.8[b]
(vi), (vii)	$k^a_{R+I\rightarrow RI}, k^a_{RI+I\rightarrow RI_2}$	4.6×10^4[b]	4.6×10^4[b]	4.6×10^4[b]
(vi), (vii)	$k^d_{RI\rightarrow R+I}, k^d_{RI_2\rightarrow RI+I}$	0.2[b]	0.2[b]	0.2[b]

[a] Units: dissociation constants in s^{-1}, association constants in $M^{-1}s^{-1}$.
[b] Kinetic data taken from: [18].
[c] Kinetic data taken from: [17].
[d] Kinetic data taken from: [14].
[e] From thermodynamic closure relations.

association and dissociation rate constants for repressor with one inducer bound were again obtained from the geometric means of the corresponding values for free repressor and two-inducer bound repressor.

2.2.3. Set 3

We use the value of $k^a_{R+O\rightarrow RO}=1.7\times10^9$/M/s and $k^d_{RO\rightarrow R+O}=1.15\times10^{-3}$/s.from Barkley et al. [17], which gives $K_D^{RO}=0.6$ pM. Using $K_{R-O}/K_{RI2-O}=500$, the RI_2–O affinity is 0.33 nM ($K_a=3.0\times10^9$). Using a value of $k_{RI_2O\rightarrow RI2+O}{}^d=0.15$/s at 0.1 M inducer [17] we have $k_{RI_2+O\rightarrow RI2O}{}^a=4.54\times10^8$/M/s. The association and dissociation rate constants for repressor with one inducer bound were again obtained from the geometric means of the corresponding values for free repressor and two-inducer bound repressor.

The three sets of rate constants are summarized in Table 1. All of the publications from which data were taken showed mono-exponential kinetic plots, or stated that kinetic curves were simple mono-exponentials, implying straightforward first order dissociation and pseudo-first order association kinetics. Set 1 corresponds to values consistent with the higher range of measured repressor–operator affinities. Set 2 corresponds to values consistent with the lower range of measured repressor–operator affinities. Set 3 is intermediate between these two. Results from the three sets provide a measure of the sensitivity of the conclusions to experimental uncertainty in rate constants.

2.3. Simulation of rates and fluxes

Using the kinetic scheme of Fig. 1, the rate of change of concentrations of the five bound species is given by the first order differential equations:

$$\frac{\partial[RO]}{\partial t}=-k^d_{RO\rightarrow R+O}[RO]+k^a_{R+O\rightarrow RO}[R][O]+k^d_{RIO\rightarrow RO+I}[RIO]-2k^a_{RO+I\rightarrow RIO}[RO][I] \tag{1}$$

$$\frac{\partial[RIO]}{\partial t}=-k^d_{RIO\rightarrow RO+I}[RIO]+2k^a_{RO+I\rightarrow RIO}[RO][I]+2k^d_{RI_2O\rightarrow RIO+I}[RI_2O]-k^a_{RIO+I\rightarrow RI_2O}[RIO][I]-k^d_{RIO\rightarrow RI+O}[RIO]+k^a_{RI+O\rightarrow RIO}[RI][O] \tag{2}$$

$$\frac{\partial[RI_2O]}{\partial t}=-2k^d_{RI_2O\rightarrow RIO+I}[RI_2O]+k^a_{RIO+I\rightarrow RI_2O}[RIO][I]-k^d_{RI_2O\rightarrow RI_2+O}[RI_2O]+k^a_{RI_2+O\rightarrow RI_2O}[RI_2][O] \tag{3}$$

$$\frac{\partial[RI]}{\partial t}=-k^d_{RI\rightarrow R+I}[RI]+2k^a_{R+I\rightarrow RI}[R][I]+2k^d_{RI_2\rightarrow RI+I}[RI_2]-k^a_{RI+I\rightarrow RI_2}[RI][I]+k^d_{RIO\rightarrow RI+O}[RIO]-k^a_{RI+O\rightarrow RIO}[RI][O] \tag{4}$$

$$\frac{\partial[RI_2]}{\partial t}=-2k^d_{RI_2\rightarrow RI+I}[RI_2]+k^a_{RI+I\rightarrow RI_2}[RI][I]+k^d_{RI_2O\rightarrow RI_2+O}[RI_2O]-k^a_{RI_2+O\rightarrow RI_2O}[RI_2][O] \tag{5}$$

The factor of 2 accounts for the two indistinguishable sub-species of repressor where one inducer is bound to either monomer. The rates of change for the unbound species are obtained using mass conservation:

$$\frac{\partial[R]}{\partial t}=-\left(\frac{\partial[RO]}{\partial t}+\frac{\partial[RIO]}{\partial t}+\frac{\partial[RI_2O]}{\partial t}+\frac{\partial[RI]}{\partial t}+\frac{\partial[RI_2]}{\partial t}\right) \tag{6}$$

$$\frac{\partial[O]}{\partial t}=-\left(\frac{\partial[RO]}{\partial t}+\frac{\partial[RIO]}{\partial t}+\frac{\partial[RI_2O]}{\partial t}\right) \tag{7}$$

$$\frac{\partial[I]}{\partial t}=-\left(\frac{\partial[RIO]}{\partial t}+2\frac{\partial[RI_2O]}{\partial t}+\frac{\partial[RI]}{\partial t}+2\frac{\partial[RI_2]}{\partial t}\right). \tag{8}$$

Initial conditions were chosen as follows. From the work of Daber et al. [8], it is known that the *in vivo* expression level in the absence of inducer is typically 4% of maximal (*i.e.* that with no repressor present in the cell). This implies that the ratio [RO]/[O] is 96/4 = 24, so the initial free repressor concentration was set to $[R(t=0)]=24K_D^{RO}$, using the dissociation constant given in Table 1 for each data set. The total concentration of operator, [Otot], was set to 0.1 nM, with 4% as O, and 96% as RO at time zero. The value of 0.1 nM is nominal: since all the rate equations are linear in the various operator species concentrations, for a given initial [RO]/[O] ratio the ratios of fluxes through different pathways are independent of the chosen value of [Otot].

The initial concentrations of R, O and RO were thus set at equilibrium for the repressed state in the absence of the inducer. The initial concentration of inducer was then set to different values corresponding to a non-equilibrium perturbation. The rate behavior as the system returns to a new equilibrium in the presence of a given concentration of inducer then describes the kinetics of induction. Inducer concentrations were varied from 1 μM to 100 μM, the range known to cover full induction *in vivo* [8]. The rate equations were integrated over time using the Euler method [31] with a time step of 0.1 s until concentration ratios were within 1% of their final equilibrium values, as determined from the rate constants in Table 1. With the parameters in Table 1 this usually required less than 150 s of total simulation time. Shorter time steps of 0.05 s and 0.01 s gave almost identical results, indicating that numerical integration errors were negligible.

Fluxes for each of the seven steps in Fig. 1 as a function of time were obtained as the difference in the forward and backward rates for each step. For step (i) for example

$$J_{RO\rightarrow R+O}(t)=k^d_{RO\rightarrow R+O}[RO(t)]-k^a_{R+O\rightarrow RO}[R(t)][O(t)]. \tag{9}$$

Analogous equations hold for the other six steps. The total number of moles that flowed through a given step is the flux integrated over time, *e.g.* for step (i):

$$N_{RO\rightarrow R+O}=\int_0^\infty J_{RO\rightarrow R+O}(t)dt. \tag{10}$$

The total amount of repressor that is dissociated by inducer is given by

$$dN=[R(\infty)]+[RI(\infty)]+[RI_2(\infty)]-[R(0)]=N_{RO\rightarrow R+O}+N_{RIO\rightarrow RI+O}+N_{RI_2O\rightarrow RI_2+O}. \tag{11}$$

The first equality of Eq. (11) gives the total amount of repressor dissociated by the inducer using equilibrium quantities, just as in the MWC model. This has no information about the path of dissociation. In the second equality, however, the three flux terms represent the contributions of the three possible paths of dissociation: (i) spontaneous dissociation (no inducer bound), (ii) via single inducer molecule binding and (iii) via binding of two inducer molecules. In terms of the alternative mechanisms discussed in the introduction, they are identified with the three paths as follows. Dissociation without inducer, (followed by inducer binding) by definition corresponds to the repressor sequestration mechanism of Choi et al. [28]. It is also identified here with the population shift or tertiary capture mechanism. The essence of this mechanism is that the protein spontaneously undergoes an energetically unfavorable transition to a low probability state, then this state is bound by the ligand. Here the uphill transition is coming off the operator, since the initial equilibrium is well towards [RO].

To go through pathways ii) or iii) I must bind R before dissociation of the latter from O. However, there are two kinetically indistinguishable variants: 1) I binds to RO, followed by a conformational change in R to the high I affinity/low O-affinity form, followed by dissociation from O. 2) R changes to the high I affinity/low O-affinity form *without* dissociation from O and *before* I binds, then I binding and dissociation occur. Strictly speaking, only the first is induced fit. A more accurate term for both is induced dissociation (ID), which will be used hereon to describe repressor dissociation following binding of one or two inducer molecules.

With these definitions, the relative size of $(N_{RIO\rightarrow RI+O}+N_{RI_2O\rightarrow RI2+O})$ *vs.* $N_{RO\rightarrow R+O}$ quantifies the contribution of the induced dissociation mechanism *vs.* the population selection/repressor sequestration mechanism. The amount of released repressor, dN defines the amount of induction. The amount of induction can be expressed as a percentage of the maximum amount of repressor that could be bound, namely $100^*dN/[O_{total}]$.

3. Results

In Fig. 2, representative plots of flux *vs.* time are shown for kinetic parameter set 1. Panel (a) shows the flux through the three possible repressor dissociation pathways i–iii of Fig. 1, for an inducer concentration of 1 μM, well below the mid-point of the induction curve. Fluxes through all three pathways rise rapidly as inducer binds to both free and bound repressor, and gradually die away after ≈200 s as the new equilibrium is established. Flux through the RIO → RI + O dissociation pathway (ii) dominates for a very short initial time, followed by the direct dissociation pathway RO → R+O (i) as free repressor is converted to RI and RI_2. Panel (b) shows the corresponding fluxes for an inducer concentration of 100 μM, near the mid-point of the induction curve. Here the overall time course of re-equilibration is about 10-fold faster. Flux through the $RI_2O \rightarrow RI_2 + O$ pathway (iii) rises rapidly and dominates throughout, with a small contribution from RIO → RI + O, and almost no contribution from direct dissociation RO → R + O. Assuming that expression is proportional to the fraction of free operator the predicted expression level, obtained from the final equilibrium concentrations, is

$$e(I) = 100[O] / ([O] + [RO] + [RIO] + [RI_2O]) \quad (12)$$

expressed as a percentage of maximal possible expression. For the conditions of Fig. 2a and b, the calculated expression levels at [I] = 1 μM and [I] = 100 μM are 5%, and 47% respectively. Given a background level of 4% expression, these values correspond to induction levels of 1% and 43% respectively, obtained from either of the equivalent relationships e(I) − e(0) or 100 dN/[O_{total}].

Fig. 3 compares expression levels calculated from Eq. (12) using purely kinetic data from sets 1 and 3 and [I] = 1–1000 μM, with expression levels measured *in vivo* by Daber et al. [8]. Experimental expression levels were measured by intensity of fluorescence of green fluorescence protein under control of the lac operon, and normalized to maximal expression (fluorescence intensity from an *E. coli* strain with no repressor gene), and correspond to equilibrium conditions. Experimental expression levels do not derive from any of the kinetic or equilibrium constants used in Table 1. Calculated values from the two parameter sets bracket the experimental data. Considering that the experimental expression data is obtained by direct readout of expression *in vivo*, while the calculated values are obtained from kinetic parameters obtained from a range of *in vitro* studies over many years, the agreement is very good. The agreement provides evidence for the consistency and reliability of the kinetic parameters used here, and hence also for the conclusions about lac repressor kinetics to be drawn from them.

From the flux time courses, the integrated flux values for the three dissociation pathways were obtained using Eq. (10). These are plotted in Fig. 4 against the amount of induction, *i.e.* the normalized amount of repressor released from the operator. For parameter set 1, the initial induction is primarily via direct dissociation of repressor followed by sequestration of free repressor by inducer. At rising levels of induction, induced dissociation via binding of a single inducer molecule exceeds spontaneous dissociation. At induction levels of 10% and greater, induced dissociation via binding of two inducer molecules increasingly dominates. This occurs because of rapid formation of RI_2O at higher concentrations of inducer. Concomitantly, the absolute amount of repressor that dissociates spontaneously decreases, because this relatively slow pathway is unaffected by inducer concentration, and it is out-competed by pathway (iii), and to a lesser extent, by pathway (ii). For parameter sets 2 and 3 the qualitative picture is the same: the small amount of induction at lower inducer concentration has significant contribution from spontaneous dissociation, but the absolute amount is never large, and it becomes unimportant at higher levels of induction.

The largest area of uncertainty in this analysis is in the repressor–operator affinity. Parameter sets 1, 3 and parameter set 2 correspond

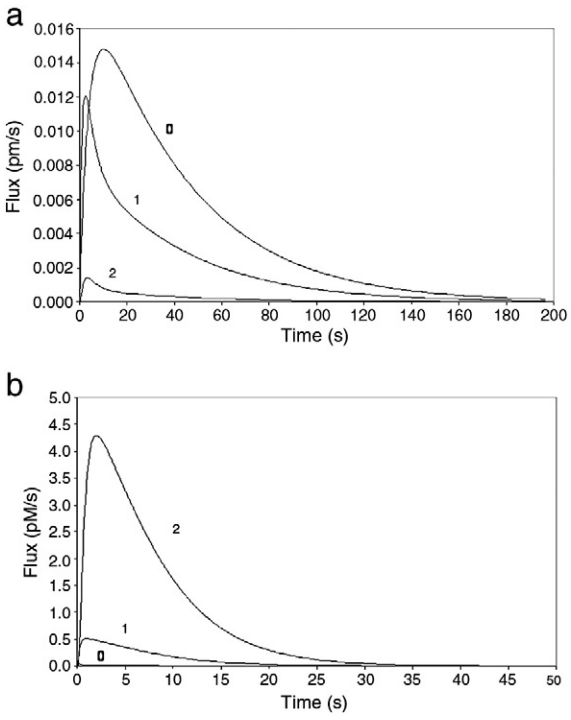

Fig. 2. Time course of fluxes for direct dissociation (no inducer bound) RO → R + O (0), via single inducer binding RIO → RI + O (1) and via double inducer binding $RI_2O \rightarrow RI_2 + O$ (2). (a) [I] = 1 μM. (b) [I] = 100 μM.

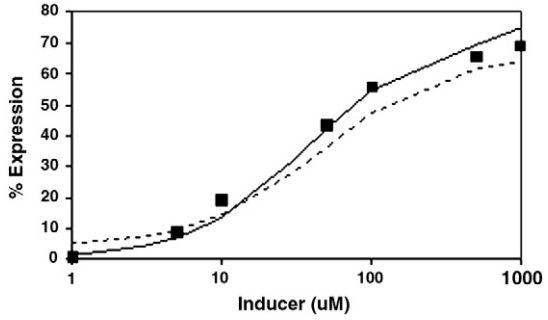

Fig. 3. Induction curve for expression from *in vivo* measurements of Daber et al. [8] (■). Calculated at equilibrium using kinetic parameter set 1 (···) and set 3 (—).

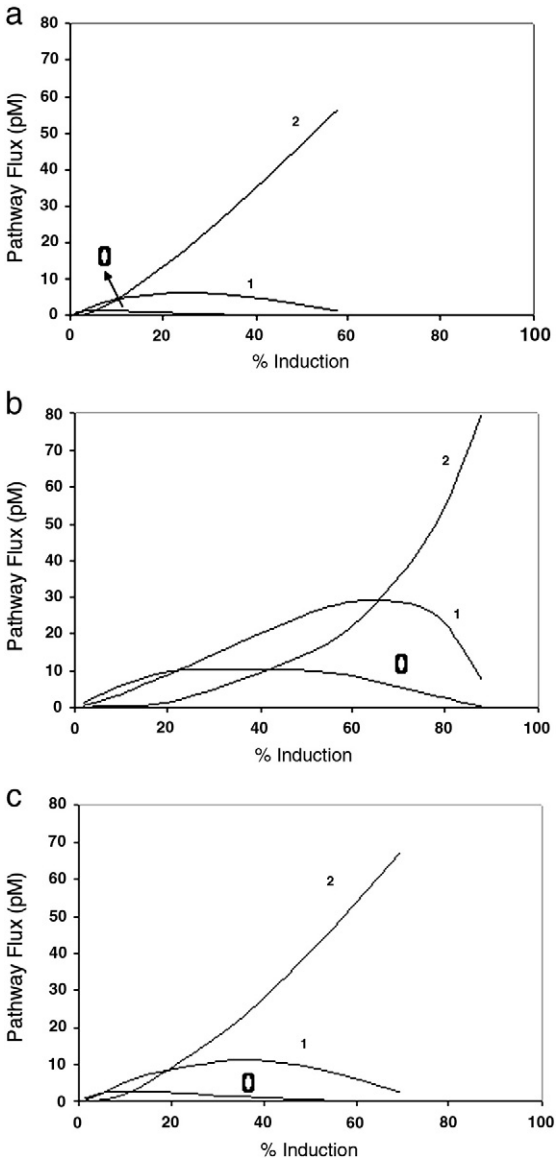

Fig. 4. Total pathway fluxes for direct dissociation (no inducer bound) (0), via single inducer binding $RIO \rightarrow RI + O$ (1) and via double inducer binding $RI_2O \rightarrow RI_2 + O$ (2) as a function of the total amount of induction produced by increasing levels of inducer. (a) Parameter set 1. (b) Parameter set 2. (c) Parameter set 3.

to the higher and lower ranges of affinities measured *in vitro*, respectively. For the lower affinity set, spontaneous dissociation of repressor makes a larger contribution throughout induction. This is because the lower affinity binding is due to a faster off rate for repressor binding (Table 1). Fig. 5 summarizes the relative importance

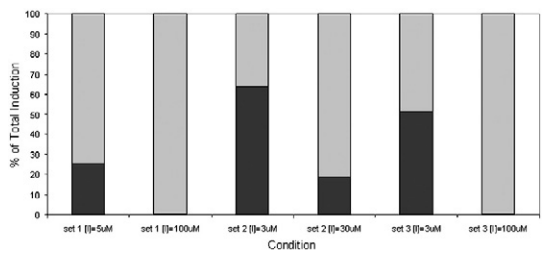

Fig. 5. Relative contributions of induced dissociation (shaded) and population selection (solid) to repressor dissociation sufficient to double background expression or produce half-maximal expression for the three kinetic parameter sets.

of the PS and ID mechanisms for the three parameter sets. The figure shows the relative contributions from the two mechanisms at two inducer levels: i) a low level of inducer sufficient to double the amount of expression over background, *i.e.* to 8%. ii) A 'mid-point' level of inducer sufficient for 50% expression. The value of [I] needed to achieve these expression levels varies slightly between parameter sets. The three parameter sets tell the same qualitative story. There is significant contribution from population selection at the lower level of expression, up to 64% for the lower affinity data set, but this mechanism is minor by the time the induction midpoint is reached.

4. Discussion

Using available kinetic and equilibrium data for the lac repressor–operator–inducer system, a detailed study of the kinetics of induction was performed in order to examine the relative importance of the induced dissociation and population selection mechanisms for induction. To account for the uncertainty in experimental data, principally in the repressor–operator affinity, three self consistent parameter sets were examined spanning a range of experimental data. Precise kinetic behavior could be obtained by straightforward integration of the first and second order rate equations for the kinetic scheme of Fig. 1 using each of the parameter sets. The relative contributions of population selection and induced dissociation were obtained by finding the total flux through the $RO \rightarrow R + O$ dissociation pathway *vs.* the $RO \rightarrow RIO \rightarrow RI_2O$, $RIO \rightarrow RI + O/RI_2O \rightarrow RI_2 + O$ pathways. For all the parameter sets, population selection was only important at low levels of induction (<20%). This is consistent with the inference of repressor sequestration by Choi et al., who studied bursts of expression from the lac operon in single cells at low inducer concentration [28]. However, even under these conditions inducer binding/induced dissociation is likely to be at least as important, since this is seen for all three parameter sets, even at micro-molar inducer (Fig. 5). At higher levels of inducer the induced dissociation mechanism dominates.

Insight into the overall qualitative features of the kinetics can be obtained by looking at the time constants for the key steps in dissociation of repressor: direct: $RO \rightarrow R + O$, and via inducer binding: $RO \rightarrow RIO \rightarrow RI_2O$ followed by $RI_2O \rightarrow RI_2 + O$. The time constants for dissociations are given by $\tau = 1/k^d$, where k^d is the appropriate first order dissociation constant from Table 1. These are invariant. The time constants for association are given by the appropriate $\tau = 1/([I]k^a)$, and these are inversely proportional to inducer concentration. These time constants are summarized in Table 2 for [I] = 1 μM and [I] = 100 μM. Very high affinity repressor operator binding is characterized by long time constant for direct dissociation. This is typically an order of magnitude longer than for the other key steps even at low inducer

Table 2
Key time constants (s)[a].

	Set 1	Set 2	Set 3
Dissociation			
$RO \rightarrow R + O$	1666	50	870
$RI_2O \rightarrow RI_2 + O$	5	2	7
$RI_2 \rightarrow RI \rightarrow R$	5	5	5
$RI_2O \rightarrow RIO \rightarrow RO$	1	1	1
Association			
$R + O \rightarrow RO$[b]	69	2	36
$RI_2 + O \rightarrow RI_2O$[b]	104	47	139
$R \rightarrow RI \rightarrow RI_2$ ([I] = 1 μM)	20	20	20
$R \rightarrow RI \rightarrow RI_2$ ([I] = 100 μM)	0.2	0.2	0.2
$RO \rightarrow RIO \rightarrow RI_2O$ ([I] = 1 μM)	112	112	112
$RO \rightarrow RIO \rightarrow RI_2O$ ([I] = 100 μM)	1	1	1

[a] Times >1 s rounded to nearest second.
[b] Time constant of pseudo first order rate for repressor binding to operator at initial free repressor concentration of $24K_D^{RO}$.

concentration. At higher inducer concentrations the ratio of time constants is even greater. Thus direct dissociation will be outcompeted under most conditions. For lesser, but still high affinity repressor operator binding, exemplified by parameter set 2, the key time constants are of the same order of magnitude at low inducer concentration, which is reflected in the fact that both PS and ID mechanisms contribute significantly here. At higher inducer concentrations, where induction is sizable, the ID mechanism still dominates. The PS mechanism contributes a fraction. This is due to the accelerating effect of inducer on the initially slow step: binding of inducer to the bound repressor.

It should also be noted that on a molecular time scale the time constants in Table 2 are long. The shortest are 0.2 s, most are seconds or longer. These time scales are comparable to those required to fold an entire protein [32]. So it is reasonable to assume that the allosteric conformational change undergone by lac, either spontaneously, or in response to (un)binding of O or I is fast compared to the association/dissociation event. Thus the PS pathway is some composite of dissociation and conformational change, but in any case it consists of the spontaneous transition of the repressor to a higher energy state before capture by the inducer.

There is one caveat to the present analysis, however. The data used here comes from the use of the gratuitous inducer, IPTG, used in the vast majority of studies of the lac repressor, not the natural inducer allolactose. So the specific amount of ID *vs.* PS mechanism found here applies to IPTG induction, whether *in vitro* or *in vivo*. The relative importance of ID and PS mechanisms might differ for allolactose if its association/dissociation constants are greatly different. However, as long as direct dissociation of repressor has a time constant ≈ 10 s or longer, and the other rate constants are similar to within an order of magnitude, and given that the PS mechanism contributes a fraction here, the ID mechanism would probably still dominate at high allolactose levels.

Finally, it is interesting to note that the lac repressor is constitutively leaky, *i.e.* that there is a low but significant level of expression in the absence of inducer [33]. This is functionally important *in vivo* for the bacteria, since the natural inducer allolactose is produced by the same enzymes controlled by the repressor. Over-tight repression effectively would result in a one way allosteric genetic switch. The leakiness occurs due to spontaneous dissociation of the repressor, so the system is predisposed, as it were, to allow capture of the unbound repressor when inducer does appear. In this regard it is not surprising in retrospect that there is significant induction via the PS mechanism at low inducer levels. One might anticipate that in other allosteric systems where the ligand-free equilibrium is further over to the P form, the PS contribution is smaller. One cannot, however, predict *a priori* the relative importance of this mechanism, nor the dominance of the ID mechanism at high ligand concentration, without a kinetic analysis of all the relevant steps such as that presented here.

Acknowledgments

The author would like to acknowledge frequent discussions of the lac repressor system with Professor Mitch Lewis, and members of his lab, Rob Daber, Leslie Milk and Matt Sochor.

References

[1] G. Varani, N. Leulliot, Current Topics in RNA–Protein Recognition: Control of Specificity and Biological Function through Induced Fit and Conformational Capture, 2001.
[2] J. Monod, J. Wyman, J.-P. Changeux, On the nature of allosteric transitions: a plausible model, J. Mol. Biol. 12 (1965) 88–118.
[3] T. Daly, K. Matthews, Allosteric regulation and operator binding to the lactose repressor, Biochemistry 25 (1986) 5479–5484.
[4] R. O'Gorman, J. Rosenberg, O. Kallai, R. Dickerson, K. Itakura, A. Riggs, K. Matthews, Equilibrium binding of inducer to lac repressor. Operator DNA, J. Biol. Chem. 255 (1980) 10107–10114.
[5] C.J. Wilson, H. Zhan, L. Swint-Kruse, K. Matthews, The lactose repressor system: paradigms for regulation, allosteric behavior and protein folding, Cell. Mol. Life Sci. 64 (2007) 3–16.
[6] L. Swint-Kruse, H. Zhan, K.S. Matthews, Integrated insights from simulation, experiment, and mutational analysis yield new details of LacI function, Biochemistry 44 (2005) 11201.
[7] R. Daber, S. Stayrook, A. Rosenberg, M. Lewis, Structural analysis of lac repressor bound to allosteric effectors, J. Mol. Biol. 370 (2007) 609–619.
[8] R. Daber, K. Sharp, M. Lewis, One is not enough, J. Mol. Biol. 392 (2009) 1133–1144.
[9] F. Jacob, J. Monod, Genetic regulatory mechanisms in the synthesis of proteins, J. Mol. Biol. 3 (1961) 318–356.
[10] C.A. Cronin, W. Gluba, H. Scrable, The lac operator–repressor system is functional in the mouse, Genes Dev. 15 (2001) 1506–1517.
[11] M.C. Hu, N. Davidson, The inducible lac operator–repressor system is functional in mammalian cells, Cell 48 (1987) 555–566.
[12] L. Caron, M. Prot, M. Rouleau, M. Rolando, F. Bost, B. Binetruy, The Lac repressor provides a reversible gene expression system in undifferentiated and differentiated embryonic stem cell, Cell. Mol. Life Sci. 62 (2005) 1605–1612.
[13] A.V. Lee, C.N. Weng, S.E. McGuire, D.M. Wolf, D. Yee, Lac repressor inducible gene expression in human breast cancer cells in vitro and in a xenograft tumor, Biotechniques 23 (1997) 1062–1068.
[14] A. Riggs, S. Bourgeois, M. Cohn, The lac repressor operator interaction III. Kinetic studies, J. Mol. Biol. 53 (1970) 401–417.
[15] A. Riggs, R. Newby, S. Bourgeois, Lac repressor–operator interaction II. Effect of galactosides and other ligands. J. Mol. Biol. 51 (1970) 303–314.
[16] A. Riggs, H. Suzuki, S. Bourgeois, Lac repressor–operator interaction I. Equilibrium studies. J. Mol. Biol. 48 (1970) 67–83.
[17] M. Barkley, A. Riggs, A. Jobe, S. Bourgeois, Interaction of effecting ligands with lac repressor and repressor–operator complex, Biochemistry 14 (1975) 1700–1712.
[18] M. Dunaway, J. Olson, J.M. Rosenberg, O.B. Kallai, R. Dickerson, K. Matthews, Kinetic studies of inducer binding to lac repressor–operator complex, J. Biol. Chem. 255 (1980) 10115–10119.
[19] K. Gunasekaran, B. Ma, R. Nussinov, Is allostery an intrinsic property of all dynamic proteins? Proteins 57 (2004) 433–443.
[20] V.J. Hilser, D. Dowdy, T.G. Oas, E. Freire, The structural distribution of cooperative interactions in proteins: analysis of the native state ensemble, PNAS 95 (1998) 9903–9908.
[21] A. Cooper, D.T.F. Dryden, Allostery without conformational change. A plausible model, Eur. Biophys. J. 11 (1984) 103–109.
[22] M.F. Perutz, G. Fermi, B. Luisi, B. Shaanan, R.C. Liddington, Stereochemistry of cooperative mechanisms in hemoglobin, Acc. Chem. Res. 20 (1987) 309–321.
[23] C.E. Bell, M. Lewis, A closer view of the conformation of the Lac repressor bound to operator, Nat. Struct. Biol. 7 (2000) 209–214.
[24] C.W. Muller, G.J. Schlauderer, J. Reinstein, G.E. Schulz, Adenylate kinase motions during catalysis: an energetic counterweight balancing substrate binding, Structure 4 (1996) 147–156.
[25] D.E.J. Koshland, The structural basis of negative cooperativity: receptors and enzymes, Curr. Opin. Struct. Biol. 6 (1996) 757761.
[26] L. Leder, Spectroscopic, calorimetric, and kinetic demonstration of conformational adaptation in peptide-antibody recognition, Biochemistry 34 (1995) 16509–16518.
[27] H. Bosshard, Molecular recognition by induced fit: how fit is the concept? News Physiol. Sci. 16 (2001) 171–173.
[28] P. Choi, L. Cai, K. Frieda, X.S. Xie, A stochastic single-molecule event triggers phenotype switching of a bacterial cell, Science 322 (2008) 442–446.
[29] K. Okazaki, S. Takada, Dynamic energy landscape view of coupled binding and protein conformational change: induced-fit versus population-shift mechanisms, PNAS 105 (2008) 11182–11187.
[30] D. Boehr, R. Nussinov, P.E. Wright, The role of dynamic conformational ensembles in biomolecular recognition, Nat. Chem. Biol. 3 (2009) 1–8.
[31] W. Press, B. Flannery, S. Teukolsky, W. Vetterling, Numerical Recipes, Cambridge University Press, Cambridge, 1986.
[32] L. Hoang, S. Bedard, M.M.G. Krishna, Y. Lin, S.W. Englander, Cytochrome c folding pathway: kinetic native-state hydrogen exchange, Proc. Natl. Acad. Sci. U. S. A. 99 (2002) 12173–12178.
[33] R. Daber, M. Lewis, A novel molecular switch, J. Mol. Biol. 391 (2009) 661–670.