GENE REGULATION IN THE *LAC* OPERON

by

Kathryn Grace Patterson

A dissertation submitted in partial fulfillment
of the requirements for the degree

of

Doctor of Philosophy

in

Mathematics

MONTANA STATE UNIVERSITY
Bozeman, Montana

July, 2009

APPROVAL

of a dissertation submitted by

Kathryn Grace Patterson

This dissertation has been read by each member of the dissertation committee and has been found to be satisfactory regarding content, English usage, format, citations, bibliographic style, and consistency, and is ready for submission to the Division of Graduate Education.

Dr. Tomáš Gedeon

Approved for the Department of Mathematical Sciences

Dr. Kenneth L. Bowers

Approved for the Division of Graduate Education

Dr. Carl A. Fox

STATEMENT OF PERMISSION TO USE

In presenting this dissertation in partial fulfillment of the requirements for a doctoral degree at Montana State University, I agree that the Library shall make it available to borrowers under rules of the Library. I further agree that copying of this dissertation is allowable only for scholarly purposes, consistent with "fair use" as prescribed in the U.S. Copyright Law. Requests for extensive copying or reproduction of this dissertation should be referred to ProQuest Information and Learning, 300 North Zeeb Road, Ann Arbor, Michigan 48106, to whom I have granted "the exclusive right to reproduce and distribute my dissertation in and from microform along with the non-exclusive right to reproduce and distribute my abstract in any format in whole or in part."

Kathryn Grace Patterson

July, 2009

ACKNOWLEDGEMENTS

I would like to thank my advisor Tomáš Gedeon for his guidance, support and all of the opportunities he has provided. I am also indebted my committee: Jack Dockery, Mark Pernarowski, Marcy Barge and Ken Bowers for their suggestions, encouragement and time reading this work. In addition, this dissertation would not have been possible without the Mathematics faculty at Montana State University. I am grateful for all of your time, patience and encouragement.

I would also like to thank Konstantin Mischaikow for his guidance, support and inviting me to visit Rutgers. My semester there was a wonderful experience.

NSF-IGERT provided funding and structure to begin my interdisciplinary studies. NSF/NIH grant W0467 and NSF grant DMS-0818785 provided funding for this project.

My family and friends provided encouragement and support throughout this process. I am particularly grateful to my husband Michael for doing everything else, so I could write.

TABLE OF CONTENTS

TABLE OF CONTENTS – CONTINUED

TABLE OF CONTENTS – CONTINUED

LIST OF TABLES

Table Page

LIST OF TABLES – CONTINUED

Table Page

LIST OF TABLES – CONTINUED

Table Page

LIST OF FIGURES

Figure Page

LIST OF FIGURES – CONTINUED

Figure Page

LIST OF FIGURES – CONTINUED

Figure Page

LIST OF FIGURES – CONTINUED

Figure Page

ABSTRACT

The *lac* operon, a jointly controlled series of genes in the bacteria E. coli, has been studied extensively since the 1940's. The *lac* operon genes are transcribed and then translated into proteins necessary for transport and digestion of lactose. The operon is activated in the presence of lactose after glucose, the preferred carbon source, has been expended.

In this thesis, we introduce a biophysical model using the Shea-Ackers framework for modeling promoter dynamics. The model spans two scales: the inputs are biophysical parameters of molecular interactions and the result is a level of gene expression - a macroscopic behavior of the cell. We include all experimentally suggested control mechanisms into the model, even though the experimental evidence is stronger for some of these mechanisms than others. We compare our model to experimental data and explore the individual contribution of the proposed mechanisms by removing them one by one and testing the reduced model's fit to the data. Finally, we find a minimal model which faithfully represents the available data, yet includes only the minimal number of control mechanisms.

1

INTRODUCTION

The central dogma of molecular biology states that the information flows from DNA[1] to messenger RNA (mRNA) to protein. In all cells, the genetic information is stored as DNA. The basic mechanism of information transfer in every living organism requires a gene, stored on the DNA, to be transcribed by RNA polymerase into mRNA, and then translated by ribosomes into protein. This process can be regulated at multiple points, the earliest of which is transcriptional regulation.

This thesis is focused on modeling transcription of the *lac* operon, a set of metabolic genes found in the prokaryote, *Escherichia coli* (E. coli). We first introduce a method for modeling transcription at the biophysical level and then introduce the biological components of the *lac* operon.

Transcription

In prokaryotes, transcription initiation can be divided into three steps [43, 65]:

- (*binding*) RNA polymerase binds to promoter DNA yielding a closed RNA polymerase-promoter complex,
- (*opening*) RNA polymerase unwinds a short segment of DNA yielding an open RNA polymerase-promoter complex,
- (*escape*) after abortive cycles of synthesis and release of short RNA products, the RNA polymerase escapes the promoter and enters into productive synthesis of RNA.

[1] As Joshua Lederberg stated in his Nobel Lecture in 1959, "No reader who recognizes deoxyribonucleic acid will need to be reminded what DNA stands for."

Considerable progress has been made in understanding the biochemistry of the various reactions in the process [7, 30, 43, 53, 66] and it is clear that while the three steps are physically coupled there is substantial freedom for independent control. Every step depends on the DNA sequence and its binding affinity to controlling agents such as protein.

Binding Control

Activation and repression of transcription initiation are primarily caused by the binding of regulatory proteins to the DNA structure in a region called a promoter or regulatory region. Assuming all chemical reactions on the DNA involving the regulatory proteins and RNA polymerase (RNAP) equilibrate on a much faster time scale than transcription, Shea and Ackers [77] have constructed a nonlinear model that describes the statistical equilibrium probability of RNAP binding to the DNA. This is a reasonable assumption in bacteria since the first time scale is on the order of milliseconds and the other on the order of minutes [6]. The inputs into this model are the binding energies of various transcription factors to DNA, their cooperative energies and the energy of DNA bending (if known). The Shea-Ackers model provides a broadly accepted quantitative framework [15] and has been experimentally validated on a variety of genes [5, 26, 71, 72, 77]. In this approach, a *state* s of the promoter is a particular configuration of transcription factors (or their absence) on the promoter, that is admissible in view of possibly overlapping binding sites. Then the probability of the occurrence of a particular state s from the set of admissible states $\mathcal{S}$ is

$$\mathbb{P}(s) = \mathbb{P}(s)([P], [r_1], \ldots, [r_m]) = \frac{K_B(s)[P]^{\alpha_s}[r_1]^{\alpha_s^1}[r_2]^{\alpha_s^2}...[r_m]^{\alpha_s^m}}{Z}, \tag{1}$$

where

$$K_B(s) := e^{-\frac{E_s}{RT}} \tag{2}$$

and the partition function Z, representing all possible states, is given by

$$Z([P],[r_1],\ldots,[r_m]) = \sum_{s\in\mathcal{S}} K_B(s)[P]^{\alpha_s}[r_1]^{\alpha_s^1}[r_2]^{\alpha_s^2}\ldots[r_m]^{\alpha_s^m}. \tag{3}$$

In these formulas E_s denotes the change in Gibbs free energy between the empty state, $s_\emptyset$, and the state $s \in \mathcal{S}$ under the normalization $E_{s_\emptyset} = 0$. The exponents α_s^i indicate the number of r_i molecules bound to the regulatory region in state s and similarly, α_s denotes the number of RNAP molecules, $[P]$ bound to the regulatory region in state s. As is standard, T denotes the temperature and R is the universal gas constant [32]. In Appendix B we derive equation (1) from principles of statistical mechanics and in Appendix C we derive the relationship between the Gibbs free energy and the association and disassociation rates of the appropriate chemical reactions.

In the original model [77] each state s which contains a bound RNAP is associated with a transcription initiation rate $k(s)$. This constant incorporates both the opening and escape processes and must be fit to data. The overall transcription rate for the gene is then

$$f([P],[r_1],\ldots,[r_m]) = \sum_{s\in\mathcal{S}} k(s)\mathbb{P}(s). \tag{4}$$

Since it is very difficult to distinguish experimentally between initiation rates corresponding to different states, it is often the case that only one rate is assumed for all states with bound RNAP. It should be emphasized that $k(s)$ and E_s in (2) are both functions of DNA structure and it is clear that binding and escape have a nonlinear correlation. For escape to occur, RNAP must break the initial bond, with energy E_s, formed with the promoter region. Therefore, as the binding energy, E_s becomes stronger, the escape rate decreases. Likewise, as E_s decreases, the escape rate increases.

Opening and escape are often modeled by simple rate constants, but the actual process is much more dynamic. There has been some effort to develop a more detailed

biophysical model for these processes. Included below is a synopsis of an existing opening model, and an escape model which is in development.

Opening

Perhaps the best current model of the opening process is due to Djordjevic and Bundschuh [24]. It uses the same chemical equilibrium assumption that we use to model the binding process. They developed a two-step model that computes the opening probability based on promoter sequence information. The first step describes the melting of the -10 box on the DNA ($\sim$5 base pairs) and the second step models the extension of the transcription bubble and the insertion of the DNA template strand into the active site channel of RNAP. When they tested the model against experimental data, they found the model to be in good agreement with experimentally determined transcription start sites found in the genome.

Escape

My collaborators (Traldi, Mischaikow, and Gedeon) are currently developing a method for modeling the escape rate, the third step of transcription initiation. Inspired by the kinetic model developed by Xue et al. [87], they incorporate the use of discrete Markov chains to represent the probability of escape given a DNA sequence. This method depends on the binding energies between base pairs spanning the DNA double helix (DNA-DNA energies) and the energies between the DNA template and the newly created mRNA strand (DNA-RNA energies).

This model is consistent with existing elongation models [8,82] which simulate the next step after transcription initiation, making it possible to combine the initiation model with an elongation model for a complete model of the entire transcription process.

Basic Biology of the *lac* Operon

We will use a framework similar to the Shea-Ackers model (equation (4)) to describe transcription in the *lac* operon. We start by describing its biology and transcriptional regulation.

The lactose operon (also known as the *lac* operon) is a set of metabolic genes found in E. coli. These genes are specific to importing and digesting the sugar lactose, and are only 'turned on,' or activated, when lactose levels outside the cell are high and glucose levels outside the cell are low. E. coli prefers the sugar glucose and as a result will not activate the genes to digest lactose until there is a sufficiently high level of external lactose and a sufficiently low level of preferred sugar. The *lac* operon, along with the phage λ, are the best understood genetic control systems. François Jacob, André Lwoff and Jacques Monod were awarded the Nobel Prize in Physiology or Medicine in 1965 "for their discoveries concerning genetic control of enzyme and virus synthesis" [2], the majority of which came from these two systems. Lwoff focused on induction of the phage λ while Monod and Jacob studied transcriptional regulation of the *lac* operon, but all three were at the Pasteur Institute at the same time, working in the same corridor discussing current experiments and hypotheses [42]. Many transcriptional control mechanisms were first described in these two systems, enabling the first genes to be copied and visualized [52].

Significance of the *lac* Operon

E. coli is a common, unicellular prokaryote (a simple organism that lacks a nucleus) which is easy to keep in the lab and whose maximal generation time is about 20 minutes in a good medium. When grown on glucose and lactose, E. coli exhibits the following phenotypical trait. When glucose is depleted, the growth rate stops for

several minutes and then resumes utilizing lactose instead of glucose. This switch in substrate utilization caught the attention of Jacques Monod and predates the discovery of DNA.

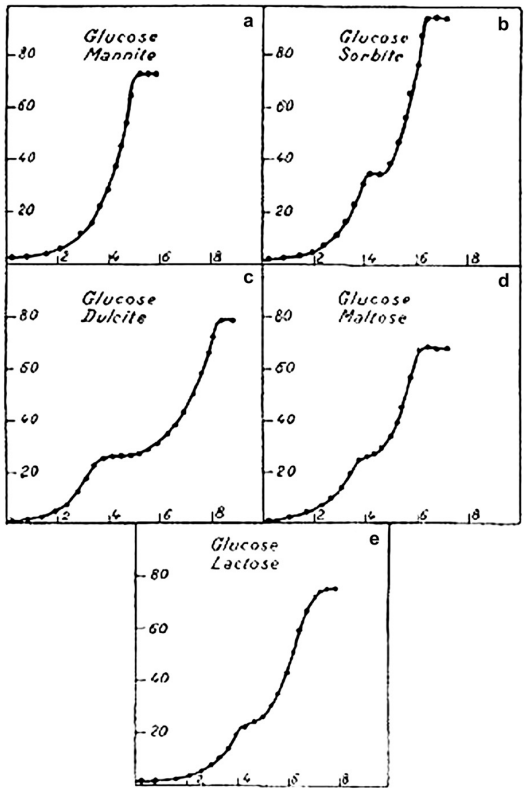


Figure 1: Diauxie from Jacques Monod's doctoral thesis [50] displaying initial exponential growth on glucose, followed by a delay in growth (the central plateau) while genes are initialized and finishing with another exponential growth phase as the second sugar is digested. Note in figure **a** that there is no diauxic growth, and how the length of time until the second growth phase varies depending on the sugar for figures **b-e**.

Near the end of Jacques Monod's thesis, he discussed with André Lwoff the following strange observation: Bacteria grown on glucose, could not continue to grow on some carbon sources (like lactose) without a plateau in the exponential growth. If the sugars were added in the other order, i.e. first grown on lactose and then glucose, growth was not affected. He named this phenomenon *diauxic growth*. Figure 1, from

Monod's thesis [50], displays bacterial population growth on the indicated sugar combinations. Diauxic growth can be observed in **b-e**, but not in **a**. Though unknown to Monod at the time of the experiment, the time needed to induce genes for digestion of the secondary sugar [52] causes the plateau phase of diauxic growth. Since there is now over 60 years of experimental data on the system, the *lac* operon is one of the best documented gene regulatory networks in molecular biology.

In addition, although the *lac* operon is a simple system, it shares common features with other systems. In fact, the study of this operon played a key role in our understanding of how universal gene control mechanisms are and how they can be transported between very different organisms. In his 1996 book, "The *lac* Operon: A Short History of a Genetic Paradigm," [52] Dr. Benno Müller-Hill describes how similar the *lac* repressor is to:

1. nine other E. coli repressors in total sequence;

2. three other binding proteins which contain similar core sequences responsible for binding the inducer; and

3. at least five other proteins (three of which are found in eukaryotes) which have the same 3-D structure even though the sequences differ.

As we compile the information from the genome projects for the various organisms (animal, plant, fungi, and so on), it is apparent that there are classes of similar protein domains whose members are likely descended from a common ancestor [17]. Consider the E. coli genome which has 4.7×10^6 base pairs. Since it takes three base pairs to encode one amino acid, E. coli is capable of producing about 1.5×10^6 amino acids resulting in about 5000 different proteins having a molecular weight of 30,000 daltons. If every protein belongs to a family as large as the *lac* repressor, then there

are only 300 protein structures in E. coli with a few possible functions each, a much more manageable number to study than 5000. By knowing and understanding the *lac* operon, its three genes and their products (four if we include the repressor), one could understand about 1% of the E. coli proteins [52].

From a mathematical standpoint, the dynamics of the *lac* operon provide a non-trivial, biological example of bistability. There are molecules (typically referred to as artificial inducers) which activate parts of the same pathway as lactose. When a population of E. coli is bathed in a mix of glucose and artificial inducer, a portion of the population will express the genes in the *lac* operon while the remaining portion will not, i.e. the population is bistable. However, when the artificial inducer is replaced with lactose, bistability is lost. How does the cell create a bistable switch? What are the benefits? Is it possible to force the *lac* operon to be bistable on lactose by changing the gene regulation? Scientists are actively trying to answer these questions.

The *lac* operon has a long, biological history, is mathematically interesting, and serves as a genetic paradigm with the potential for broad applications.

System Description

In order for a prokaryote such as E. coli to make a protein, RNA polymerase (RNAP) must bind the promoter region, separate the two strands of DNA (form an open complex), and then break away from the initial promoter region (escape) to begin transcription. Since E. coli does not have a nucleus, ribosomes bind the nascent mRNA strand and immediately begin translating the mRNA into protein. Consequently, transcription and translation (making protein from mRNA) occur side by side in the cytoplasm.

The *lac* operon is composed of a set of three genes (lacZ, lacY, and lacA) and their regulatory region. Immediately preceding the *lac* operon on the E. coli genome

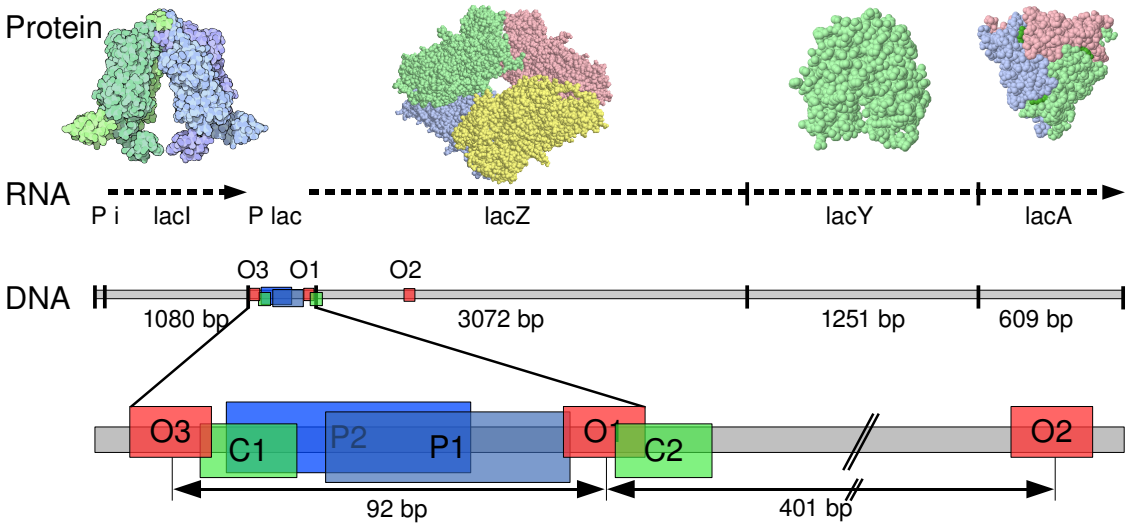


Figure 2: Cartoon image describing the *lac* repressor coding region and the *lac* operon. In the E. coli genome the DNA coding for a *lac* repressor subunit is preceded by a promoter region, P_i and immediately followed by the *lac* operon. The *lac* operon consists of a regulatory region and the lacZ, lacY, and lacA genes. As shown at the bottom of the figure, the regulatory region is composed of multiple binding sites: the P1 and P2 promoter regions bind RNAP (there are also at least two other sites P3 and P4, but these bind RNAP very weakly); C1 and C2 bind CAP; and O1, O2 and O3 bind the *lac* repressor. O1 and O3 are separated by 92 base pairs, and O1 and O2 are separated by 401 base pairs. Once RNAP binds a promoter region and escapes, it produces a RNA copy of all three genes. The RNA is then translated into a polypeptide chain which is then folded into protein. The 3-D structure of each protein is displayed above its respective gene, and has been generated from the RCSB Protein Data Bank [14]. From left to right are the *lac* repressor tetramer [28], β-galactosidase [64] (also a tetramer), lactose permease [4], and galactoside acetyltransferase [3] (a trimer).

is the promoter region and gene for the *lac* repressor (see Figure 2). A wildtype *lac* repressor protein is a tetramer which means it is composed of four monomer units of the protein.

The genes of the *lac* operon, shown on the DNA in Figure 2, store the genetic information for the proteins involved in transporting and metabolizing the sugar lactose. LacZ, codes for β-galactosidase, an enzyme which converts lactose to allolactose (the natural inducer for the operon), or cleaves lactose into glucose and galactose.

(Allolactose, glucose and galactose are metabolized in the cell.) The next gene, lacY codes for permease, a membrane protein which transports external lactose into the cell. The final gene in the *lac* operon, lacA, codes for transacetylase, an enzyme not directly involved lactose metabolism. Since lacA and transacetylase play no role in the self regulation of the *lac* operon, they will not be included in further discussion. If enough external lactose is present, permease will pump it in, β-galactosidase will convert it to allolactose which will bind the repressor, removing the repressor from the regulatory region of the gene, thus creating more permease and β-galactosidase.

lac Regulation

Consider the regulatory region as shown in Figure 2. In order for transcription to take place, RNAP must bind the promoter at either P1 or P2 as shown in blue, and then form an open complex and escape the promoter region to begin transcription. There are two regulatory proteins which can activate or repress transcription by binding to the regulatory region of the operon: **c**atabolite gene **a**ctivator **p**rotein (CAP) and the *lac* repressor. Note that the **c**atabolite gene **a**ctivator **p**rotein (CAP) and the **c**AMP **r**eceptor **p**rotein (CRP) are used interchangeably in the literature. To create a notational distinction between the protein without cyclic adenosine monophosphate (cAMP) bound (which does not bind the regulatory region), and the complex with cAMP bound, from here on CRP will refer to the cAMP receptor protein without cAMP, and CAP will refer to the protein complex.

Activator CAP: cAMP is a commonly found signalling molecule occurring in many species. In E. coli, glucose transport is one pathway which affects intracellular cAMP concentration. The enzyme IIA^{Glc}, involved in bringing glucose into the cell, will accept a phosphate group. When an E. coli cell is growing in an environment

abundant in glucose, the IIA^{Glc} enzyme is dephosphorylated (a phosphate group is removed) as glucose is transported into the cell. When glucose is depleted from the environment, phosphorylated IIA^{Glc} builds up in the cell and up-regulates adenylate cyclase. An increase in adenylate cyclase will increase cAMP synthesis triggering the formation of CAP complexes [23, 29].

A CAP complex is formed when cAMP binds the CRP protein, and is involved in activating more than one hundred promoters [10, 46] including the *lac* operon. Although it is not a part of the *lac* operon, CAP has a large effect on the rate at which the non-repressed operon is transcribed. Once bound to the C1 binding site (in green in Figure 2) on the DNA, CAP activates the *lac* operon in two ways:

1. it reduces the dissociation constant of RNAP, which enhances binding by about 4-fold and

2. it helps stabilize the RNAP-DNA open complex at the start of the process of transcription, leading to a 15-20 fold [48] total increase in transcription.

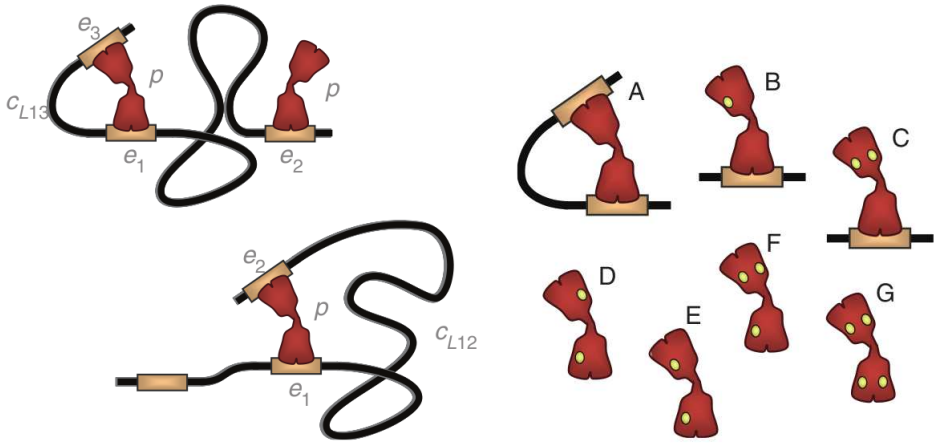
On the other hand, if CAP binds C2, it competes with the repressor for position on the DNA. However, CAP binds C2 weaker than the repressor binds O1 and there is little evidence to suggest that C2 plays a major regulatory role in this system [39,46]. Similarly, there are promoter sites P3 and P4 whose existence has been inferred from the sequence data, but have not been shown to affect the system. C2, P3 and P4 will not be mentioned further.

Repressor: As the name suggests, the *lac* repressor is the regulatory protein which represses the *lac* operon by binding the DNA. It is constitutively expressed, that is to say, the level of protein production is constant and independent of the *lac* operon genes. The primary operator or binding site for the *lac* repressor is O1, but

there are also two additional binding sites, O2 and O3 (all repressor binding sites are in red in Figure 2) which the repressor binds with decreasing affinity. If a repressor is bound to O1, transcription cannot occur, but a repressor bound to O2 or O3 is unable to block transcription. Why then do O2 and O3 exist? The *lac* repressor is composed of two DNA binding domains where each domain is a dimer. Thus, each repressor can bind two operators at once, forming a DNA loop, greatly reducing transcription and increasing repression.

In the presence of the inducer allolactose, each monomeric unit of the repressor tetramer can bind one inducer molecule. This causes a conformal change which greatly reduces or completely removes the DNA binding affinity. Figure 3a and Figure 3b, taken from [69], illustrate repressor-DNA looping conformations and the various types of inducer impaired repressors, respectively. Each yellow dot in Figure 3b represents an inducer molecule bound to a repressor subunit. For the impaired inducers in Figure 3b (B-G), as long as one binding domain remains free of inducer, the repressor is able to bind an operator as displayed in B and C. The inducer-repressor combinations in Figure 3b (D-G) have a greatly reduced affinity (or no affinity) for the DNA operators.

The *lac* Repressor and DNA Looping: Although originally discovered in the E. coli *ara* operon [68], the effect of looping on gene regulation has been observed in many regulatory processes in eukaryotes (cells with a nucleus) and prokaryotes. Two examples of systems utilizing DNA looping are the *lac* operon and the λ phage in E. coli. They each utilize looping to increase repression, but the distance in base pairs between two looped operator sites is much further in the phage (2400 base pairs) than in the *lac* operon (401 and 92 base pairs). In the phage, the energy to loop the DNA decreases the total free energy of the protein-DNA complex. Interestingly, in the *lac*



(a) Repressor mediated DNA looping (b) Repressor-Inducer Binding

Figure 3: Repression is greatly increased by repressor mediated DNA looping as shown on the left, however looping and repression can be abolished by the addition of an inducer. The yellow dots represent inducer molecules bound to repressor subunits. These figures are from Saiz and Vilar [69], and they represent the energy of a repressor binding O1, O2 and O3 by e_1, e_2 and e_3, respectively.

operon, the free energy needed to bind two operators with one repressor (i.e. form a loop) is higher than the energy needed to bind two operators with two repressors.

Why then would the system ever loop the DNA? There are only about ten repressor tetramers in the cell. If a repressor manages to bind the DNA and form a loop, the local concentration of repressor increases. Once looped, if one domain of the repressor falls off the DNA, the loop keeps the binding domain in proximity of this binding site, increasing the probability of re-binding the DNA. This effect increases repression (>1000-fold) [61, 62]. In combination, the low concentration of the repressor and the increase in local concentration due to loop formation cause the looped conformation to be the favored state for typical *in vivo* repressor concentrations.

DNA loop formation is important for gene regulation as indicated by the huge increase in *lac* repression. Yet, the formation of short loops requires energy. In fact, naive models predict short loops require such large amounts of energy that they

are unlikely to occur. However, short loops have been observed *in vivo* much more frequently than can be predicted by naive models. Recently, nonspecific binding proteins (proteins which do not require a specific site to bind the DNA), have been shown to assist in bending the DNA [12,13,21]. One example of a nonspecific binding protein is HU. Figure 4 from [21] displays the DNA length along the x-axis and the J factor along the y-axis. The J factor describes how easily DNA loops form; a high J factor indicates loops form easily. The solid line represents the ease with which HU-bound DNA forms a loop of a given length whereas the dotted curve represents naked DNA loop formation. Nonspecific binding proteins like HU may explain why short loops form more frequently than the DNA structure models predict.

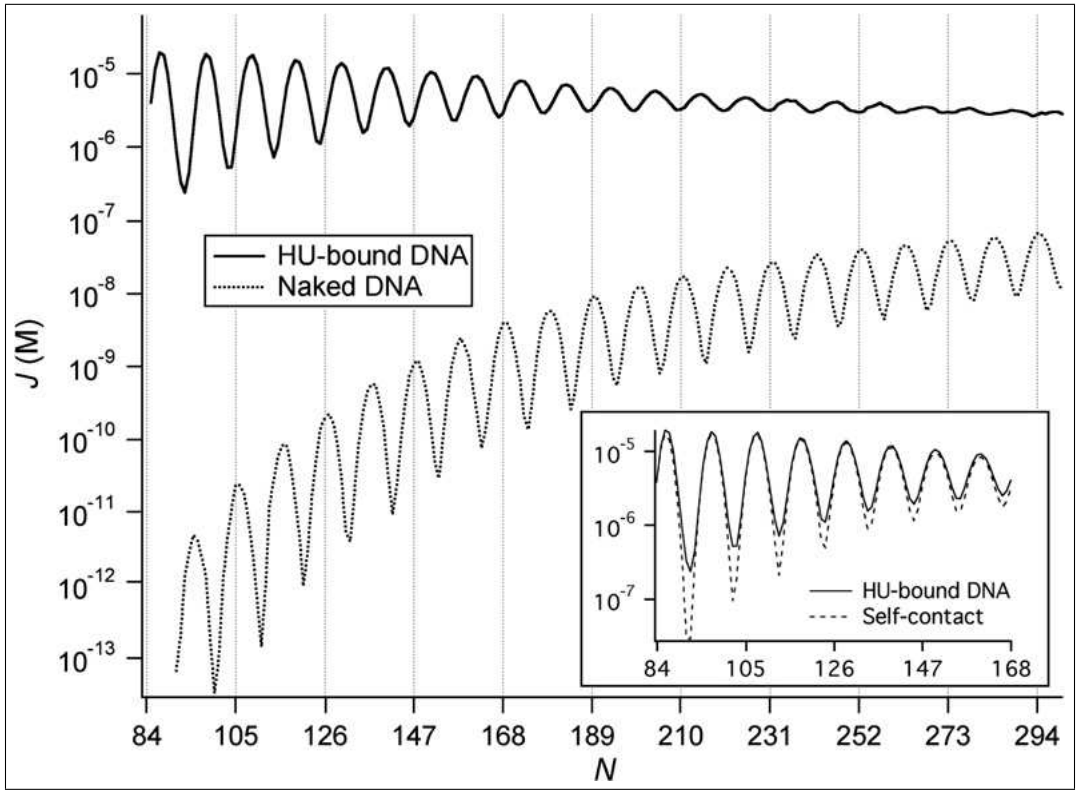


Figure 4: J factor for naked DNA and HU-bound DNA. HU is a protein known to bind and help bend the DNA. [21]

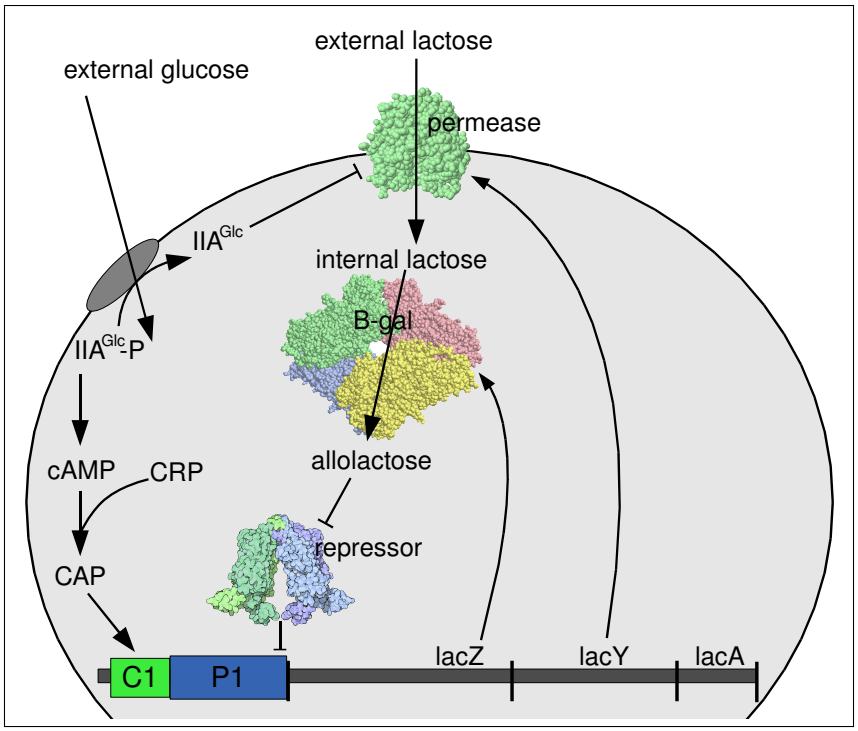


Figure 5: Positive feedback loop of the *lac* operon as described in the literature. [70,74]

Feedback Regulation: The positive feedback loop is at the heart of the *lac* regulation (see Figure 5). When the cellular medium contains low glucose and high lactose concentrations, the permease protein transports external lactose into the cell. The internal lactose is converted by β-galactosidase to allolactose which then binds to individual monomers of the repressor tetramer causing a conformal change in the subunit. This conformal change impairs the repressor binding domain and removes the repression from the *lac* operon allowing transcription and translation to begin. The resulting subunits of β-galactosidase protein assemble to form an active tetramer, while permease, the protein monomer encoded by lacY, moves to the cellular membrane. Once inserted into the membrane, permease increases the transport of lactose and β-galactosidase increases the production of allolactose. The entire cycle forms a

strong positive feedback loop which allows E. coli to quickly switch from glucose to lactose utilization.

It has been suggested [31,70,74,86] that under conditions of high external glucose, the cellular cAMP concentration is low and under low external glucose, the cAMP concentration increases to reinforce the positive feedback. It is well known that cAMP levels rise when glucose levels are low [40, 41, 58]. This, however, may in fact occur under more general conditions as well; the concentration of intracellular cAMP is similar for cells growing on glucose or lactose [29, 36, 40, 52, 79]. A spike in the internal cAMP concentration is observed as glucose runs out and the *lac* operon is induced. However, once the cell is able to digest lactose, the internal cAMP levels return to values similar to those of a cell growing on glucose [40]. We can speculate the temporary spike is a signal for a lack of resources rather than the lack of a specific resource, glucose.

Inducer Exclusion: If external glucose concentrations are sufficiently high, glucose interacts with the permease proteins in the cell membrane and blocks the transport of lactose into the cell. Since this prevents the formation of allolactose, it effectively shuts down the *lac* operon. Even if a lactose molecule enters the cell, there are so few β-galactosidase enzymes when the operon is not induced that the lactose molecule would likely not be converted to allolactose [52]. This process has been named *inducer exclusion*.

In combination, E. coli has developed a switch so effective that every experiment containing a wildtype population of E. coli grown on some fixed level of glucose and lactose results in a homogeneous population. Either all are induced and therefore expressing the *lac* genes, or none are.

lac Operon on Artificial Inducers: An artificial inducer is a molecule which has the same regulatory effects as allolactose without being diluted through cell growth and digestion. Similar to allolactose, once in the cell, artificial inducers bind and impair *lac* repressor subunits. However, an artificial inducer does not need to be converted by β-galactosidase to become active and is not digested by the cell. This allows experimentalists to separate induction from the dilution of the inducer caused by digestion and cell growth. The artificial inducers enter the cell, disable the repressor, and induce the *lac* operon (see Figure 7).

Interestingly, a bistable E. coli population has been observed when an artificial inducer like methyl-1-thio-β-D-galactoside (TMG) or isopropyl-1-thio-β-D-galactoside (IPTG) has been used in place of lactose (see Figure 6a). Although others have observed bistability, Ozbudak et al. [63] collected single cell data measurements rather than measuring the population as a whole. This allowed the authors to show hysteresis (Figure 6b) for a fixed concentration of glucose and varying TMG concentration. Furthermore, they were able to generate 14 data points and fit a bifurcation diagram to the data (Figure 6c as published by Ozbudak et al. [63]). Under an artificial inducer, the positive feedback loop eliminates the need for β-galactosidase since both TMG and IPTG are able to bind the repressor without a conformation change (see Figure 7).

Why is bistability so easily attainable on an artificial inducer but not on lactose? The structure of interaction is different, compare Figures 7 and 8. First, the artificial inducer does not need to go through a conformal change via β-galactosidase to bind the repressor. Second, when lactose is the inducer, β-galactosidase digests about half of the internal lactose and eventually all of the allolactose into glucose and galactose reducing the concentration of available inducer. Third, the cell grows on lactose (but not on an artificial inducer) increasing the volume of the cell and diluting

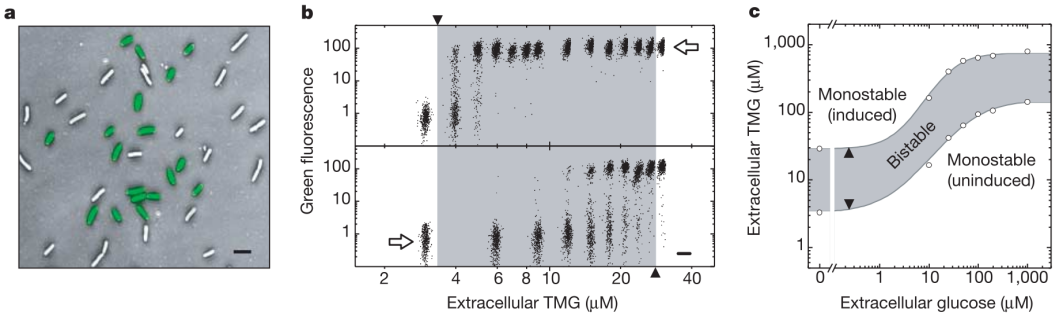


Figure 6: Bistable population in E. coli [63]. In figure **a**, each oval represents an individual E. coli cell. The cells were grown on a mixture of glucose and TMG, turning green when the cells express the *lac* genes or remaining white when repressing the *lac* operon. In figure **b**, varying concentrations of TMG are introduced in the absence of glucose. In the upper pane of figure **b** all cells were grown in 1mM TMG (this is a high concentration) overnight, and then transferred into the reported concentration of TMG for an additional 20 hours. It is assumed that this time is long enough for the cells to reach equilibrium. Each dot in **b** refers to one cell and green fluorescence >10 implies the *lac* operon is being expressed in the cell. In the lower pane, the same procedure was followed except the initial TMG concentration was 0mM. If we were to overlay the top pane with the bottom pane, hysteresis would become clear. In figure **c** the glucose and TMG concentrations are both varying. Each open circle represents recorded data, the grey region indicates a region of bistability, and the white monostability.

the concentration of available inducer. Finally, lactose regulates its own uptake by dephosphorylating an enzyme involved in glucose transport called the IIA^{Glc} enzyme. When glucose is abundant, as glucose is transported into the cell, the IIA^{Glc} enzyme is dephosphorylated, causing inducer exclusion [35]. When glucose is depleted and lactose is first utilized by the cell, there is an initial excess of phosphorylated IIA^{Glc} enzyme. This frees permease and as a result, the rate of lactose transport into the cell is fast. After a few minutes however, the internal lactose has caused about 50% of the IIA^{Glc} enzyme to be dephosphorylated, slowing the uptake rate of external lactose. In comparison, no dephosphorylation of IIA^{Glc} has been observed when using TMG as the inducer [35]. Testing for IIA^{Glc} dephosphorylation in the presence of IPTG has only been reported in the presence of lactose, resulting in a much higher

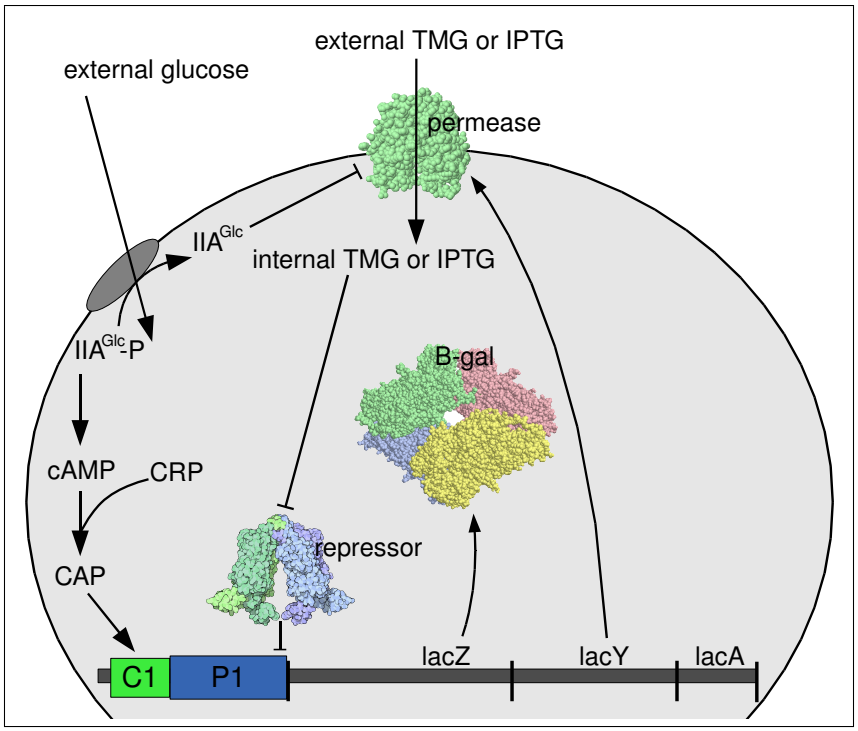


Figure 7: Positive feedback loop of the *lac* operon when using an artificial inducer rather than the natural inducer lactose.

dephosphorylation. Unfortunately, because lactose is also involved, it is impossible to tell if the IPTG caused the dephosphorylation, or if the excess of internal lactose due to the over expression of the *lac* operon induced by the addition of IPTG caused the dephosphorylation. Recall from the CAP section that the phosphorylated state of this same enzyme activates cAMP and thus CAP (the activator for the *lac* operon). It should be noted that the phosphorylation state of enzyme IIA^{Glc} is not the only factor influencing the intracellular cAMP concentrations [34].

Results

Experimentalists have collected data on two levels: at the biophysical level where they determine the energy required to form a specific complex, and at the expression level where they measure the rate at which a protein is expressed. Many biolog-

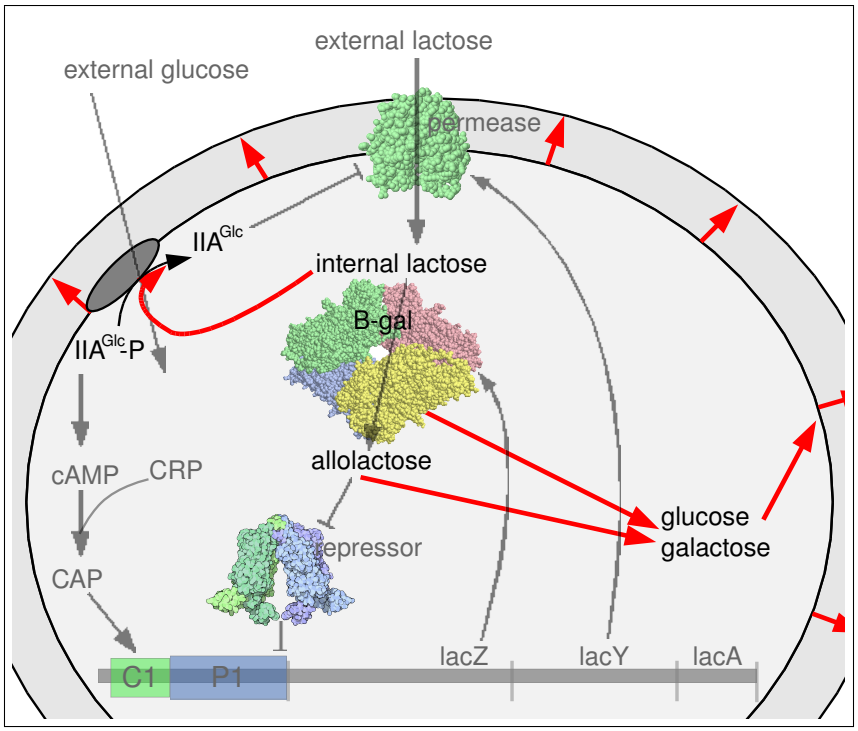


Figure 8: The factors that weaken the positive feedback overlaying the positive feedback loop of the *lac* operon on lactose. Not only is the natural inducer lactose digested into glucose and galactose, but only about half of the internal lactose is converted to allolactose. The remaining lactose is also digested to glucose and galactose. The resulting glucose and galactose are digested and in turn cause cell growth, diluting the internal concentration. In addition, lactose has been shown to auto-regulate its own uptake [35]. However, the artificial inducers TMG and IPTG are neither digested nor cause cell growth. Furthermore, it is known that TMG does not slow the uptake of external inducer by dephosphorylating the enzyme IIA^{Glc} [35].

ical models are phenomenological. They directly model expression levels with no representation of the underlying biophysics. These models use ordinary differential equations and approximate complex cellular mechanisms with Michalis-Menten or Hill functions.

We have introduced a Shea-Ackers framework which allows us to model binding control on the biophysical level. With this type of model we are able to predict expression levels from the underlying biophysical measurements. Our goal is to construct such a model of the *lac* operon. The first step is to develop a highly detailed, biophys-

ical model of *lac* regulation. To do so we will fit model parameters to experimental data which recorded repression values. Each repression value represents a ratio of protein expression under different cellular conditions.

Can we accurately define a biophysical model to represent the *lac* operon repression data? Are there some constraints which prevent this model from correctly predicting the data? Can the model pinpoint areas where the biology is incomplete?

We have defined a biophysical model with parameters derived from published data. We are able to accurately represent repression data, but clear constraints are revealed. The most obvious is the disparity between the energy predicted from structural models of short DNA loops, and the energy that our biophysical model requires to match the data. Our model requires the looping energy to be $\sim$10 to 13 kcal/mol lower than the structural models predict. This suggests a component, which changes the energy of looping, is missing from one of the models. One possibility is a nonspecific binding protein which binds the DNA and lowers the energy to form a loop [21].

Additionally, we test several components which have been suggested, but not confirmed, as contributors to the *lac* operon behavior. We find the model fails to match a portion of the repression data if we do not increase the energy of the state which has CAP bound to C1 at the same time as a repressor bound to O3. Furthermore, we show that if the model allows repressors to bind experimentally deleted sites at a reduced rate, we are able to match the repression data and keep the binding energies for the O1, O2 and O3 operators within their experimentally predicted free energy values. If we do not allow the repressors to bind the experimentally deleted sites, we must decrease the binding energy for repressor to the operators by 2 kcal/mol. This increase pushes the free energy of the repressor to the O1 operator 1 kcal/mol beyond the strongest reported biological measurement [37] to -15.01 kcal/mol.

MODEL OVERVIEW

The central role in transcription of the *lac* operon is played by RNAP binding to the P1 binding site. This binding is either enhanced by CAP binding the C1 binding site, or repressed by the *lac* repressor binding the O1 binding site. The repression is enhanced further when the free domain of the repressor binds to O2 or O3. When both binding domains of a repressor tetramer are bound, the DNA is looped.

To provide the most complete model of transcription initiation, we will consider several additional control mechanisms: the additional repressor binding site O3*, the additional RNAP binding site P2, CAP assisted DNA looping, steric interference between proteins bound to the C1 and O3 sites, reduced binding of the inducer impaired *lac* repressor and low specific binding of the repressor to experimentally deleted sites. We will now describe each of these control mechanisms in greater detail.

Primary Components

The key components of transcription in the *lac* operon are:

- the dual activation of transcription by CAP,
- repression of transcription by the *lac* repressor,
- increased repression when the DNA becomes looped between two operators and
- repressor inactivation by the inducer.

CAP Activation

Transcription of the *lac* operon is initiated when RNAP binds the DNA, forms an open complex and escapes to make mRNA. CAP has two distinct ways in which it activates transcription. First, CAP bound to the C1 site assists in the formation of a RNAP-DNA open complex when RNAP is bound to P1. Thus, the transcription initiation rate $k_{fC1}(s)$ that corresponds to the state where CAP is bound to C1 will be greater than the rate $k_f(s)$ that corresponds to states where CAP is not bound. Second, a CAP protein bound to C1 also lowers the total binding energy necessary for RNAP to bind P1. In particular, the free energy required for binding both CAP and RNAP to C1 and P1, respectively, is less than the free energy for binding each individually.

$$\Delta G_{P1} + \Delta G_{C1} + \Delta G_{coop} < \Delta G_{P1} + \Delta G_{C1}.$$

Here $\Delta G_{P1} < 0$ is the free energy for binding RNAP to P1, $\Delta G_{C1} < 0$ is the free energy for binding CAP to C1 and ΔG_{coop} is the so called cooperative energy. When $\Delta G_{coop} < 0$ there is a favorable cooperative interaction between CAP bound to C1 and RNAP bound to P1. This binding enhancement is reflected in the change in free energy E_s and is distinctly different from changing the transcription initiation rate $k_f(s)$. See equation (1) in the Introduction.

The *lac* Repressor and DNA Looping

To down-regulate or repress transcription, a repressor tetramer must bind to operator O1 on the DNA. This prevents transcription by blocking RNAP from binding. Two additional operators, O2 and O3, help to increase the repression. The O2 operator binds the repressor about 10 times weaker than O1 and lies in the lacZ gene,

401 base pairs from O1. The O3 operator, which lies in front of the P1 promoter, binds the repressor about 300 times weaker than O1 and is 92 base pairs from O1 (see figure 2 in the Introduction). In 1990, Oehler et al. [62] performed a series of deletion experiments and found a 70-fold decrease in repression when both O2 and O3 were deleted, but only a 2- to 3-fold decrease in repression when one of these operators was deleted. These studies helped confirm that the DNA is looped and that this looping is integral for the full repression of the *lac* operon.

We can model the contribution of loop formation to the repression in the same framework as the binding control of transcription by assigning free energy E_s to each looped state of the operon. The free energy of looping can be determined by evaluating the free energy difference between the looped state, and the non-looped state. For example, consider the lower looped conformation in Figure 3a where a single repressor binds O1 and O2. The free energy of this state E_s includes the free energy for binding the repressor to O1, which we denote by ΔG_{O1}, and the free energy for binding the repressor to O2, which we denote by ΔG_{O2}. Then the free energy required to form the loop between O1 and O2 is $\Delta G_{O12} = E_s - \Delta G_{O1} - \Delta G_{O2}$. Note that the more negative ΔG_{O12} is, the more probable this state is. In general, the repressor affinity for each operator can be expressed as a free energy ΔG_{Oj} where $j \in \{1, 2, 3\}$. The more negative ΔG_{Oj}, the stronger the binding affinity. If a repressor binds two operators, Oi and Oj, the energy of the state s includes the term ΔG_{Oij} to describe the additional energy needed to loop the DNA. Thus, $E_s = \Delta G_i + \Delta G_j + \Delta G_{Oij}$. The loop formation is described in greater detail in Chapter 4.

The free energy of a loop can vary greatly and depends on the DNA sequence, the number of base pairs in the loop, or even other proteins involved. In the case of the *lac* operon, the loops are so short that it costs energy ($\Delta G_{Oij} > 0$) to bend the DNA in order to form the loops. Surprisingly, even with energy cost of the short loop,

the total energy of the looped complex is lower than the non-looped complex (two repressors bound to two different operators) for low concentrations. We will discuss this in further detail in Chapter 4.

The Inducer

When an inducer like lactose, IPTG or TMG is absent, the repressor blocks transcription by binding one domain of the repressor to O1 and looping the DNA to bind the other repressor domain to one of the other operators. If a cell is exposed to a high external concentration of inducer, in the absence of a preferred sugar like glucose, the inducer will enter the cell and bind the repressor resulting in a conformal change. Each monomer of the repressor has a binding region for one unit of inducer. A binding domain, which is formed by two monomers, is less effective when bound by the inducer. Thus the inducer removes the repression, allowing transcription of the *lac* genes.

In the presence of the inducer there are three functionally different types of repressor: R represents repressor free of inducer, R' represents repressors with one impaired binding domain and R'' represents repressors with two impaired binding domains. (See Figure 3 in the Introduction.) Since a binding domain must be free of inducer to bind an operator, the concentration of repressors available to effectively bind the DNA decreases as inducer is added to the system. R can bind and loop the DNA, R' cannot effectively loop the DNA, and R'' is unable to effectively bind the DNA.

Secondary Components

As will be demonstrated in the Results section, the model based only on the dominant components of the *lac* operon control is unable to match the Oehler et al. [61, 62] repression data. Only when we extended our model to include additional features, are we able to match the data. These additional components are:

- the additional RNAP binding site P2;
- the additional binding site for the *lac* repressor O3*;
- CAP assisted DNA looping;
- O3-CAP steric interference;
- reduced binding of the inducer impaired *lac* repressor and
- low (but non-zero) specific binding of the *lac* repressor to sites that are assumed to be deleted in the Oehler et al. [61, 62] experiments.

An in depth discussion of these secondary components follows.

Additional RNAP Binding Sites

At least four promoter regions, where RNAP can bind the DNA, have been identified in the *lac* operon [22, 44]. We have already discussed the primary RNAP binding site, P1, for *lac* gene expression. The P2 biding site overlaps the P1 binding site and RNAP can only bind to one of them at a time. P3 and P4 are positioned on the DNA near P2, and they also overlap P1 and the CAP binding site C1. Figure 2 in the Introduction shows this overlap. We now describe the evidence [49] which suggests that P2 binds RNAP more effectively than P1 in the absence of CAP, see

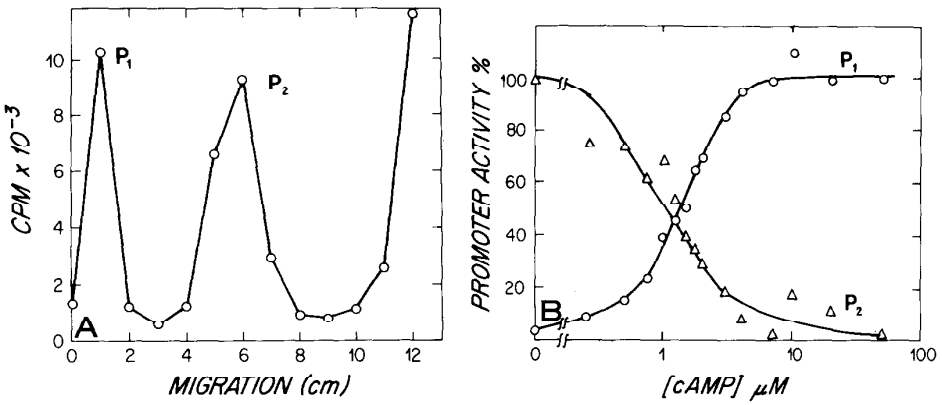


Figure 9: As described in [49], using 2nM of a wildtype *lac* promoter fragment along with 50 nM RNAP and 200 nM CRP, the left panel shows the two abortive initiation products, short pieces of mRNA released during transcription initiation, from the P1 and P2 promoters and the right panel displays the relative activity of each promoters as a function of cAMP concentration. As cAMP increases, more CAP is available to bind C1, activate P1 and repress P2.

figure 9. The left panel of this figure displays the relative concentration of abortive transcripts from the P1 and P2 promoters in the *lac* operon. Abortive transcripts are short pieces of mRNA released during transcription initiation while RNAP is bound to the promoter. The relative occupancy of the two promoters, as a function of cAMP concentration, is described in the right panel. The promoter occupancy is determined from the activity in the abortive transcript assay.

During the *in vitro* experiment CRP and RNAP are in solution with the DNA. As the cAMP concentration increases, the *lac* operon activator, CAP (cAMP bound to CRP), also increases in concentration. High concentration of CAP will push the promoter activity in favor of P1. CAP affects the promoter activity in multiple ways. First, the C1 site and the P2 promoter overlap so that both sites cannot be bound at the same time. Additionally, CAP binds the DNA stronger than RNAP binds P2. Therefore, if the concentration of CAP and RNAP are similar, CAP is more

likely to bind C1 than RNAP is to bind P2. Finally, once CAP binds C1 it recruits RNAP to the P1 promoter. At the time of this experiment, it was thought that the internal concentration of cAMP was higher when E. coli was digesting the sugar lactose in comparison to a more favorable sugar like glucose. In 1996 Inada et al. [40] determined internal cAMP concentrations on lactose were similar to internal cAMP concentrations on glucose. What functional purpose, if any, does P2 provide now that we know that internal cAMP concentrations are similar on glucose and lactose?

In our model, the P2 binding region of the DNA is represented with its own transcription rate function, see equation (4) in the Introduction. The competition between P1 and P2 is expressed through their binding affinity, $K_B(s)$, to the DNA, their transcription initiation rates, $k_f(s)$, and the partition function (equation (3) in the Introduction) which includes both promoters.

Additional Repressor Binding Site O3*

There is support from two independent sources for the existence of an additional repressor binding site named O3*. This site is five to six base pairs further from O1 than the O3 operator, and is only occupied by the repressor when CAP binds to C1. The first supporting evidence for the O3* operator came from the work of Hudson and Fried [25, 39].

They analyzed enzyme cutting patterns for the *lac* operon DNA to observe the interaction between CAP and the *lac* repressor. The procedure combines CAP and the repressor in solution with DNA and allows sufficient time for the solution to come to equilibrium before introducing an enzyme which splices unbound DNA. When the solution is analyzed, the segments of DNA which have protein bound remain intact. Hudson and Fried found the repressor would not bind O3 in the presence of CAP unless a high concentration of repressor was introduced, and then the repressor would

bind to a site shifted five to six base pairs upstream of O3 [25]. These experiments suggest that although O3 and C1 are 11 base pairs apart, the sites cannot be bound concurrently. However, they can bind together if the repressor shifts to bind a new site.

The corroborating evidence that this new site is involved in *lac* operon function comes from the work of Swigon and Olson [81]. Their work focused on determining the energy required to loop the DNA between the O1 and O3 operators of the *lac* operon. Experimentally, short loops are formed with a much higher frequency than the predicted looping energy would indicate. This suggests the models predicting the looping energy are missing some component which lowers the looping energy. To reduce the total energy of the state, Swigon and Olson searched the DNA around O3 for a shifted site which would bind the *lac* repressor and explain the Hudson and Fried observations. They found an additional repressor binding site, O3*, five base pairs from O3. Furthermore, their model predicts that the looped complex, consisting of a repressor bound to O1 and O3* and CAP bound to C1 is energetically more favorable (by 1.3 kcal/mol) than the same complex where the repressor is bound to O3.

The addition of the O3* site introduces several more possible states into the partition function (see equation (3) in the Introduction).

CAP Assisted DNA Looping

It is well known that CAP binding to the C1 site bends the DNA ([46, 81]). Furthermore, the models for DNA loop formation struggle to accurately predict the likelihood of short loop formation (like the O1-O3 loop). Swigon and Olson model the O1-O3 loop with and without CAP bound to the DNA as well as the O1-O3* loop with CAP bound to the DNA. Several loop conformations were examined for each state. For the state with an O1-O3 loop and CAP bound, the predicted free energy

for seven of the eight loop conformations exceeds the free energy of the O1-O3 looped state without CAP bound. The remaining conformation has comparable energy with or without CAP bound at about 16 kcal/mol [81]. In contrast, a looping conformation formed when the repressor binds O1 and O3* becomes very favorable [81] with CAP bound at an energy of ∼14.8 kcal/mol. Since C1 is between O3 and O1, we suggest that CAP bound to C1 may also assist in O1-O3*, O2-O3 and O2-O3* loop formation and include these in our model.

CAP assisted DNA looping is reflected in the free energy of the state, E_s. When CAP is bound to C1 and a DNA loop is formed across the C1 binding site, a cooperative term ($\Delta G_{C1loop} < 0$) is included in E_s. This term represents the decrease in energy resulting from CAP assisted DNA loop formation.

Steric Interference between CAP Bound to C1 and a Repressor Bound to O3

The O3 operator and the C1 binding site are 11 base pairs apart on the DNA suggesting there is ample room to bind both the *lac* repressor and CAP at the same time [81]. However, when Swigon and Olson [81] analyzed the O1-O3 DNA loop, they also considered the O1-O3-CAP loop. Their analysis of the O1-O3-CAP loop determined that steric interference can occur between CAP and the repressor or CAP and the DNA for many of the minimum energy configurations. Steric interference (or negative steric interaction) means the surface of the CAP protein interacts negatively with the DNA or the repressor. In support of this, Hudson and Fried [25] found that CAP and the repressor would not bind concurrently without the repressor shifting five to six base pairs.

The steric interference between CAP bound to C1 and a repressor to O3 is represented by $\Delta G_{C1-O3} > 0$. This free energy is included in the free energy of the state,

E_s, for all states in which CAP is bound to C1 and a repressor is bound to the O3 operator.

Binding of the Inducer-Impaired *lac* Repressor to the Operators

Horton et al. [37] find that the IPTG impaired *lac* repressor continues to bind the O1 operator, but at a reduced binding affinity. This opposes the Oehler et al. [60] prediction which states that impaired repressors binding domains do not bind the DNA at all.

Inducer-impaired binding domains of the repressor are modeled by tracking the concentration of R'' (the repressors with doubly impaired domains), and introducing R'' as a regulatory protein in equation (1) in the Introduction. In order to include R'' as a binding repressor we need to determine binding affinities for R'' to each operator.

Binding of the Repressor to Deleted Operators

The experimental removal of a binding site is achieved by mutating some of the DNA nucleotides. This often only lowers the likelihood of binding of the repressor to this site, rather then completely precluding this binding.

For example, the mutant (0,0,3) studied in Oehler et al. [61] has the O1 and O2 sites experimentally deleted. However, Horton et al. [37] demonstrated in their experiments with the O1 site, that each base pair contributes a certain partial energy to the binding energy of the entire O1 site. The O1 deleted and the O2 deleted sites differs from the O1 site by 4 and 11 base pairs, respectively (Figure 2 in the Introduction). We can use this data from Horton et al. [37] to predict the decrease in binding energy of the repressor to the deleted O1 and O2 sites with these 4 and 11 base pair modifications. The deleted O1 and O2 sites may still bind the repressor with this lowered affinity.

TRANSCRIPTION OF THE *LAC* OPERON

Since the *lac* operon has been studied since the 1940's [50], and its bistability since 1957 [19, 59], data is abundant. This chapter will begin with a brief review of experimental data. We first describe the data collected by Ozbudak et al. [63] which documents the bistability of the *lac* operon. Then we will describe the repression data collected by Oehler et al. [61, 62] and lastly we report on the change in free energy required to loop the DNA as reported by Swigon and Olson [81].

The second half of the chapter will introduce our model and give a brief review of existing models of the *lac* operon.

Experimental Data

Bistability Data

Although other groups have observed a bistable population of E. coli [19], the Ozbudak et al. paper [63] provided a beautiful display of hysteresis as well as a bifurcation diagram, shown in Figure 6 in the Introduction. To collect this data, the green fluorescent protein (gfp) reporter gene was placed into the chromosome of E. coli MG1655 cells under the control of the *lac* promoter. Additionally, the cells contained a plasmid carrying a red fluorescent reporter (HcRed) under the control of the galactitol (*gat*) promoter. The *gat* promoter has a CAP binding site, and as a result, transcription at the *gat* promoter, measured by red fluorescence, is a direct measurement of CAP levels. Using these reporter genes, Ozbudak et al. [63] collected single cell measurements, rather than population averages.

If RNAP binds the *lac* promoter, this gfp reporter protein will turn the cell green allowing for easy recognition of *lac* expression. The assumption here is that if gfp is

expressed, then both the inserted *lac* promoter and the native *lac* operon are induced. An image of a bistable population at 0μM glucose and 18μM TMG is in the left frame of Figure 6 (in the Introduction). In the central image of Figure 6, each dot represents the green fluorescence of a single cell at the indicated external TMG concentration while glucose is fixed at 0μM. All of the cells recorded in the top pane were induced overnight, before being introduced to the indicated concentration of TMG with 0mM glucose. The lower pane is similar except the cells were not induced prior to introducing TMG. The open circles in the right pane are recorded from data and then Ozbudak et al. [63] fit the curves to create the bifurcation diagram with the two parameters, external glucose and external TMG. To generate this figure, Ozbudak et al. [63] created two master cultures of E.coli, one with the *lac* operon induced using a high level of TMG and the other with the *lac* operon uninduced by using a TMG concentration of 0mM. Each culture was allowed to grow overnight to reach equilibrium. Sub-populations of approximately 1000 cells were then exposed for 20 hours to the TMG and glucose concentrations specified in **b** and **c**. An example of a bistable population is shown in **a** where the cells expressing the *lac* operon have over 100 times the green fluorescence of the white, uninduced cells. For Figure 6**a**, the cellular medium contained 18μM TMG and 0mM glucose. Figure 6**b** displays the green fluorescence of the sub-populations at the indicated TMG concentration while glucose is constant at 0mM. The cells in the upper pane originated from the induced master culture, while the lower pane cells originated from the uninduced master culture. Each column represents single cell measurements of green fluorescence collected from sub-populations grown at the indicated TMG concentration. If the upper and lower panes were superimposed, hysteresis would be visible. Bistability was achieved on both IPTG and TMG (see Figure 6 in the Introduction for TMG bistability), but not lactose.

In another experiment, bistability was examined in the presence of four plasmids and 25 plasmids. With four plasmids, a range of bistability was still visible, but with 25 plasmids, only a graded response to increasing inducer was observed. Ozbudak et al. [63] conclude that the additional operators, three for each plasmid, act to titrate the repressor concentration. With 25 plasmids, there are 78 operators and only about 10 repressor tetramers per cell, leaving promoters free to bind RNAP for all inducer concentrations. These experiments provide important quantitative results to strive for when modeling the *lac* operon.

Repression Data

Due to some unique properties of β-galactosidase, the lacZ gene has been used throughout microbiology as a reporter gene. The function of the β-galactosidase enzyme is to cleave lactose into glucose and galactose which are easier for the cell to digest. The bond that must be cleaved is a β-galactoside bond and other compounds containing this bond can be recognized as a substrate for β-galactosidase. The discovery of the cheap, colorless substrate o-nitrophenyl-β-galactoside (ONPG) has resulted in huge advances in the understanding of cellular processes. When β-galactosidase cleaves ONPG, colorless galactose and yellow o-nitrophenol are produced. The resulting yellow can be measured accurately enough to detect one molecule of β-galactosidase per cell in a few milliliters of cell suspension [52]. By inserting the lacZ gene after a gene of interest and then introducing ONPG, biologists can determine if the gene in question is currently active. To help quantify this information Jeffrey Miller published a protocol in 1972 which is now referred to as a "Miller" assay, and sometimes the resulting units are also referred to as "Miller Units."

The "Miller" assay of a cellular population uses spectroscopy to measure the absorbance of light for the population at 420, 500, and 600 nm so that

$$1\text{Miller Unit} = 1000 \times \frac{Abs_{420} - (1.75 \times Abs_{550})}{t \times \nu \times Abs_{600}}$$

where

- Abs_{420} is the absorbance of the yellow o-nitrophenol,
- $1.75 \times Abs_{550}$ approximates the scatter from cell debris at 420nm,
- t is the reaction time in minutes,
- ν is the volume,
- Abs_{600} reflects the cell density [1].

Absorbance is defined as $Abs_\lambda = -\log_{10}(I/I_0)$ where I is the light intensity at wavelength λ after passing through the cell sample and I_0 is the light intensity before passing through the sample.

It should be noted that Abs_{600} is different for each spectrophotometer (used to measure the absorbance). Each one should be calibrated by plating dilutions of known Abs_{600} cultures to determine the number of viable cells per Abs_{600}. Also, this formula was originally designed so that uninduced E. coli populations produce about 1 Miller Unit (MU), and induced E. coli populations produce about 1000 MU (for cells grown on lactose or IPTG). However, in one protocol, 1500-1800 MU were reported for the E. coli strain MG1655 on 1mM IPTG. I mention this so the reader is aware of data variation.

When comparing different strains for their ability to repress the *lac* operon, a ratio of β-galactosidase activity under induced conditions (1mM IPTG) to the activity

under uninduced conditions (no IPTG) is computed. Repression is defined as

$$\mathcal{R} = \frac{\beta\text{-galactosidase(1mM)}}{\beta\text{-galactosidase(0mM)}}, \tag{5}$$

where the β-galactosidase activity is determined by a Miller assay. On the assumption that in the presence of excess ONPG, the production of o-nitrophenol per unit time is proportional to the concentration of β-galactosidase, a higher repression ratio indicates better repression. This measurement of repression has been used to compare the effects of deleting operators as well as different loop lengths.

Of particular interest are data from Oehler et al. [60–62]. By using an E. coli mutant which had the *lac* operon deleted and individually reintroducing combinations of the *lac* operators (O1, O2, and O3) using a phage λ vector, Oehler et al. measured the repression for almost all possible combinations of operators at various levels of repressor concentration. Measurements were also recorded with a mutated repressor which could only form dimers in solution rather than a tetramer like the wildtype.

To represent the various mutants, we need to keep track of both the original position of the operators as well as the original DNA sequence of the operators. Let a 3-tuple ((i, j, k), $i, j, k \in \{0, 1, 2, 3\}$) represent a particular mutant where i holds the position of the O1 operator on the DNA; j holds the position of the O2 operator and k holds the position of the O3 operator. We will use 0 to represent a deleted site while 1, 2 and 3 will denote the DNA sequences for O1, O2 and O3 sites, respectively. For example, $(1, 2, 3)$ represents the wildtype *lac* operon; $(1, 0, 0)$ has the O2 and O3 operators deleted; and $(3, 0, 1)$ has the DNA for the O3 operator in the O1 position, O2 deleted, and the DNA for O1 in the O3 position.

Oehler et al. published papers based on these types of mutations in 1990 [62] and again in 1994 [61] with repressor tetramers and mutated repressor dimers at varying concentrations. The three experimental concentrations were: wildtype (WT) levels

Table 1: Repression measurements reported by Oehler et al. [61, 62] for various mutants, as described by the 3-tuple (i, j, k). Each vector represents a particular mutant: (1,2,3) is the wildtype; (1,0,0) has O2 and O3 deleted; (3,0,1) has O3 DNA in the O1 position, O2 deleted, and O1 DNA in the O3 position and so on. For the column headers: WT describes the repression measured in the presence of wildtype repressor concentration, $\sim$5 $\times$ WT implies the cellular concentration of repressor was approximately 5 times larger than the wildtype, and similarly for 90 times wildtype. The first column of the 1994 data groups similar mutants: A. is the wildtype, B. contains mutants with a deletion in the O2 position, C. contains mutants with a deletion in the O3 position, positions O2 and O3 are deleted in D. and E. has deletions in the O1 and O2 positions.

	1994	$\sim$ 5$\times$WT	$\sim$ 90$\times$WT
A.	(1,2,3)	8100	$\geq$19,000
B.	(1,0,1)	$\geq$ 12,000	$\geq$ 46,000
	(1,0,3)	6200	$\geq$21,000
	(3,0,1)	890	3900
	(3,0,3)	38	960
C.	(1,2,0)	2300	6800
	(2,2,0)	360	560
	(3,2,0)	6.8	15
D.	(1,0,0)	200	4700
	(2,0,0)	21	320
	(3,0,0)	1.3	16
E.	(0,0,1)	18	28
	(0,0,3)	1.7	25

1990	WT	$\sim$ 5$\times$WT	$\sim$ 90$\times$WT
(1,2,3)	1300	6700	16,000
(1,0,3)	440	3900	15,000
(1,2,0)	700	1400	3600
(1,0,0)	18	140	2700
(0,2,3)	1.9	4.4	29
(0,0,3)	1.0	1.9	21
(0,2,0)	1.0	1.1	1.2
(0,0,0)	1.0	1.1	1.3

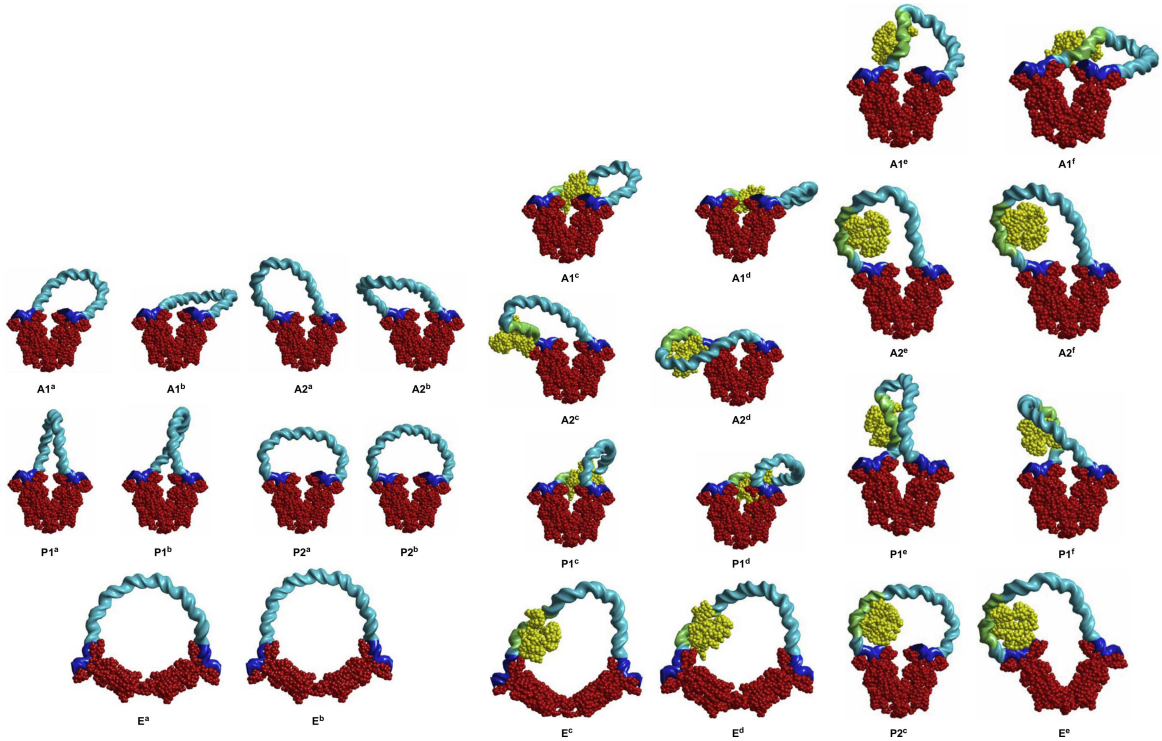
which amounts to about 10 tetramers or 40 monomers per cell, 5 $\times$ WT and 90 $\times$ WT. The later paper was partially published to correct some dimer data from the first paper. Given this notation, see Table 1 for various mutants and repression measurements.

Observe how the measured repression differs between the 1990 paper and the 1994 paper. For example, the repression for (1,0,0) was measured to be 140 in 1990 and 200 in 1994 at 5 $\times$ WT concentration, and repression was measured to be 2700 in 1990 and 4700 in 1994 at 90 $\times$ WT concentration. It should be noted that when

these experiments were run, the 1990 DNA deletion sequences were different from the 1994 deletion sequences as displayed in Table 2. In this table, the first column either contains the year the mutation was used, or WT to indicate it is the wildtype sequence. The '$^-$' indicates a deletion sequence, and O1$^-$a and O1$^-$b are the O1 DNA deletions used in 1994 for (0,0,3) and (0,0,1), respectively. The slightly higher repression in 1994 suggests that the plasmids they used may bind the repressor slightly better than those used in 1990.

Looping Data

Swigon and Olson modeled the structure of multiprotein-DNA assemblies using a base-pair level theory of DNA elasticity. In particular, they studied the O1-O3 loop in the *lac* operon [80, 81] in the presence and absence of the activator CAP. Furthermore, they discovered that the incorporation of O3*, a shifted operator which overlaps O3 in all but 5 base pairs, considerably lowers the predicted energy of loop formation when CAP is bound to C1. For each complex (O1-O3, O1-O3 with CAP, and O1-O3* with CAP) they calculated the energy cost of forming the DNA loop and presented a sample of the minimum energy configurations. We include Figure 10 from Swigon et al. [81]. (a) displays the DNA looped by a repressor bound to the O1 and O3 operators, (b) displays the DNA looped by a repressor bound to the O1 and O3 operators with CAP bound to C1 and in (c) the DNA is looped by a repressor bound to the O1 and O3* operators with CAP bound to C1. The DNA is represented in light blue, the operators in dark blue, the repressor in red, and CAP in yellow. These DNA loops represent the minimum energy configurations, and Table 3 displays their relative looping energies.



(a) O1-O3 loop, no CAP (b) O1-O3 loop (c) O1-O3* loop

Figure 10: Minimum energy configurations of (a) the O1-repressor-O3 loop, (b) the O1-repressor-O3 loop with CAP bound and (c) the O1-repressor-O3* loop with CAP bound for the *lac* operon as predicted by Swigon et al. [81].

Table 2: DNA sequences for wildtype and deleted operators as reported in the 1990 and 1994 Oehler et al. papers [61,62]. WT is the wildtype, 1990 and 1994 denote the year the subsequent deletion sequence was implemented. O1, O2, O3 and O3* denote the operator with '$^-$' representing a deleted operator. For example, in the table above, 1994 O2$^-$ is followed by the DNA sequence for the deleted O2 operator used in the experiments published in 1994 [61] for the mutants: (1,0,1), (1,0,3), (3,0,1), (3,0,3), (1,0,0), (2,0,0), (3,0,0), (0,0,1) and (0,0,3), i.e. for all mutants where the O2 operator is deleted. Each wildtype sequence is in capital letters and the deletion sequences have a capital if the base pair matches the wildtype and a lower case letter for a base pair substitution differing from the wildtype sequence. The 1994 O1$^-$a and O1$^-$b are the O1 DNA deletions used in (0,0,3) and (0,0,1), respectively. The O3* sequences are for the shifted O3 operator which has been proposed by Swigon and Olson [81]. O3*(O1) is the resulting sequence if O1 DNA is inserted into the O3 position.

WT	O1	5'- A A T T G T G A G C G G A T A A C A A T T -3'
1990	O1$^-$	5'- A A T T G T t A G C G G A g A A g A A T T -3'
1994	O1$^-$a	5'- g A T T G T t A G C t t A T A A t A A T T -3'
1994	O1$^-$b	5'- g A a c g a c A t C c t c c g c t A g g T -3'
WT	O2	5'- A A A T T G T A G C G A G T A A C A A C C -3'
1990	O2$^-$	5'- g A A g g t T A a t G A a T A g C A c C C -3'
1994	O2$^-$	5'- g A A T g t T A a t G A a T A g C A c C C -3'
WT	O3	5'- G G C A G T G A G C G C A A C G C A A T T -3'
1990	O3$^-$	5'- t c g A t c G A G C T C A A C G C A A T T -3'
1994	O3$^-$	5'- a a C c t c G A G C T C A A C G C A A T T -3'
WT	O3*	5'- A A G C G G G C A G T G A G G G C A A C G -3'
1990	O3*$^-$	5'- A A G C G t c g A t c G A G c t C A A C G -3'
1994	O3*$^-$	5'- A A G C G a a C c t g G A G c t C A A C G -3'
1994	O3*(O1)	5'- A A G C G a a t t G T G A G c G g A t a a -3'.

Table 3: The predicted free energy of looping from Swigon et al. [81]. The first column correlates to the loop of the same name in the Figure 10. The second column is the free energy of forming the loop in k_BT (Boltzmann's constant times temperature) and the third column is the free energy in kcal/mol. We highlight in blue the lowest energy for each looping complex. See Appendix B for more on the relation between k_BT and kcal/mol.

Figure 10(a)			Figure 10(b)			Figure 10(c)		
loop	k_BT	kcal/mol	loop	k_BT	kcal/mol	loop	k_BT	kcal/mol
$A1^a$	36.5	22.52	$A1^c$	42.2	26.04	$A1^e$	70.2	43.31
$A1^b$	41.4	25.54	$A1^d$	52.4	32.33	$A1^f$	49.4	30.48
$A2^a$	36.7	22.64	$A2^c$	82.9	51.15	$A2^e$	68.8	42.45
$A2^b$	44.1	27.21	$A2^d$	87.3	53.86	$A2^f$	23.9	14.75
$P1^a$	42.2	26.04	$P1^c$	48.5	29.92	$P1^e$	99.0	61.08
$P1^b$	64.1	39.55	$P1^d$	88.0	54.30	$P1^f$	58.3	35.9
$P2^a$	77.4	47.76	E^c	40.4	24.93	$P2^c$	38.2	23.57
$P2^b$	49.1	30.29	E^d	26.2	16.17	E^e	59.8	36.90
E^a	45.9	28.32						
E^b	26.1	16.1						

Differential Equation Model of the *lac* Operon

In this section we first introduce a differential equation model of the *lac* operon and from this model we derive an expression for the repression value $\mathcal{R}$. It is this value that we wish to compare to the experimental values of Oehler et al. [61]. We then explain in detail all the terms of this expression.

We will follow the model structure of Santillán and Mackey [74]. Dynamical variables will represent the mRNA concentration M, protein concentration P and intracellular inducer concentration I. Note that I could represent lactose, IPTG or TMG when we try to simulate experiments with different inducers. The equations take the form

$$\dot{M} = DF(I) - \gamma_M M \tag{6}$$

$$\dot{P} = k_P M - \gamma_P P \tag{7}$$

$$\dot{I} = k_I \beta_I \beta_G P - 2\phi_M \mathcal{M}(I) P/4 - \gamma_I I, \tag{8}$$

We will now explain all the terms and parameters in these equations.

mRNA Concentration

In equation (6), D represents the average copy number of the *lac* operon in an E.coli cell and the function $F(I)$ describes the rate of mRNA production as a function of the inducer. In order for mRNA to be produced, RNAP must bind the DNA and transcribe the *lac* genes. This function $F(I)$ represents transcription from both the primary promoter P1 as well as the secondary promoter P2 and is defined

$$F(I) = f_1(I) + f_2(I) + f_{1C1}(I).$$

Here $f_1(I)$, $f_2(I)$ and $f_{1C1}(I)$ are transcription rate functions, see equation (4) in the Introduction. Transcription from RNAP bound to P1 without CAP is represented by

$f_1(I)$, and transcription from RNAP bound to P2 is represented by $f_2(I)$. Since CAP assists in open-complex formation, we define the function $f_{1C1}(I)$ to represent transcription from RNAP bound to P1 with CAP bound to C1. The term γ_M represents the dilution rate of mRNA due to cell growth and degradation.

Protein Concentration

The *lac* genes are transcribed together as one strand of mRNA. At the beginning of each transcribed gene there is a binding site for a ribosome. Ribosomes bind and translate each gene from mRNA into a chain of amino acids producing permease, β-galactosidase and transacetylase monomers. Permease is a monomer, but it takes four copies of the β-galactosidase monomer to make one β-galactosidase enzyme. P represents the concentration of amino acid chains of each gene in the cell.

k_P is the translation initiation rate. This represents the time it takes ribosomes to bind the mRNA strand and make a protein. Similar to mRNA, the term γ_P describes the degradation and dilution of amino acid chains due to cell growth.

Internal Inducer Concentration

Equation (8) has three components. The first component, $k_I\beta_I\beta_G P$, represents the import of extracellular inducer. The maximal rate of inducer uptake per permease is represented by k_I. $\beta_I = \frac{extI}{\kappa_I + extI}$ represents the normalized inducer uptake rate per active permease as a function of external inducer ($extI$) concentration. The decrease per active permease as a function of external glucose ($extG$) is represented by $\beta_G = 1 - \phi_G \frac{extG}{\kappa_G + extG}$. This is also known as inducer exclusion.

The term $2\phi_M \mathcal{M}(I)P/4$ is specific to lactose and represents the digestion of inducer per β-galactosidase enzyme. Approximately half of the internal lactose is turned into allolactose (the isomer of lactose which bind the repressor) and the other half

is digested into glucose and galactose. $\phi_M \mathcal{M}(I)$ where $\mathcal{M}(I) = \frac{I}{\kappa_M + I}$ represents the lactose metabolism rate per β-galactosidase. ϕ_M is the lactose metabolism rate and κ_M is the saturation constant. When the operon is induced with an artificial inducer $2\phi_M \mathcal{M}(I)$ is zero due to the fact that artificial inducers like IPTG and TMG are not digested. The term, $P/4$, represents the concentration β-galactosidase tetramers.

The last term $\gamma_I I$ accounts for the dilution of internal lactose due to cell growth.

Table 4: Estimated parameter values for equations (6)-(8) from Santillán et al. [74]. mbp stands for molecules per bacterium.

k_P	$\simeq 18.8$ min^{-1}	κ_M	$\simeq 7.0 \times 10^5$ mpb
k_I	$\simeq 6.0 \times 10^4$ min^{-1}	κ_I	$\simeq 680$ μM
κ_G	$\simeq 1.0$ μM	ϕ_G	$\simeq 0.35$
ϕ_M	$\in [0, 4.0 \times 10^4]$ min^{-1}	γ_M	$\simeq 0.48$min^{-1}
γ_P	$\simeq 0.03$min^{-1}	γ_I	$\simeq 0.02$min^{-1}
D	$\simeq 2$ mpb		

Model Predicted Repression Value

To derive an expression for the repression value measured by Oehler et al. [61] we assume that the inducer I is IPTG and thus the term $2\phi_M \mathcal{M}(I)$, describing the rate of lactose metabolism by a β-galactosidase enzyme, drops out of the inducer equation (see equation (8)). When the system is in equilibrium, equations (6)-(8) can be solved for equilibrium concentrations (M^*, P^*, I^*)

$$M^* = DF(I^*)/\gamma_M$$

$$P^* = k_P M^*/\gamma_P$$

$$I^* = k_I \beta_I \beta_G P^*/\gamma_I.$$

We assume that the β-galactosidase activity is proportional to the protein concentration. Then the repression ratio (equation (5) in Chapter 3) can be represented as

follows

$$\begin{aligned}\mathcal{R} &= \frac{\beta\text{-galactosidase(1mM)}}{\beta\text{-galactosidase(0mM)}} \\ &\simeq \frac{P^*|_{(1\text{mM})}}{P^*|_{(0\text{mM})}} \\ &= \frac{\frac{k_P M^*}{\gamma_P}|_{(1\text{mM})}}{\frac{k_P M^*}{\gamma_P}|_{(0\text{mM})}} \\ &= \frac{M^*|_{(1\text{mM})}}{M^*|_{(0\text{mM})}} \\ &= \frac{(DF(I^*)/\gamma_M)|_{(1\text{mM})}}{(DF(I^*)/\gamma_M)|_{(0\text{mM})}} \\ &= \frac{F(I^*)|_{(1\text{mM})}}{F(I^*)|_{(0\text{mM})}},\end{aligned}$$

where

$$F(I^*) = f_1(I^*) + f_2(I^*) + f_{1C1}(I^*)$$

is the production rate of mRNA. Therefore

$$\begin{aligned}\mathcal{R} &= \frac{F(I^*)|_{(1\text{mM})}}{F(I^*)|_{(0\text{mM})}} \\ &= \frac{(f_1(I^*) + f_2(I^*) + f_{1C1}(I^*))\,|_{(1\text{mM})}}{(f_1(I^*) + f_2(I^*) + f_{1C1}(I^*))\,|_{(0\text{mM})}}. \qquad (9)\end{aligned}$$

In the next section we describe in detail expressions $f_1(I)$, $f_2(I)$ and $f_{1C1}(I)$ and the parameters that enter into these expressions.

mRNA Transcription

In 1982, Shea and Ackers [5] introduced statistical mechanics as a tool to represent protein-DNA interactions in the cell and the method has since been validated in several other models [26, 27, 69, 71, 72, 77]. We present here our complete model with all primary and secondary components. Each transcription function f_* is defined by:

$$f_* = \sum_{s\in\mathcal{S}_*} k(s)\mathbb{P}(s) = \sum_{s\in\mathcal{S}_*} k(s)\frac{K_B(s)[RNAP]^{\alpha_s}[CAP]^{\alpha_s^1}[R]^{\alpha_s^2}[R']^{\alpha_s^3}[R'']^{\alpha_s^4}}{Z}; \qquad (10)$$

comparing to equation (4) in the Introduction we have $m = 4$ with [RNAP] = $[P]$, $[CAP] = [r_1]$, $[R] = [r_2]$, $[R'] = [r_3]$ and $[R''] = [r_4]$ ([.] denotes concentration). R, R' and R'' represent repressor tetramers with zero, one and two inducer impaired binding domains as in Figure 3b in the Introduction. (If we do not allow impaired repressor domains to bind the DNA, $m = 3$ and $[R'']^{\alpha_s^4}$ is removed from f_*.) The partition function Z remains the same for f_1, f_2 and f_{1C1}. However, f_{1C1} has a faster transcription initiation rate $k_{fC1}(s)$ since CAP assists in open complex formation. Furthermore, the set of states which leads to transcription (the numerator of f_*) differs for f_1, f_2, and f_{1C1}. For example, consider the transcription rate function $f_2 = f_2([\text{RNAP}], [CAP], [R], [R'], [R''])$ as an example. The states $\mathcal{S}_2$ for f_2 are all possible combinations of RNAP, CAP, repressor and DNA, where RNAP is bound to P2 and a repressor is not bound to O1. In addition, since P2 overlaps the C1 binding site, CAP cannot bind when RNAP is bound to P2. f_1 and f_{1C1} are similar in concept with their own states and restrictions. The full representation of F contains hundreds of terms and can be viewed in Appendix D.

To compute F, when all primary and secondary components are included, we must know the energy of the individual states and the concentrations of RNAP, CAP, R, R', and R''.

States of F: Each state s is used to describe a particular binding configuration of the species (RNAP, CAP, R, R' and R'') and the DNA. For each bound protein there is an associated free energy. As an example, if CAP binds C1, there is a free energy ΔG_{C1} representing the energy change between the empty DNA strand and the CAP bound DNA strand. This is true for each species. Furthermore, some species interact to increase or decrease the total energy of the DNA-species complex. For example, if CAP binds C1 and RNAP binds P1, CAP interacts with RNAP to make the complex

more stable. The stability is observed in the form of ΔG_{coop} in the free energy of the state $E_s = \Delta G_{C1} + \Delta G_{P1} + \Delta G_{coop}$. If $\Delta G_{coop} < 0$ the interaction increases the state stability, if $\Delta G_{coop} > 0$ the interaction is unfavorable and if $\Delta G_{coop} = 0$ there is no benefit or deficit to the interaction. This interaction is sometimes called cooperativity and is discussed in greater detail in Appendix C.

In the primary components of the model, there are two occurrences of cooperativity. The first is between CAP and RNAP as described above, and the second occurs when a repressor tetramer loops the DNA by binding two operators. With CAP and RNAP, the interaction decreases the total energy of the state. When a repressor loops the DNA, energy must be put into the system to bend the DNA. To model the primary components we must determine the following free energies: the binding of the repressor to each operator (ΔG_{O1}, ΔG_{O2}, and ΔG_{O3}); the energy of looping between the O1 and O2 operators (ΔG_{O12}), the energy of looping between the O1 and O3 operators (ΔG_{O13}) and the energy of looping between the O2 and O3 operators (ΔG_{O23}); the dissociation constant for RNAP to P1 (K_{P1}); the dissociation constant for CAP to C1 (K_{C1}) and the cooperativity K_{coop} between RNAP bound to P1 and CAP bound to C1. (The free energy ΔG is proportional to the dissociation constant K by the equation $K = \exp(-\Delta G/RT)$ as discussed in Appendix C.)

To incorporate the secondary components there are several additional free energies to be determined. These are listed below.

- To include P2, we must determine a value for the dissociation constant, K_{P2}.
- By introducing O3*, the free energy between the repressor and O3*, ΔG_{O3*} as well as the energy for the additional looping configurations: ΔG_{O13*} and ΔG_{O23*} must be represented. Recall, O3* overlaps the O3 position in all but five base pairs. Since we want to represent all mutations performed by Oehler

et al. [61], we must also determine a free energy for the sequence which is in the O3* position when the O1 DNA is inserted into the O3 position. We denote this free energy $\Delta G_{O3*(O1)}$.

- CAP assisted DNA looping introduces the term, $\Delta G_{C1loop} < 0$. When CAP is bound to C1, ΔG_{C1loop} represents a decrease in the energy of loop formation for O1-O3, O1-O3*, O2-O3 and O2-O3*.

- C1-O3 steric interference, ΔG_{C1-O3}, is present when CAP is bound to C1 and a repressor is bound to O3 and increases the energy of the complex.

- Allowing impaired repressors to bind the DNA at a reduced affinity requires that we determine their free energies: ΔG^I_{O1}, ΔG^I_{O2}, ΔG^I_{O3} and ΔG^I_{O3*}.

- The assumption that a repressor still has some affinity for an experimentally deleted site requires that we determine the free energy of binding a free repressor to the deleted site, and the free energy of binding an impaired repressor to the deleted site. The deleted operator free energies will be denoted by a '-' such as ΔG_{O2^-}. Since Oehler et al. [61] performed two different O1 operator deletions in 1994, we also include ΔG_{O1^-a} and ΔG_{O1^-b}.

Protein Concentrations: We now describe the concentrations of RNAP, CAP and the repressor as well as the exponents α_s, α^1_s, α^2_s, α^3_s and α^4_s in equation (10). Each exponent represents the number of each species bound to the DNA for a state s. For example, if RNAP is bound to P1 and CAP is bound to C1 but no repressor is bound to the DNA, the exponents are $\alpha_s = 1$, $\alpha^1_s = 1$, $\alpha^2_s = 0$, $\alpha^3_s = 0$ and $\alpha^4_s = 0$. A zero indicates no molecules of that species are bound for state s.

In the general formulation, the exponent α_s refers to the number of bound RNAP. In the case of the *lac* operon, only one RNAP can bind at a time. Thus, $\alpha_s \in \{0, 1\}$

where 0 indicates RNAP is not bound in the state s and 1 indicates RNAP is bound for the state s. We will assume a constant concentration of active RNAP.

The concentration of the activator CAP is a function of CRP and cAMP. CAP is the CRP-cAMP complex and becomes active when one molecule of cAMP is bound to CRP. Thus we represent

$$[\text{CAP}] = [\text{CRP}]\frac{[\text{cAMP}]}{[\text{cAMP}] + K_{\text{cAMP}}}$$

where K_{cAMP} is the dissociation constant for cAMP to CRP [45]. Since the only binding site for CAP is C1, the exponent $\alpha_s^1 \in \{0, 1\}$.

The Oehler et al. [61] experiments are performed at three different values of repressor concentration: ∼10 tetramers, ∼50 tetramers, and ∼900 tetramers per cell. This value represents the total concentration, $[R_t]$, of repressor. When the inducer is absent, $[R] = [R_t]$. However, when the inducer is present $[R_t] = [R] + [R'] + [R'']$ and the concentrations of $[R]$, $[R']$ and $[R'']$ are dependent on the concentration of inducer. The dissociation constant, K_I, between a repressor monomer and an inducer molecule has been calculated by Oehler et al. [60]. Using the total repressor concentration, the total inducer concentration and K_I we can represent $[R]$, $[R']$ and $[R'']$ as follows:

$$[R] = \frac{K_I^4[R_t]}{K_I^4 + [I_f]K_I^3 + [I_f]^2K_I^2 + [I_f]^3K_I + [I_f]^4} \tag{11}$$

$$[R'] = \frac{[I_f][R]}{K_I} + \frac{[I_f]^2[R]}{3K_I^2} \tag{12}$$

$$[R''] = [R_t] - [R] - [R']. \tag{13}$$

We derive these equations for $[R]$, $[R']$ and $[R'']$ in the next chapter. In Figure 3b (in the Introduction), subfigure A represents the free repressor R, while B and C represent the repressors with one impaired binding domain, R'. D-G represent the doubly impaired binding domains included in R''.

The exponents associated with the repressor are α_s^2, α_s^3 and α_s^4. The binding sites for the repressor are O1, O2, O3 and O3*. Because O3 and O3* overlap, only one of these two operators can be bound at a time. If no repressor tetramers are bound to the DNA, α_s^2, α_s^3 and α_s^4 are zero. The highest value α_s^2, α_s^3 and α_s^4 can attain is three which represents three operators bound by three different repressors. However, since two types of repressor tetramer (impaired or not) cannot bind the same operator concurrently, it must be the case that $\alpha_s^2 + \alpha_s^3 + \alpha_s^4 \leq 3$. Additionally, if the state s includes a DNA loop, the exponent only counts the one repressor that is in the loop, not the number of operators bound.

We derive these parameter values in detail in the next chapter. We include a list of all of the parameters necessary to determine $F(I)$ in Table 5.

Table 5: A full list of free energy values and concentrations necessary to compute F. Note that the concentrations of $[R]$, $[R']$ and $[R'']$ are dependent on the concentration of free inducer ($[I_f]$) and total repressor ($[R_t]$).

ΔG_{O1}	ΔG^I_{O1}	ΔG_{O1^-a}	[RNAP]	$[R_t]$
ΔG_{O2}	ΔG^I_{O2}	ΔG_{O1^-b}	[CAP]	$[I_f]$
ΔG_{O3}	ΔG^I_{O3}	ΔG_{O2^-}	[cAMP]	[CRP]
ΔG_{O3*}	ΔG^I_{O3*}	ΔG_{O3^-}	ΔG_{O3*^-}	$[R]$
ΔG_{O12}	ΔG_{O13}	ΔG_{O13*}	$\Delta G_{O3*(O1)}$	$[R']$
ΔG_{O23}	ΔG_{O23*}	ΔG_{C1loop}	ΔG_{C1-O3}	$[R'']$
K_{C1}	K_{P1}	K_{P2}	K_{coop}	K_{cAMP}

$$[\text{CAP}] = [\text{CRP}]\frac{[\text{cAMP}]}{[\text{cAMP}] + K_{cAMP}}$$

$$[R] = \frac{K_I^4[R_t]}{K_I^4 + [I_f]K_I^3 + [I_f]^2K_I^2 + [I_f]^3K_I + [I_f]^4}$$

$$[R'] = \frac{[I_f][R]}{K_I} + \frac{[I_f]^2[R]}{3K_I^2}$$

$$[R''] = [R_t] - [R] - [R']$$

A brief literature review of existing models of the *lac* operon follows.

Existing Models of the *lac* Operon

Bistability Models

Although many models have been created to model the *lac* operon, data released in 2004 [63] showing bistability on the artificial inducer TMG provide a new target for quantitative reproduction by a model. The recent models from Santillán et al. [70, 74] and van Hoek et al. [83, 84] have attempted to match the bistability data in Figure 6 (in the Introduction).

In 1997 Wong, Gladney, and Keasling [86] published a model including an impressive number of parameter values collected from various reference papers. This was published before Ozbudak et al. [63] released their experimentally measured bistability figure, and subsequently, some of the values used in the Wong paper have been ignored in order to match the figure published by Ozbudak.

When Wong et al. [86] created their model, they were unsure of the mechanism controlling cAMP synthesis, so they wrote the model with two different options. The first assumed cAMP synthesis was inhibited by extracellular glucose, and the second assumed cAMP was inhibited by the rate at which glucose was transported through the phosphoenolpyruvate: sugar phosphotransferase system (PTS). For these two model choices, they report, "the simulation results were virtually identical."

There is also some question as to what happens to the glucose forming from lactose and allolactose digestion. What is known is that galactose is converted to a phosphorylated form of glucose, Glu6P, which is the form of glucose produced by import of external glucose by the PTS system. It is possible that the glucose first diffuses out of the cell and then re-enters through the PTS system. However, since Ozbudak et al. [63] did not see bistability on lactose, it seems more likely that the glucose is phosphorylated and digested within the cell.

In addition to their experimental bistability data, Ozbudak et al. [63] created a model for the *lac* system (see Figure 6 in the Introduction and the section Bistability Data). In the model they match three parameters to data: ρ, a repression factor; β, the TMG uptake rate per lacY molecule and α, the maximal *lac* activity if every repressor molecule were inactive. Their resulting bifurcation diagram is a function of extracellular glucose and extracellular TMG concentrations see Figure 6 in the Introduction, taken from [63]. The region of bistability is enclosed in grey. In figure **c**, the open circles indicate the saddle nodes, as implied by the data, for the indicated external concentrations of glucose and TMG. Ozbudak et al. [63] then fit the bifurcation diagram using those data points, indicating the bistable region with grey, the lower white region indicates only the lower equilibrium exists, and the upper white region indicates only the upper equilibrium exists. These experiments were also performed using IPTG which induced bistability, and lactose, where bistability was absent.

A new generation of models have been developed to incorporate this beautiful data. Using many of the same values as Wong et al. [86], van Hoek and Hogeweg [84] developed a 10-dimensional differential equation system with the idea of addressing two key problems in previous models. The first is that the model outcome depends highly on parameter values used. Though each parameter choice is supported by experimental data, the values come from different experiments with varying cell strains, temperatures, chemicals, and other factors. The same biological experiment under different conditions may provide different results. These results affect the model parameter values, with the potential of greatly changing the model predictions. Secondly, since existing models lack an analytical solution, the key terms which lead to bistability are difficult to determine.

To address the first issue, they incorporated the promoter activity function from Setty et al. [76] for mRNA production. This function results from solving a set of equilibrium reactions representing a simplified *lac* operon. In particular, the operon has only five possible states: free, repressed, bound to RNAP, bound to CAP, or bound with both RNAP and CAP. This corresponds to a subset of the primary components since they do not consider repressor mediated DNA looping. Transcription is assumed when the operon has both RNAP and CAP bound. Then they evolve the 10-dimensional system *in silico* under fluctuating glucose and lactose conditions, allowing several parameters in the Setty function to change with a fixed probability. Each change is designed to represent a possible mutation.

To address the second issue, they find an analytical expression to predict conditions during which bistability occurs. Their results showed that bistability was avoided on the natural inducer, but not on the artificial inducers.

In 2007 van Hoek and Hogeweg [83] took the same DE model and introduced noise in gene expression. Again the model parameters were evolved, but in the stochastic model [83] the final parameters predicted a higher transcription rate in comparison to the earlier model. They interpreted this result to mean the parameters evolve to minimize stochasticity in gene expression. Additionally they conclude that the dynamics of the *lac* operon are well represented by a deterministic model when on lactose, and that a stochastic representation is much more important for artificial inducers due to the increased positive feedback [83].

Narang et al. [54, 55] published several papers discussing bistability and found that dilution due to growth from lactose and/or glucose forces the system away from bistability. In the 2008 paper, Narang and Pilyugin [55] designed a model with two goals, first to update a model by Chung and Stephanopoulos [18] with a more accurate operon representation and second, to understand the absence of bistability during

growth on lactose. After updating the model, they conclude the dynamics of the new model quantitatively match the data. As for the bistability, they determine bistability is suppressed by dilution. For example, if cells are in the presence of non-galactosidic carbon sources (TMG, IPTG, etc.), their growth is independent of concentration since the cells cannot metabolize non-galactosidic sources. In the most recent paper [57] they discuss effects of inducer transport across the cell membrane, specifically diffusive influx (import without passing through permease) and carrier efflux (export through permease rather than diffusion). Though their model is biochemically based and they closely analyze repression from DNA looping, they minimize the role of CAP and assume it is independent of looping.

For many years Michael Mackey has been involved in modeling the *lac* operon. The first model his group published was a five-dimensional deterministic system representing mRNA, β-galactosidase, permease, allolactose and lactose. Another couple of deterministic models were published in 2004 [71, 89]: one was three-dimensional with mRNA, β-galactosidase, and allolactose represented; and the other six-dimensional with mRNA for β-galactosidase, mRNA for permease, β-galactosidase, permease, allolactose and cAMP. The six-dimensional system improved upon the other two models by replacing the Hill type function of mRNA production with a biophysically based representation and incorporating all three operators with DNA looping. Even though all three models used parameters from biological data, each one predicted the *lac* operon to be bistable for realistic extracellular glucose and lactose concentrations.

After the Ozbudak et al. [63] paper showed bistability for a range of external glucose and external TMG (but not on external lactose) Santillán, Mackey and Zeron published a new model in 2007 [74]. This model shared similar concepts with their prior models: inducer uptake, lactose metabolism, and simplified mRNA production functions. Results reasonably matched the 2004 Ozbudak et al. data [63] and San-

tillán et al. [74] determined the region of bistability on external inducer increased as external glucose increased, and shifted to higher inducer concentrations as the inducer metabolism ranged from 0 min^{-1}, representing artificial inducers, to 4.0×10^4 min^{-1}, for lactose. They conclude Ozbudak et al. [63] did not see bistability because the region of bistability is small and thus hard to detect for the tested concentration values. In 2008 Santillán published his most recent model [70] incorporating variable growth, the binding of partially impaired repressors, and CAP cooperativity in looping between the O1 and O3 operators. Unfortunately, in the 2008 model, after deriving parameters from data, Santillán was unable to recover the Ozbudak [63] data until several of the key parameters were reduced by 85%.

Santillán and Mackey published [73] a thorough review of bistability models for the *lac* operon in 2008.

Repression Ratio Models

In the *lac* operon, repression is primarily accomplished by a combination of competitive binding between RNAP and the *lac* repressor at the P1 and O1 sites, or by blocking transcription at the P2 site. However, looping between two of the three operator sites will cause the local concentration of repressor to be higher, increasing the probability of binding an operator, and increasing repression. Many of the models [45, 69, 80, 90] are striving to match repression data rather than exploring bistability. Repression data is recorded as ratio of measured β-galactosidase activity at 1mM IPTG over the β-galactosidase activity at 0mM IPTG:

$$\mathcal{R} = \frac{\beta\text{-galactosidase(1mM)}}{\beta\text{-galactosidase(0mM)}}.$$

This ratio is designed to capture the relative repression and since IPTG removes repression, a larger ratio indicates greater repression. Typically this data is used to learn how mutations change operon effectiveness.

Saiz and Vilar have published several models [68, 69, 85] focused on DNA looping. In their 2008 paper [69], they derive a transcription model containing a complete thermodynamic representation for repressor binding and looping. Also, they suggest that a biologically deleted operator site may still bind the repressor, but at a reduced affinity as supported by Horton et al. [37]. Unfortunately, when representing RNAP and CAP, they use phenomenological parameters rather than using a biophysical model. If CAP does not assist in loop formation, this simplification might be acceptable, but results from [39, 45, 81] suggest CAP is involved in forming short DNA loops.

In 2007 Kuhlman et al. [45] published a biophysical model and repression data they recorded. The repression data was recorded for several mutants which lacked either permease or one of the genes involved in cAMP production. The model was designed using statistical mechanics and then rearranged into a Hill function. After fitting the Hill function to the data they were able to analyze, with greater clarity, how each mutation affected the data. They concluded that induction is sensitive to repressor mediated looping and that looping is significantly enhanced by CAP.

Loop Formation Models

Zhang et al. [90] and Swigon et al. [80, 81] have created models which analyze the energy needed for loop closure. These models are interested in short to medium loops (50-180 bp) and compute the free energy required to close the loop. Early models, which accurately capture DNA behavior for longer loops by representing DNA as a stiff polymer, become grossly inaccurate when trying to match data collected for

short loops [67]. The new models allow for protein flexibility as well as DNA base pair interactions (tilt, roll, twist, etc.) to determine DNA elasticity.

PARAMETERIZING THE MODEL AND FITTING THE DATA

Parameter Selection and Estimation

We will first describe the selection of parameters involved in RNAP binding, followed by RNAP activation and CAP binding and finishing with DNA looping and the binding of the *lac* repressor.

Our model considers two binding sites P1 and P2 for RNAP on the DNA. For both sites we need to determine the binding energy of the RNAP-DNA interaction and concentration of RNAP. Further, we need to determine transcription rates for the states where RNAP is bound to P1 and to P2.

RNAP Concentration ([RNAP])

In E. coli, the active number of RNAP molecules is ∼1500 molecules per cell [16]. For a cell size of 1×10^{-15}L, the calculation in molar is:

$$\frac{\text{RNAP}}{\text{cell}} = \frac{1500 \text{ molecules}}{\text{cell}} \frac{1\text{cell}}{2.0 \times 10^{-15}\text{L}} \frac{1\text{M}}{N_A} = 1.25 \times 10^{-6}\text{M}, \tag{14}$$

where Avogadro's number is $N_A = 6.022 \times 10^{23}$ molecules/mol and the average cell size is 2.0×10^{-15}L [52]. Thus the concentration of [RNAP] = 1.25×10^{-6}M (equation (14)).

Binding of RNAP to P1 (K_{P1})

When RNAP binds a promoter, multiple steps (open complex formation and escape) must occur before transcription can begin. In 2003, Liu et al. [48] analyzed the kinetics of open complex formation for RNAP bound to P1 with and without CAP using stopped-flow kinetics and manual mixing. These two techniques record the change in fluorescence which occurs as the DNA is separated to form the open

complex. To calculate the dissociation constant K_d, the fluorescence data is fit with a curve of the form

$$k_{obs} = \frac{k_2[\text{RNAP}]}{K_d + [\text{RNAP}]} + k_{-2},$$

where k_{obs} is the recorded data. This form follows from the assumption that the following two step chemical reaction between the DNA and RNAP takes place:

$$[\text{RNAP}] + [\text{DNA}] \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} [\text{RNAP:DNA}_c] \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} [\text{RNAP:DNA}_1], \tag{15}$$

where [RNAP:DNA$_c$] represents RNAP bound to the DNA as a closed complex, and [RNAP:DNA$_1$] is the first step of open complex formation. The experimentalists observe multiple steps during open complex formation and propose the following kinetic scheme for this process:

$$[\text{RNAP}] + [\text{DNA}] \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} [\text{RNAP:DNA}_c] \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} [\text{RNAP:DNA}_1] \underset{k_{-3}}{\overset{k_3}{\rightleftharpoons}} [\text{RNAP:DNA}_2] \overset{k_4}{\rightleftharpoons} [\text{RNAP:DNA}_3].$$

Using this scheme they matched the data and determined the kinetic rates in Table 6. (k_{-3} was not reported.) The K_d value listed in the $-$CAP column represents the binding of RNAP to the DNA in the absence of CAP and is defined as:

$$K_d := \frac{k_{-1}}{k_1} = 5.6 \times 10^{-6} \text{ M}.$$

This value represents our $K_{P1} = 5.6 \times 10^{-6}$M. We conclude the dissociation rate between RNAP and the P1 promoter is $K_{P1} = k_{-1}/k_1 = K_d$ where K_d is the dissociation constant in the $-$CAP column of Table 6. Thus $K_{P1} = 5.6 \times 10^{-6}$M.

RNAP-DNA Interaction in Presence of CAP (K_{coop})

We will next discuss the K_d value represented in the +CAP column of Table 6 and our calculation for determining what portion of K_d is the cooperativity between CAP and RNAP (K_{coop}).

Table 6: Dissociation constants and kinetic rates for RNAP bound to the P1 site on the DNA with and without CAP, determined by stopped-flow kinetics and manual mixing as published by Liu et al. [48]. The K_d value in the $-$CAP column represents the binding of RNAP to P1. We denote this by K_{P1} rather than K_d. Note, the experimentalists either were not able to determine k_{-3} or did not report the value.

Rate constants	+CAP	$-$CAP
K_d (μM)	$0.15 \pm 0,03$	0.56 ± 0.09
$k_2(\mathrm{s}^{-1})$	15.14 ± 0.14	9.10 ± 0.23
$k_{-2}(\mathrm{s}^{-1})$	1.78 ± 0.16	7.86 ± 0.32
$k_3(\mathrm{s}^{-1})$	2.5	3.3
$k_4(\mathrm{s}^{-1})$	3.04×10^{-3}	1.7×10^{-3}

When calculating the kinetic constants for RNAP bound with CAP, Liu et al. [48] use the same chemical equation with the assumption that all DNA is bound by CAP prior to introducing RNAP:

$$[\mathrm{RNAP}] + [\mathrm{CAP{:}DNA}] \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} [\mathrm{RNAP{:}CAP{:}DNA_c}] \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} [\mathrm{RNAP{:}CAP{:}DNA_1}] \underset{k_{-3}}{\overset{k_3}{\rightleftharpoons}} [\mathrm{RNAP{:}CAP{:}DNA_2}] \overset{k_4}{\rightleftharpoons} [\mathrm{RNAP{:}CAP{:}DNA_3}].$$

Here $K_d := \frac{k_{-1}}{k_1}$ and is the dissociation constant for RNAP and the CAP-DNA complex. To fit the fluorescence data, the concentration of CAP bound DNA is taken to be the concentration of DNA in the system. However, when CAP is added to the system, the reaction

$$[\mathrm{CAP}] + [\mathrm{DNA}] \overset{K_{C1}}{\rightleftharpoons} [\mathrm{CAP{:}DNA}] \tag{16}$$

with dissociation constant K_{C1} also occurs, implying some fraction of the DNA will not be bound by CAP. For this mixed solution of DNA with and without CAP, it should be assumed that some fluorescence occurs from RNAP bound to P1 without CAP. To fit the fluorescence data, k_{obs}, the model should include the chemical equation (16) representing the mix of activated and not activated RNAP fluorescence. The reported value for activated RNAP is $K_d = 0.15 \times 10^{-6}$ M listed in Table 6,

and though this fits the curve, the assumed form is wrong as it does not include the interaction between CAP and the DNA.

To recover the actual dissociation constant between RNAP and CAP bound DNA, K_1, we assume DNA and CAP come to equilibrium before introducing RNAP.

Rather than assuming the form of equation (15), the first step should include the CAP and DNA interaction:

$$[\text{CAP:DNA}] + [\text{RNAP}] \underset{r_2}{\overset{r_1}{\rightleftharpoons}} [\text{CAP:RNAP:DNA}_c] \underset{r_4}{\overset{r_3}{\rightleftharpoons}} [\text{CAP:RNAP:DNA}_1] \tag{17}$$

with $K_1 = \frac{r_2}{r_1}$. To simplify the notation we let

C:=CAP, R:=RNAP, D:=DNA, CD:=CAP:DNA, CRD:=CAP:RNAP:DNA. Further, we use subscript c to denote the closed complex and 1 denotes a complex in an intermediate step toward open complex. If we apply the steady-state condition to $[CRD_c]$ in equation (17) we have:

$$r_1[R][CD] + r_4[CRD_1] = (r_2 + r_3)[CRD_c]. \tag{18}$$

From equation (16) we know the equilibrium constant $K_{C1} = \frac{[C][D]}{[CD]}$. Rearranging this equation we have:

$$[CD] = \frac{[C][D]}{K_{C1}}. \tag{19}$$

We now substitute for $[CD]$ in equation (18) to get

$$\frac{r_1[R][C][D]}{K_{C1}} + r_4[CRD_1] = (r_2 + r_3)[CRD_c]. \tag{20}$$

We also use the fact that the total DNA ($[D_o]$) is preserved. The DNA conservation law is

$$[D] = [D_o] - [CD] - [CRD_c] - [CRD_1]. \tag{21}$$

If we substitute equation (19) into the DNA conservation law, $[D]$ can be rewritten as:

$$[D] = \frac{[D_o] - [CRD_c] - [CRD_1]}{1 + [C]/K_{C1}}. \tag{22}$$

Equations (22) and (20) can be combined and simplified in a few steps as follows:

$$\begin{aligned}
\frac{r_1[R][C]}{K_{C1}}\frac{[D_o]-[CRD_c]-[CRD_1]}{1+[C]/K_{C1}}+r_4[CRD_1] &= (r_2+r_3)[CRD_c] \\
\frac{r_1[R][C]([D_o]-[CRD_1])}{K_{C1}+[C]}-\frac{r_1[R][C][CRD_c]}{K_{C1}+[C]}+r_4[CRD_1] &= (r_2+r_3)[CRD_c] \\
\frac{\frac{r_1[R][C]([D_o]-[CRD_1])}{K_{C1}+[C]}+r_4[CRD_1]}{\left(r_2+r_3+\frac{r_1[R][C]}{K_{C1}+[C]}\right)} &= [CRD_c]. \qquad (23)
\end{aligned}$$

Let

$$\gamma = \frac{[C]}{K_{C1}+[C]}. \qquad (24)$$

Then, equation (23) can be written as:

$$[CRD_c] = \frac{r_1[R]\gamma([D_o]-[CRD_1])+r_4[CRD_1]}{(r_2+r_3+r_1[R]\gamma)}. \qquad (25)$$

Liu et al. [48] record the rate at which $[CRD_1]$ appears. To match this data with our model we derive the rate of $[CRD_1]$ production from equation (17). We then substitute equation (25) into the differential equation and simplify.

$$\begin{aligned}
\frac{d[CRD_1]}{dt} &= r_3[CRD_c]-r_4[CRD_1] \\
&= r_3\frac{r_1[R]\gamma([D_o]-[CRD_1])+r_4[CRD_1]}{(r_2+r_3+r_1[R]\gamma)}-r_4[CRD_1] \\
&= \frac{r_1r_3[R]\gamma[D_o]-r_1r_3[R]\gamma[CRD_1]+r_3r_4[CRD_1]}{(r_2+r_3+r_1[R]\gamma)}-r_4[CRD_1] \\
&= \frac{r_1r_3[R]\gamma[D_o]}{(r_2+r_3+r_1[R]\gamma)}-\left(\frac{r_1r_3[R]\gamma-r_3r_4}{r_2+r_3+r_1[R]\gamma}+r_4\right)[CRD_1] \\
&= \beta-\alpha[CRD_1] \qquad (26)
\end{aligned}$$

where we set

$$\alpha := \frac{r_1r_3[R]\gamma-r_3r_4}{r_2+r_3+r_1[R]\gamma}+r_4 \qquad (27)$$

and

$$\beta := \frac{r_1r_3[R]\gamma[D_o]}{(r_2+r_3+r_1[R]\gamma)}.$$

Thus,

$$\frac{d[CRD_1]}{dt} = \beta - \alpha[CRD_1] \tag{28}$$

with solution $[CRD_1] = \frac{-\beta}{\alpha}e^{-\alpha t} + \frac{\beta}{\alpha}$.

When the experimentalists record the data, they are able to track the rate at which $[CRD_1]$ appears. We would like to fit our rate of $[CRD_1]$ production, α, to their observed rate of production, k_{obs}. Thus, we assume

$$k_{obs} = \alpha = \frac{r_1 r_3[R]\gamma - r_3 r_4}{r_2 + r_3 + r_1[R]\gamma} + r_4$$

which can be rewritten as

$$\begin{aligned}
\alpha &= \frac{r_1 r_3[R]\gamma - r_3 r_4 + r_4 r_2 + r_3 r_4 + r_1 r_4[R]\gamma}{r_2 + r_3 + r_1[R]\gamma} \\
&= \frac{r_1 r_3[R]\gamma + r_4(r_2 + r_1[R]\gamma)}{r_2 + r_3 + r_1[R]\gamma} \\
&= \frac{r_3[R]\gamma + r_4(r_2/r_1 + [R]\gamma)}{r_2/r_1 + r_3/r_1 + [R]\gamma} \\
&= \frac{r_3[R]\gamma + r_4(K_1 + [R]\gamma)}{K_1 + r_3/r_1 + [R]\gamma}
\end{aligned} \tag{29}$$

with $K_1 = r_2/r_1$.

If $r_3 << r_2$, then $\frac{r_3}{r_1} << \frac{r_2}{r_1} = K_1$ so r_3/r_1 may be ignored in the denominator. This assumption implies RNAP falls off the DNA much faster than the RNAP:CAP:DNA complex moves on to the first step of open complex formation. This assumption is standard when fitting rapid reaction kinetic data [48, 78]. Thus

$$\begin{aligned}
\alpha &= \frac{r_3[R]\gamma + r_4(K_1 + [R]\gamma)}{K_1 + [R]\gamma} \\
&= \frac{r_3[R]\gamma}{K_1 + [R]\gamma} + r_4 \\
&= \frac{r_3[R]}{K_1/\gamma + [R]} + r_4.
\end{aligned} \tag{30}$$

When Liu et al. [48] fit their data, they use the equation

$$k_{obs} = \frac{k_2[R]}{K_d + [R]} + k_{-2},$$

which has the form of equation (30) with $k_2 = r_3$, $K_d = K_1/\gamma$ and $k_{-2} = r_4$. Thus, their reported value K_d in Table 6 in the + CAP column corresponds to the value K_1/γ in our model:

$$K_d = K_1/\gamma. \tag{31}$$

Recall, we would like to determine the cooperativity between CAP bound to C1 and RNAP bound to P1. K_1 can also be related to the change in free energy between the state with CAP bound to C1 and RNAP bound to P1 and the state with CAP bound to the C1 by:

$$K_1 = \exp(-(\Delta G_{C1} + \Delta G_{P1} + \Delta G_{coop} - \Delta G_{C1})/RT). \tag{32}$$

Here, the free energy for the state with CAP bound to C1 and RNAP bound to P1 is $\Delta G_{C1} + \Delta G_{P1} + \Delta G_{coop}$. The free energy for the state with CAP bound to C1 is ΔG_{C1}. This can be rearranged as follows:

$$\begin{aligned} K_1 &= \exp(-(\Delta G_{P1} + \Delta G_{coop})/RT) \\ &= \exp(-\Delta G_{P1}/RT)\exp(-\Delta G_{coop}/RT) \\ &= K_{P1}K_{coop}. \end{aligned} \tag{33}$$

Refer to Appendix C for more detail on the relationship between the dissociation constant and the free energy.

Combining equations (31) and (33) we have:

$$\begin{aligned} K_d\gamma = K_1 &= K_{P1}K_{coop} \\ K_d\gamma &= K_{P1}K_{coop}. \end{aligned} \tag{34}$$

In the prior section, we showed the dissociation constant for RNAP to P1 without CAP is $K_{P1} = 5.62 \times 10^{-7}$ M. γ is defined in equation (24) as $\gamma = \frac{[C]}{K_{C1}+[C]}$ and the dissociation constant $K_d = 1.52 \times 10^{-7}$ M is reported in the +CAP column of Table 6.

We would like to determine K_{coop}. Substituting these values into equation (34) we have:

$$K_d\gamma = K_{P1}K_{coop} \tag{35}$$

$$1.52 \times 10^{-7}\frac{[C]}{[C]+K_{C1}} = 5.62 \times 10^{-7}K_{coop}. \tag{36}$$

Therefore, the cooperative factor by which CAP increases the binding affinity for RNAP to P1 when CAP is bound to C1 is

$$K_{coop} = \frac{1.5}{5.6}\frac{[C]}{[C]+K_{C1}} \tag{37}$$

where $[C]$ is fixed at the initial (but unknown) CAP concentration used in the Liu et al. experiments [48].

Binding of RNAP to P2 (K_{P2})

Malan and McClure [49] measured the ratio of promoter activity between P1 and P2 in the absence of cAMP. To determine a dissociation constant for K_{P2} the ratio of promoter activity was multiplied by K_{P1}. This provides us with a relative binding energy:

$$K_{P2} \simeq \frac{2.2}{5}K_{P1} \text{ M}.$$

Transcription Initiation Rate Constants (k_f and k_{fC1})

The transcription initiation rate should represent the average time required for a bound RNAP to form an open complex and escape the promoter. Based on the values from Liu et al. [48] the transcription initiation rates should be

$$k_f \simeq \left(\frac{k_2}{k_{-2}}\right)\left(\frac{k_3}{k_{-3}}\right)(k_4) \text{ s}^{-1} \text{ (from the } -\text{CAP column of Table 6)} \tag{38}$$

$$k_{fC1} \simeq \left(\frac{k_2}{k_{-2}}\right)\left(\frac{k_3}{k_{-3}}\right)(k_4) \text{ s}^{-1} \text{ (from the } +\text{CAP column of Table 6)} \tag{39}$$

but k_{-3} was not recovered from the data. Because k_4 is more limiting we assume $k_3/k_{-3} = 1$. Thus inserting the values from Table 6 and multiplying by 60 sec/1 min:

$$k_{fC1} \simeq \left(\frac{15.14}{1.78}\right)\left(3.04 \times 10^{-3}\right)(60)\ \text{min}^{-1},$$
$$k_f \simeq \left(\frac{9.1}{7.86}\right)\left(1.75 \times 10^{-3}\right)(60)\ \text{min}^{-1},$$

are the transcription initiation rates at P1 with and without CAP bound, respectively.

The transcription initiation rate for P2 was assumed to be the same as the non-activated P1 rate, k_f. We conclude $k_{fC1} \simeq 1.55$ min^{-1} and $k_f \simeq 0.12$ min^{-1}.

Activator CAP

The activator of the *lac* operon is the CRP protein dimer with one cAMP molecule bound, forming the CAP complex. Ozbudak et al. [63] found CAP levels to decrease as glucose increased, but to be independent of TMG. Inada et al. [40] reported an increase in intracellular cAMP concentration between the first and second growth phase as the cell switched from glucose to lactose, but similar levels during growth on either sugar. The analysis here will consider a fixed glucose concentration and a fixed inducer concentration at steady state, thus the concentrations for cAMP and CRP will be held constant. The active concentration of CAP depends on the internal cAMP and CRP concentrations. We model this dependence in the same way as Kuhlman et al. [45]:

$$[\text{CAP}] = [\text{CRP}]\frac{[\text{cAMP}]}{[\text{cAMP}] + K_{cAMP}}.$$

Here [CAP] is the concentration of the activator, [CRP]$= 1.5 \times 10^{-6}$M ([20]) is the concentration of dimers without cAMP bound, [cAMP]$= 3.0 \times 10^{-6}$M ([40]) is the internal cAMP concentration and the value $K_{cAMP} = 3.0 \times 10^{-6}$ M ([9]). This allows us to calculate the concentration of CAP. For the dissociation constant between CAP and C1 on the DNA we use $K_{C1} \simeq 1.2 \times 10^{-8}$ M ([47]).

To conclude, for the parameters related to the activator CAP, we will use the values listed in Table 7.

Table 7: Parameter values determined in the section Activator CAP.

[CAP]	$= [\text{CRP}]\frac{[\text{cAMP}]}{[\text{cAMP}]+K_{cAMP}}$
[CRP]	$= 1.5 \times 10^{-6}$ M
[cAMP]	$= 3.0 \times 10^{-6}$ M
K_{cAMP}	$= 3.0 \times 10^{-6}$ M
K_{C1}	$\simeq 1.2 \times 10^{-8}$ M

lac Repressor

The *lac* repressor is the most complicated component of the transcriptional regulation. The following parameters are required for the model:

1. the concentration of free repressor ($[R]$), repressor with one impaired binding domain ($[R']$) and repressor with two impaired binding domains ($[R'']$);

2. the binding energy for a free binding domain of the repressor to each operator: O1, O2, O3, O3* and O3*(O1) for mutants (1,0,1) and (3,0,1) (ΔG_{O1}, ΔG_{O2}, ΔG_{O3}, ΔG_{O3^*} and $\Delta G_{O3^*(O1)}$);

3. the binding energy for an inducer impaired binding domain of the repressor to each operator: O1, O2, O3, O3* and O3* for mutants (1,0,1) and (3,0,1) (ΔG^I_{O1}, ΔG^I_{O2}, ΔG^I_{O3}, $\Delta G^I_{O3^*}$ and $\Delta G^I_{O3^*(O1)}$);

4. the energy of looping the DNA between any pair of the operators: O1, O2, O3 or O3* (ΔG_{O12}, ΔG_{O13}, ΔG_{O23}, ΔG_{O13^*}, and ΔG_{O23^*});

5. the increase in energy required to bind O1 with repressor and C1 with CAP at the same time (ΔG_{C1-O3});

6. the specific binding energy of the repressor to an experimentally deleted site (ΔG_{O1^-a}, ΔG_{O1^-b}, ΔG_{O2^-}, ΔG_{O3^-}, $\Delta G_{O3^{*-}}$, $\Delta G^I_{O1^-a}$, $\Delta G^I_{O1^-b}$, $\Delta G^I_{O2^-}$, $\Delta G^I_{O3^-}$ and $\Delta G^I_{O3^{*-}}$).

To estimate the concentration of repressor, we use the values published with the repression data [61, 62]: wildtype levels (WT), five times WT, and 90 times WT. In molar, these translate to: WT $\simeq 8.3 \times 10^{-9}$ M, $5 \times$ WT $\simeq 42 \times 10^{-9}$ M, and $90 \times$ WT $\simeq 747 \times 10^{-9}$ M for an average cell size of 2.0×10^{-15}L ([52]).

Using mass-action kinetics, the total concentration of inducer, $[I_t]$; repressor, $[R_t]$; and the dissociation constant for inducer to repressor will determine the repressor concentrations $[R]$, $[R']$, and $[R'']$. Oehler et al. [60] calculate K_I, the dissociation constant for the inducer IPTG to repressor by matching a curve to repression data. A brief review of their calculations are below, followed by a derivation for the concentration of non-inducer bound (free) repressor.

The calculations from Oehler et al. [60] use the fact that the repressor tetramer is composed of two binding domains and each binding domain is a dimer. In the notation, they only consider the one binding domain which is interacting with the O1 operator. The binding domain can have zero, one or two molecules of IPTG bound. They are trying to determine how many molecules of IPTG must be bound before

the binding domain ineffectively binds the operator. They propose that

$$\begin{aligned}\mathcal{R} &= \frac{\text{occupied operator}}{\text{free operator}} \\ &= \frac{[D_t]}{K_{O1}}\frac{K_I^2}{(K_I+[I_t])^2} + \frac{[D_t]}{K_{O1-I}}\frac{2K_I}{(K_I+[I_t])^2} + \frac{[D_t]}{K_{O1-2I}}\frac{[I_t]^2}{(K_I+[I_t])^2}\end{aligned} \tag{40}$$

where $[D_t]$ is the total concentration of binding domains. K_{O1} is the equilibrium dissociation constant for the repressor-O1 complex, K_{O1-I} is the equilibrium dissociation constant when the repressor is bound to O1 and one inducer is also bound to the binding domain, and K_{O1-2I} is the equilibrium dissociation constant when the repressor binding domain is bound to O1 with two inducer molecules bound. K_I is the equilibrium dissociation constant for the inducer IPTG binding one monomer of the repressor. The affinity of the *lac* repressor binding domain to the operator is $\sim$1000-fold lower when saturated with IPTG than that of *lac* repressor in the absence of inducer [11]. Due to this fact, it is assumed the repressor binding domain with two inducer molecules bound contributes very little to repression at low experimental concentrations of inducer. Therefore equation (40) can be reduced to

$$\mathcal{R} = \frac{[D_t]}{K_{O1}}\frac{K_I^2}{(K_I+[I_t])^2} + \frac{[D_t]}{K_{O1-I}}\frac{2K_I}{(K_I+[I_t])^2}. \tag{41}$$

Using non-linear regression analysis Oehler et al. [60] determined that once a binding domain is bound by one inducer, the repressor is no longer able to contribute to repression. Therefore equation (41) can be reduced further to

$$\mathcal{R} = \frac{[D_t]}{K_{O1}}\frac{K_I^2}{(K_I+[I_t])^2} \tag{42}$$

with

$$K_I = 6.7(\pm 0.4) \times 10^{-6}\ \text{M}. \tag{43}$$

Using this value of K_I we compute the probability that a repressor is free from inducer. First consider the probability that a repressor monomer is free from inducer

(IPTG):

$$P_f = \frac{K_I}{K_I + [I_t]},$$

where $[I_t]$ is the total concentration of IPTG. Then, the concentration of free repressor $[R]$ is calculated using mass action kinetics. Consider these four reactions where the dissociation constant $K_I = \frac{k_-}{k_+}$:

$$\begin{aligned}
[I_f] + [R] &\underset{k_-}{\overset{k_+}{\rightleftharpoons}} [I_1R] \\
[I_f] + [I_1R] &\underset{k_-}{\overset{k_+}{\rightleftharpoons}} [I_2R] \\
[I_f] + [I_2R] &\underset{k_-}{\overset{k_+}{\rightleftharpoons}} [I_3R] \\
[I_f] + [I_3R] &\underset{k_-}{\overset{k_+}{\rightleftharpoons}} [I_4R].
\end{aligned} \tag{44}$$

There are four ways to add one inducer to one repressor, each of which behaves as equation (44). The rate of change of free inducer, I_f can be represented as:

$$[\dot{I_f}] = 4k_-[I_1R] - 4k_+[I_f][R]. \tag{45}$$

Also, since the inducer is conserved, the conservation law for inducer is:

$$[I_t] = [I_f] + [I_1R] + 2[I_2R] + 3[I_3R] + 4[I_4R]. \tag{46}$$

If we are at equilibrium we can solve equation (45) for I_1R as follows:

$$\begin{aligned}
0 &= 4k_-[I_1R] - 4k_+[I_f][R] \\
[I_1R] &= \frac{[I_f][R]}{K_I}.
\end{aligned} \tag{47}$$

Equation (47) can be inserted into the inducer conservation law (equation (46)) and

$$[I_t] = [I_f] + \frac{[I_f][R]}{K_I} + 2[I_2R] + 3[I_3R] + 4[I_4R].$$

Similarly, if one inducer is bound there are three possible positions able to accept an inducer, in addition to the choice of losing the inducer currently bound. The differential equation for I_1R is:

$$\dot{[I_1R]} = 3k_-[I_2R] - 3k_+[I_f][I_1R] + k_+[I_f][R] - k_-[I_1R].$$

Assuming steady state conditions and substituting in equation (47) we find a representation for I_2R:

$$\begin{aligned} 0 &= 3k_-[I_2R] - 3k_+[I_f][I_1R] + k_+[I_f][R] - k_-[I_1R] \\ 0 &= 3k_-[I_2R] - 3k_+[I_f][I_1R] \\ [I_2R] &= \frac{[I_f][I_1R]}{K_I} \\ &= \frac{[I_f]}{K_I}\frac{[I_f][R]}{K_I} \\ &= \frac{[I_f]^2[R]}{K_I^2}. \end{aligned} \tag{48}$$

The equation for $[I_3R]$ follows similarly. The differential equation for I_2R is represented as:

$$\dot{[I_2R]} = 2k_-[I_3R] - 2k_+[I_f][I_2R] + 2k_+[I_f][I_1R] - 2k_-[I_2R].$$

Assuming steady state leads to:

$$0 = 2k_-[I_3R] - 2k_+[I_f][I_2R] + 2k_+[I_f][I_1R] - 2k_-[I_2R]$$

which can be simplified since $2k_+[I_f][I_1R] = 2k_-[I_2R]$ in steady state. Thus,

$$\begin{aligned} 0 &= k_-[I_3R] - k_+[I_f][I_2R] \\ [I_3R] &= \frac{[I_f][I_2R]}{K_I} \\ &= \frac{[I_f]^3[R]}{K_I^3}. \end{aligned} \tag{49}$$

The final differential equation is represented as:

$$\dot{[I_3R]} = k_-[I_4R] - k_+[I_f][I_3R] + 4k_+[I_f][I_2R] - 4k_-[I_3R].$$

Again, assuming we are in steady state, this equals zero, and we solve for $[I_4R]$:

$$\begin{aligned} 0 &= k_-[I_4R] - k_+[I_f][I_3R] + 4k_+[I_f][I_2R] - 4k_-[I_3R] \\ 0 &= k_-[I_4R] - k_+[I_f][I_3R] \\ [I_4R] &= \frac{[I_f][I_3R]}{K_I} \\ &= \frac{[I_f]^4[R]}{K_I^4} \end{aligned} \tag{50}$$

Thus, substituting equations (48)-(50) into the conservation law for inducer (equation (46)) we have

$$[I_t] = [I_f] + \tfrac{[I_f][R]}{K_I} + 2\tfrac{[I_f]^2[R]}{K_I^2} + 3\tfrac{[I_f]^3[R]}{K_I^3} + 4\tfrac{[I_f]^4[R]}{K_I^4}.$$

To determine the total repressor $[R_t]$ concentration, we begin with the conservation law for the repressor:

$$[R_t] = [R] + [I_1R] + [I_2R] + [I_3R] + [I_4R]. \tag{51}$$

Substituting the equations (47)-(50) into equation (51) leads to:

$$\begin{aligned} [R_t] &= [R] + \frac{[I_f][R]}{K_I} + \frac{[I_f]^2[R]}{K_I^2} + \frac{[I_f]^3[R]}{K_I^3} + \frac{[I_f]^4[R]}{K_I^4} \\ &= [R]\left(1 + \frac{[I_f]}{K_I} + \frac{[I_f]^2}{K_I^2} + \frac{[I_f]^3}{K_I^3} + \frac{[I_f]^4}{K_I^4}\right). \end{aligned}$$

Thus, the concentration of free repressor can be calculated as:

$$\begin{aligned} [R] &= \frac{[R_t]}{\left(1 + \frac{[I_f]}{K_I} + \frac{[I_f]^2}{K_I^2} + \frac{[I_f]^3}{K_I^3} + \frac{[I_f]^4}{K_I^4}\right)} \\ &= \frac{K_I^4[R_t]}{K_I^4 + [I_f]K_I^3 + [I_f]^2K_I^2 + [I_f]^3K_I + [I_f]^4}. \end{aligned}$$

The concentration of repressors with one impaired binding domain can be computed as the sum of equation (47) plus one-third of equation (48):

$$\begin{aligned}[R'] &= [I_1R] + [I_2R]/3 \\ &= \frac{[I_f][R]}{K_I} + \frac{[I_f]^2[R]}{3K_I^2}\end{aligned}$$

and using the conservation law (equation (51)):

$$[R''] = [R_t] - [R] - [R'].$$

To summarize, we use these expressions to determine the concentration of repressors with no, one and two impaired binding domains (R, R', and R'' respectively). The parameters introduced in this section are listed in Table 8.

Table 8: Total concentration of repressor at wildtype (WT), 5 × WT and 90 × WT concentrations, the IPTG-repressor dissociation constant, K_I, the average E. coli cell size and concentrations for repressor tetramers with no, one and two impaired binding domains.

WT $[R_t]$	$\simeq 8.3 \times 10^{-9}$ M
$5 \times$ WT $[R_t]$	$\simeq 42 \times 10^{-9}$ M
$90 \times$ WT $[R_t]$	$\simeq 747 \times 10^{-9}$ M
K_I	$= 6.7(\pm 0.4) \times 10^{-6}$ M
cell size	$= 2.0 \times 10^{-15}$ L
$[R]$	$= \dfrac{K_I^4[R_t]}{K_I^4 + [I_f]K_I^3 + [I_f]^2K_I^2 + [I_f]^3K_I + [I_f]^4}$
$[R']$	$= \dfrac{[I_f][R]}{K_I} + \dfrac{[I_f]^2[R]}{3K_I^2}$
$[R'']$	$= [R_t] - [R] - [R']$

Loop Formation

One of the key components of repression occurs when one repressor binds two operators forming a loop in the DNA. This can be observed in the experimental results from Oehler et al. [61, 62] in which they design a repressor mutant capable of

form dimers, but not tetramers. The repressor dimers still bind the operators with the same affinity, but the ability to loop is removed. Using the mutated repressor, they measure the repression at two concentrations: 5×WT and 90×WT, observing repression levels comparable to those with two deleted operators. For comparison, at a concentration of 5×WT the tetramers improve repression between (1,0,0) and (1,2,0) or (1,0,3) by about an order of magnitude. This increase in repression in the presence of tetramers results from having two repressor binding sites in (1,2,0) and (1,0,3), allowing a repressor to loop the DNA by binding both operators. When the repressor is mutated so it only forms dimers, the (1,2,0) and (1,0,3) repression levels are similar to (1,0,0) since no loop can form to increase the repression.

For repression to occur, the O1 position must be bound by a repressor. Interestingly, even though the free energy of looping the DNA between any of the *lac* operators is positive (requires energy to form), the looped configuration seems to be preferred at the WT and 5×WT concentrations of repressor, but not for 90×WT. To understand this behavior, one must compare the total free energy of the looped and non-looped conformations. Consider the operators sites O1 and O2. Recall that ΔG_{O1} and ΔG_{O2} are the free energies for binding a repressor to O1 and O2, respectively. Also, the binding of a repressor to an operator results in a decrease in energy ($\Delta G_{O1} < 0$ and $\Delta G_{O2} < 0$), but to form a DNA loop between two operators increases the free energy, $\Delta G_{O12} > 0$. Then,

$$\Delta G_{O1} + \Delta G_{O2} < \Delta G_{O1} + \Delta G_{O2} + \Delta G_{O12} < \Delta G_{O1}.$$

However, when a protein binds the DNA, the concentration of protein also affects the probability of complex formation. Recall from equation (10) in the section, Differential Equation Model of the *lac* Operon, the probability of a state s is represented as

$\mathbb{P}(s)$ where

$$\mathbb{P}(s) = \frac{K_B(s)[RNAP]^{\alpha_s}[CAP]^{\alpha_s^1}[R]^{\alpha_s^2}[R']^{\alpha_s^3}[R'']^{\alpha_s^4}}{Z}. \tag{52}$$

We would like to compare the probability of forming a DNA loop between O1 and O2. Let s_1 be the state formed when one repressor binds operator O1 and another repressor binds operator O2. The probability of the state s_1 is

$$\mathbb{P}(s_1) = \frac{K_B(s_1)[R]^2}{Z}$$

with

$$K_B(s_1) = \exp(-(\Delta G_{O1} + \Delta G_{O2})/RT),$$

where $[R]$ is the repressor concentration, and the exponent represents that two different repressors are bound in the state s_1. Let s_2 be the state in which one repressor is bound to operators O1 and O2. The probability of the state s_2 is

$$\mathbb{P}(s_2) = \frac{K_B(s_2)[R]}{Z}$$

where

$$K_B(s_2) = \exp(-(\Delta G_{O1} + \Delta G_{O2} + \Delta G_{O12})/RT).$$

At low concentrations ($[R] \sim 10^{-9}$ M), the probability of the state, s_2, with the O1-O2 loop is higher than the probability of the state, s_1, with two repressors binding the O1 and O2 operators without looping,

$$\mathbb{P}(s_2) > \mathbb{P}(s_1)$$
$$\frac{K_B(s_2)[R]}{Z} > \frac{K_B(s_1)[R]^2}{Z}.$$

However, as the repressor concentration increases, $[R]$ increases and the inequality switches (see figure 11)

$$\mathbb{P}(s_2) < \mathbb{P}(s_1)$$
$$\frac{K_B(s_2)[R]}{Z} < \frac{K_B(s_1)[R]^2}{Z}.$$

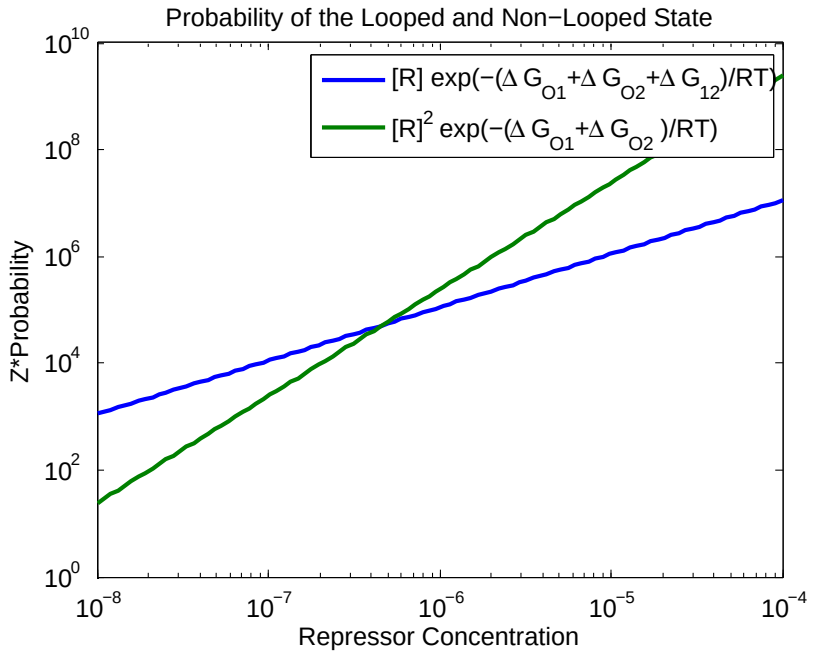


Figure 11: Repressor concentration is along the x-axis and the product of the partition function Z and the probability of the state is represented along the y-axis ($Z\mathbb{P}(s)$). The blue curve represents the looped conformation, the green curve represents the non-looped state. Notice how the cooperativity is favorable only up to some repressor concentration threshold. For this figure $\Delta G_{O12} = 9$ kcal/mol.

More generally, the switching point which determines the favorable state occurs when the probability of the non-looped state with two repressors bound ($[R]^2$) equals the probability of the looped state with one repressor bound:

$$\frac{K_B(s_1)[R]^2}{Z} = \frac{K_B(s_2)[R]}{Z}$$
$$\frac{e^{-(\Delta G_i+\Delta G_j)/RT}[R]^2}{Z} = \frac{e^{-(\Delta G_i+\Delta G_j+\Delta G_{Oij})/RT}[R]}{Z}.$$

This simplifies to,

$$[R] = e^{-\Delta G_{Oij}/RT},$$

where $i \neq j$, i and j are operators, $[R]$ is the repressor concentration, and RT is the gas constant times temperature. As ΔG_{Oij} increases (which implies it is harder

to loop), the critical concentration of repressors above which looping is favored, is decreased.

To directly determine the energy of looping is difficult; therefore it is often predicted through modeling. Using a base-pair level theory of DNA elasticity, the Olson group published two papers modeling the DNA loop between operators O1 and O3 [80,81]. They describe several types of loop formation and determine the likelihood of occurrence based on the predicted free energy. Because the distance between the operators is so short, 92 base pairs, the predicted energy to form the loop is large.

The looping of the DNA is a well accepted component of the *lac* operon, but the results from the Olson group encourage the additional inclusion of O3*, CAP looping assistance, and C1-O3 steric interference into the model. We estimated the the looping energies based on the values published by Swigon and Olson [81]. However, our model requires considerably smaller values to match the repression data. Thus, ΔG_{O12}, ΔG_{O13}, ΔG_{O13*}, ΔG_{O23} and ΔG_{O23*} are treated as free parameters with the constraint that the values must be positive. For ΔG_{C1loop}, Swigon and Olson suggest for the O1-O3 loop this value should be around -3.1kcal/mol. Due to lack of information about CAP looping assistance and C1-O3 interference, $\Delta G_{C1loop} \leq 0$ and $\Delta G_{C1-O3} \geq 0$ are also free parameters.

Binding Energies of the Repressor

We must find dissociation constants for each operator (ΔG_{O1}, ΔG_{O2}, ΔG_{O3}, ΔG_{O3*}, $\Delta G_{O3*(O1)}$, ΔG^I_{O1}, ΔG^I_{O2}, ΔG^I_{O3}, ΔG^I_{O3*}, $\Delta G^I_{O3*(O1)}$, ΔG_{O1^-a}, ΔG_{O1^-b}, ΔG_{O2^-}, ΔG_{O3^-}, ΔG_{O3*^-}, $\Delta G^I_{O1^-a}$, $\Delta G^I_{O1^-b}$, $\Delta G^I_{O2^-}$, $\Delta G^I_{O3^-}$, and $\Delta G^I_{O3*^-}$). The free energy for O1 has been measured repeatedly, but the free energy for the other operators, in particular any experimentally deleted site as well as O3*, has not been published. For O1, the equilibrium constant (K_{O1}) and the free energy (ΔG_{O1}) for

repressor to operator are related by the equation $K_{O1} = \exp(-\Delta G_{O1}/RT)$ where ΔG_{O1} is the free energy in kcal/mol, and RT is the product of the universal gas constant and temperature. See Appendix C for additional information on the equilibrium constant as a function of free energy. In terms of free energy, the binding energy for repressor to O1 ranges from -12.9 kcal/mol $\geq \Delta G_{O1} \geq -14.1$ kcal/mol [37, 60]. It has been reported that a repressor binds O1 $\sim$10 times stronger than O2 and $\sim$300 times stronger than O3 [51]. These ranges will be used to restrict the possible binding affinity for repressor to operator.

To determine the free energy for inducer impaired repressor to operators, O3*, and low specific binding to deleted sites, the incremental binding measurements for the *lac* repressor to O1 from Horton et al. [37] will be used. When Oehler et al. [61, 62] mutated the *lac* operon, they published the central 21 base pairs for each operator mutation sequence. Horton et al. [37] report the free energy contribution of each base pair to the whole binding affinity for repressor to O1, as well as the difference per base pair in the binding affinity for the IPTG impaired *lac* repressor to O1. The reported energies are in addition to a basal level of binding energy between the *lac* repressor and the DNA. We denote the basal level of binding ΔG_b. In Figure 12, the reported data for each base pair is displayed on the left, and the values determined from the figure for the central 21 base pairs are listed on the right. Each row in Figure 12B represents one base pair, its incremental binding affinity $\pm$ error followed by the incremental binding affinity in the presence of IPTG $\pm$ error in kcal/mol.

We now describe how we use Figure 12B to compute the binding energies of the repressor to various operators. Let H, H_e, H^I and H_e^I be vectors representing the data in Figure 12B as follows.

- H, the first column of Figure 12B, is the incremental energy vector for repressor to O1.
- H_e, computed from the second column of Figure 12B is the absolute value of the error for H.
- H^I is the change in incremental energy for a repressor to O1 in the presence of IPTG (third column).
- H_e^I is computed from the forth column and is the absolute value of the error for H^I.

We calculate the approximate free energy for binding free repressor to the DNA sequence O1 using the formula:

$$A_{O1} \simeq \Delta G_b + \vec{\mathbf{1}} \cdot H,$$

where A_{O1} is the sum of the incremental binding energies in H and ΔG_b is the basal binding energy. To compute the error for ΔG_{O1} we sum the values of H_e. Thus, the absolute value of the error for ΔG_{O1} is

$$E_{O1} = \vec{\mathbf{1}} \cdot H_e.$$

Therefore the free energy for binding free repressor to the operator O1 is in the range:

$$\Delta G_{O1} \in [A_{O1} - E_{O1}, A_{O1} + E_{O1}].$$

When IPTG is bound, we approximate the free energy for binding impaired repressor to the DNA sequence O1 by adding the sum of H^I to ΔG_{O1} as follows:

$$A_{O1}^I \simeq \Delta G_{O1} + \vec{\mathbf{1}} \cdot H^I.$$

Recall, the more negative the free energy is, the stronger the protein binds the DNA. Then to calculate the error we sum the values of H_e^I. Thus the absolute value of the error for ΔG_{O1}^I is

$$E_{O1}^I = \vec{\mathbf{1}} \cdot H_e^I,$$

and the free energy for binding an impaired repressor binding domain to the O1 operator is in the range:

$$\Delta G_{O1}^I \in [A_{O1}^I - E_{O1}^I, A_{O1}^I + E_{O1}^I].$$

Next we determine the range of free energy values for binding the repressor to each of the sequences listed in Table 2 (in Chapter 3). We would like to reproduce the 1994 data from Oehler et al. [61]. Therefore, when Table 2 lists two different deletions for the same operator, we choose the sequence from 1994. Let V_i, $i \in$ {O2,O3,O3*,O1$^-$a,O1$^-$b,O2$^-$,O3$^-$,O3*$^-$,O3*(O1)} be a vector of length 21 of zeros and ones. We compare each base pair of the sequence i to the O1 sequence. For every base pair which matches the O1 sequence, the vector V_i will have a 1 in the corresponding position. We represent every base pair which differs from the O1 sequence with a zero in V_i. Define $\overline{V_i} := \vec{\mathbf{1}} - V_i$. The vector V_i will be referred to as the similarity vector for sequence i.

As an example, consider the wildtype sequence for O1 and O2 from Table 2:

WT O1 5'-`AATTGTGAGCGGATAACAATT`-3'

WT O2 5'-`AAATTGTAGCGAGTAACAACC`-3'

V_{O2} `(11010001111001111100)`.

The last row contains the similarity vector (V_{O2}) of ones and zeros.

To calculate the approximate free energy of a repressor binding operator i, we take the dot product between the similarity vector V_i and H:

$$A_i \simeq \Delta G_b + V_i \cdot H.$$

When a base pair does not match the O1 sequence, we assume the energy increases from the energy recorded for O1. However, it is unlikely the energy contribution is zero. Thus, to compute the error estimate for ΔG_i, we take the dot product between the similarity vector and the error, $V_i \cdot H_e$, which determines the error in the data measurements, and add to it the absolute value of the energy from H for every base pair mismatch:

$$E_i = V_i \cdot H_e + |\overline{V_i} \cdot H|.$$

This gives us

$$\Delta G_i \in [A_i - E_i, A_i + E_i].$$

Finally, we compute the free energy of an impaired repressor binding operator i, ΔG_i^I by adding the sum of the measured impairment, $V_i \cdot H^I$, to ΔG_i so that:

$$A_i^I \simeq \Delta G_i + V_i \cdot H^I.$$

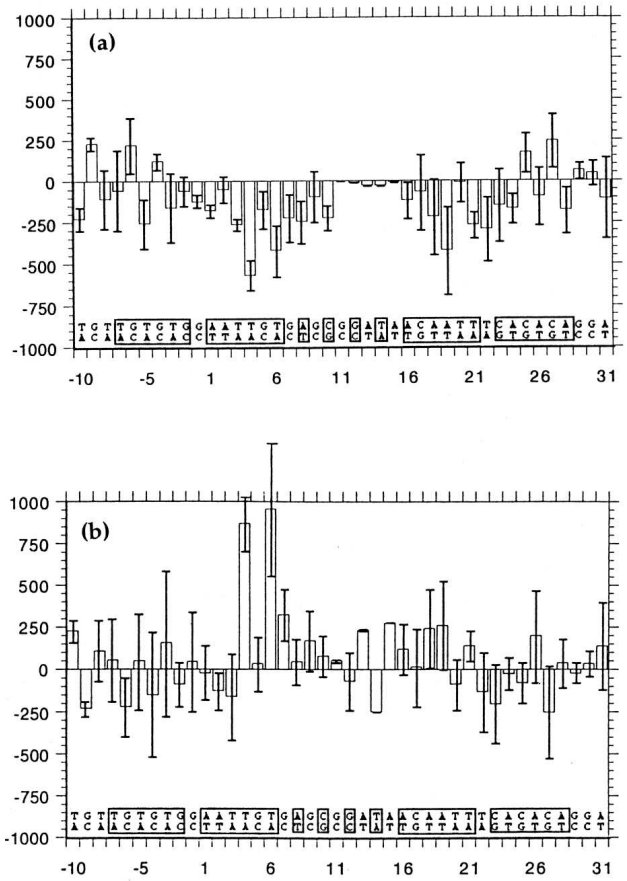
We compute the error as the sum of the error for base pairs which match O1, $V_i \cdot H_e^I$, and increase it by the sum of the absolute value of base pair mismatches. Thus, the absolute value of the error for an impaired repressor binding domain bind the operator i is

$$E_i^I = V_i \cdot H_e^I + |\overline{V_i} \cdot H^I|.$$

Therefore the energy for an impaired repressor binding domain to bind the operator i is in the range

$$\Delta G_i^I \in [A_i^I - E_i^I, A_i^I + E_i^I].$$

The relative estimates from the literature and the incremental binding energies from Horton et al. are the restrictions imposed when fitting the model to the data. See Table 10 for the predicted free energy of binding for the various DNA sequences.



A

O1 DNA	no IPTG O1± error cal/mol		1 mM IPTG O1± error cal/mol	
	H	H_e	H^I	H_e^I
A	−175	50	−25	175
A	−50	75	−125	125
T	−250	30	−175	275
T	−575	100	+875	175
G	−175	120	+50	160
T	−400	150	+950	400
G	−225	150	+325	150
A	−250	125	+50	150
G	−100	150	+175	190
C	−225	75	+750	125
G	0	0	+50	0
G	−15	0	−75	180
A	−25	0	+250	0
T	−25	0	−275	0
A	0	0	+290	0
A	−125	120	+125	160
C	−75	250	+20	250
A	−225	250	+250	250
A	−425	250	+275	275
T	0	150	−100	175
T	−275	80	+150	100

B

Figure 12: A. Incremental binding affinity (per base pair in cal/mol at 298 K) as published by Horton, Lewis and Lu [37] for (a) the *lac* repressor to operator 1 (O1) and (b) the change in incremental binding affinity of the *lac* repressor to O1 in the presence of 10^{-3}M IPTG. To determine the free energy for the repressor to operator, the sum of the incremental energies must be added to some basal energy. B. The incremental energies in cal/mol for the central 21 base pairs as determined from A.

RESULTS

While the role of mathematics in general and mathematical models in particular is well established in physics, chemistry, engineering and perhaps to some degree in ecology, this role is less clear in the emerging field of systems biology. In many influential papers, in addition to experimental results, a model is presented as well. More often then not, however, the model is phenomenological. The reaction terms in ordinary differential equations are arbitrary Hill functions; some models include only protein concentrations, some both protein and mRNA concentrations; some include noise and some are deterministic. The parameters in the Hill functions are chosen arbitrarily and if the resulting curves fit the experimental data, there is very little discussion of appropriateness of the model. In these cases the model provides an illustration, a metaphor of the underlying biology. Clearly, one of the causes of this state of affairs is the complexity and multi-scale character of modeled systems. Gene regulation is a result of molecular interactions with DNA, which in turn is governed by the biophysics of bond and loop formation. Clearly an appropriate modeling level is problem dependent. In neuroscience the problem of the multi-scale character of the reality has been to some extent abated by Hodgkin and Huxley, whose fundamental model of signal propagation is well accepted. This model predicts macroscopic behavior (signal propagation) based on information about channel dynamics. What is remarkable is that it is more than a model: it is a modeling platform that is generalizable to different tissues by simply changing the set of channels that enter the model and measuring for each channel a well defined set of characteristics.

In molecular biology we do not have a Hodgkin-Huxley model. The closest analog is the Shea-Ackers framework for modeling the promoter dynamics. It can be used

for any gene; there is a well defined set of inputs (in this case binding energies) and the result is a gene expression level. As in the case of the Hodgkin-Huxley equations, the model spans two scales: the inputs are biophysical parameters of molecular interactions and the result is an expression - a macroscopic behavior of the cell.

The most important test for every modeling scheme is a comparison to data. In systems biology the expression level data is much more abundant than the biophysical data. In fact, there are only a handful of systems where the latter data is rich enough to even attempt to develop a full scale model based on biophysical parameters. As we have documented in previous pages, the *lac* operon is one of them. By developing the most complete, biophysical, data driven model of the *lac* operon, we attempt to answer the following list of questions. Is such model capable of matching the data collected on a different level, i.e. expression data? Does the model point out deficiencies in our understanding of the *lac* operon control by pointing to a specific parameter that needs to have a different value from those predicted by current biology in order to match the data? What parts of the regulation are most important for proper functioning of the *lac* operon? Which seem to be redundant? To answer the last two questions we initially include all potential control mechanisms in the model. Then we remove them one by one and test the fit to the model. In all the instances this leads to a worse fit to the data. However, since many parameters have only ranges of possible values, it is possible that the fit can be improved by subsequent optimization. We deem a control mechanism not important for the *lac* operon if this second optimization results in a good fit to the data. We realize that additional data would very likely affect this decision, since it would put more constraints on the model, which in turn would require the inclusion of more mechanisms. Our ultimate goal then is to develop a so called *minimal* model that would faithfully represent the available data, yet would include only the minimal number of control mechanisms.

Initially, the model contains all of the primary components: CAP activation, blocked transcription when the *lac* repressor binds O1, increased repression via DNA looping and deactivation of the repressor by inducer. We have also included several additional components: the P2 promoter, the O3* operator, CAP assisted looping, C1-O3 steric interference, reduced binding of the inducer impaired *lac* repressor and low specific binding to deleted sites. This model represents 684 possible protein-DNA configuration for the *lac* operon, see Appendix D. All concentrations of proteins and dissociation constants have been measured experimentally or are approximated from experimental data.

We find that using this model, we are able to accurately reproduce the Oehler et al. [61] data for all mutants except (3,0,3). Therefore this model is capable of predicting the expression data based on biophysics.

On the other hand, the model also points out a possible gap in biology. In order to achieve this fit to the data, the free energy of looping has to be greatly reduced ($\sim$10 to 13 kcal/mol) from the predicted values based on the DNA structure [81]. This is indicative of a missing component, possibly a non-specific binding protein such as HU [13, 21]. See Table 10 in the following pages and Table 3 (from Chapter 3) for a comparison between the DNA looping values used for the complete model, as shown in Figure 13, and the predicted values based on the DNA structure from Swigon and Olson [81].

Next we begin testing the relative contributions of the secondary components. We discover that the secondary components fall into three classes: those which can be removed and the resulting system fits the data with no additional change in other parameters (impaired repressor binding domains and the P2 promoter), those which fit the data after changing other parameters within their experimentally predicted

bounds (the O3* operator and CAP assisted DNA looping) and those whose absence precludes a good fit even if other parameters are changed (C1-O3 steric interference).

We will also point out that components in the first and second categories may have very different behavior for the repression curves than the full model. If these changes occur outside of the concentrations where the repression data is available, we are unable to distinguish between the full model and the reduced model.

Table 9: Repression measurements reported by Oehler et al. [61] for various mutants, as described by the 3-tuple (i, j, k). Each vector represents a particular mutant: (1,2,3) is the wildtype; (1,0,0) has O2 and O3 deleted; (3,0,1) has O3 DNA in the O1 position, O2 deleted, and O1 DNA in the O3 position and so on. In the case of an inequality, the experimental procedure could only determine a lower bound; the actual repression value may be much higher than the reported value. For the column headers: WT describes the repression measured in the presence of wildtype repressor concentration, $\sim$5$\times$WT implies the cellular concentration of the repressor was approximately 5 times larger than the wildtype, and similarly for 90 times wildtype. The first column of the data groups similar mutants: A. is the wildtype, B. contains mutants with a deletion in the O2 position, C. contains mutants with a deletion in the O3 position, D. contains mutants with O2 and O3 deletions and E. has deletions in the O1 and O2 positions.

	1994	$\sim$ 5$\times$WT	$\sim$ 90$\times$WT
A.	(1,2,3)	8100	$\geq$19,000
B.	(1,0,1)	$\geq$ 12,000	$\geq$ 46,000
	(1,0,3)	6200	$\geq$21,000
	(3,0,1)	890	3900
	(3,0,3)	38	960
C.	(1,2,0)	2300	6800
	(2,2,0)	360	560
	(3,2,0)	6.8	15
D.	(1,0,0)	200	4700
	(2,0,0)	21	320
	(3,0,0)	1.3	16
E.	(0,0,1)	18	28
	(0,0,3)	1.7	25

Finally, we observe that if we do not allow repressors to bind experimentally deleted sites, the binding energy for the repressor to the operators must be increased

2 kcal/mol. This increase forces the free energy, ΔG_{O1}, 1 kcal/mol stronger than the biologically reported range of values $[37, 60]$ to -15.01 kcal/mol. If this value is biologically acceptable, we find the minimal model sufficiently represents the Oehler et al. [61] data. This minimal model retains the steric interaction between CAP at C1 and the repressor at O3, but it removes all other secondary components and it does not allow the repressor to bind deleted sites. These simplifications reduce the model to 92 terms rather than the 684 terms of the complete model.

One of the main advantages of having vast amounts of data available for the *lac* operon is that we have estimates for all of the parameter values related to the primary components and some of the secondary components. However, the disadvantage is that these values vary from one E. coli strain to another, from experiment to experiment and from laboratory to laboratory. Additionally, if the experimentalists make an incorrect assumption about the behavior they are observing when the data is collected, the reported parameter values can be affected. As an illustration, recall the calculation of K_{coop} in Chapter 4. When fitting the data, the experimentalists incorrectly assumed that all of the DNA would be bound by CAP before introducing RNAP. This assumption resulted in an incorrect reported value for the dissociation constant between RNAP and CAP bound DNA. We had to fix this error. Because of the variation in the data, we expect some parameter adjustments are necessary. Therefore, we begin with the complete model and adjust the parameters, within the range of estimates, to best fit the model to the data. Similarly, when we remove a component from the model, we again refit the parameters to best fit the data.

Once a component is removed from the model, the set of possible states of the promoter is recalculated. The change in states can alter the repression curves to the point that the model is no longer able to accurately reproduce the data. We have two methods we use to refit the model to the data. First we determine which states

dominate the repression ratio and methodically adjust the cooperative free energies to increase or decrease the curves as needed. Note that as long as the numerical experiment allows repressors to bind delete sites, the repression calculation for every mutant contains states with DNA looping. Therefore it is often the case that the cooperative free energies affect all of the repression curves.

Second, when additional parameter adjustments are necessary, ΔG_{b}, the basal energy of the repressor to the DNA, is modified to fit the mutants (1,0,0), (2,0,0) and (3,0,0). A change in the ΔG_{b} parameter uniformly changes the free energies for binding a repressor to an operator but keeps the relative binding energies unchanged. For example, a repressor should bind O1 about 10 times stronger than O2 and 300 times stronger than O3. Changing ΔG_{b} will not alter these relative binding energies. Therefore, if the ΔG_{b} parameter is increased all of the curves shift lower, and if ΔG_{b} is decreased all of the curves shift higher.

In one numerical experiment we found it necessary to adjust the free energy of binding for $\Delta G_{O3*(O1)}$, $\Delta G^{I}_{O3*(O1)}$, ΔG_{O3*} and ΔG^{I}_{O3*}. Recall, the O3* operator overlaps the O3 operator in all but five base pairs, therefore any change in the O3 operator changes the O3* operator. Thus we use O3*(O1) to denote the case when the O1 operator DNA is substituted into the O3 position. O3* denotes the O3 operator is in its wildtype position. The remaining parameters are held constant at biologically relevant values.

For each experiment, we present our results in a set of figures and a table. Each figure plots the repressor concentration in molar along the x-axis and the repression value along the y-axis. The left figure contains the repression curves for all mutants with the O3 operator deleted (groups C and D from Table 9). With the inclusion of the wildtype operator, the right figure contains all of the mutants with the O2 operator deleted (groups A, B and E from Table 9). The boxes mark data values

from Table 9 as reported by Oehler et al. [61] at 200 subunits (4.2×10^{-8}M) and 3600 subunits (7.473×10^{-7}M). When the experimentalists [61] collected this data, there were few instances in which the experimental method could only record a lower bound. It is likely that the actual repression value is higher than the reported value. The associated mutants are: (1,2,3), (1,0,1) and (1,0,3) at a repressor concentration of 7.47×10^{-7} M (3600 subunits) and (1,0,1) at a repressor concentration of 4.3×10^{-8} M (200 subunits). The lower bounds are denoted by an inequality in Table 9 and by a box with an 'x' in the figures.

Each set of figures has a corresponding table containing the relevant parameter values. For the complete model the table is presented in three parts. Part A contains the repressor-operator binding energy for the operator (k) in the absence (ΔG_k) and presence (ΔG_k^I) of IPTG as determined from the Horton et al. [37] data (see Figure 12 in Chapter 4). Part B contains the free energy of looping, the basal energy and the steric interaction between C1 and O3. Finally, part C contains the remaining parameter values which are always held constant (i.e. RNAP concentration, CAP concentration, RNAP binding affinities, and so on). For the numerical experiments, part C is withheld. When parameters are changed from the set presented in the complete model they will be denoted by colored text. If the parameter is part of the experiment, the value will be denoted by red. If the parameter was altered to refit the data it will be denoted by blue followed by an arrow. If the arrow points up, we had to increase the parameter value in comparison to the value we use in the complete model. If the arrow points down, we had to decrease the value in comparison to the value we use in the complete model.

To quantify how well the model fits to the data, we calculate ε for the complete model and each numerical experiment.

$$\varepsilon = \sum_{m \in \mathcal{M}} |\log(D_1(m)) - \log(P_1(m))| + |\log(D_2(m)) - \log(P_2(m))|, \qquad (53)$$

where $\mathcal{M}$ is the set of all mutants; $D_1(m)$ and $D_2(m)$ are the repression values from Oehler et al. [61] for mutant m at 200 and 3600 subunits respectively and $P_1(m)$ and $P_2(m)$ are the model predicted repression values for mutant m at 200 and 3600 subunits respectively. The data points which were only reported as a lower bound are not included, as these inequalities are satisfied in all our numerical experiments. We use this error calculation since the repression values range from 1.3 to 8100 and because we want to calculate fold differences. For example, when a data point is 20 and the model prediction is 10, this difference should contribute to the error in the same way as if the data point is 200 and the model prediction is 100. We will compute ε for each model.

We now discuss the numerical experiments. We will number the experiments NE1 through NE8, where NE stands for numerical experiment.

NE1 The P2 promoter: the P2 promoter is not included in the model.

NE2 The O3* operator: the O3* operator is not included in the model.

NE3 CAP assisted DNA loop formation: CAP does not assist in bending the DNA.

NE4 C1-O3 steric interaction: CAP bound to C1 and a repressor bound to O3 do not interact.

NE5 Impaired binding domains: we do not allow inducer impaired binding domains to bind to operators.

NE6 Repressor bound to O2 blocks transcription: if O2 is bound, transcription does not occur.

NE7 Low-specific binding to deleted sites: repressors are not allowed to bind deleted sites.

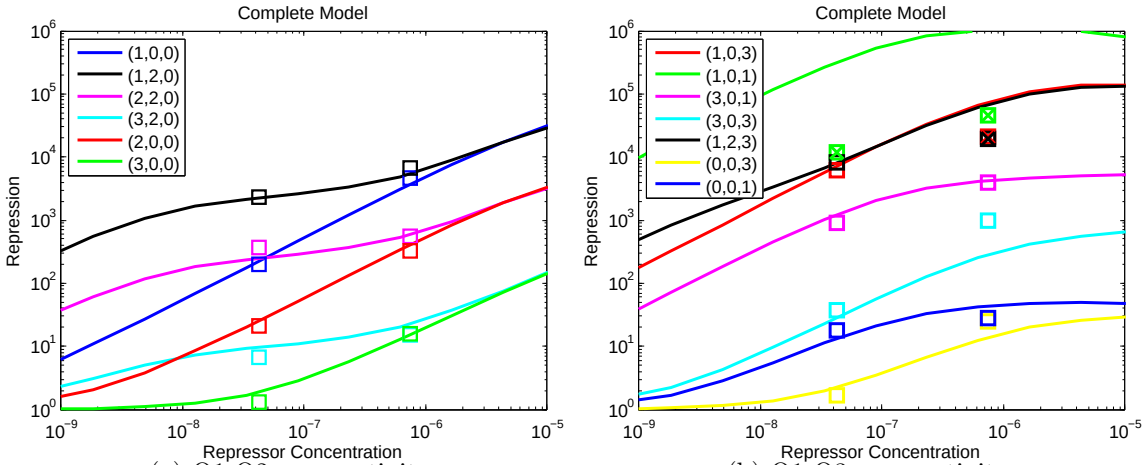
NE8 Minimal model: we remove the P2 promoter, the O3* operator, CAP assisted DNA loop formation, we do not allow inducer impaired binding domains to bind the DNA, and we remove low-specific binding to deleted sites.

We have noticed that the fit to the repression data containing non-deleted operators in the O1 and O2 positions is much better in comparison to the curves with non-deleted operators in the O1 and O3 positions. The O1-O2 cooperativity figures match the data easily until O2 is allowed to block transcription (NE6, Figure 19) or the low specific binding to deleted sites is removed (NE7, Figure 20). The parameters which recover the O1-O2 fit are primarily the ΔG_{b} value underlying the Horton incremental binding affinities and ΔG_{O12}, the energy required to form a loop of DNA between the O1 and the O2 operator positions. In contrast, the curves containing an O1-O3 loop are easily perturbed and recovery involves all other looping parameters (see Table 10B).

The Complete Model

The complete model implements all of the primary components of the *lac* operon. In addition, after reviewing the literature we determined several secondary components which may also be important to accurately represent the *lac* operon behavior. The secondary components included in the complete model are the additional promoter P2, the shifted binding site O3*, the steric interaction between CAP bound

Complete Model, the associated parameters are in Table 10



(a) O1-O2 cooperativity (b) O1-O3 cooperativity

Figure 13: Repression level curves as a function of the repressor concentration for the complete model. The parameter values we used to attain these curves are listed in Table 10.

C1 and a repressor bound O3, and the assistance in loop formation by CAP. We also allow low specific binding to deleted sites.

We implement a model that includes the components described in the previous chapters. Table 10A reports the free energy for binding a repressor to an operator. These free energy values are accurate relative to each other meaning a repressor binds O1 about 10 times stronger than O2, and a repressor binds O1 about 300 times stronger than O3. Moreover, every free energy value is within the error bound determined from the Horton data [37] except for O1$^-$a. This deleted operator is only used in the mutant (0,0,3). The reported repression values for (0,0,3) are not similar to the values for (3,0,3), see Figure 13. However, if we decrease the free energy of the deleted O1 site, O1$^-$a, so that it is within the error bounds, the (0,0,3) repression curve parallels the (3,0,3) repression curve. This is because $\Delta G_{O3} = -9.6$ kcal/mol while the free energy for the deleted site is predicted to be $\Delta G_{O1^-a} \leq -9.19$ kcal/mol for a basal energy of -8.25 kcal/mol. The predicted binding energy value is so close

Table 10: Parameter values for Figure 13, the complete model.

A. Free energy of binding without				and with IPTG			
ΔG_{O1}	$\simeq$	-13.03	kcal/mol	ΔG^I_{O1}	$\simeq$	-6.905	kcal/mol
ΔG_{O2}	$\simeq$	-11.665	kcal/mol	ΔG^I_{O2}	$\simeq$	-7.12	kcal/mol
ΔG_{O3}	$\simeq$	-9.6	kcal/mol	ΔG^I_{O3}	$\simeq$	-4.18	kcal/mol
ΔG_{O3^*}	$\simeq$	-8.485	kcal/mol	$\Delta G^I_{O3^*}$	$\simeq$	-4.075	kcal/mol
ΔG_{O1^-a}	$\simeq$	-7.68	kcal/mol	$\Delta G^I_{O1^-a}$	$\simeq$	-2.175	kcal/mol
ΔG_{O1^-b}	$\simeq$	-6.72	kcal/mol	$\Delta G^I_{O1^-b}$	$\simeq$	-5.895	kcal/mol
ΔG_{O2^-}	$\simeq$	-7.945	kcal/mol	$\Delta G^I_{O2^-}$	$\simeq$	-4.2	kcal/mol
ΔG_{O3^-}	$\simeq$	-7.205	kcal/mol	$\Delta G^I_{O3^-}$	$\simeq$	-3.245	kcal/mol
ΔG_{O3^*-}	$\simeq$	-6.21	kcal/mol	$\Delta G^I_{O3^*-}$	$\simeq$	-4.225	kcal/mol
$\Delta G_{O3^*(O1)}$	$\simeq$	-8.45	kcal/mol	$\Delta G^I_{O3^*(O1)}$	$\simeq$	-3.135	kcal/mol
B. Basal and cooperative free energies							
ΔG_{b}	$\simeq$	-8.25	kcal/mol	ΔG_{C1loop}	$\simeq$	-2.47	kcal/mol
ΔG_{C1-O3}	$\simeq$	3.5	kcal/mol	ΔG_{O12}	$\simeq$	6.89	kcal/mol
ΔG_{O23}	$\simeq$	12.01	kcal/mol	ΔG_{O23_c}	$\simeq$	9.54	kcal/mol
ΔG_{O23*}	$\simeq$	12.01	kcal/mol	ΔG_{O23*_c}	$\simeq$	9.54	kcal/mol
ΔG_{O13}	$\simeq$	5.41	kcal/mol	ΔG_{O13_c}	$\simeq$	2.94	kcal/mol
ΔG_{O13*}	$\simeq$	6.77	kcal/mol	$\Delta G_{O13^*_c}$	$\simeq$	4.3	kcal/mol

C. Fixed parameter values							
RT	$\simeq$	0.617	kcal/mol	[RNAP]	$\simeq$	1.25×10^{-6}	M
K_{P1}	$\simeq$	5.6×10^{-7}	M	K_{P2}	$\simeq$	$K_{P1} * 2.2/5$	M
[C]	$\simeq$	5×10^{-7}	M	K_{CAP}	$\simeq$	1.2×10^{-8}	M
K_{coop}	$\simeq$	$1.5/5.6 * \frac{[\mathrm{C}]}{[\mathrm{C}]+K_{CAP}}$		[cAMP]	$\simeq$	3.3×10^{-6}	M
[CRP]	$\simeq$	1.5×10^{-6}	M	K_{cAMP}	$\simeq$	3.03×10^{-6}	M
[CAP]	$\simeq$	$[\mathrm{CRP}] * \frac{[\mathrm{cAMP}]}{[\mathrm{cAMP}]+K_{cAMP}}$	M	K_I	$\simeq$	6.7×10^{-6}	M
$[R_t]$	$\simeq$	8.3×10^{-9}	M	cell size	$\simeq$	2.0×10^{-15}	L

to ΔG_{O3}, that the same behavior is observed. Therefore, we increase ΔG_{O1^-a} 1.51 kcal/mol. The error bounds and free energy values determined from Horton [37] are in Table 11. These values are determined assuming a basal energy of $\Delta G_b = -8.25$.

If we use the structurally predicted DNA looping energies [81] we are unable to reproduce the data. However, if we reduce the energy required for DNA looping by about 11 kcal/mol, we accurately reproduce the data observations for all mutants with the exception of (3,0,3). The (3,0,3) mutant was reported [61] to have higher

Table 11: Free energy values for each operator i in kcal/mol as predicted from the Horton et al. [37] data for the basal energy $\Delta G_b = -8.25$ kcal/mol. Notice the repressor-operator free energies ΔG_i (also in Table 10) are all within $A_i \pm E_i$ in the absence of IPTG and ΔG_i^I are within $A_i^I \pm E_i^I$ in the presence of IPTG, with the exception of O1$^-$a.

$\Delta G_b =$	no IPTG			1 mM IPTG		
-8.25	A_i	E_i	ΔG_i	A_i^I	E_i^I	ΔG_i^I
O1$^-$a	-11.365	2.175	-7.680	-3.915	3.385	-2.175
O1$^-$b	-9.275	3.320	-6.720	-5.645	4.785	-5.895
O2$^-$	-9.475	3.360	-7.945	-6.435	4.385	-4.200
O3$^-$	-10.525	3.320	-7.205	-5.210	4.830	-3.245
O3*$^-$	-9.415	3.445	-6.210	-5.540	5.655	-4.225
O1	-11.865	2.125	-13.030	-9.220	3.315	-6.905
O2	-11.300	2.430	-11.665	-7.905	3.335	-7.120
O3	-10.875	3.115	-9.600	-6.405	4.190	-4.180
O3*	-9.640	3.520	-8.485	-7.540	5.530	-4.075
O3*(O1)	-8.915	3.445	-8.450	-8.125	5.665	-3.135

repression at the high data point (7.47×10^{-7}M or 3600 subunits) than we are able to achieve. To increase the (3,0,3) curve further requires decreasing the energy required to form the O1-O3* loop. Unfortunately, this parameter is also included as the leading term of the repression for several other mutants. Thus, increasing the curve for the (3,0,3) mutant will force other curves above their respective data points. Table 10B contains the free energy of looping the DNA along with the basal component, the energy of CAP assisted looping and the energy of the steric interaction between a CAP bound to C1 and a repressor bound to O3 for the complete model.

Table 10C contains the experimentally determined fixed parameters.

The complete model accurately reproduces the data [61] as long as the free energy of looping is greatly reduced from the values based on the DNA structure [81] (see Table 3 in Chapter 3). This indicates that a component is still missing from the model, possibly a non-specific binding protein able to assist in looping the DNA, such as HU [13, 21].

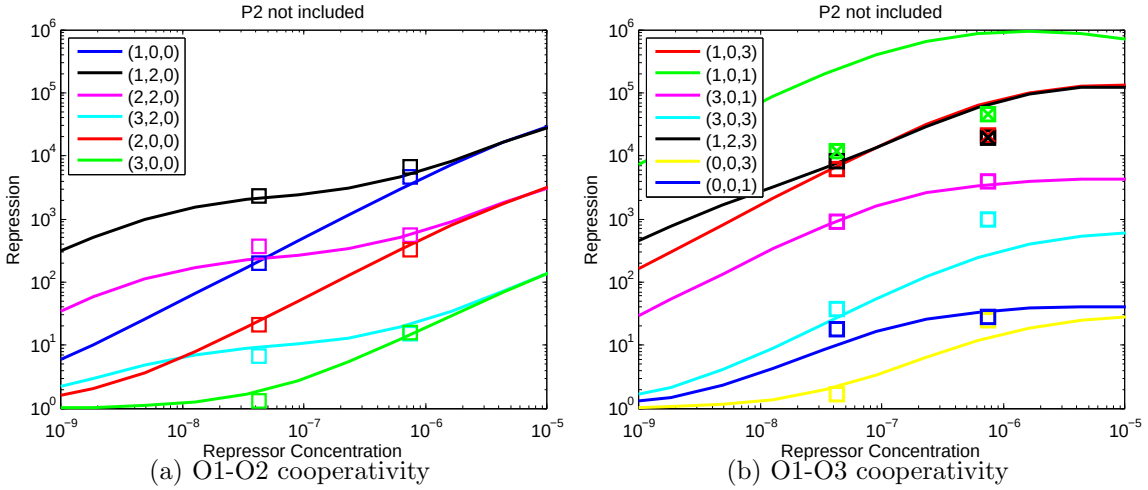
To determine how well the complete model reproduces the data, we calculate ε from Equation (53). We find $\varepsilon = 2.7247$.

Secondary Component Experiments

It is well established that the primary components are key to proper functioning of the *lac* operon. In this section we test the relative contribution of the secondary components to the functioning of the *lac* operon. In each numerical experiment we remove a secondary component and then attempt to refit the data by changing other parameters. The resulting model and parameter values are displayed.

NE1: The P2 Promoter

Removal of the P2 promoter



(a) O1-O2 cooperativity (b) O1-O3 cooperativity

Figure 14: Repression level curves as a function of the repressor concentration as determined when the P2 promoter site is not included in the model. The associated parameters are in Table 12.

Malan and McClure [49] observed higher activity from the P2 promoter than the primary promoter, P1, under low cAMP concentrations in vitro. We included P2 in our complete model and here we investigate how the repression values change once P2 is removed from the model.

Table 12: Parameter values for Figure 14 where P2 is removed from the complete model.

A. Free energy of binding without				and with IPTG			
ΔG_{O1}	$\simeq$	-13.03	kcal/mol	ΔG^I_{O1}	$\simeq$	-6.905	kcal/mol
ΔG_{O2}	$\simeq$	-11.665	kcal/mol	ΔG^I_{O2}	$\simeq$	-7.12	kcal/mol
ΔG_{O3}	$\simeq$	-9.6	kcal/mol	ΔG^I_{O3}	$\simeq$	-4.18	kcal/mol
ΔG_{O3^*}	$\simeq$	-8.485	kcal/mol	$\Delta G^I_{O3^*}$	$\simeq$	-4.075	kcal/mol
ΔG_{O1^-a}	$\simeq$	-7.68	kcal/mol	$\Delta G^I_{O1^-a}$	$\simeq$	-2.175	kcal/mol
ΔG_{O1^-b}	$\simeq$	-6.72	kcal/mol	$\Delta G^I_{O1^-b}$	$\simeq$	-5.895	kcal/mol
ΔG_{O2^-}	$\simeq$	-7.945	kcal/mol	$\Delta G^I_{O2^-}$	$\simeq$	-4.2	kcal/mol
ΔG_{O3^-}	$\simeq$	-7.205	kcal/mol	$\Delta G^I_{O3^-}$	$\simeq$	-3.245	kcal/mol
ΔG_{O3^*-}	$\simeq$	-6.21	kcal/mol	$\Delta G^I_{O3^*-}$	$\simeq$	-4.225	kcal/mol
$\Delta G_{O3^*(O1)}$	$\simeq$	-8.45	kcal/mol	$\Delta G^I_{O3^*(O1)}$	$\simeq$	-3.135	kcal/mol
B. Basal and cooperative free energies							
ΔG_{b}	$\simeq$	-8.25	kcal/mol	ΔG_{C1loop}	$\simeq$	-2.47	kcal/mol
ΔG_{C1-O3}	$\simeq$	3.5	kcal/mol	ΔG_{O12}	$\simeq$	6.89	kcal/mol
ΔG_{O23}	$\simeq$	12.01	kcal/mol	ΔG_{O23_c}	$\simeq$	9.54	kcal/mol
ΔG_{O23*}	$\simeq$	12.01	kcal/mol	ΔG_{O23*_c}	$\simeq$	9.54	kcal/mol
ΔG_{O13}	$\simeq$	5.41	kcal/mol	ΔG_{O13_c}	$\simeq$	2.94	kcal/mol
ΔG_{O13*}	$\simeq$	6.77	kcal/mol	$\Delta G_{O13^*_c}$	$\simeq$	4.3	kcal/mol

The resulting repression curves, shown in Figure 14, display a minute decrease in repression using the same energy values as the complete model with $\varepsilon = 2.6730$. Table 12 contains the parameter values for Figure 14. For the operon in steady state the concentration of internal cAMP is similar for cells digesting lactose or glucose, thus the contribution of P2 should not change under varying glucose and lactose.

We conclude the P2 promoter need not be included to accurately represent the rate of transcription (and therefore the repression) from the *lac* operon. However, P2 may be necessary to model the dynamic behavior during diauxic growth as E.coli shift from digesting glucose to lactose.

NE2: The O3* Operator

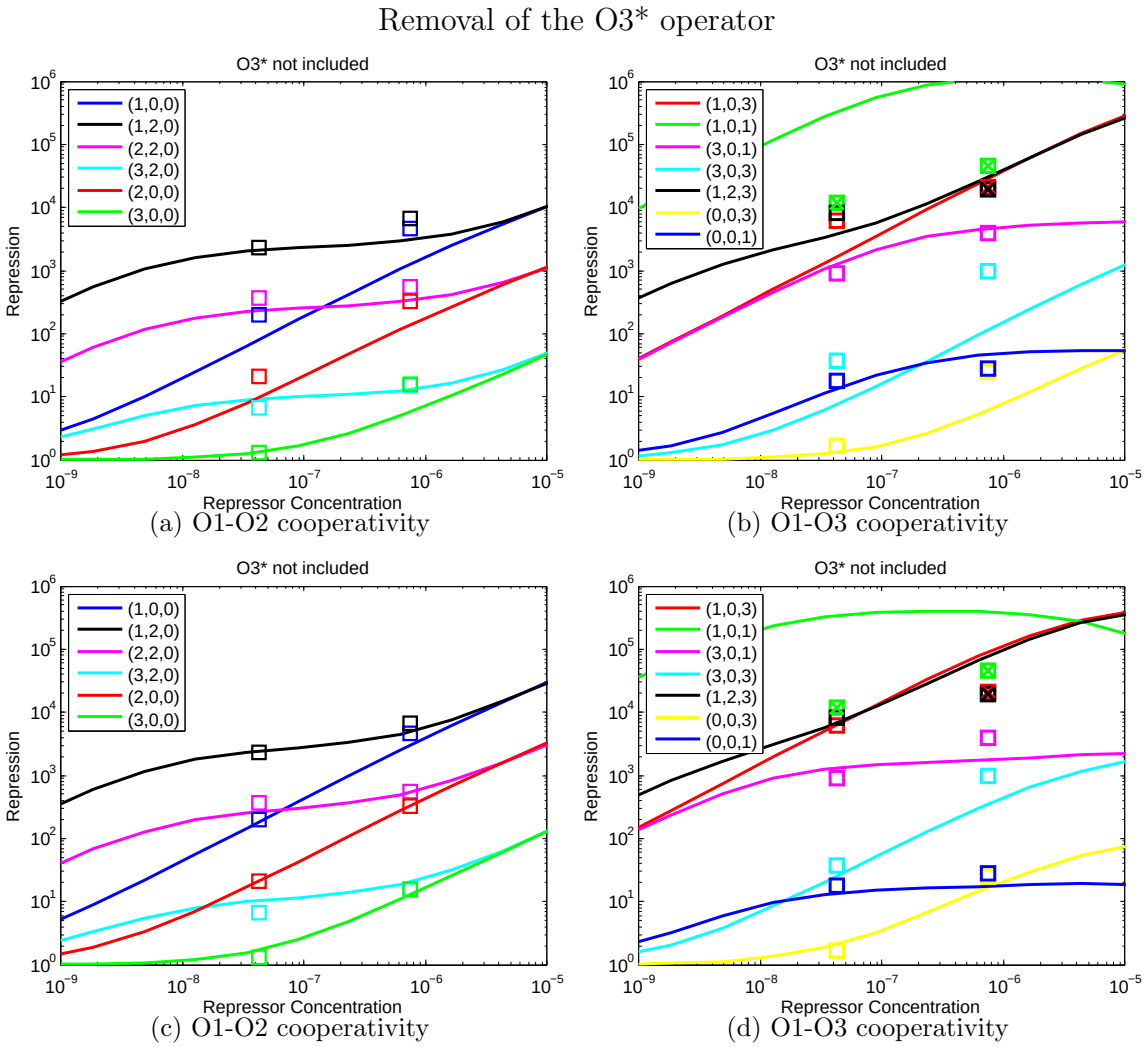


Figure 15: Repression level curves as a function of the repressor concentration when the O3* operator is removed from the model. (a) and (b) are the repression curves with the parameters from the complete model (Table 10). (c) and (d) are the repression curves once the parameters have been adjusted to refit the curves to the data. The new parameters are listed in Table 13.

Hudson and Fried [25] observed that the O3 operator shifts five to six base pairs in the presence of high concentrations repressor and CAP. Swigon and Olson [81] have called this shifted operator site O3*. Swigon and Olson [81] predict the looping energy based on the DNA and protein structure for a repressor bound to O1 and O3,

Table 13: Parameter values for Figure 15 where O3* is removed from the complete model.

A. Free energy of binding without				and with IPTG			
ΔG_{O1}	$\simeq$	-13.03	kcal/mol	ΔG^I_{O1}	$\simeq$	-6.905	kcal/mol
ΔG_{O2}	$\simeq$	-11.665	kcal/mol	ΔG^I_{O2}	$\simeq$	-7.12	kcal/mol
ΔG_{O3}	$\simeq$	-9.6	kcal/mol	ΔG^I_{O3}	$\simeq$	-4.18	kcal/mol
ΔG_{O3^*}			N/A	$\Delta G^I_{O3^*}$			N/A
ΔG_{O1-a}	$\simeq$	-7.68	kcal/mol	ΔG^I_{O1-a}	$\simeq$	-2.175	kcal/mol
ΔG_{O1-b}	$\simeq$	-6.72	kcal/mol	ΔG^I_{O1-b}	$\simeq$	-5.895	kcal/mol
ΔG_{O2-}	$\simeq$	-7.945	kcal/mol	ΔG^I_{O2-}	$\simeq$	-4.2	kcal/mol
ΔG_{O3-}	$\simeq$	-7.205	kcal/mol	ΔG^I_{O3-}	$\simeq$	-3.245	kcal/mol
ΔG_{O3^*-}			N/A	$\Delta G^I_{O3^*-}$			N/A
$\Delta G_{O3^*(O1)}$			N/A	$\Delta G^I_{O3^*(O1)}$			N/A
B. Basal and cooperative free energies							
ΔG_{b}	$\simeq$	-8.25	kcal/mol	ΔG_{C1loop}	$\simeq$	-2.47	kcal/mol
ΔG_{C1-O3}	$\simeq$	$2\downarrow$	kcal/mol	ΔG_{O12}	$\simeq$	$6.83\downarrow$	kcal/mol
ΔG_{O23}	$\simeq$	$10.78\downarrow$	kcal/mol	ΔG_{O23_c}	$\simeq$	$8.31\downarrow$	kcal/mol
ΔG_{O23*}			N/A	ΔG_{O23*_c}			N/A
ΔG_{O13}	$\simeq$	$5.84\uparrow$	kcal/mol	ΔG_{O13_c}	$\simeq$	$3.37\uparrow$	kcal/mol
ΔG_{O13*}			N/A	$\Delta G_{O13^*_c}$			N/A

a repressor bound to O1 and O3 with CAP bound to C1, and a repressor bound to O1 and O3* with CAP bound to C1, see Table 3. They determine that this shifted site, O3*, allows CAP to bind C1 and a repressor to bind O1 and O3* at a lower energy than if the repressor binds O1 and O3. Prior models of the *lac* operon have not included the O3* operator.

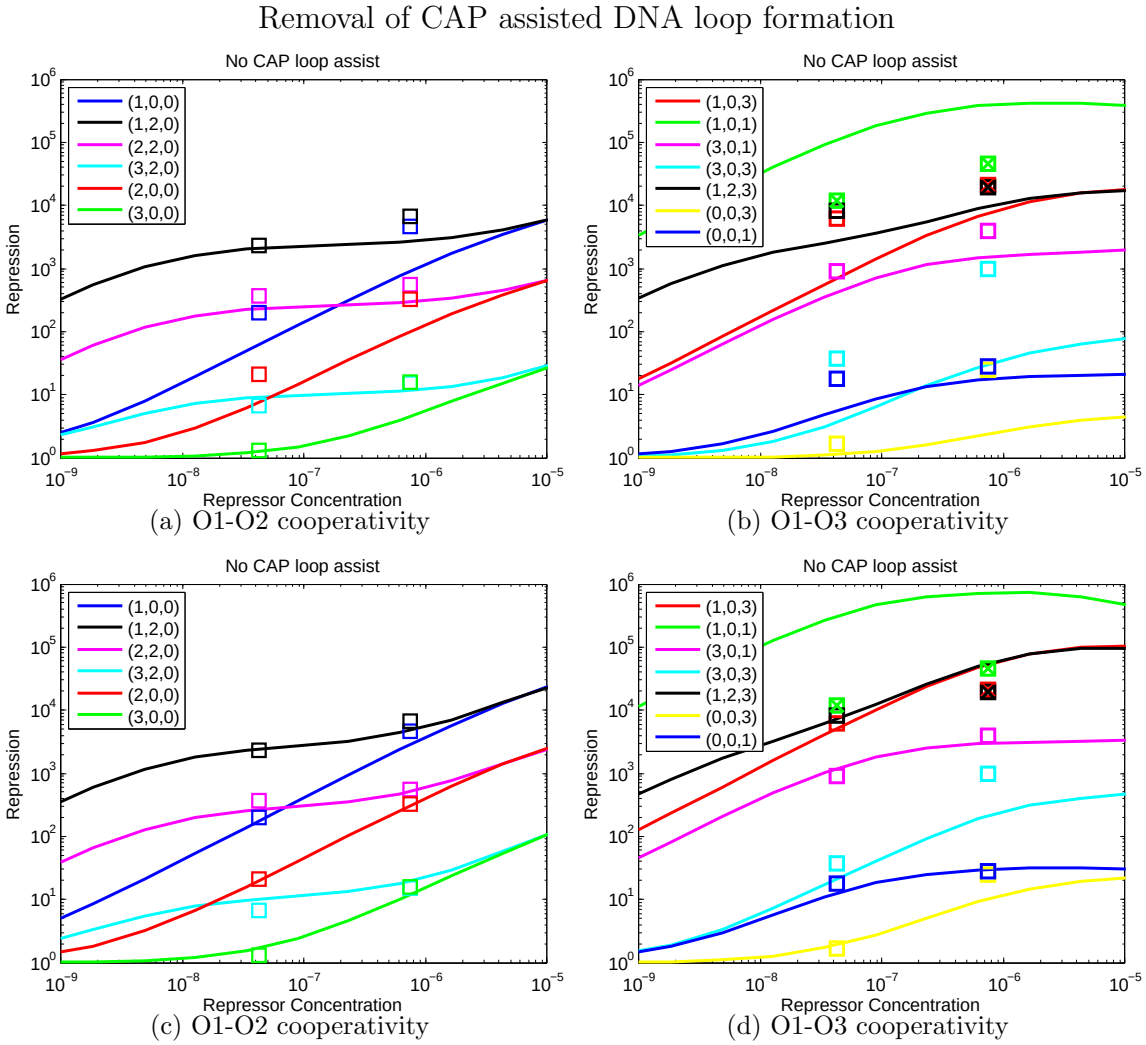
We want to investigate the effect of the O3* operator on the repression values. We remove the O3* operator from the model and plot the repressor concentration along the x-axis against the repression curves along the y-axis in Figure 15. In Figures 15(a)-(b), O3* has been removed, but the parameters have not been adjusted from the complete model. The removal of O3* results in $\varepsilon = 6.9815$.

In Figures 15(c)-(d), the repressor-operator free energies are conserved from the complete model but the cooperative parameters (ΔG_{C1-O3}, ΔG_{O12}, ΔG_{O13}, ΔG_{O23},

ΔG_{O13c}, and ΔG_{O23c}) must be adjusted to best fit the model to the data. Figures 15(c)-(d) result from a reduction in the steric interaction between CAP at C1 and a repressor at O3, an increase in the energy of forming a loop between O1 and O3 and a decrease in all other looping energies. These changes are listed in Table 13. The resulting repression curves for the (3,0,1) and (3,0,3) mutants are little low when the repressor concentration is 7.47×10^{-7} (3600 subunits), but otherwise the data is well fit with $\varepsilon = 2.8828$. Notice the change in the (1,0,1), (3,0,1) and (0,0,1) curves. In comparison to the complete model, all three have a higher repression at low repressor concentrations and a lower repression at high repressor concentrations. The wildtype concentration of repressor is $\sim 10 \times 10^{-9}$. If new experimental repression values were determined at the wildtype (WT) concentration as well 5 $\times$ WT and 90 $\times$ WT, we may be able to determine the contribution of O3* to the repression.

Comparing Figure 15(a) to Figure 15(c) and Figure 15(b) to Figure 15(d), we observe O3* does affect the model, but we are able to compensate for its removal by changing the looping parameters. The current data is not sufficient to support the inclusion of the O3* operator in the model of the *lac* operon.

NE3: CAP assisted DNA loop formation



(a) O1-O2 cooperativity

(b) O1-O3 cooperativity

(c) O1-O2 cooperativity

(d) O1-O3 cooperativity

Figure 16: Repression level curves as a function of the repressor concentration. Here CAP no longer assists in repressor-DNA loop formation. (a) and (b) use the parameters from the complete model (Table 10). In (c) and (d) the parameters have been readjusted to fit the data. The new parameter values are in Table 14.

The protein CAP is known to kink (or bend) the DNA [45, 46]. When Swigon and Olson [81] include CAP into their energy predictions of repressor mediated DNA loop formation they find that, in certain configurations, CAP causes a decrease in the energy required to form the loop. Since the CAP binding site, C1, is between

Table 14: Parameter values for Figure 16 where CAP is not allowed to assist in loop formation.

A. Free energy of binding without				and with IPTG			
ΔG_{O1}	$\simeq$	-13.03	kcal/mol	ΔG^I_{O1}	$\simeq$	-6.905	kcal/mol
ΔG_{O2}	$\simeq$	-11.665	kcal/mol	ΔG^I_{O2}	$\simeq$	-7.12	kcal/mol
ΔG_{O3}	$\simeq$	-9.6	kcal/mol	ΔG^I_{O3}	$\simeq$	-4.18	kcal/mol
ΔG_{O3^*}	$\simeq$	-8.485	kcal/mol	$\Delta G^I_{O3^*}$	$\simeq$	-4.075	kcal/mol
ΔG_{O1^-a}	$\simeq$	-7.68	kcal/mol	$\Delta G^I_{O1^-a}$	$\simeq$	-2.175	kcal/mol
ΔG_{O1^-b}	$\simeq$	-6.72	kcal/mol	$\Delta G^I_{O1^-b}$	$\simeq$	-5.895	kcal/mol
ΔG_{O2^-}	$\simeq$	-7.945	kcal/mol	$\Delta G^I_{O2^-}$	$\simeq$	-4.2	kcal/mol
ΔG_{O3^-}	$\simeq$	-7.205	kcal/mol	$\Delta G^I_{O3^-}$	$\simeq$	-3.245	kcal/mol
$\Delta G_{O3^{*-}}$	$\simeq$	-6.21	kcal/mol	$\Delta G^I_{O3^{*-}}$	$\simeq$	-4.225	kcal/mol
$\Delta G_{O3^*(O1)}$	$\simeq$	-8.45	kcal/mol	$\Delta G^I_{O3^*(O1)}$	$\simeq$	-3.135	kcal/mol
B. Basal and cooperative free energies							
ΔG_{b}	$\simeq$	-8.25	kcal/mol	ΔG_{C1loop}	$\simeq$	0	kcal/mol
ΔG_{C1-O3}	$\simeq$	$3\downarrow$	kcal/mol	ΔG_{O12}	$\simeq$	$6.84\downarrow$	kcal/mol
ΔG_{O23}	$\simeq$	$10.54\downarrow$	kcal/mol	ΔG_{O23_c}	$\simeq$	ΔG_{O23}	kcal/mol
ΔG_{O23*}	$\simeq$	$10.54\downarrow$	kcal/mol	ΔG_{O23*_c}	$\simeq$	ΔG_{O23*}	kcal/mol
ΔG_{O13}	$\simeq$	$4.68\downarrow$	kcal/mol	ΔG_{O13_c}	$\simeq$	ΔG_{O13}	kcal/mol
ΔG_{O13*}	$\simeq$	$4.68\downarrow$	kcal/mol	$\Delta G_{O13^*_c}$	$\simeq$	ΔG_{O13^*}	kcal/mol

the O1 and O3 operators, these configurations always include the repressor binding either the O3 or O3* operator. This seems to work opposite the CAP-repressor steric interaction at C1 and O3. However, these contributions do not simply cancel, since the steric interaction occurs regardless of loop formation. We denote the assistance in loop formation by CAP, ΔG_{C1loop}, and investigate its contribution to the repression of the *lac* operon.

In the complete model, we represent the DNA loop formation in the absence of CAP by ΔG_{Oij} and in the presence of CAP by ΔG_{Oij_c} for $i \in \{1,2\}$ and $j \in \{3,3*\}$. In Figures 16(a)-(b), the CAP assisted DNA loop formation has been removed from the model, but the parameters have not yet been adjusted. We are able to recover the fit to the data by decreasing the looping free energies which involve the O3 or O3* operator (ΔG_{C1-O3}, ΔG_{O12}, ΔG_{O13}, ΔG_{O23}, ΔG_{O13*} and ΔG_{O23*}), as displayed in

Figures 16(c)-(d) and Table 14. We are able to fit the current data with $\varepsilon = 3.1026$. However, as with O3*, the curves outside the current data are qualitatively different than the complete model. If new repression values were recorded for the (3,0,1), (1,0,1) and (0,0,1) mutants at WT, 5 $\times$ WT and 90 $\times$ WT repressor concentrations, our conclusions may differ. Data for 5 $\times$ WT and 90 $\times$ WT repressor concentrations does exist. However, because of data variation, repression would ideally be measured under the same conditions at all three repressor concentrations. We conclude that the current data does not support the inclusion of CAP assisted loop formation into the model of the repression of the *lac* operon.

NE4: C1-O3 Steric Interaction

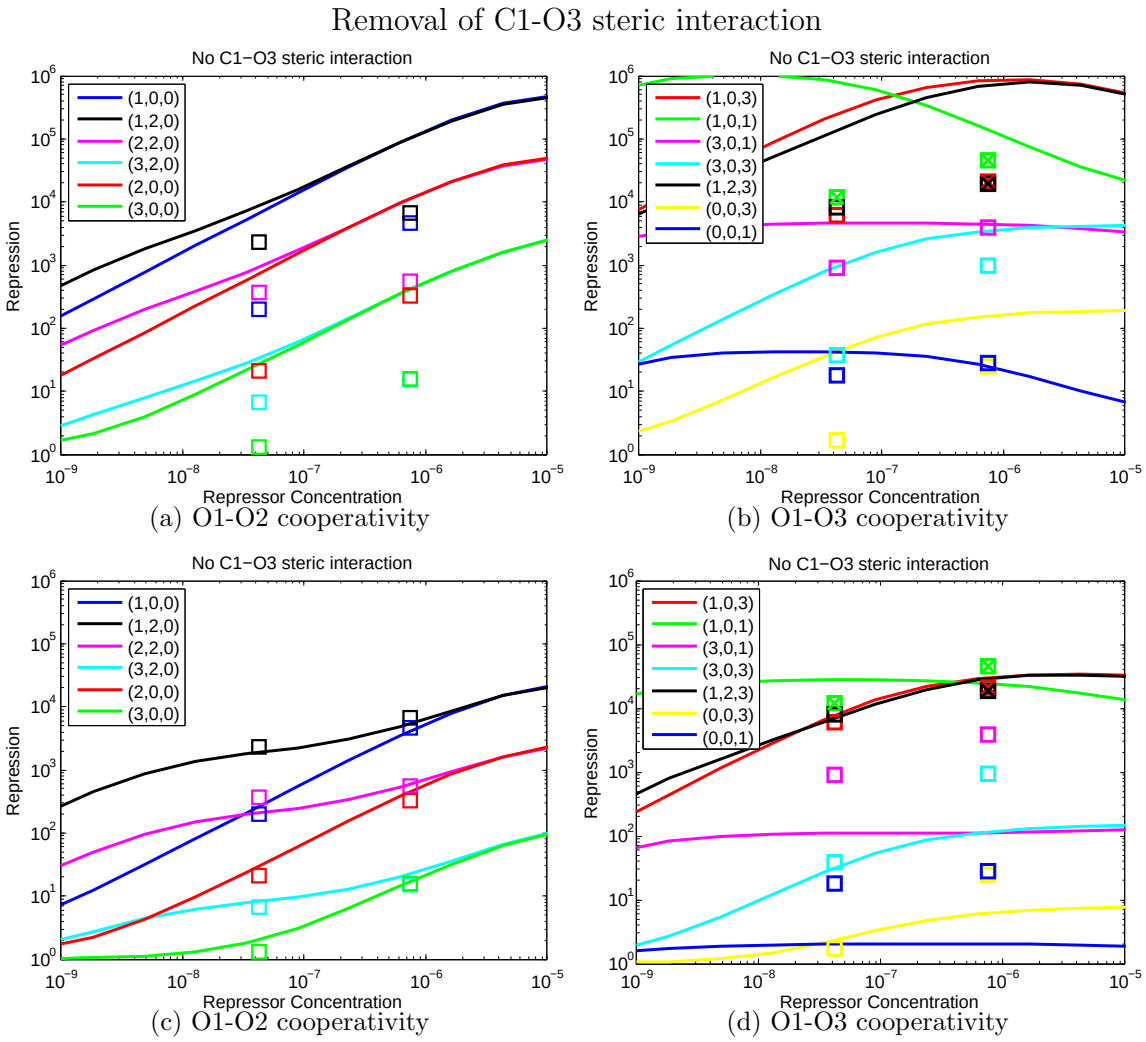


Figure 17: Repression level curves as a function of the repressor concentration. Here the steric interaction is removed, between a repressor bound to O3 and CAP bound to C1, from the model. (a) and (b) are the curves with parameters from the complete model (Table 10). (c) and (d) are the repression curves once the parameters have been readjusted to fit the model to the data. The new parameters are listed in Table 15.

Although the C1 binding site is 11 base pairs away from the O3 binding site [81], Hudson and Fried observed that CAP and the repressor will not concurrently bind [25, 39]. Swigon and Olson studied the O1-O3 loop with and without CAP bound to C1 and found the outer surface of the CAP protein makes unfavorable steric contacts

Table 15: Parameter values for Figure 17 where $\Delta G_{C1-O3} = 0$.

A. Free energy of binding without				and with IPTG			
ΔG_{O1}	$\simeq$	-13.03	kcal/mol	ΔG^I_{O1}	$\simeq$	-6.905	kcal/mol
ΔG_{O2}	$\simeq$	-11.665	kcal/mol	ΔG^I_{O2}	$\simeq$	-7.12	kcal/mol
ΔG_{O3}	$\simeq$	-9.6	kcal/mol	ΔG^I_{O3}	$\simeq$	-4.18	kcal/mol
ΔG_{O3^*}	$\simeq$	-8.485	kcal/mol	$\Delta G^I_{O3^*}$	$\simeq$	-4.075	kcal/mol
ΔG_{O1^-a}	$\simeq$	-7.68	kcal/mol	$\Delta G^I_{O1^-a}$	$\simeq$	-2.175	kcal/mol
ΔG_{O1^-b}	$\simeq$	-6.72	kcal/mol	$\Delta G^I_{O1^-b}$	$\simeq$	-5.895	kcal/mol
ΔG_{O2^-}	$\simeq$	-7.945	kcal/mol	$\Delta G^I_{O2^-}$	$\simeq$	-4.2	kcal/mol
ΔG_{O3^-}	$\simeq$	-7.205	kcal/mol	$\Delta G^I_{O3^-}$	$\simeq$	-3.245	kcal/mol
ΔG_{O3^*-}	$\simeq$	-6.21	kcal/mol	$\Delta G^I_{O3^*-}$	$\simeq$	-4.225	kcal/mol
$\Delta G_{O3^*(O1)}$	$\simeq$	-8.45	kcal/mol	$\Delta G^I_{O3^*(O1)}$	$\simeq$	-3.135	kcal/mol
B. Basal and cooperative free energies							
ΔG_{b}	$\simeq$	-8.25	kcal/mol	ΔG_{C1loop}	$\simeq$	-2.47	kcal/mol
ΔG_{C1-O3}	$\simeq$	0	kcal/mol	ΔG_{O12}	$\simeq$	$7.02\uparrow$	kcal/mol
ΔG_{O23}	$\simeq$	$12.69\uparrow$	kcal/mol	ΔG_{O23_c}	$\simeq$	$10.23\uparrow$	kcal/mol
ΔG_{O23*}	$\simeq$	$12.08\uparrow$	kcal/mol	ΔG_{O23*_c}	$\simeq$	$9.61\uparrow$	kcal/mol
ΔG_{O13}	$\simeq$	$7.76\uparrow$	kcal/mol	ΔG_{O13_c}	$\simeq$	$5.29\uparrow$	kcal/mol
ΔG_{O13*}	$\simeq$	$7.14\uparrow$	kcal/mol	$\Delta G_{O13^*_c}$	$\simeq$	$4.67\uparrow$	kcal/mol

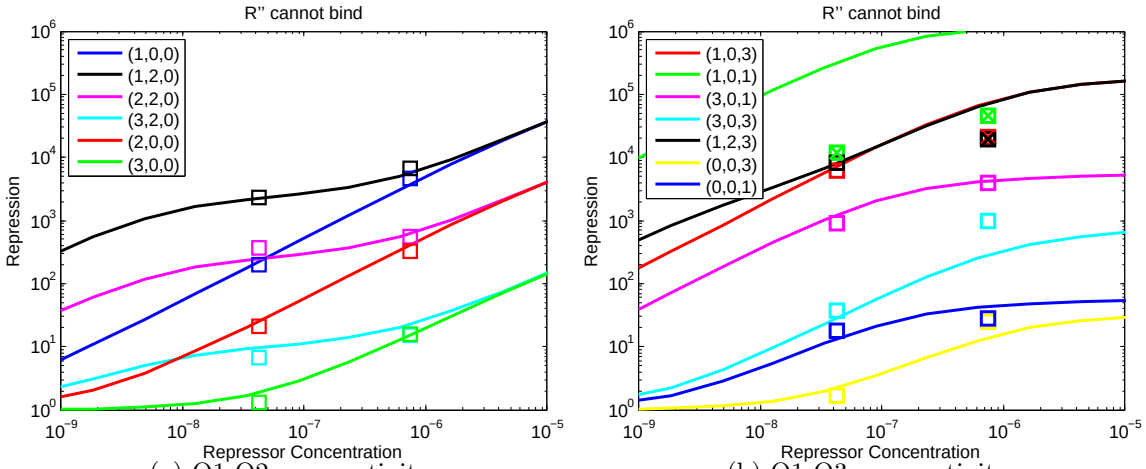
with the repressor or the DNA in several looping configurations. For the complete model in Figure 13, this parameter (ΔG_{C1-O3}) increases the energy required to bind a repressor to O3 and CAP to C1 by 3.5 kcal/mol, see Table 10. We investigate the contribution of ΔG_{C1-O3} to the repression values by removing the steric interaction, by setting $\Delta G_{C1-O3} = 0$.

When the steric interaction is removed between CAP bound to C1 and a repressor bound to O3, the model predictions for operators containing an operator in the O3 position drastically change. See Figures 17(a)-(b). To compensate, the energy to form any looped state (ΔG_{O12}, ΔG_{O13}, ΔG_{O23}, ΔG_{O13*} and ΔG_{O23*}) is increased. The new values are displayed in Figures 17(c)-(d) and Table 15. Even with these changes, removing the steric interference causes the states with O1 DNA in the O3 position to flatten considerably. In particular, (3,0,1) and (0,0,1) are considerably lower than the recorded values, and (3,0,3) fails to attain the recorded value at high concentration.

We find $\varepsilon = 7.8477$, much worse than the complete model. We conclude the C1-O3 steric interaction is vital to correctly represent the repression from the O1-O3 loop.

NE5: Impaired Binding Domains

Impaired Binding Domains are removed, associated parameters are in Table 16



(a) O1-O2 cooperativity (b) O1-O3 cooperativity

Figure 18: Repression level curves as a function of the repressor concentration as determined from the model when impaired repressor binding domains are not allowed to bind the DNA.

Each subunit of a repressor tetramer accepts one inducer molecule. Each binding domain is composed of two subunits and becomes impaired when one or two inducer molecules are bound to it. In the complete model, an impaired binding domain is allowed to bind at a reduced affinity as described in Table 10 by ΔG_k^I, where k is any operator DNA. These energies lie within the error bounds computed from the incremental base pair energies determined by Horton et al. [37] in the presence of IPTG. On the other hand, Oehler et al. [60] measure the concentration of free operator as a function of increasing IPTG concentration and find that their model describes the data substantially better if it does not allow impaired domains to bind the DNA.

Therefore we test the model under the assumption that impaired repressor binding domains are unable to bind the DNA. This change eliminates 464 states from the repression ratio, each of which contain a doubly impaired repressor bound to the DNA.

Table 16: Parameter values for Figure 18 where an impaired repressor domain is not allowed to bind DNA.

A. Free energy of binding without				and with IPTG			
ΔG_{O1}	$\simeq$	-13.03	kcal/mol	ΔG^I_{O1}			N/A
ΔG_{O2}	$\simeq$	-11.665	kcal/mol	ΔG^I_{O2}			N/A
ΔG_{O3}	$\simeq$	-9.6	kcal/mol	ΔG^I_{O3}			N/A
ΔG_{O3^*}	$\simeq$	-8.485	kcal/mol	$\Delta G^I_{O3^*}$			N/A
ΔG_{O1-a}	$\simeq$	-7.68	kcal/mol	ΔG^I_{O1-a}			N/A
ΔG_{O1-b}	$\simeq$	-6.72	kcal/mol	ΔG^I_{O1-b}			N/A
ΔG_{O2-}	$\simeq$	-7.945	kcal/mol	ΔG^I_{O2-}			N/A
ΔG_{O3-}	$\simeq$	-7.205	kcal/mol	ΔG^I_{O3-}			N/A
ΔG_{O3^*-}	$\simeq$	-6.21	kcal/mol	$\Delta G^I_{O3^*-}$			N/A
$\Delta G_{O3^*(O1)}$	$\simeq$	-8.45	kcal/mol	$\Delta G^I_{O3^*(O1)}$			N/A
B. Basal and cooperative free energies							
ΔG_{b}	$\simeq$	-8.25	kcal/mol	ΔG_{C1loop}	$\simeq$	-2.47	kcal/mol
ΔG_{C1-O3}	$\simeq$	3.5	kcal/mol	ΔG_{O12}	$\simeq$	6.89	kcal/mol
ΔG_{O23}	$\simeq$	12.01	kcal/mol	ΔG_{O23_c}	$\simeq$	9.54	kcal/mol
ΔG_{O23*}	$\simeq$	12.01	kcal/mol	ΔG_{O23*_c}	$\simeq$	9.54	kcal/mol
ΔG_{O13}	$\simeq$	5.41	kcal/mol	ΔG_{O13_c}	$\simeq$	2.94	kcal/mol
ΔG_{O13*}	$\simeq$	6.77	kcal/mol	$\Delta G_{O13^*_c}$	$\simeq$	4.3	kcal/mol

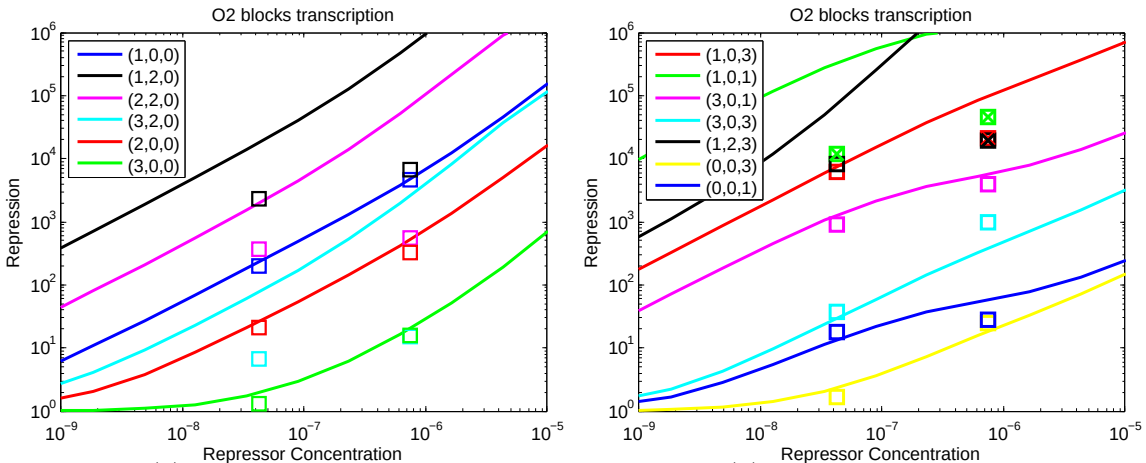
However, the model still includes all the primary components as well as P2, O3*, CAP assisted DNA looping, O3-CAP negative cooperativity due to steric interactions and low specific binding to deleted sites.

Surprisingly, the model predictions fit the data without any subsequent refitting of the parameter values. The repression curves are presented in Figure 18 and the parameters are in Table 16. The only visible change is a slight increase in repression at concentrations of repressor greater than 10^{-6}M. We find $\varepsilon = 2.7303$, only slightly larger than the complete model. Thus, our results support the conclusions of Oehler et al. [60]; if impaired repressors do bind the DNA, the effect is negligible.

Additional Experiments

NE6: Repressor bound to O2 Blocks Transcription

Repressor bound to O2 blocks transcription, associated parameters in Table 17



(a) O1-O2 cooperativity (b) O1-O3 cooperativity

Figure 19: Repression level curves as a function of the repressor concentration as determined from the model when a repressor bound to O2 completely blocks transcription.

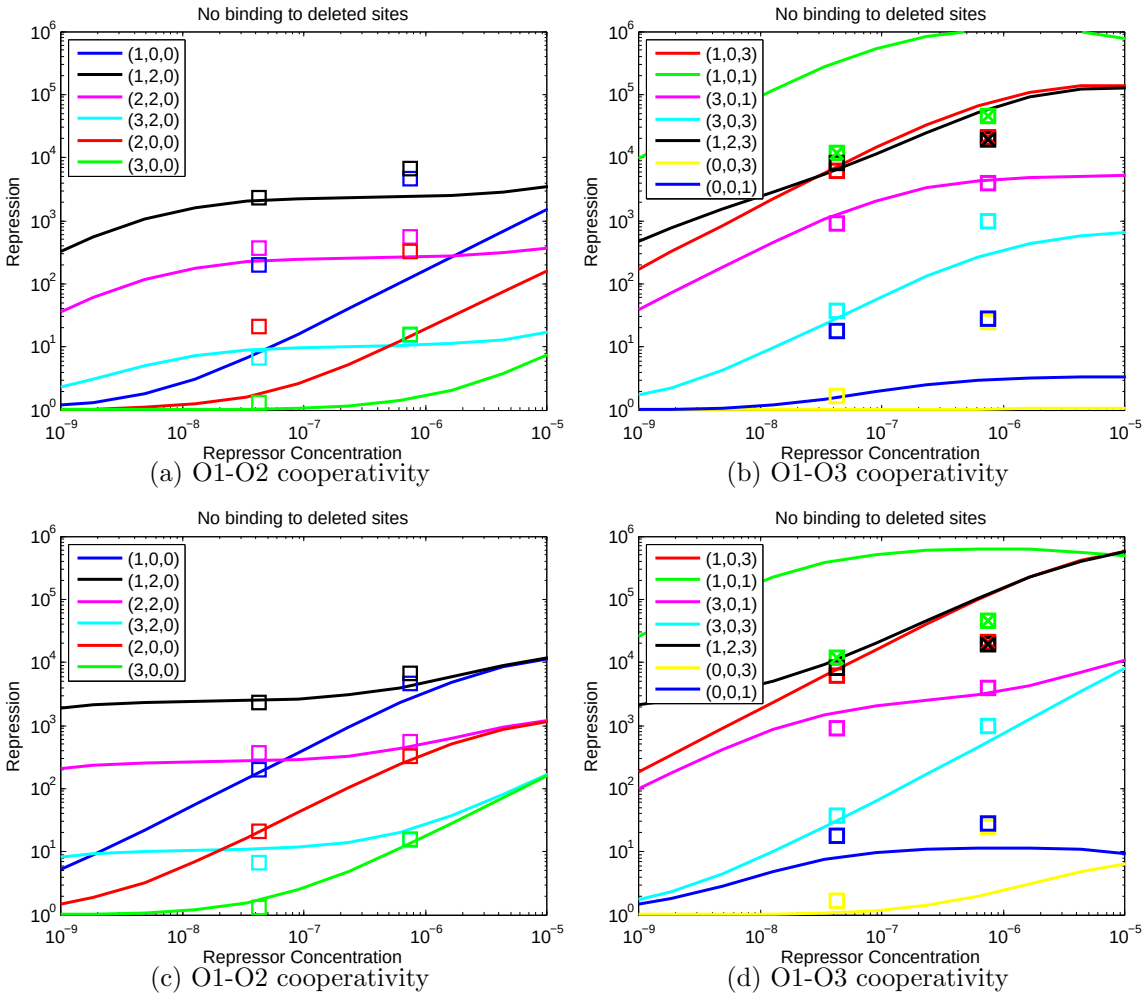
Though the *lac* repressor is not known to block transcription when bound to the O2 operator, the wild type O2 operator is located within the lacZ gene. If the model assumes a repressor bound to the O2 operator blocks transcription, the effect is described by Figure 19. The change is most apparent in the mutants with a non-deleted operator in the O2 position. The resulting repression curves for the wild type operator and mutants of the form $(i, j, 0)$ no longer level off between the two data points. Instead they continue to grow as the repressor concentration is increased. There were no parameter values which could recover the fit to the data, therefore Figure 19 was generated with the values from the complete model. We conclude that a repressor bound to O2 does not block transcription of the *lac* genes.

Table 17: Parameter values for Figure 19 the model which allows a repressor bound to O2 to block transcription.

A. Free energy of binding without				and with IPTG			
ΔG_{O1}	$\simeq$	-13.03	kcal/mol	ΔG^I_{O1}	$\simeq$	-6.905	kcal/mol
ΔG_{O2}	$\simeq$	-11.665	kcal/mol	ΔG^I_{O2}	$\simeq$	-7.12	kcal/mol
ΔG_{O3}	$\simeq$	-9.6	kcal/mol	ΔG^I_{O3}	$\simeq$	-4.18	kcal/mol
ΔG_{O3^*}	$\simeq$	-8.485	kcal/mol	$\Delta G^I_{O3^*}$	$\simeq$	-4.075	kcal/mol
ΔG_{O1^-a}	$\simeq$	-7.68	kcal/mol	$\Delta G^I_{O1^-a}$	$\simeq$	-2.175	kcal/mol
ΔG_{O1^-b}	$\simeq$	-6.72	kcal/mol	$\Delta G^I_{O1^-b}$	$\simeq$	-5.895	kcal/mol
ΔG_{O2^-}	$\simeq$	-7.945	kcal/mol	$\Delta G^I_{O2^-}$	$\simeq$	-4.2	kcal/mol
ΔG_{O3^-}	$\simeq$	-7.205	kcal/mol	$\Delta G^I_{O3^-}$	$\simeq$	-3.245	kcal/mol
$\Delta G_{O3^{*}-}$	$\simeq$	-6.21	kcal/mol	$\Delta G^I_{O3^{*}-}$	$\simeq$	-4.225	kcal/mol
$\Delta G_{O3^*(O1)}$	$\simeq$	-8.45	kcal/mol	$\Delta G^I_{O3^*(O1)}$	$\simeq$	-3.135	kcal/mol
B. Basal and cooperative free energies							
ΔG_{b}	$\simeq$	-8.25	kcal/mol	ΔG_{C1loop}	$\simeq$	2.47	kcal/mol
ΔG_{C1-O3}	$\simeq$	3.5	kcal/mol	ΔG_{O12}	$\simeq$	6.89	kcal/mol
ΔG_{O23}	$\simeq$	12.01	kcal/mol	ΔG_{O23_c}	$\simeq$	9.54	kcal/mol
ΔG_{O23*}	$\simeq$	12.01	kcal/mol	ΔG_{O23*_c}	$\simeq$	9.54	kcal/mol
ΔG_{O13}	$\simeq$	5.41	kcal/mol	ΔG_{O13_c}	$\simeq$	2.94	kcal/mol
ΔG_{O13*}	$\simeq$	6.77	kcal/mol	$\Delta G_{O13^*_c}$	$\simeq$	4.3	kcal/mol

NE7: Low-Specific Binding to Deleted Sites

Repressor is not allowed to bind deleted sites



(a) O1-O2 cooperativity

(b) O1-O3 cooperativity

(c) O1-O2 cooperativity

(d) O1-O3 cooperativity

Figure 20: Repression level curves as a function of the repressor concentration as determined from the model when repressor tetramers are not allowed to bind deleted sites. (a) and (b) are the repression curves with the parameters from the complete model (Table 10). (c) and (d) display the repression curves after the parameters have been readjusted to fit the data. The new parameters are in Table 18.

A protein binding site on the DNA can be experimentally deleted by changing some of the base pairs within the binding region. Rather than the site being deleted, the mutated site has a reduced binding for the protein, possibly low enough to be undetected experimentally. Our complete model allows a repressor to bind a deleted

Table 18: Free energy estimates for Figure 20 between a repressor binding domain and an operator sequence on the DNA. This model does not allow repressors to bind to deleted operators.

A. Free energy of binding without				and with IPTG			
ΔG_{O1}	$\simeq$	-15.01	kcal/mol	ΔG^I_{O1}	$\simeq$	-8.905	kcal/mol
ΔG_{O2}	$\simeq$	-13.665	kcal/mol	ΔG^I_{O2}	$\simeq$	-9.12	kcal/mol
ΔG_{O3}	$\simeq$	-11.6	kcal/mol	ΔG^I_{O3}	$\simeq$	-9.12	kcal/mol
ΔG_{O3^*}	$\simeq$	$-6.885\uparrow$	kcal/mol	$\Delta G^I_{O3^*}$	$\simeq$	$-2.475\uparrow$	kcal/mol
ΔG_{O1^-a}			N/A	$\Delta G^I_{O1^-a}$			N/A
ΔG_{O1^-b}			N/A	$\Delta G^I_{O1^-b}$			N/A
ΔG_{O2^-}			N/A	$\Delta G^I_{O2^-}$			N/A
ΔG_{O3^-}			N/A	$\Delta G^I_{O3^-}$			N/A
$\Delta G_{O3^{*-}}$			N/A	$\Delta G^I_{O3^{*-}}$			N/A
$\Delta G_{O3^*(O1)}$	$\simeq$	$-7.45\uparrow$	kcal/mol	$\Delta G^I_{O3^*(O1)}$	$\simeq$	$-2.135\uparrow$	kcal/mol
B. Basal and cooperative free energies							
ΔG_{b}	$\simeq$	$-10.25\downarrow$	kcal/mol	ΔG_{C1loop}	$\simeq$	2.47	kcal/mol
ΔG_{C1-O3}	$\simeq$	$6\uparrow$	kcal/mol	ΔG_{O12}	$\simeq$	$8.88\uparrow$	kcal/mol
ΔG_{O23}	$\simeq$	$11.84\downarrow$	kcal/mol	ΔG_{O23_c}	$\simeq$	$9.37\downarrow$	kcal/mol
ΔG_{O23*}	$\simeq$	$11.84\downarrow$	kcal/mol	ΔG_{O23*_c}	$\simeq$	$9.37\downarrow$	kcal/mol
ΔG_{O13}	$\simeq$	$8.14\uparrow$	kcal/mol	ΔG_{O13_c}	$\simeq$	$5.67\uparrow$	kcal/mol
ΔG_{O13*}	$\simeq$	$7.52\uparrow$	kcal/mol	$\Delta G_{O13^*_c}$	$\simeq$	$5.06\uparrow$	kcal/mol

operator with a reduced affinity. In this experiment we explore the difference between allowing repressors to bind deleted sites and completely removing the deleted sites from the model.

In Figures 20(a)-(b), we show the repression curves for the model which does not include the binding energy for a repressor to any of the deleted sites with the complete model parameter values listed in Table 10. Note the repression curves with two deletions drop significantly while the other mutants are relatively unchanged.

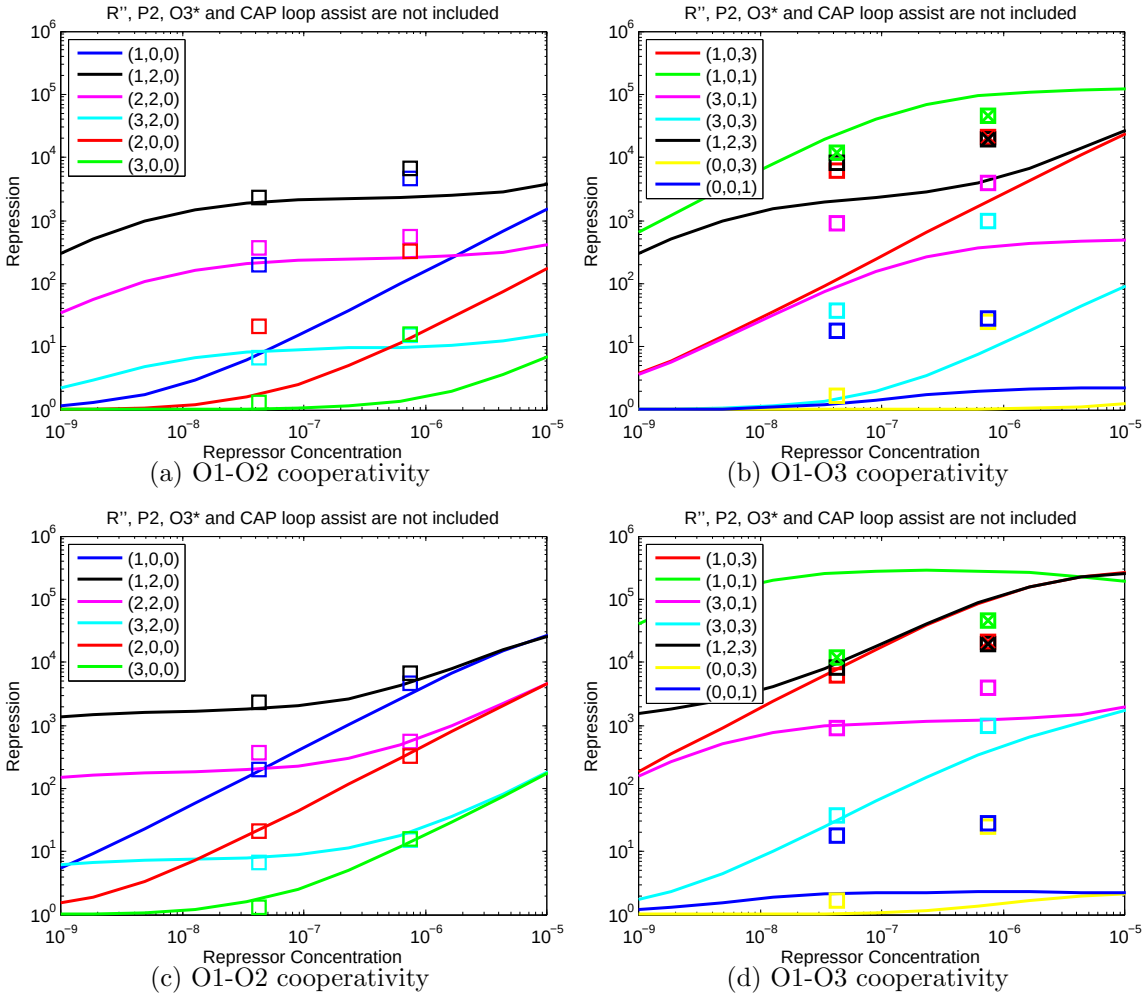
We were unable to recover the fit to the data until we decreased the basal energy 2 kcal/mol to $\Delta G_b = -10.25$ kcal/mol. This adjustment uniformly increases all binding affinities for the repressor binding domains to the operators, see Table 18A. Next we adjusted the cooperativity levels in Table 18B to get the fit in Figures 20(c)-(d).

Once we refit the curves to the data, we determine $\varepsilon = 4.5248$. When comparing the figures for the complete model and the model which allows low-specific binding to deleted sites, we see they differ for low concentrations in the mutants which have the O3 operator deleted. If experimentalists generated a new data set for the mutants with the O3 operator deleted at WT, 5 $\times$ WT and 90 $\times$ WT repressor concentrations, it would be possible to determine the importance of modeling the deleted sites as binding regions for the repressor.

Both models maintain the relative binding energies for repressor to operator, and both models require substantially lower looping energies than are predicted structurally [81]. However, the decrease in basal energy to -10.25 kcal/mol may prove to be biologically inaccurate. When $\Delta G_{\mathrm{b}} = -10.25$ kcal/mol, the free energy for the repressor to O1 is $\Delta G_{O1} = -15.01$ kcal/mol. This value is outside the range predicted by the literature (-14.1 kcal/mol [37] to -12.9 kcal/mol [60]). Furthermore, these parameter changes caused the terms with O3* and O3*(O1) to dominate the repression ratios in several of the mutants. As a result, we had to increase the free energy of ΔG_{O3*} and $\Delta G_{O3*(O1)}$ an additional 3 kcal/mol to attain the fit in Figures 20(c)-(d). The resulting energies are so weak that the O3* and O3*(O1) operators contribute very little to the repression. See Table 18. The existing data does not support the inclusion of low-specific binding to the deleted sites as long as the decrease in free energy is biologically reasonable.

NE8: Minimal Model

Minimal model, associated parameters in Table 19



(a) O1-O2 cooperativity (b) O1-O3 cooperativity

(c) O1-O2 cooperativity (d) O1-O3 cooperativity

Figure 21: Repression level curves as a function of the repressor concentration as determined from the minimal model. (a) and (b) are the repression curves with parameters from the complete model (Table 10). (c) and (d) are the repression curves after the parameters have been adjusted to fit the curves to the data. The new parameters are in Table 19.

We now remove all components which have been determined unnecessary for an accurate representation of the *lac* operon. It is possible that the deletion of each component individually can be compensated for by adjusting other parameters, but these adjustments are incompatible. If we can remove all of the unnecessary compo-

Table 19: Parameter values for Figure 21, the minimal model.

A. Free energy of binding without				and with IPTG			
ΔG_{O1}	$\simeq$	-15.1	kcal/mol	ΔG^I_{O1}			N/A
ΔG_{O2}	$\simeq$	-13.735	kcal/mol	ΔG^I_{O2}			N/A
ΔG_{O3}	$\simeq$	-11.67	kcal/mol	ΔG^I_{O3}			N/A
ΔG_{O3^*}			N/A	$\Delta G^I_{O3^*}$			N/A
ΔG_{O1-a}			N/A	ΔG^I_{O1-a}			N/A
ΔG_{O1-b}			N/A	ΔG^I_{O1-b}			N/A
ΔG_{O2-}			N/A	ΔG^I_{O2-}			N/A
ΔG_{O3-}			N/A	ΔG^I_{O3-}			N/A
ΔG_{O3^*-}			N/A	$\Delta G^I_{O3^*-}$			N/A
$\Delta G_{O3^*(O1)}$			N/A	$\Delta G^I_{O3^*(O1)}$			N/A
B. Basal and cooperative free energies							
ΔG_{b}	$\simeq$	$-10.32\downarrow$	kcal/mol	ΔG_{C1loop}	$\simeq$	0	kcal/mol
ΔG_{C1-O3}	$\simeq$	3.5	kcal/mol	ΔG_{O12}	$\simeq$	$9.14\uparrow$	kcal/mol
ΔG_{O23}	$\simeq$	$11.73\downarrow$	kcal/mol	ΔG_{O23_c}	$\simeq$	ΔG_{O23}	
ΔG_{O23*}			N/A	ΔG_{O23*_c}			N/A
ΔG_{O13}	$\simeq$	$6.92\uparrow$	kcal/mol	ΔG_{O13_c}	$\simeq$	ΔG_{O13}	
ΔG_{O13*}			N/A	$\Delta G_{O13^*_c}$			N/A

nents and still attain the fit, this model, which we call minimal, can then be used in subsequent modeling studies.

In Figure 21 we remove the O3* operator, the P2 promoter, and the CAP assisted DNA loop formation. In addition, we do not allow repressors to bind deleted sites nor do we allow impaired repressors to bind the DNA. These changes reduce the number of states in the (1,2,3) wild type operon from 684 in the complete model to 92. Because we do not allow the repressor to bind to deleted sites, any mutant with a deleted operator has even fewer states. The reduced model with parameter values from the complete model is displayed in Figures 21(a)-(b).

By decreasing the basal energy to $\Delta G_b = -10.32$ kcal/mol, and adjusting the looping cooperativity we are able to recover the fit to the data almost as well as the complete model. See Table 19 and Figures 21(c)-(d). We determine $\varepsilon = 5.1908$ for

minimal model and conclude that the minimal model should be sufficient to model the general behavior of the *lac* operon.

Conclusions

We introduced several, scientifically supported components into a model of the *lac* operon to derive the most complete biological model available. After fitting the model to the data, we systematically begin removing certain components. We determine the P2 promoter and the binding of impaired repressors contribute negligible amounts of repression. Additionally, although the O3* operator and the CAP assisted DNA looping change the qualitative behavior of the repression curves, these components can be removed and the model is still able to reasonably match the Oehler [61] repression data. For all of these changes, the binding affinity for the repressor to O1 stays within the biologically determined range of energy. The only component we determine to be vital to predicting *lac* operon repression is the CAP-repressor steric interference at O3 and C1.

For the minimal model we remove all non-vital components and we do not allow repressor tetramers to bind deleted operators. These changes reduce the number of states for the wildtype operon (1,2,3) from 684 states in the complete model to 92 states in the minimal model. The minimal model provides an acceptable representation of the Oehler data [61] and is significantly less complicated. However, to attain the fit, the free energy of binding a repressor to the DNA must decrease 2 kcal/mol. This decrease in energy forces ΔG_{O1} about 1 kcal/mol more negative than the range of biological measurements for repressor to the O1 operator.

However, some curves do change outside the range of measurements. We mention that additional repression data could distinguish between the models. For some

mutants Oehler et al. [62] recorded repression at wildtype repressor concentrations. However, the cells were of a different type and the operon deletion sequences were different from the 1994 deletions. The variation in data can be observed between the 1990 and 1994 data at the higher repressor concentrations. Therefore, additional repression data at multiple repressor concentrations is required before we can determine if further components can be deemed necessary for the model.

REFERENCES CITED

[1] Beta-galatcosidase assay (a better Miller) http://openwetwear.org.

[2] Nobelprize.org.

[3] PDB ID: 1krv, XG Wang, LR Olsen, and SL Roderick. Structure of the *lac* operon galactoside acetyltransferase. *Structure*, 10:581–8, 2002.

[4] PDB ID: 1pv6, J Abramson, I Smirnova, V Kasho, G Verner, HR Kaback, and S Iwata. Structure and mechanism of the lactose permease of *Escherichia coli*. *Science*, 301:610–5, 2003.

[5] GK Ackers, AD Johnson, and MA Shea. Qualitative model for gene regulation by λ phage repressor. *Proc Nat Acad Sci USA*, 79(4):1129–33, February 1982.

[6] U Alon. *An introduction to systems biology*. Chapman & Hall/CRC, 2007.

[7] L Bai, TJ Santangelo, and MD Wang. Single-molecule analysis of RNA polymerase transcription. *Ann Rev Biophys Biomol Struct*, 35:343–60, 2006.

[8] L Bai, A Shundrovsky, and MD Wang. Sequence-dependent kinetic model for transcription elongation by RNA polymerase. *J Mol Biol*, 344:335–49, 2004.

[9] CH Baker, ST Tomlinson, AE García, and JG Harman. Amino acid substitution at position 99 affects the rate of CRP subunit exchange. *Biochemistry*, 40(41):12329–38, 2001.

[10] C Balsalobre, J Johansson, and BE Uhlin. Cyclic AMP-dependent osmoregulation of *crp* gene expression in *Escherichia coli*. *J Bacteriol*, 188:5935–44, 2006.

[11] BD Barkley, AD Riggs, A Jobe, and S Burgeois. Interaction of effecting ligands with *lac* repressor and repressor-operator complex. *Biochemistry*, 14(8):1700–12, 1975.

[12] NA Becker, JD Kahn, and LJ Maher III. Effects of nucleoid proteins on DNA repression loop formation in *Escherichia coli*. *Nucleic Acids Res*, 35(12):3988–4000, 2007.

[13] NA Becker, JD Kahn, and LJ Maher III. Eukaryotic HMGB proteins as replacements for HU in *E. coli* repression loop formation. *Nucleic Acids Res*, 36(12):4009–21, 2008.

[14] HM Berman, K Henrick, and H Nakamura. Announcing the worldwide Protein Data Bank. *Nat Strut Biol*, 10(12):980, 2003.

[15] L Bintu, NE Buchler, HG Garcia, U Gerland, T Hwa, J Kondev, and R Phillips. Transcriptional regulation by the numbers: models. *Curr Opin Genet Dev*, 15(2):116–24, 2002.

[16] H Bremer and PP Dennis. Modulation of chemical composition and other parameters of the cell by growth rate. In F.C. Neidhart et. al, editor, *In Escherichia coli and Salmonella thyphymurium: Cellular and Molecular Biology*, volume 2, pages 1553–1569. American Society for Microbiology, Washington DC, 1996.

[17] C Chothia and J Gough. Genomic and structural aspects of protein evolution. *Biochem J*, 419(1):15–28, 2009.

[18] JD Chung and G Stephanopoulos. On physiological multiplicity and population heterogeneity of biological systems. *Chem Eng Sci*, 51(9):109–21, 1996.

[19] M Cohn and K Horibata. Inhibition by glucose of the induced synthesis of the β-galactosidase-enzyme system of *Escherichia coli*: Analysis of maintenance. *J Bacteriol*, 78:601–12, 1959.

[20] P Cossart and B Gicquel-Sanzey. Regulation of expression of the crp gene of *Escherichia coli* K-12: in vivo study. *J Bacteriol*, 161(1):454–7, 1985.

[21] L Czapla, D Swigon, and WK Olson. Effects of the nucleoid protein HU on the structure, flexibility, and ring-closure properties of DNA deduced from Monte Carlo simulations. *J Mol Biol*, 382:353–70, 2008.

[22] D Czarniecki, RJ Noel, Jr, and WS Reznikoff. The -45 region of the *Escherichia coli lac* promoter: CAP-dependent and CAP-independent transcription. *J Bacteriol*, 179(2):423–9, 1997.

[23] J Deutscher, C Francke, and PW Postma. How phosphotransferase system-related protein phosphorylation regulates carbohydrate metabolism in bacteria. *Microbiol Mol Biol Rev*, 70(4):939–1031, 2006.

[24] M Djordjevic and R Bundschuh. Formation of the open complex by bacterial RNA polymerase–a quantitative model. *Biophys J*, 94:4233–48, 2008.

[25] MG Fried and JM Hudson. DNA looping and *lac* repressor-CAP interaction. *Science*, 274(5294):1930–1, 1996.

[26] T Gedeon, K Mischaikow, K Patterson, and E Traldi. Binding cooperativity in phage lambda is not sufficient to produce an effective switch. *Biophys J*, 94(9):3384–92, 2008.

[27] T Gedeon, K Mischaikow, K Patterson, and E Traldi. When activators repress and repressors activate: a qualitative analysis of the Shea-Ackers model. *Bull Math Biol*, 70(6):1660–83, 2008.

[28] DS Goodsell. lac Repressor. http://dx.doi.org/10.2210/rcsb_pdb/mom_2003_3.

[29] B Görke and J Stülke. Carbon catabolite repression in bacteria: many ways to make the most out of nutrients. *Nat Rev Microbiol*, 6(8):613–24, 2008.

[30] SJ Greive and PH von Hippel. Thinking quantitatively about transcriptional regulation. *Nature Rev Mol Cell Bio*, 6:221–32, 2005.

[31] AJF Griffiths, JH Miller, DT Suzuki, RC Lewontin, and WM Gelbart. *An Introduction to Genetic Analysis*. WH Freeman and Company, 2000.

[32] TL Hill. *Introduction to Statistical Thermodynamics*. Addison Wesley, 1960.

[33] BM Hogema, JC Arents, R Bader, K Eijkemans, T Inada, H Aiba, and PW Postma. Inducer exclusion by glucose 6-phosphate in *Escherichia coli*. *Mol Microbiol*, 28(4):755–65, 1998.

[34] BM Hogema, JC Arents, R Bader, K Eijkemans, H Yoshida, H Takahashi, H Aiba, and PW Postma. Inducer exclusion in *Escherichia coli* by non-PTS substrates: the role of the PEP to pyruvate ratio in determining the phosphorylation state of enzyme IIA^{Glc}. *Mol Microbiol*, 30(3):487–98, 1998.

[35] BM Hogema, JC Arents, R Bader, and PW Postma. Autoregulation of lactose uptake through the LacY permease by enzyme IIA^{Glc} of the PTS in *Escherichia coli* K-12. *Mol Microbiol*, 31(6):1825–33, 1999.

[36] BM Hogema, JC Arents, T Inada, H Aiba, K van Dam, and PW Postma. Catabolite repression by glucose-6-phosphate, gluconate and lactose in *Escherichia coli*. *Mol Microbiol*, 24(4):857–67, 1997.

[37] N Horton, M Lewis, and P Lu. *Escherichia coli* lac repressor-*lac* operator interaction and the influence of allosteric effectors. *J Mol Biol*, 265:1–7, 1997.

[38] http://www.wikipedia.org/.

[39] JM Hudson and MG Fried. Co-operative interactions between the catabolite gene activator protein and the *lac* repressor at the lactose promoter. *J Mol Biol*, 214:381–96, 1990.

[40] T Inada, K Kimata, and H Aiba. Mechanism responsible for glucose-lactose diauxie in Escherichia coli: challenge to the cAMP model. *Genes Cells*, 1:293–301, 1996.

[41] H Ishizuka, A Hanamura, T Kunimura, and H Aiba. Mechanism of the down-regulation of cAMP receptor protein by glucose in *Escherichia coli*: role of autoregulation of the *crp* gene. *EMBO J*, 13(13):3077–82, 1994.

[42] F Jacob. Genetics of the bacterial cell. Nobel Lecture, Dec 1965.

[43] AN Kapanidis, E Margeat, S Ho, E Kortkhonjia, S Weiss, and RH Ebright. Initial transcription by RNA polymerase proceeds through a DNA-scrunching mechanism. *Science*, 314(5802):1144–7, November 2006.

[44] IM Keseler, C Bonavides-Martinez, J Collado-Vides, S Gama-Castro, RP Gunsalus, DA Johnson, M Krummenacker, LM Nolan, S Paley, IT Paulsen, M Peralta-Gil, A Santos-Zavaleta, AG Shearer, and PD Karp. Ecocyc: A comprehensive view of *Escherichia coli* biology. *Nucleic Acids Res*, 37:D464–70, 2009.

[45] T Kuhlman, Z Zhang, MH Saier, Jr, and T Hwa. Combinatorial transcriptional control of the lactose operon of *Escherichia coli*. *Proc Nat Acad Sci USA*, 104(14):6043–8, 2007.

[46] CL Lawson, D Swigon, KS Murakami, SA Darst, HM Berman, and RH Ebright. Catabolite activator protein: DNA binding and transcription activation. *Curr Opin Struct Biol*, 14:10–20, 2004.

[47] S Lindemose, PE Nielsen, and NE Mollegaard. Dissecting direct and indirect readout of cAMP receptor protein DNA binding using an inosine and 2,6-diaminopurine in vitro selection system. *Nucleic Acids Res*, 36(14):4797–807, 2008.

[48] M Liu, G Gupte, S Roy, RP Bandwar, SS Patel, and S Garges. Kinetics of transcription initiation at lacP1. Multiple roles of cyclic AMP receptor protein. *J Biol Chem*, 278(41):39755–61, 2003.

[49] TP Malan and WR McClure. Dual promoter control of the *Escherichia coli* lactose operon. *Cell*, 39:173–80, 1984.

[50] J Monod. *Recherche sur la croissance des cultures bactériennes*. PhD thesis, Paris: Hermann ed, 1942.

[51] J Müller, S Oehler, and B Müller-Hill. Repression of *lac* promoter as a function of distance, phase and quality of an auxiliary *lac* operator. *J Mol Biol*, 257:21–9, 1996.

[52] B Müller-Hill. *The* lac *Operon: A Short History of a Genetic Paradigm*. de Gruyter, 1996.

[53] KS Murakami and SA Darst. Bacterial RNA polymerase: the whole story. *Curr Opin Struct Biol*, 13:31–9, 2003.

[54] A Narang. Effect of DNA looping on the induction kinetics of the *lac* operon. *J Theor Biol*, 247:695–712, 2007.

[55] A Narang and SS Pilyugin. Bistability of the *lac* operon during growth of *Escherichia coli* on lactose and lactose + glucose. *Bull Math Biol*, 70:1032–64, 2008.

[56] P Nelson. *Biological Physics: Energy, Information, Life*. WH Freeman, 2002.

[57] JT Noel, SS Pilyugin, and A Narang. The diffusive influx and carrier efflux have a strong effect on the bistability of the *lac* operon in *Escherichia coli*. *J Theor Biol*, 256:14–28, 2009.

[58] L Notley-McRobb, A Death, and T Ferenci. The relationship between external glucose concentration and cAMP levels inside *Escherichia coli*: implications for models of phosphotransferase-mediated regulation of adenylate cyclase. *Microbiology*, 143:1909–18, 1997.

[59] A Novick and M Weiner. Enzyme induction as an all-or-none phenomenon. *Proc Nat Acad Sci USA*, 43:553–66, 1957.

[60] S Oehler, S Alberti, and B Müller-Hill. Induction of the *lac* promoter in the absence of DNA loops and the stoichiometry of induction. *Nucleic Acids Res*, 34(2):606–12, 2006.

[61] S Oehler, M Amouyal, P Kolkhof, B von Wilcken-Bergmann, and B Müller-Hill. Quality and position of the three *lac* operators of *E. coli* define efficiency of repression. *EMBO J*, 13(14):3348–55, 1994.

[62] S Oehler, ER Eismann, H Krämer, and B Müller-Hill. The three operators of the *lac* operon cooperate in repression. *EMBO J*, 9(4):973–9, 1990.

[63] EM Ozbudak, M Thattai, HN Lim, BI Shraiman, and A Van Oudenaarden. Multistability in the lactose utilization network of Escherichia coli. *Nature*, 427(6976):737–40, 2004.

[64] PDB ID: 1bgl, RH Jacobson, XJ Zhang, RF Dubose, and BW Matthews. Three-dimensional structure of β-galactosidase from E. coli. *Nature*, 369:761–6, 1994.

[65] A Revyakin, C Liu, RH Ebright, and TR Strick. Abortive initiation and productive initiation by RNA polymerase involve DNA scrunching. *Science*, 314(5802):1139–43, November 2006.

[66] S Roy, S Garges, and S Adhya. Activation and repression of transcription by differential contact: two sides of a coin. *J Biol Chem*, 273:14059–62, 1998.

[67] L Saiz and JMG Vilar. DNA looping: the consequences and its control. *Curr Opin Struct Biol*, 16(3):344–50, 2006.

[68] L Saiz and JMG Vilar. Stochastic dynamics of macromolecular-assembly networks. *Mol Syst Biol*, 2:2006.0024, 2006.

[69] L Saiz and JMG Vilar. *Ab initio* thermodynamic modeling of distal multisite transcription regulation. *Nucleic Acids Res*, 36(3):726–31, 2008.

[70] M Santillán. Bistable behavior in a model of the *lac* operon in *Escherichia coli* with variable growth rate. *Biophys J*, 94(6):2065–81, 2008.

[71] M Santillán and MC Mackey. Influence of catabolite repression and inducer exclusion on the bistable behavior of the *lac* operon. *Biophys J*, 86:1282–92, March 2004.

[72] M Santillán and MC Mackey. Why the lysogenic state of phage λ is so stable: A mathematical modeling approach. *Biophys J*, 86:75–84, January 2004.

[73] M Santillán and MC Mackey. Quantitative approaches to the study of bistability in the *lac* operon of *Escherichia coli*. *J R Soc Interface*, 5 Suppl 1:S29–39, 2008.

[74] M Santillán, MC Mackey, and ES Zeron. Origin of Bistability in the *lac* Operon. *Biophys J*, 92(11):3830–42, 2007.

[75] DV Schroeder. *An Introduction to Thermal Physics*. Addison Wesley Longman, 2000.

[76] Y Setty, AE Mayo, MG Surette, and U Alon. Detailed map of a cis-regulatory input function. *Proc Nat Acad Sci USA*, 100(13):7702–7, 2003.

[77] MA Shea and GK Ackers. The O_R control system of bacteriophage lambda : A physical-chemical model for gene regulation. *J Mol Bio*, 181(2):211–30, 1985.

[78] S Strickland, G Palmer, and V Massey. Determination of dissociation constants and specific rate constants of enzyme-substrate (or protein-ligand) interactions from rapid reaction kinetic data. *J Biol Chem*, 250(11):4048–52, 1975.

[79] J Stülke and W Hillen. Carbon catabolite repression in bacteria. *Curr Opin Microbiol*, 2:195–201, 1999.

[80] D Swigon, BD Coleman, and WK Olson. Modeling the Lac repressor-operator assembly: The influence of DNA looping on Lac repressor conformation. *Proc Nat Acad Sci USA*, 103(26):9879–84, 2006.

[81] D Swigon and WK Olson. Mesoscale modeling of multi-protein-DNA assemblies: The role of the catabolic activator protein in Lac-repressor-mediated looping. *Int J Non Linear Mech*, 43(10):1082–93, 2008.

[82] VR Tadigotla, DO Maoiléidigh, AM Sengupta, V Epshtein, RH Ebright, E Nudler, and AE Ruckenstein. Thermodynamic and kinetic modeling of transcriptional pausing. *Proc Nat Acad Sci USA*, 103(12):4439–44, 2006.

[83] M van Hoek and P Hogeweg. The effect of stochasticity on the *lac* operon: and evolutionary perspective. *PLoS Comp Biol*, 3:1071–82, June 2007.

[84] MJA van Hoek and P Hogeweg. In silico evolved *lac* operons exhibit bistability for artificial inducers, but not for lactose. *Biophys J*, 91:2833–43, October 2006.

[85] JMG Vilar and L Saiz. DNA looping in gene regulation: from the assembly of macromolecular complexes to the control of transcriptional noise. *Curr Opin Genet Dev*, 15(2):136–44, 2005.

[86] P Wong, S Gladney, and JD Keasling. Mathematical model of the *lac* operon: Inducer exclusion, catabolite repression, and diauxic growth on glucose and lactose. *Biotechnol Prog*, 13:132–43, 1997.

[87] XC Xue, F Liu, and ZC Ou-Yang. A kinetic model of transcription initiation by RNA polymerase. *J Mol Biol*, 378(3):520–9, May 2 2008.

[88] N Yildirim and MC Mackey. Feedback regulation in the lactose operon: a mathematical modeling study and comparison with experimental data. *Biophys J*, 84:2841–51, 2003.

[89] N Yildirim, M Santillán, D Horike, and MC Mackey. Dynamics and bistability in a reduced model of the *lac* operon. *Chaos*, 14:279–92, 2004.

[90] Y Zhang, AE McEwen, DM Crothers, and SD Levene. Analysis of *in-vivo* LacR-mediated gene repression based on the mechanics of DNA looping. *PLoS ONE*, 1:e136, 2006.

APPENDICES

APPENDIX A

MAIN DEFINITIONS

Definition A.1. **Eukaryote:** A phylogenetic domain of organisms, characterized by the presence of a nucleus. Animals, plants and fungi are all eukaryotes.

Definition A.2. **Inducer Exclusion:** The process of blocking the non-PTS carbon sources (in our case lactose) from being transported into the cell. Historically this was thought to be only caused by glucose, but it has also been observed with glucose 6-phosphate [33] and to a small extent with lactose [35].

Definition A.3. **Plasmid:** A plasmid is a small closed loop of DNA often used by biologists to introduce genes of interest into a cell, or by cells to perform horizontal gene transfer (sharing of genes), or the result of a viral infection.

Definition A.4. **Prokaryote:** A phylogenetic domain of organisms, characterized by the absence of a cell nucleus. Members include bacteria like *Escherichia coli*.

Definition A.5. **Transcription:** The process of creating an RNA copy of DNA, performed by RNA polymerase.

Definition A.6. **Translation:** The process of producing protein from mRNA, performed by ribosomes which bind the mRNA and match amino acids to the mRNA template.

Definition A.7. **Wildtype:** This refers to the form of a gene, protein, or organism as found in nature.

APPENDIX B

STATISTICAL MECHANICS

To model the process of transcription we assume that if the operon is not repressed, transcription of that gene occurs at some given rate, k_f, provided that RNA polymerase (RNAP) is bound to the promoter. Thus transcription occurs when the regulatory region is bound by any possible combination of protein and DNA where RNAP is bound and the promoter is not repressed. We assume that all chemical reactions occur at thermodynamic equilibrium. Under these assumptions, the rate of mRNA production can be modeled as a product of a rate of production, k_f, and the probability that a set of favorable states occurs. To account for gene regulation by proteins, the probability a particular protein (with fixed concentration) binds to a fixed site must be determined. This requires some statistical mechanics.

The fundamental assumption for statistical mechanics is the following postulate:

THEOREM B.1 (Fundamental Postulate). *Given an isolated system in equilibrium, it is found with equal probability in each of its accessible microstates. [38, 56]*

A microstate describes detailed microscopic properties, such as all positions and momenta of all atoms for a gas in a container, whereas a macrostate refers to macroscopic properties like temperature, pressure, and volume.

Entropy is a measure of the statistical "disorder" of the thermodynamic system; the amount of uncertainty that remains about the exact microscopic state of the system, given a description of its macroscopic properties [38]. Entropy satisfies the following relation:

$$S(s) := k_B \ln \Omega(s)$$

where

- k_B is Boltzmann's constant, and

- $\Omega(s)$ is the number of microstates accessible to the system for the given macrostate s.

Solving for $\Omega(s)$,

$$\Omega(s) = \exp(S(s)/k_B).$$

Consider a system in contact with a reservoir, where energy can flow between the system and the reservoir, and the reservoir has an infinitely large heat capacity to maintain constant temperature for the combined system. The combined system is isolated, so by the Fundamental Postulate, all microstates are equally probable. Then the probability of a particular state s occurring is directly proportional to the number of microstates accessible to the reservoir. A ratio of probabilities for two different macrostates, s_1 and s_2, can be calculated as follows:

$$\begin{aligned}\frac{\mathbb{P}(s_1)}{\mathbb{P}(s_2)} &= \frac{\Omega(s_1)}{\Omega(s_2)} \\ &= \frac{e^{S(s_1)/k_B}}{e^{S(s_2)/k_B}} \\ &= e^{[S(s_1)-S(s_2)]/k_B} \\ &= e^{\Delta S/k_B} \end{aligned} \tag{54}$$

The entropy change in the system from s_1 to s_2 is small in comparison to the entropy of the reservoir. Thus the generalized thermodynamic identity from the 1$^{\text{st}}$ law of thermodynamics can be applied to the reservoir:

$$\Delta U = T\Delta S - P\Delta V + \sum_i \mu_i \Delta N_i \tag{55}$$

where Δ is "change in", U = energy, T = degrees Kelvin, S = entropy, P = pressure, V = volume, μ_i = chemical potential of the i$^{\text{th}}$ species and N_i is the number of moles of the of the i$^{\text{th}}$ species, all with respect to the reservoir.

Solving for ΔS in (55) gives:

$$\Delta S \approx \frac{1}{T}\left(\Delta U - \sum_i \mu_i \Delta N_i\right). \tag{56}$$

The $P\Delta V$ term is ignored since the change in volume is usually zero or at least minimal in comparison to ΔU for biochemical reactions. Since any change in energy or particles gained by the reservoir must be lost by the system, equation (56) becomes:

$$\Delta S \approx \frac{1}{T}\left(-\Delta E + \sum_i \mu_i \Delta n_i\right), \tag{57}$$

where E and n_i refer to the small system.

Substitute equation (57) into (54) to get:

$$\begin{aligned}
\frac{\mathbb{P}(s_1)}{\mathbb{P}(s_2)} &= e^{\Delta S/k_B} \\
&= \exp\left(\frac{-1}{k_B T}(\Delta E - \sum_i \mu_i \Delta n_i)\right) \\
&= \exp(-[E(s_1) - E(s_2) - \sum_i \mu_i n_i(s_1) + \sum_i \mu_i n_i(s_2)]/k_B T) \\
&= \frac{\exp\left(\frac{-1}{k_B T}\left[E(s_1) - \sum_i \mu_i n_i(s_1)\right]\right)}{\exp\left(\frac{-1}{k_B T}\left[E(s_2) - \sum_i \mu_i n_i(s_2)\right]\right)}.
\end{aligned}$$

This ratio of probabilities is a ratio of exponential factors called Gibbs factors.

$$Gibbs\ factor := \exp\left(\frac{-\left[E(s) - \sum_i \mu_i n_i(s)\right]}{k_B T}\right)$$

To obtain an absolute probability, take $s_1 = s$ and $s_2 =$ all states. Then

$$\frac{\mathbb{P}(s_1)}{\mathbb{P}(s_2)} = \frac{\mathbb{P}_s}{1} = \frac{\exp\left(\frac{1}{k_B T}\left[E(s) - \sum_i \mu_i n_i(s)\right]\right)}{\sum_j \exp\left(\frac{1}{k_B T}\left[E(s_j) - \sum_i \mu_i n_i(s_j)\right]\right)}$$

$$\mathbb{P}_s = \frac{\exp\left(\frac{1}{k_B T}\left[E(s) - \sum_i \mu_i n_i(s)\right]\right)}{Z}$$

where $\mathbb{P}_s = \mathbb{P}(s)$ and Z is the grand partition function or Gibbs sum

$$Z := \sum_s \mathbb{P}_s$$

a sum over all possible states, s, allowing the probabilities sum to one. With some rearrangement, it can shown

$$\begin{aligned}\mathbb{P}_s &= \frac{1}{Z}\exp\left(\frac{-E(s)}{k_B T}\right)\exp\left(\frac{\mu_1 n_1(s)}{k_B T}\right)\exp\left(\frac{\mu_2 n_2(s)}{k_B T}\right)\ldots\exp\left(\frac{\mu_p n_p(s)}{k_B T}\right) \\ &= \frac{1}{Z}\exp\left(\frac{-E(s)}{k_B T}\right)[species_1]^{n_1(s)}[species_2]^{n_2(s)}\ldots[species_p]^{n_p(s)},\end{aligned}$$

where

$$\mu_i = -k_B T \ln([species_i]) \text{ ([75])}$$

and $[species_i]$ is the ith molecule involved in the system.

Converting between kcal/mol and $k_B T$

Sometimes, rather than using joules (the SI unit) for energy E(s), kcal/mol is used. The energies are related as follows. The gas constant (R) and Boltzmann's constant (k_B) have the relationship:

$$R/N_A = k_B,$$

where R is in J mol^{-1} K^{-1}, and k_B is in J K^{-1} ($k_B = 1.3806504 \times 10^{-23}$J K^{-1}). Also, R can be converted to kcal mol^{-1} K^{-1}. Denote R_j the gas constant in J mol^{-1} K^{-1} and R_k the gas constant in kcal mol^{-1} K^{-1}. Then,

$$\begin{aligned} R_j/N_A &= k_B \\ R_j &= k_B * N_A \\ &= 1.3806504 \times 10^{-23}\text{J K}^{-1} * 6.022 \times 10^{23}\text{mol}^{-1} \\ &= 8.31\text{J K}^{-1}\text{mol}^{-1}. \end{aligned}$$

Now, convert both sides to kcal mol^{-1} K^{-1}:

$$\begin{aligned} R_j * 2.39 \times 10^{-4} &= 8.31\text{J K}^{-1}\text{mol}^{-1} * 2.39 \times 10^{-4}\text{kcal/J} \\ R_k &= 1.986 \times 10^{-3}\text{kcal K}^{-1}\text{mol}^{-1} \\ &= .6/T \text{ kcal/mol if T=310.15 Kelvin which is 37°C} \end{aligned}$$

Thus, if we assume ΔG is in kcal/mol but want to convert to k_BT we must convert from R_k to R_j and then use the equation $R_j/N_A = k_B$:

$$\begin{aligned} \frac{\Delta G(\text{kcal/mol})}{R_kT(\text{kcal/mol})} &= \frac{\Delta G(\text{kcal/mol})}{R_kT(\text{kcal/mol})} * \frac{4184(\text{J/kcal})}{4184(\text{J/kcal})} \\ &= \frac{\Delta G * 4184 \text{ (J/mol)}}{R_jT \text{ (J/mol)}} \end{aligned}$$

Substituting $R_j = k_B * N_A$ in the denominator gives:

$$\begin{aligned} \frac{\Delta G(\text{kcal/mol})}{R_kT(\text{kcal/mol})} &= \frac{\Delta G * 4184(\text{J/mol})}{k_BN_AT(\text{J/mol})} \\ &= \frac{\Delta G * 4184/N_A \text{ (Joules)}}{k_BT \text{ (Joules)}} \\ \frac{\Delta G_{new}}{k_BT} &= \frac{\Delta G * 4184/N_A}{k_BT}. \end{aligned}$$

The units of ΔG_{new} are k_BT.

For example, if the free energy is 2 kcal/mol, the dissociation constant is $K_d = \exp(-2/RT)$, see Appendix C. We convert the free energy to k_BT as follows:

$$\begin{aligned}\frac{2(\text{kcal/mol})}{R_kT(\text{kcal/mol})} &= \frac{2 * 4184/N_A}{k_BT} \\ &= \frac{8368(\text{J/mol})}{(1.38 \times 10^{-23}\text{JK}^{-1})(6.022 \times 10^{23}\text{mol}^{-1})(310.15\text{K})} \\ &= 3.247.\end{aligned}$$

We started with $2/RT$ with $K_d = \exp^{-2/RT}$. Now after our conversion, K_d is also equal to $e^{-3.247} = e^{-3.247k_BT/k_BT}$ and the free energy is $\Delta G_{new} = 3.247k_BT$.

Transcription rate function

Thus, to model transcription we define $\mathbb{P}_s$ as the probability of a state s, and k_s as the transcription initiation rate for a state s. The probability of a particular state s is:

$$\mathbb{P}_s := \frac{K_s[RNAP]^{\alpha_s}[CAP]^{\alpha_s^1}[R]^{\alpha_s^2}[R]^{\alpha_s^3}[R'']^{\alpha_s^4}}{\sum_i K_i[RNAP]^{\alpha_i}[CAP]^{\alpha_i^1}[R]^{\alpha_i^2}[R]^{\alpha_i^3}[R'']^{\alpha_i^4}}$$

where $K_s = \exp(-E(s)/RT)$ is the equilibrium constant (see Appendix C), and α_s, α_s^1, α_s^2, and α_s^3 represent the number of bound molecules of RNAP, CAP, R, R', and R'', respectively. Then,

$$f := \sum_s k_s\mathbb{P}_s = \frac{\sum_s K_s[RNAP]^{\alpha_s}[CAP]^{\alpha_s^1}[R]^{\alpha_s^2}[R]^{\alpha_s^3}[R'']^{\alpha_s^4}}{\sum_i K_i[RNAP]^{\alpha_i}[CAP]^{\alpha_i^1}[R]^{\alpha_i^2}[R]^{\alpha_i^3}[R'']^{\alpha_i^4}}$$

where k_s is the transcription initiation rate and the denominator is the sum of probabilities of all possible states Z.

APPENDIX C

EQUILIBRIUM CONSTANT

The equilibrium constant for a state s, K_s, represents the ratio of products and reactants involved in a reaction to create the state s. Consider the chemical reaction:

$$[A] + [B] \underset{k_-}{\overset{k_+}{\rightleftharpoons}} [AB]$$

where k_+ is the forward reaction rate and k_- is the backwards reaction rate. Then, the differential equations which represent this chemical reaction are

$$\dot{[A]} = k_-[AB] - k_+[A][B]$$
$$\dot{[B]} = k_-[AB] - k_+[A][B]$$
$$\dot{[AB]} = k_+[A][B] - k_-[AB].$$

If we assume we are at steady state, the reaction becomes

$$0 = k_-[AB] - k_+[A][B]$$
$$0 = k_-[AB] - k_+[A][B]$$
$$0 = k_+[A][B] - k_-[AB].$$

Thus all three equations provide us with the relation

$$k_-[AB] = k_+[A][B]$$

which can be rewritten as:

$$\frac{k_-}{k_+} = \frac{[A][B]}{[AB]}.$$

The the equilibrium constant is defined as:

$$K_{AB} := \frac{[A][B]}{[AB]}$$

where

$$K_{AB} = \frac{k_-}{k_+}.$$

The equilibrium constant depends on temperature and pressure (or volume), so often the change in Gibbs free energy is more useful. The free energy relates to the equilibrium constant with following equation:

$$K_{AB} = \alpha \exp\left(\frac{-\Delta G_{AB}}{RT}\right)$$

where α is a constant which carries the units, ΔG_{AB} represents the change in free energy in kcal/mol, R is the universal gas constant, and T is temperature. Note that α can always be scaled to 1 by changing the value of ΔG_{AB}.

We can observe the importance of α when changing units. For example, Let $K_{AB} = 10 \times 10^{-6}$M. Then if $\alpha = 1$M, the free energy $\Delta G_{AB} = -RT \ln(10 \times 10^{-6}) \approx -7.1$ kcal/mol. However, say we represent K_{AB} in micromolar so that $K_{AB} = 10\mu$M. If forget to adjust α, the free energy $\Delta G_{AB} = -RT \ln(10) \approx 1.42$ kcal/mol.

Free energy for cooperativity

In equation (1) in the Introduction, $K_B(s) = \exp(-E_s/RT)$ and E_s represents the change in free energy between the empty DNA and the particular state s. When the state includes more than one protein or RNAP bound to the DNA, E_s can be decomposed into the individual free energies of binding. Consider the state which has a single RNAP bound and a single regulatory protein, R, bound. The state energy E_s can be represented as the sum of the free energy of binding RNAP to the DNA, ΔG_P, and the free energy of binding the regulatory protein to the DNA, ΔG_R. An additional term, ΔG_{coop} is included if the regulatory protein directly affects RNAP binding. Thus,

$$E_s = \Delta G_P + \Delta G_R + \Delta G_{coop}$$

where $\Delta G_{coop} < 0$ implies the regulatory protein enhances RNAP binding, $\Delta G_{coop} = 0$ implies no interaction and when the regulatory protein decreases the probability of RNAP binding, $\Delta G_{coop} > 0$.

APPENDIX D

PROBABILITY FUNCTION

The terms of the complete model follow. For space, we have written all free energies as equilibrium constants. For example, we will not write $\exp(-\Delta G_{O1}/RT)$ as K_{O1} and $\exp(-\Delta G^I_{O1}/RT)$ as K_{O1_I} Recall:

$$\mathcal{R} = \frac{\beta\text{-galactosidase(1mM)}}{\beta\text{-galactosidase(0mM)}} = \frac{F(I)|_{(1\text{mM})}}{F(I)|_{(0\text{mM})}} = \frac{(f_1(I) + f_2(I) + f_{1C1}(I))\,|_{(1\text{mM})}}{(f_1(I) + f_2(I) + f_{1C1}(I))\,|_{(0\text{mM})}},$$

and each

$$f_* = \sum_{s\in\mathcal{S}_*} k(s)\mathbb{P}(s) = \sum_{s\in\mathcal{S}_*} k(s)\frac{K_B(s)[RNAP]^{\alpha_s}[CAP]^{\alpha_s^1}[R]^{\alpha_s^2}[R']^{\alpha_s^3}[R'']^{\alpha_s^4}}{Z}$$

where R, R' and R'' are dependent on I.

For low IPTG, the numerator of $F(I)$ is:

$F(0mM)_{num} = [R][RNAP][CAP]K_{O2}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}+$
$[RNAP][CAP]K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}+ [R][RNAP][CAP]K_{O2}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*}K_{23*c}+$
$[R][RNAP]K_{O2}K_{P2}^{-1}+ [RNAP]K_{P2}^{-1}+ [R][RNAP]K_{O2}K_{P1}^{-1}+$
$[RNAP]K_{P1}^{-1}+ [R][RNAP][CAP]K_{CAPO3}K_{O2}K_{O3}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{23c}+$
$[R][RNAP]K_{O2}K_{O3}K_{23}K_{P2}^{-1}+ [R]^2[RNAP][CAP]K_{O2}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*}+$
$[R][RNAP][CAP]K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*}+ [R][RNAP]K_{O2}K_{O3}K_{23}K_{P1}^{-1}+$
$[R]^2[RNAP]K_{O2}K_{O3}K_{P2}^{-1}+ [R][RNAP]K_{O3}K_{P2}^{-1}+ [R]^2[RNAP]K_{O2}K_{O3}K_{P1}^{-1}+$
$[R][RNAP]K_{O3}K_{P1}^{-1}+ [R]^2[RNAP][CAP]K_{CAPO3}K_{O2}K_{O3}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}+$
$[R][RNAP][CAP]K_{CAPO3}K_{O3}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}+ [R]^2[RNAP]K_{O2}K_{P2}^{-1}K_{O3*}+$
$[R][RNAP]K_{O2}K_{P2}^{-1}K_{O3*}K_{23*}+ [R][RNAP]K_{P2}^{-1}K_{O3*}+ [R]^2[RNAP]K_{O2}K_{P1}^{-1}K_{O3*}+$
$[R][RNAP]K_{O2}K_{P1}^{-1}K_{O3*}K_{23*}+ [R][RNAP]K_{P1}^{-1}K_{O3*}.$

For low IPTG, the denominator of $F(I)$ is:

$F(0mM)_{den} = [R][RNAP][CAP]K_{O1}K_{O2}K_{12}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}+$
$[R]^2[RNAP][CAP]K_{O1}K_{O2}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*}K_{13*c}+$
$[R][RNAP][CAP]K_{O1}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*}K_{13*c}+ [R][CAP]K_{O1}K_{O2}K_{12}K_{CAP}^{-1}+$

$$[R]^2[CAP]K_{O1}K_{O2}K_{CAP}^{-1}K_{O3*}K_{13*c}+ [R][CAP]K_{O1}K_{CAP}^{-1}K_{O3*}K_{13*c}+$$
$$[R][RNAP]K_{O1}K_{O2}K_{12}K_{P2}^{-1}+ [R]^2[RNAP][CAP]K_{O1}K_{O2}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}+$$
$$[R][RNAP][CAP]K_{O1}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}+ [R][RNAP]K_{O1}K_{O2}K_{12}K_{P1}^{-1}+$$
$$[R]K_{O1}K_{O2}K_{12}+ [R]^2[RNAP][CAP]K_{CAPO3}K_{O1}K_{O2}K_{O3}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{13c}+$$
$$[R][RNAP][CAP]K_{CAPO3}K_{O1}K_{O3}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{13c}+$$
$$[R]^2[RNAP][CAP]K_{O1}K_{O2}K_{12}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*}+$$
$$[R][RNAP][CAP]K_{O2}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}+$$
$$[R]^2[RNAP][CAP]K_{O1}K_{O2}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*}K_{23*c}+$$
$$[RNAP][CAP]K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}+ [R]^2[RNAP]K_{O1}K_{O2}K_{O3}K_{12}K_{P2}^{-1}+$$
$$[R]^2[CAP]K_{O1}K_{O2}K_{CAP}^{-1}+ [R][CAP]K_{O1}K_{CAP}^{-1}+ [R]^2[RNAP]K_{O1}K_{O2}K_{O3}K_{13}K_{P2}^{-1}+$$
$$[R][RNAP]K_{O1}K_{O3}K_{13}K_{P2}^{-1}+ [R]^2[RNAP]K_{O1}K_{O2}K_{O3}K_{12}K_{P1}^{-1}+$$
$$[R]^2[RNAP]K_{O1}K_{O2}K_{O3}K_{13}K_{P1}^{-1}+ [R][RNAP]K_{O1}K_{O3}K_{13}K_{P1}^{-1}+$$
$$[R]^2[CAP]K_{CAPO3}K_{O1}K_{O2}K_{O3}K_{CAP}^{-1}K_{13c}+ [R][CAP]K_{CAPO3}K_{O1}K_{O3}K_{CAP}^{-1}K_{13c}+$$
$$[R]^2[RNAP]K_{O1}K_{O2}K_{P2}^{-1}+ [R][RNAP][CAP]K_{O2}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*}K_{23*c}+$$
$$[R]^2[CAP]K_{O1}K_{O2}K_{12}K_{CAP}^{-1}K_{O3*}+ [R][RNAP]K_{O1}K_{P2}^{-1}+ [R]^2K_{O1}K_{O2}K_{O3}K_{12}+$$
$$[R][CAP]K_{O2}K_{CAP}^{-1}+ [R]^2[CAP]K_{O1}K_{O2}K_{CAP}^{-1}K_{O3*}K_{23*c}+$$
$$[CAP]K_{CAP}^{-1}+ [R]^2K_{O1}K_{O2}K_{O3}K_{13}+ [R]K_{O1}K_{O3}K_{13}+$$
$$[R]^2[RNAP]K_{O1}K_{O2}K_{P1}^{-1}+ [R][RNAP]K_{O1}K_{P1}^{-1}+$$
$$[R]^2[RNAP][CAP]K_{CAPO3}K_{O1}K_{O2}K_{O3}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{23c}+$$
$$[R]^2[RNAP]K_{O1}K_{O2}K_{P2}^{-1}K_{O3*}K_{13*}+ [R][RNAP]K_{O1}K_{P2}^{-1}K_{O3*}K_{13*}+$$
$$[R]^2[RNAP][CAP]K_{CAPO3}K_{O1}K_{O2}K_{O3}K_{12}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}+$$
$$[R]^2[RNAP]K_{O1}K_{O2}K_{P1}^{-1}K_{O3*}K_{13*}+ [R]^2[RNAP]K_{O1}K_{O2}K_{12}K_{P2}^{-1}K_{O3*}+$$
$$[R][RNAP]K_{O1}K_{P1}^{-1}K_{O3*}K_{13*}+ [R]^2K_{O1}K_{O2}+ [R][RNAP]K_{O2}K_{P2}^{-1}+$$
$$[R]^2[RNAP]K_{O1}K_{O2}K_{O3}K_{23}K_{P2}^{-1}+ [R]K_{O1}+ [RNAP]K_{P2}^{-1}+$$
$$[RNAP][CAP]K_{O1}K_{O2}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*}+$$
$$[R]^2[RNAP][CAP]K_{O1}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*}+$$

$$[R]^2[RNAP]K_{O1}K_{O2}K_{12}K_{P1}^{-1}K_{O3*}+ [R][CAP]K_{O2}K_{CAP}^{-1}K_{O3*}K_{23*c}+$$
$$[R][RNAP]K_{O2}K_{P1}^{-1}+ [R]^2[RNAP]K_{O1}K_{O2}K_{O3}K_{23}K_{P1}^{-1}+ [RNAP]K_{P1}^{-1}+$$
$$[RNAP]K_{O1}K_{O2}K_{O3}K_{P2}^{-1}+ [R][RNAP][CAP]K_{CAPO3}K_{O2}K_{O3}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{23c}+$$
$$[R]^2K_{O1}K_{O2}K_{O3*}K_{13*}+ [R]^2[RNAP]K_{O1}K_{O3}K_{P2}^{-1}+ [R]K_{O1}K_{O3*}K_{13*}+$$
$$[R]^2[CAP]K_{CAPO3}K_{O1}K_{O2}K_{O3}K_{CAP}^{-1}K_{23c}+ [RNAP]K_{O1}K_{O2}K_{O3}K_{P1}^{-1}+$$
$$[R]^2[RNAP]K_{O1}K_{O3}K_{P1}^{-1}+ [R]^2K_{O1}K_{O2}K_{12}K_{O3*}+$$
$$[R]^2[CAP]K_{CAPO3}K_{O1}K_{O2}K_{O3}K_{12}K_{CAP}^{-1}+ [R]K_{O2}+ [R]^2K_{O1}K_{O2}K_{O3}K_{23}+$$
$$[R][RNAP]K_{O2}K_{O3}K_{23}K_{P2}^{-1}+ 1+ [R]^2[RNAP][CAP]K_{O2}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*}+$$
$$[R][RNAP][CAP]K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*}+ [CAP]K_{O1}K_{O2}K_{CAP}^{-1}K_{O3*}+$$
$$[R][RNAP]K_{O2}K_{O3}K_{23}K_{P1}^{-1}+ [R]^2[CAP]K_{O1}K_{CAP}^{-1}K_{O3*}+ K_{O1}K_{O2}K_{O3}+$$
$$[R]^2[RNAP]K_{O2}K_{O3}K_{P2}^{-1}+ [R]^2K_{O1}K_{O3}+ [R][RNAP]K_{O3}K_{P2}^{-1}+$$
$$[R][CAP]K_{CAPO3}K_{O2}K_{O3}K_{CAP}^{-1}K_{23c}+ [R]^2[RNAP]K_{O2}K_{O3}K_{P1}^{-1}+$$
$$[R][RNAP]K_{O3}K_{P1}^{-1}+ [RNAP][CAP]K_{CAPO3}K_{O1}K_{O2}K_{O3}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}+$$
$$[R]K_{O2}K_{O3}K_{23}+ [R]^2[RNAP][CAP]K_{CAPO3}K_{O1}K_{O3}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}+$$
$$[RNAP]K_{O1}K_{O2}K_{P2}^{-1}K_{O3*}+ [R]^2[RNAP]K_{O1}K_{O2}K_{P2}^{-1}K_{O3*}K_{23*}+$$
$$[R]^2[RNAP]K_{O1}K_{P2}^{-1}K_{O3*}+ [R]^2[CAP]K_{O2}K_{CAP}^{-1}K_{O3*}+ [R][CAP]K_{CAP}^{-1}K_{O3*}+$$
$$[R]^2K_{O2}K_{O3}+ [R]K_{O3}+ [RNAP]K_{O1}K_{O2}K_{P1}^{-1}K_{O3*}+$$
$$[R]^2[RNAP]K_{O1}K_{O2}K_{P1}^{-1}K_{O3*}K_{23*}+ [R]^2[RNAP]K_{O1}K_{P1}^{-1}K_{O3*}+$$
$$[R]^2[RNAP][CAP]K_{CAPO3}K_{O2}K_{O3}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}+$$
$$[R][RNAP][CAP]K_{CAPO3}K_{O3}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}+ K_{O1}K_{O2}K_{O3*}+$$
$$[R]^2K_{O1}K_{O2}K_{O3*}K_{23*}+ [R]^2[RNAP]K_{O2}K_{P2}^{-1}K_{O3*}+ [R][RNAP]K_{O2}K_{P2}^{-1}K_{O3*}K_{23*}+$$
$$[R]^2K_{O1}K_{O3*}+ [R][RNAP]K_{P2}^{-1}K_{O3*}+ [CAP]K_{CAPO3}K_{O1}K_{O2}K_{O3}K_{CAP}^{-1}+$$
$$[R]^2[CAP]K_{CAPO3}K_{O1}K_{O3}K_{CAP}^{-1}+ [R]^2[RNAP]K_{O2}K_{P1}^{-1}K_{O3*}+$$
$$[R][RNAP]K_{O2}K_{P1}^{-1}K_{O3*}K_{23*}+ [R][RNAP]K_{P1}^{-1}K_{O3*}+ [R]^2K_{O2}K_{O3*}+$$
$$[R]K_{O2}K_{O3*}K_{23*}+ [R]K_{O3*}+ [R]^2[CAP]K_{CAPO3}K_{O2}K_{O3}K_{CAP}^{-1}+$$
$$[R][CAP]K_{CAPO3}K_{O3}K_{CAP}^{-1}.$$

For high IPTG, the numerator of $F(I)$ is:

$$F(1mM)_{num} = [RNAP][CAP]K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1} + [RNAP]K_{P2}^{-1} + [RNAP]K_{P1}^{-1} +$$
$$[R'][RNAP][CAP]K_{O2_I}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1} + [R'][RNAP][CAP]K_{O2}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1} +$$
$$[R'][RNAP][CAP]K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*_I} + [R'][RNAP]K_{O2_I}K_{P2}^{-1} +$$
$$[R'][RNAP]K_{O2_I}K_{P1}^{-1} + [R'][RNAP][CAP]K_{O2}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*}K_{23*c} +$$
$$[R'][RNAP]K_{O2}K_{P2}^{-1} + [R'][RNAP]K_{P2}^{-1}K_{O3*_I} + [R'][RNAP]K_{O2}K_{P1}^{-1} +$$
$$[R'][RNAP]K_{P1}^{-1}K_{O3*_I} + [R'][RNAP]K_{O3_I}K_{P2}^{-1} +$$
$$[R'][RNAP][CAP]K_{O2}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*}K_{23*c} + [R'][RNAP]K_{O3_I}K_{P1}^{-1} +$$
$$[R'][RNAP][CAP]K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*} + [R'][RNAP]K_{O3}K_{P2}^{-1} +$$
$$[R']^2[RNAP][CAP]K_{O2_I}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*_I} +$$
$$[R][RNAP][CAP]K_{O2}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1} + [R'][RNAP]K_{O3}K_{P1}^{-1} +$$
$$[R'][RNAP][CAP]K_{CAPO3}K_{O3_I}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1} +$$
$$[R'][RNAP][CAP]K_{O2}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*}K_{23*c} +$$
$$[R][RNAP][CAP]K_{O2}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*}K_{23*c} +$$
$$[R'][R'][RNAP][CAP]K_{O2}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*_I} +$$
$$[R'][RNAP][CAP]K_{CAPO3}K_{O3}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1} + [R'][RNAP]K_{P2}^{-1}K_{O3*} +$$
$$[R'][RNAP][CAP]K_{CAPO3}K_{O2}K_{O3}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{23c} + [R'][RNAP]K_{P1}^{-1}K_{O3*} +$$
$$[R']^2[RNAP]K_{O2_I}K_{P2}^{-1}K_{O3*_I} + [R'][RNAP]K_{O2}K_{P2}^{-1}K_{O3*}K_{23*} +$$
$$[R'][RNAP]K_{23}K_{O2}K_{O3}K_{P2}^{-1} + [R][RNAP]K_{O2}K_{P2}^{-1} +$$
$$[R'][RNAP][CAP]K_{CAPO3}K_{O3}K_{O2}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{23c} +$$
$$[R']^2[RNAP]K_{O2_I}K_{P1}^{-1}K_{O3*_I} + [R'][RNAP]K_{O2}K_{P1}^{-1}K_{O3*}K_{23*} +$$
$$[R'][RNAP]K_{23}K_{O2}K_{O3}K_{P1}^{-1} + [R][RNAP]K_{O2}K_{P1}^{-1} + [R']^2[RNAP]K_{O2_I}K_{O3_I}K_{P2}^{-1} +$$
$$[R][RNAP][CAP]K_{CAPO3}K_{O2}K_{O3}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{23c} +$$
$$[R'][RNAP][CAP]K_{CAPO3}K_{O2}K_{O3}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{23c} +$$
$$[R']^2[RNAP]K_{O2_I}K_{O3_I}K_{P1}^{-1} + [R'][RNAP]K_{O3}K_{23}K_{O2}K_{P2}^{-1} +$$
$$[R'][R'][RNAP][CAP]K_{O2_I}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*} + [R][RNAP]K_{O2}K_{O3}K_{23}K_{P2}^{-1} +$$

$$[R'][R''][RNAP]K_{O2}K_{P2}^{-1}K_{O3*_I}+ [R'][RNAP]K_{O2}K_{P2}^{-1}K_{O3*}K_{23*}+$$
$$[R'][RNAP]K_{O2}K_{23}K_{O3}K_{P2}^{-1}+ [R][RNAP][CAP]K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*}+$$
$$[R'][RNAP]K_{O3}K_{23}K_{O2}K_{P1}^{-1}+ [R'][R''][RNAP]K_{O3}K_{O2_I}K_{P2}^{-1}+$$
$$[R][RNAP]K_{O2}K_{O3}K_{23}K_{P1}^{-1}+ [R'][R''][RNAP]K_{O2}K_{P1}^{-1}K_{O3*_I}+$$
$$[R'][RNAP]K_{O2}K_{P1}^{-1}K_{O3*}K_{23*}+ [R'][RNAP]K_{O2}K_{23}K_{O3}K_{P1}^{-1}+$$
$$[R][RNAP]K_{O3}K_{P2}^{-1}+ [R'][R''][RNAP]K_{O2}K_{O3_I}K_{P2}^{-1}+$$
$$[R'][R''][RNAP]K_{O3}K_{O2_I}K_{P1}^{-1}+ [R][RNAP]K_{O3}K_{P1}^{-1}+$$
$$[R'][R''][RNAP]K_{O2}K_{O3_I}K_{P1}^{-1}+$$
$$[R'']^2[RNAP][CAP]K_{CAPO3}K_{O2_I}K_{O3_I}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}+$$
$$[R']^2[RNAP][CAP]K_{O2}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*}+ [R']^2[RNAP]K_{O2}K_{O3}K_{P2}^{-1}+$$
$$[R']^2[RNAP]K_{O2}K_{O3}K_{P1}^{-1}+$$
$$[R'][R''][RNAP][CAP]K_{CAPO3}K_{O3}K_{O2_I}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}+$$
$$[R][R''][RNAP][CAP]K_{O2}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*_I}+ [R'][R''][RNAP]K_{O2_I}K_{P2}^{-1}K_{O3*}+$$
$$[R'][RNAP]K_{O2}K_{P2}^{-1}K_{O3*}K_{23*}+ [R][RNAP][CAP]K_{CAPO3}K_{O3}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}+$$
$$[R'][R''][RNAP][CAP]K_{CAPO3}K_{O2}K_{O3_I}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}+$$
$$[R][RNAP]K_{O2}K_{P2}^{-1}K_{O3*}K_{23*}+ [R][RNAP]K_{P2}^{-1}K_{O3*}+$$
$$[R'][R''][RNAP]K_{O2_I}K_{P1}^{-1}K_{O3*}+ [R'][RNAP]K_{O2}K_{P1}^{-1}K_{O3*}K_{23*}+$$
$$[R][RNAP]K_{O2}K_{P1}^{-1}K_{O3*}K_{23*}+ [R][RNAP]K_{P1}^{-1}K_{O3*}+$$
$$[R']^2[RNAP][CAP]K_{CAPO3}K_{O2}K_{O3}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}+ [R']^2[RNAP]K_{O2}K_{P2}^{-1}K_{O3*}+$$
$$[R']^2[RNAP]K_{O2}K_{P1}^{-1}K_{O3*}+ [R][R''][RNAP][CAP]K_{O2_I}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*}+$$
$$[R][R''][RNAP]K_{O2}K_{P2}^{-1}K_{O3*_I}+ [R][R''][RNAP]K_{O3}K_{O2_I}K_{P2}^{-1}+$$
$$[R][R''][RNAP]K_{O2}K_{P1}^{-1}K_{O3*_I}+ [R][R''][RNAP]K_{O2}K_{O3_I}K_{P2}^{-1}+$$
$$[R][R''][RNAP]K_{O3}K_{O2_I}K_{P1}^{-1}+ [R][R''][RNAP]K_{O2}K_{O3_I}K_{P1}^{-1}+$$
$$[R][R'][RNAP][CAP]K_{O2}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*}+$$
$$[R][R'][RNAP][CAP]K_{O2}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*}+ [R][R'][RNAP]K_{O3}K_{O2}K_{P2}^{-1}+$$
$$[R][R'][RNAP]K_{O2}K_{O3}K_{P2}^{-1}+ [R][R'][RNAP]K_{O3}K_{O2}K_{P1}^{-1}+$$

$$[R][R'][RNAP]K_{O2}K_{O3}K_{P1}^{-1}+$$
$$[R][R''][RNAP][CAP]K_{CAPO3}K_{O3}K_{O2_I}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}+$$
$$[R][R''][RNAP]K_{O2_I}K_{P2}^{-1}K_{O3*}+$$
$$[R][R''][RNAP][CAP]K_{CAPO3}K_{O2}K_{O3_I}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}+$$
$$[R][R''][RNAP]K_{O2_I}K_{P1}^{-1}K_{O3*}+$$
$$[R][R'][RNAP][CAP]K_{CAPO3}K_{O3}K_{O2}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}+$$
$$[R][R'][RNAP][CAP]K_{CAPO3}K_{O2}K_{O3}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}+$$
$$[R][R'][RNAP]K_{O2}K_{P2}^{-1}K_{O3*}+\ [R][R'][RNAP]K_{O2}K_{P2}^{-1}K_{O3*}+$$
$$[R][R'][RNAP]K_{O2}K_{P1}^{-1}K_{O3*}+\ [R][R'][RNAP]K_{O2}K_{P1}^{-1}K_{O3*}+$$
$$[R]^2[RNAP][CAP]K_{O2}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*}+\ [R]^2[RNAP]K_{O2}K_{O3}K_{P2}^{-1}+$$
$$[R]^2[RNAP]K_{O2}K_{O3}K_{P1}^{-1}+\ [R]^2[RNAP][CAP]K_{CAPO3}K_{O2}K_{O3}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}+$$
$$[R]^2[RNAP]K_{O2}K_{P2}^{-1}K_{O3*}+\ [R]^2[RNAP]K_{O2}K_{P1}^{-1}K_{O3*}.$$

For high IPTG, the denominator of $F(I)$ is:

$$F(1mM)_{den} = [RNAP][CAP]K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}+\ [CAP]K_{CAP}^{-1}+$$
$$[RNAP]K_{P2}^{-1}+\ [RNAP]K_{P1}^{-1}+\ 1+\ [R''][RNAP][CAP]K_{O2_I}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}+$$
$$[R'][RNAP][CAP]K_{O1}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}+$$
$$[R'][RNAP][CAP]K_{O1}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*}K_{13*c}+$$
$$[R''][RNAP][CAP]K_{O1_I}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}+$$
$$[R''][RNAP][CAP]K_{O1}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*}K_{13*c}+$$
$$[R'][RNAP][CAP]K_{O2}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}+\ [R''][CAP]K_{O2_I}K_{CAP}^{-1}+$$
$$[R''][RNAP][CAP]K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*_I}+\ [R'][CAP]K_{O1}K_{CAP}^{-1}+$$
$$[R'][CAP]K_{O1}K_{CAP}^{-1}K_{O3*}K_{13*c}+\ [R'][RNAP][CAP]K_{O1}K_{12}K_{O2}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}+$$
$$[R][RNAP][CAP]K_{O1}K_{O2}K_{12}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}+\ [R''][CAP]K_{O1_I}K_{CAP}^{-1}+$$
$$[R''][CAP]K_{O1}K_{CAP}^{-1}K_{O3*}K_{13*c}+\ [R''][RNAP][CAP]K_{12}K_{O1}K_{O2}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}+$$

$$[R''][RNAP]K_{O2_I}K_{P2}^{-1}+ [R'][RNAP]K_{O1}K_{P2}^{-1}+$$
$$[R][RNAP][CAP]K_{O1}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*}K_{13*c}+ [R'][CAP]K_{O2}K_{CAP}^{-1}+$$
$$[R''][RNAP]K_{O2_I}K_{P1}^{-1}+ [R''][CAP]K_{CAP}^{-1}K_{O3*_I}+ [R'][RNAP]K_{O1}K_{P1}^{-1}+$$
$$[R''][RNAP]K_{O1_I}K_{P2}^{-1}+ [R'][CAP]K_{O1}K_{12}K_{O2}K_{CAP}^{-1}+$$
$$[R''][RNAP][CAP]K_{O2}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*}K_{23*c}+$$
$$[R][CAP]K_{O1}K_{O2}K_{12}K_{CAP}^{-1}+ [R''][RNAP]K_{O1_I}K_{P1}^{-1}+$$
$$[R'][RNAP][CAP]K_{O2}K_{12}K_{O1}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}+ [R'']K_{O2_I}+$$
$$[R''][CAP]K_{12}K_{O1}K_{O2}K_{CAP}^{-1}+ [R'][RNAP]K_{O2}K_{P2}^{-1}+ [R']K_{O1}+$$
$$[R''][RNAP]K_{P2}^{-1}K_{O3*_I}+ [R'][RNAP]K_{O2}K_{P1}^{-1}+ [R'][RNAP]K_{O1}K_{12}K_{O2}K_{P2}^{-1}+$$
$$[R][CAP]K_{O1}K_{CAP}^{-1}K_{O3*}K_{13*c}+ [R'']K_{O1_I}+$$
$$[R'][RNAP][CAP]K_{O1}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*}K_{13*c}+ [R''][RNAP]K_{P1}^{-1}K_{O3*_I}+$$
$$[R][RNAP]K_{O1}K_{O2}K_{12}K_{P2}^{-1}+ [R''][RNAP]K_{O3_I}K_{P2}^{-1}+$$
$$[R'][R''][RNAP][CAP]K_{O1}K_{O2_I}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}+$$
$$[R'][R''][RNAP][CAP]K_{O1}K_{O2_I}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*}K_{13*c}+$$
$$[R'][RNAP]K_{O1}K_{12}K_{O2}K_{P1}^{-1}+ [R'][RNAP][CAP]K_{O2}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*}K_{23*c}+$$
$$[R''][RNAP]K_{12}K_{O1}K_{O2}K_{P2}^{-1}+ [R][RNAP][CAP]K_{O1}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}+$$
$$[R''][CAP]K_{O2}K_{CAP}^{-1}K_{O3*}K_{23*c}+ [R][RNAP]K_{O1}K_{O2}K_{12}K_{P1}^{-1}+$$
$$[R''][RNAP]K_{O3_I}K_{P1}^{-1}+ [R']K_{O2}+ [R'][CAP]K_{O2}K_{12}K_{O1}K_{CAP}^{-1}+$$
$$[R'']^2[RNAP][CAP]K_{O1_I}K_{O2_I}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}+$$
$$[R'']^2[RNAP][CAP]K_{O1}K_{O2_I}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*}K_{13*c}+$$
$$[R''][RNAP]K_{12}K_{O1}K_{O2}K_{P1}^{-1}+ [R'']K_{O3*_I}+ [R'][RNAP][CAP]K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*}+$$
$$[R'][RNAP]K_{O1}K_{P2}^{-1}K_{O3*}K_{13*}+ [R']K_{O1}K_{12}K_{O2}+ [R'][RNAP]K_{O3}K_{P2}^{-1}+$$
$$[R]K_{O1}K_{O2}K_{12}+ [R'][RNAP]K_{O1}K_{P1}^{-1}K_{O3*}K_{13*}+$$
$$[R][RNAP][CAP]K_{CAPO3}K_{O1}K_{O3}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{13c}+$$
$$[R'']K_{O3_I}+ [R''][RNAP]K_{O1}K_{P2}^{-1}K_{O3*}K_{13*}+$$
$$[R'][RNAP][CAP]K_{CAPO3}K_{O1}K_{O3}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{13c}+$$

$$[R'][CAP]K_{O1}K_{CAP}^{-1}K_{O3*}K_{13*c}+ [R']^2[RNAP][CAP]K_{O1}K_{O2}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}+$$
$$[R'']K_{12}K_{O1}K_{O2}+ [R']^2[RNAP][CAP]K_{O1}K_{O2}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*}K_{13*c}+$$
$$[R'']^2[RNAP][CAP]K_{O2_I}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*_I}+$$
$$[R][RNAP][CAP]K_{O2}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}+ [R'][RNAP]K_{O3}K_{P1}^{-1}+$$
$$[R'][RNAP]K_{O2}K_{12}K_{O1}K_{P2}^{-1}+ [R'][R''][CAP]K_{O1}K_{O2_I}K_{CAP}^{-1}+$$
$$[R'][R''][RNAP][CAP]K_{O1}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*_I}+$$
$$[R'][R''][CAP]K_{O1}K_{O2_I}K_{CAP}^{-1}K_{O3*}K_{13*c}+ [R'][CAP]K_{O2}K_{CAP}^{-1}K_{O3*}K_{23*c}+$$
$$[R''][RNAP]K_{O1}K_{P1}^{-1}K_{O3*}K_{13*}+$$
$$[R''][RNAP][CAP]K_{CAPO3}K_{O1}K_{O3}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{13c}+ [R][CAP]K_{O1}K_{CAP}^{-1}+$$
$$[R'][R''][RNAP][CAP]K_{O2}K_{O1_I}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}+$$
$$[R''][RNAP][CAP]K_{CAPO3}K_{O3_I}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}+ [R][RNAP]K_{O1}K_{O3}K_{13}K_{P2}^{-1}+$$
$$[R'][R''][RNAP][CAP]K_{O1}K_{O2}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*}K_{13*c}+$$
$$[R'][RNAP]K_{O2}K_{12}K_{O1}K_{P1}^{-1}+ [R']K_{O1}K_{O3*}K_{13*}+ [R'][RNAP]K_{O1}K_{13}K_{O3}K_{P2}^{-1}+$$
$$[R'']^2[CAP]K_{O1_I}K_{O2_I}K_{CAP}^{-1}+ [R'']^2[RNAP][CAP]K_{O1_I}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*_I}+$$
$$[R'']^2[CAP]K_{O1}K_{O2_I}K_{CAP}^{-1}K_{O3*}K_{13*c}+ [R'][CAP]K_{CAP}^{-1}K_{O3*}+ [R']K_{O3}+$$
$$[R][RNAP]K_{O1}K_{O3}K_{13}K_{P1}^{-1}+ [R'][RNAP]K_{O1}K_{13}K_{O3}K_{P1}^{-1}+$$
$$[R'']K_{O1}K_{O3*}K_{13*}+ [R''][RNAP]K_{13}K_{O1}K_{O3}K_{P2}^{-1}+ [R'][R''][RNAP]K_{O1}K_{O2_I}K_{P2}^{-1}+$$
$$[R'][RNAP][CAP]K_{O2}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*}K_{23*c}+$$
$$[R][CAP]K_{CAPO3}K_{O1}K_{O3}K_{CAP}^{-1}K_{13c}+$$
$$[R][R''][RNAP][CAP]K_{O1}K_{O2_I}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*}K_{13*c}+$$
$$[R][RNAP][CAP]K_{O2}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*}K_{23*c}+$$
$$[R'][CAP]K_{CAPO3}K_{O1}K_{O3}K_{CAP}^{-1}K_{13c}+ [R']K_{O2}K_{12}K_{O1}+$$
$$[R'][R''][RNAP][CAP]K_{O2}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*_I}+ [R][RNAP]K_{O1}K_{P2}^{-1}+$$
$$[R'][RNAP][CAP]K_{CAPO3}K_{O3}K_{O1}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{13c}+$$
$$[R']^2[CAP]K_{O1}K_{O2}K_{CAP}^{-1}+ [R''][RNAP]K_{13}K_{O1}K_{O3}K_{P1}^{-1}+$$
$$[R']^2[CAP]K_{O1}K_{O2}K_{CAP}^{-1}K_{O3*}K_{13*c}+ [R'']^2[CAP]K_{O2_I}K_{CAP}^{-1}K_{O3*_I}+$$

$$[R'][R''][RNAP]K_{O1}K_{O2_I}K_{P1}^{-1}+ [R][CAP]K_{O2}K_{CAP}^{-1}+$$
$$[R'][RNAP][CAP]K_{CAPO3}K_{O3}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}+ [R'']^2[RNAP]K_{O1_I}K_{O2_I}K_{P2}^{-1}+$$
$$[R'][R''][CAP]K_{O1}K_{CAP}^{-1}K_{O3*_I}+ [R]K_{O1}K_{O3}K_{13}+ [R'][RNAP]K_{P2}^{-1}K_{O3*}+$$
$$[R']K_{O1}K_{13}K_{O3}+ [R][RNAP]K_{O1}K_{P1}^{-1}+$$
$$[R'][R''][RNAP][CAP]K_{O1}K_{12}K_{O2}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*_I}+$$
$$[R''][RNAP][CAP]K_{CAPO3}K_{O2}K_{O3}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{23c}+$$
$$[R''][CAP]K_{CAPO3}K_{O1}K_{O3}K_{CAP}^{-1}K_{13c}+ [R][RNAP]K_{O1}K_{P2}^{-1}K_{O3*}K_{13*}+$$
$$[R'][R''][CAP]K_{O2}K_{O1_I}K_{CAP}^{-1}+ [R''][CAP]K_{CAPO3}K_{O3_I}K_{CAP}^{-1}+$$
$$[R][R''][RNAP][CAP]K_{O1}K_{O2}K_{12}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*_I}+$$
$$[R'][R''][CAP]K_{O1}K_{O2}K_{CAP}^{-1}K_{O3*}K_{13*c}+$$
$$[R'][R''][RNAP][CAP]K_{O1}K_{O2}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*}K_{23*c}+$$
$$[R'']^2[RNAP]K_{O1_I}K_{O2_I}K_{P1}^{-1}+ [R'][RNAP]K_{O3}K_{13}K_{O1}K_{P2}^{-1}+$$
$$[R'']^2[CAP]K_{O1_I}K_{CAP}^{-1}K_{O3*_I}+ [R'][RNAP]K_{P1}^{-1}K_{O3*}+$$
$$[R'']K_{13}K_{O1}K_{O3}+ [R'']^2[RNAP][CAP]K_{12}K_{O1}K_{O2}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*_I}+$$
$$[R'][R'']K_{O1}K_{O2_I}+ [R][RNAP]K_{O1}K_{P1}^{-1}K_{O3*}K_{13*}+ [R']^2[RNAP]K_{O1}K_{O2}K_{P2}^{-1}+$$
$$[R'']^2[RNAP]K_{O2_I}K_{P2}^{-1}K_{O3*_I}+$$
$$[R'']^2[RNAP][CAP]K_{O2}K_{O1_I}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*}K_{23*c}+$$
$$[R''][RNAP]K_{O2}K_{P2}^{-1}K_{O3*}K_{23*}+ [R''][RNAP]K_{23}K_{O2}K_{O3}K_{P2}^{-1}+$$
$$[R][RNAP]K_{O2}K_{P2}^{-1}+ [R'][RNAP]K_{O3}K_{13}K_{O1}K_{P1}^{-1}+ [R]K_{O1}+$$
$$[R][R'][RNAP][CAP]K_{O1}K_{O2}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*}K_{13*c}+$$
$$[R'][CAP]K_{O2}K_{CAP}^{-1}K_{O3*}K_{23*c}+ [R'][R''][RNAP]K_{O1}K_{P2}^{-1}K_{O3*_I}+$$
$$[R][R''][CAP]K_{O1}K_{O2_I}K_{CAP}^{-1}K_{O3*}K_{13*c}+ [R][CAP]K_{O2}K_{CAP}^{-1}K_{O3*}K_{23*c}+$$
$$[R'']^2K_{O1_I}K_{O2_I}+ [R'][R''][CAP]K_{O2}K_{CAP}^{-1}K_{O3*_I}+ [R']^2[RNAP]K_{O1}K_{O2}K_{P1}^{-1}+$$
$$[R'][RNAP][CAP]K_{CAPO3}K_{O3}K_{O2}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{23c}+$$
$$[R'][CAP]K_{CAPO3}K_{O3}K_{O1}K_{CAP}^{-1}K_{13c}+$$
$$[R'][R''][RNAP][CAP]K_{O1}K_{O2_I}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*}K_{13*c}+$$

$$[R'']^2[RNAP]K_{O2_I}K_{P1}^{-1}K_{O3*_I}+ [R''][RNAP]K_{O2}K_{P1}^{-1}K_{O3*}K_{23*}+$$
$$[R'][R''][RNAP]K_{O2}K_{O1_I}K_{P2}^{-1}+ [R''][RNAP]K_{23}K_{O2}K_{O3}K_{P1}^{-1}+ [R']K_{O3*}+$$
$$[R][RNAP]K_{O2}K_{P1}^{-1}+ [R'']^2[RNAP]K_{O2_I}K_{O3_I}K_{P2}^{-1}+$$
$$[R'][CAP]K_{CAPO3}K_{O3}K_{CAP}^{-1}+ [R'][R''][RNAP]K_{O1}K_{P1}^{-1}K_{O3*_I}+$$
$$[R][RNAP][CAP]K_{CAPO3}K_{O2}K_{O3}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{23c}+$$
$$[R'][RNAP][CAP]K_{CAPO3}K_{O2}K_{O3}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{23c}+$$
$$[R'']^2[RNAP]K_{O1_I}K_{P2}^{-1}K_{O3*_I}+ [R]K_{O1}K_{O3*}K_{13*}+ [R'][R''][RNAP]K_{O1}K_{O3_I}K_{P2}^{-1}+$$
$$[R'][R''][CAP]K_{O1}K_{12}K_{O2}K_{CAP}^{-1}K_{O3*_I}+ [R''][CAP]K_{CAPO3}K_{O2}K_{O3}K_{CAP}^{-1}K_{23c}+$$
$$[R']^2[RNAP][CAP]K_{O2}K_{O1}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*}K_{23*c}+$$
$$[R][R''][RNAP][CAP]K_{O1}K_{O2_I}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}+ [R']K_{O3}K_{13}K_{O1}+$$
$$[R'][R''][RNAP]K_{O2}K_{O1_I}K_{P1}^{-1}+ [R'']^2[RNAP]K_{O2_I}K_{O3_I}K_{P1}^{-1}+$$
$$[R][R''][CAP]K_{O1}K_{O2}K_{12}K_{CAP}^{-1}K_{O3*_I}+ [R'][R''][CAP]K_{O1}K_{O2}K_{CAP}^{-1}K_{O3*}K_{23*c}+$$
$$[R'][RNAP]K_{O1}K_{P2}^{-1}K_{O3*}K_{13*}+ [R'']^2[RNAP]K_{O1_I}K_{P1}^{-1}K_{O3*_I}+$$
$$[R'][RNAP]K_{O3}K_{23}K_{O2}K_{P2}^{-1}+ [R'][R''][RNAP]K_{O1}K_{O3_I}K_{P1}^{-1}+$$
$$[R'][R''][RNAP][CAP]K_{O2}K_{12}K_{O1}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*_I}+ [R']^2K_{O1}K_{O2}+$$
$$[R'']^2[RNAP]K_{O1_I}K_{O3_I}K_{P2}^{-1}+ [R'']^2K_{O2_I}K_{O3*_I}+ [R'']K_{O2}K_{O3*}K_{23*}+$$
$$[R'']K_{23}K_{O2}K_{O3}+ [R'][R''][RNAP][CAP]K_{O2_I}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*}+$$
$$[R]K_{O2}+ [R'']^2[CAP]K_{12}K_{O1}K_{O2}K_{CAP}^{-1}K_{O3*_I}+$$
$$[R'][R''][RNAP]K_{O1}K_{O2_I}K_{P2}^{-1}K_{O3*}K_{13*}+ [R][RNAP]K_{O2}K_{O3}K_{23}K_{P2}^{-1}+$$
$$[R'][R''][RNAP]K_{O2}K_{P2}^{-1}K_{O3*_I}+$$
$$[R'][R''][RNAP][CAP]K_{O2}K_{O1_I}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*}K_{23*c}+$$
$$[R'][RNAP]K_{O2}K_{P2}^{-1}K_{O3*}K_{23*}+ [R'][RNAP]K_{O2}K_{23}K_{O3}K_{P2}^{-1}+$$
$$[R'][R'']K_{O1}K_{O3*_I}+ [R']^2[RNAP][CAP]K_{O1}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*}+$$
$$[R'][RNAP]K_{O1}K_{P1}^{-1}K_{O3*}K_{13*}+ [R][RNAP][CAP]K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*}+$$
$$[R'']^2[CAP]K_{O2}K_{O1_I}K_{CAP}^{-1}K_{O3*}K_{23*c}+ [R'][RNAP]K_{O3}K_{23}K_{O2}K_{P1}^{-1}+$$
$$[R][R'][CAP]K_{O1}K_{O2}K_{CAP}^{-1}K_{O3*}K_{13*c}+ [R'']^2[RNAP]K_{O1_I}K_{O3_I}K_{P1}^{-1}+$$

$$[R'][R''][RNAP]K_{O3}K_{O2_I}K_{P2}^{-1}+ [R'][R'']K_{O2}K_{O1_I}+$$
$$[R'][R''][RNAP]K_{O1}K_{O2_I}K_{P1}^{-1}K_{O3*}K_{13*}+ [R][RNAP]K_{O2}K_{O3}K_{23}K_{P1}^{-1}+$$
$$[R']^2[RNAP][CAP]K_{O1}K_{O2}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*}K_{13*c}+$$
$$[R][R''][RNAP][CAP]K_{CAPO3}K_{O1}K_{O3}K_{O2_I}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{13c}+$$
$$[R'']^2K_{O2_I}K_{O3_I}+ [R'][R''][RNAP]K_{O2}K_{P1}^{-1}K_{O3*_I}+ [R']^2[RNAP]K_{O1}K_{O3}K_{P2}^{-1}+$$
$$[R'][R''][RNAP]K_{O1}K_{12}K_{O2}K_{P2}^{-1}K_{O3*_I}+ [R'][RNAP]K_{O2}K_{P1}^{-1}K_{O3*}K_{23*}+$$
$$[R'][RNAP]K_{O2}K_{23}K_{O3}K_{P1}^{-1}+ [R'']^2[RNAP]K_{O1}K_{O2_I}K_{P2}^{-1}K_{O3*}K_{13*}+$$
$$[R'][R''][RNAP][CAP]K_{CAPO3}K_{O1}K_{O3}K_{O2_I}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{13c}+$$
$$[R'']^2K_{O1_I}K_{O3*_I}+ [R][RNAP]K_{O3}K_{P2}^{-1}+ [R'][R''][RNAP]K_{O2}K_{O3_I}K_{P2}^{-1}+$$
$$[R'][CAP]K_{CAPO3}K_{O3}K_{O2}K_{CAP}^{-1}K_{23c}+$$
$$[R'][R''][RNAP][CAP]K_{O1_I}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*}+ [R'][R'']K_{O1}K_{O3_I}+$$
$$[R'][R''][CAP]K_{O1}K_{O2_I}K_{CAP}^{-1}K_{O3*}K_{13*c}+ [R][R''][RNAP]K_{O1}K_{O2}K_{12}K_{P2}^{-1}K_{O3*_I}+$$
$$[R'][R''][RNAP]K_{O3}K_{O2_I}K_{P1}^{-1}+ [R][CAP]K_{CAPO3}K_{O2}K_{O3}K_{CAP}^{-1}K_{23c}+$$
$$[R'][CAP]K_{CAPO3}K_{O2}K_{O3}K_{CAP}^{-1}K_{23c}+$$
$$[R'']^2[R'][RNAP][CAP]K_{O1}K_{O2_I}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*_I}+$$
$$[R][R'][RNAP][CAP]K_{O2}K_{O1}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}+$$
$$[R][R'][RNAP][CAP]K_{O1}K_{O2}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}+ [R']^2[RNAP]K_{O1}K_{O3}K_{P1}^{-1}+$$
$$[R'][R''][RNAP]K_{O1}K_{12}K_{O2}K_{P1}^{-1}K_{O3*_I}+$$
$$[R][R'][RNAP][CAP]K_{O1}K_{O2}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*}K_{13*c}+$$
$$[R'']^2[RNAP]K_{O1}K_{O2_I}K_{P1}^{-1}K_{O3*}K_{13*}+ [R'][R''][RNAP]K_{O3}K_{O1_I}K_{P2}^{-1}+$$
$$[R'']^2[RNAP]K_{12}K_{O1}K_{O2}K_{P2}^{-1}K_{O3*_I}+ [R']K_{O1}K_{O3*}K_{13*}+$$
$$[R][RNAP]K_{O3}K_{P1}^{-1}+$$
$$[R'']^2[RNAP][CAP]K_{CAPO3}K_{O1}K_{O3}K_{O2_I}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{13c}+$$
$$[R'][R''][RNAP]K_{O2}K_{O3_I}K_{P1}^{-1}+ [R'][R''][RNAP]K_{O1}K_{12}K_{O2}K_{O3_I}K_{P2}^{-1}+$$
$$[R']K_{O3}K_{23}K_{O2}+ [R']^2[CAP]K_{O2}K_{O1}K_{CAP}^{-1}K_{O3*}K_{23*c}+$$
$$[R][R''][CAP]K_{O1}K_{O2_I}K_{CAP}^{-1}+ [R][R''][RNAP][CAP]K_{O1}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*_I}+$$

$$[R][R'']RNAP]K_{O1}K_{O2}K_{12}K_{P1}^{-1}K_{O3*_I}+ [R'']^2K_{O1_I}K_{O3_I}+$$
$$[R'']^2[RNAP][CAP]K_{CAPO3}K_{O2_I}K_{O3_I}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}+$$
$$[R][R'']RNAP]K_{O1}K_{O3}K_{13}K_{O2_I}K_{P2}^{-1}+ [R][R'']RNAP]K_{O1}K_{O2}K_{12}K_{O3_I}K_{P2}^{-1}+$$
$$[R'][R'']K_{O1}K_{O2_I}K_{O3*}K_{13*}+ [R]K_{O2}K_{O3}K_{23}+$$
$$[R'][R'']RNAP]K_{O1}K_{13}K_{O3}K_{O2_I}K_{P2}^{-1}+ [R'][R'']K_{O2}K_{O3*_I}+ [R']K_{O2}K_{O3*}K_{23*}+$$
$$[R']K_{O2}K_{23}K_{O3}+ [R']^2[RNAP][CAP]K_{O2}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*}+$$
$$[R'][R'']RNAP][CAP]K_{CAPO3}K_{O1}K_{O3_I}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}+$$
$$[R'][R'']CAP]K_{O2}K_{12}K_{O1}K_{CAP}^{-1}K_{O3*_I}+ [R']^2[RNAP]K_{O1}K_{O2}K_{P2}^{-1}K_{O3*}K_{13*}+$$
$$[RNAP][CAP]K_{O1_I}K_{O2_I}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*_I}+$$
$$[R][R'']RNAP][CAP]K_{O2}K_{O1_I}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}+ [R'][R'']RNAP]K_{O3}K_{O1_I}K_{P1}^{-1}+$$
$$[R'']^2[RNAP]K_{12}K_{O1}K_{O2}K_{P1}^{-1}K_{O3*_I}+$$
$$[R][R'']RNAP][CAP]K_{O2}K_{O1}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*}K_{13*c}+$$
$$[R'][R'']RNAP]K_{O1}K_{12}K_{O2}K_{O3_I}K_{P1}^{-1}+ [R'][R'']CAP]K_{O2_I}K_{CAP}^{-1}K_{O3*}+$$
$$[R'']^2[RNAP]K_{12}K_{O1}K_{O2}K_{O3_I}K_{P2}^{-1}+ [R'][R'']CAP]K_{O2}K_{O1_I}K_{CAP}^{-1}K_{O3*}K_{23*c}+$$
$$[R'][R'']K_{O3}K_{O2_I}+ [R']^2[CAP]K_{O1}K_{CAP}^{-1}K_{O3*}+$$
$$[R][R'']RNAP]K_{O1}K_{O3}K_{13}K_{O2_I}K_{P1}^{-1}+ [R][R'']RNAP]K_{O1}K_{O2}K_{12}K_{O3_I}K_{P1}^{-1}+$$
$$[R'][R'']RNAP]K_{O1}K_{13}K_{O3}K_{O2_I}K_{P1}^{-1}+ [R']^2[RNAP]K_{O2}K_{O3}K_{P2}^{-1}+$$
$$[R']^2K_{O1}K_{O3}+ [R][CAP]K_{CAP}^{-1}K_{O3*}+ [R'][R'']K_{O1}K_{12}K_{O2}K_{O3*_I}+$$
$$[R'']^2K_{O1}K_{O2_I}K_{O3*}K_{13*}+ [R'']^2[RNAP]K_{13}K_{O1}K_{O3}K_{O2_I}K_{P2}^{-1}+$$
$$[R']^2[RNAP][CAP]K_{O1}K_{12}K_{O2}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*}+$$
$$[R']^2[RNAP]K_{O1}K_{O2}K_{P1}^{-1}K_{O3*}K_{13*}+$$
$$[R][R'][RNAP][CAP]K_{CAPO3}K_{O1}K_{O3}K_{O2}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{13c}+$$
$$[R]K_{O3}+ [R'][R'']K_{O2}K_{O3_I}+ [R'']^2[RNAP][CAP]K_{CAPO3}K_{O1_I}K_{O3_I}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}+$$
$$[R'][R'']RNAP]K_{O1}K_{O2}K_{P2}^{-1}K_{O3*}K_{13*}+$$
$$[R']^2[RNAP][CAP]K_{CAPO3}K_{O1}K_{O2}K_{O3}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{13c}+$$
$$[R][R'']K_{O1}K_{O2}K_{12}K_{O3*_I}+ [R][R'][RNAP][CAP]K_{O1}K_{O2}K_{12}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*}+$$

$$[R']^2[CAP]K_{O1}K_{O2}K_{CAP}^{-1}K_{O3*}K_{13*c}+$$

$$[R']^2[RNAP][CAP]K_{O1}K_{O2}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*}K_{23*c}+$$

$$[R'']^2[RNAP]K_{12}K_{O1}K_{O2}K_{O3_I}K_{P1}^{-1}+ [R][R'']^2[CAP]K_{CAPO3}K_{O1}K_{O3}K_{O2_I}K_{CAP}^{-1}K_{13c}+$$

$$[R'][R'']^2[CAP]K_{CAPO3}K_{O1}K_{O3}K_{O2_I}K_{CAP}^{-1}K_{13c}+ [R][R'']^2[RNAP]K_{O1}K_{O2_I}K_{P2}^{-1}+$$

$$[R'][R'']^2[CAP]K_{O1_I}K_{CAP}^{-1}K_{O3*}+$$

$$[R'][R'']^2[RNAP][CAP]K_{CAPO3}K_{O3}K_{O1}K_{O2_I}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{13c}+$$

$$[R']^2[RNAP]K_{O2}K_{O3}K_{P1}^{-1}+ [R']^2[RNAP]K_{O1}K_{12}K_{O2}K_{O3}K_{P2}^{-1}+$$

$$[R][R'][RNAP][CAP]K_{O2}K_{O1}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*}K_{23*c}+$$

$$[R'']^2[RNAP]K_{13}K_{O1}K_{O3}K_{O2_I}K_{P1}^{-1}+$$

$$[R']^2[R''][RNAP][CAP]K_{O1}K_{O2}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*_I}+$$

$$[R'][R'']K_{O3}K_{O1_I}+ [R'']^2K_{12}K_{O1}K_{O2}K_{O3*_I}+$$

$$[R'][R''][RNAP][CAP]K_{12}K_{O1}K_{O2}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*}+ [R'][R'']K_{O1}K_{12}K_{O2}K_{O3_I}+$$

$$[R'][R''][RNAP][CAP]K_{CAPO3}K_{O3}K_{O2_I}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}+$$

$$[R'][R''][RNAP]K_{O1}K_{O2}K_{P1}^{-1}K_{O3*}K_{13*}+$$

$$[R][R''][RNAP][CAP]K_{O2}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*_I}+$$

$$[R][R'][RNAP]K_{O1}K_{O2}K_{12}K_{O3}K_{P2}^{-1}+ [R'][R''][RNAP]K_{O2}K_{12}K_{O1}K_{P2}^{-1}K_{O3*_I}+$$

$$[R'']^2[R'][CAP]K_{O1}K_{O2_I}K_{CAP}^{-1}K_{O3*_I}+$$

$$[R'][R''][RNAP][CAP]K_{CAPO3}K_{O1}K_{O2}K_{O3}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{13c}+$$

$$[R][R'][CAP]K_{O2}K_{O1}K_{CAP}^{-1}+ [R][R'][CAP]K_{O1}K_{O2}K_{CAP}^{-1}+$$

$$[R']^2[RNAP][CAP]K_{CAPO3}K_{O1}K_{O3}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}+ [R][R'']K_{O1}K_{O3}K_{13}K_{O2_I}+$$

$$[R][R'][CAP]K_{O1}K_{O2}K_{CAP}^{-1}K_{O3*}K_{13*c}+ [R][R'']K_{O1}K_{O2}K_{12}K_{O3_I}+$$

$$[R'][R''][RNAP]K_{O2_I}K_{P2}^{-1}K_{O3*}+$$

$$[R'][R''][RNAP][CAP]K_{O2}K_{O1_I}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*}K_{23*c}+$$

$$[R'][R'']K_{O1}K_{13}K_{O3}K_{O2_I}+ [R'][RNAP]K_{O2}K_{P2}^{-1}K_{O3*}K_{23*}+$$

$$[R][R''][RNAP]K_{O1}K_{O2_I}K_{P1}^{-1}+ [R][RNAP][CAP]K_{CAPO3}K_{O3}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}+$$

$$[R'][R''][RNAP][CAP]K_{CAPO3}K_{O2}K_{O3_I}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}+$$

$$[R'']^2[CAP]K_{CAPO3}K_{O1}K_{O3}K_{O2_I}K_{CAP}^{-1}K_{13c}+ [R][R'][RNAP]K_{O1}K_{O3}K_{13}K_{O2}K_{P2}^{-1}+$$
$$[R']^2[RNAP]K_{O1}K_{12}K_{O2}K_{O3}K_{P1}^{-1}+ [R']^2K_{O1}K_{O2}K_{O3*}K_{13*}+$$
$$[R']^2[RNAP]K_{O1}K_{13}K_{O2}K_{O3}K_{P2}^{-1}+ [R][R''][CAP]K_{O1}K_{CAP}^{-1}K_{O3*_I}+$$
$$[R']^2[RNAP]K_{O1}K_{P2}^{-1}K_{O3*}+ [R][R''][RNAP]K_{O1}K_{O2_I}K_{P2}^{-1}K_{O3*}K_{13*}+$$
$$[R'][R''][RNAP]K_{12}K_{O1}K_{O2}K_{O3}K_{P2}^{-1}+$$
$$[R'][R''][RNAP][CAP]K_{CAPO3}K_{O1}K_{O2}K_{O3}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{23c}+$$
$$[R][R''][RNAP][CAP]K_{O2}K_{O1_I}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*}K_{23*c}+$$
$$[R][RNAP]K_{O2}K_{P2}^{-1}K_{O3*}K_{23*}+ [R'']^2[CAP]K_{CAPO3}K_{O2_I}K_{O3_I}K_{CAP}^{-1}+$$
$$[R'']^2[R'][RNAP][CAP]K_{O2}K_{O1_I}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*_I}+$$
$$[R][RNAP]K_{P2}^{-1}K_{O3*}+ [R][R'][RNAP]K_{O1}K_{O2}K_{12}K_{O3}K_{P1}^{-1}+$$
$$[R'']^2K_{12}K_{O1}K_{O2}K_{O3_I}+ [R'][R''][RNAP]K_{O2}K_{12}K_{O1}K_{P1}^{-1}K_{O3*_I}+$$
$$[R']^2[CAP]K_{O2}K_{CAP}^{-1}K_{O3*}+ [R'][R''][CAP]K_{CAPO3}K_{O1}K_{O3_I}K_{CAP}^{-1}+$$
$$[R'][R''][RNAP]K_{O3}K_{13}K_{O1}K_{O2_I}K_{P2}^{-1}+ [CAP]K_{O1_I}K_{O2_I}K_{CAP}^{-1}K_{O3*_I}+$$
$$[R'][R''][RNAP]K_{O2}K_{12}K_{O1}K_{O3_I}K_{P2}^{-1}+$$
$$[R'][R''][RNAP]K_{O2_I}K_{P1}^{-1}K_{O3*}+ [R']^2K_{O2}K_{O3}+$$
$$[R'][RNAP]K_{O2}K_{P1}^{-1}K_{O3*}K_{23*}+ [R][R''][CAP]K_{O2}K_{O1_I}K_{CAP}^{-1}+$$
$$[R'][R''][RNAP][CAP]K_{CAPO3}K_{O3}K_{O1_I}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}+$$
$$[R][R''][CAP]K_{O2}K_{O1}K_{CAP}^{-1}K_{O3*}K_{13*c}+$$
$$[R][R''][RNAP][CAP]K_{O1}K_{O2}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*}K_{23*c}+$$
$$[R'][R''][RNAP][CAP]K_{CAPO3}K_{O1}K_{12}K_{O2}K_{O3_I}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}+$$
$$[R][R'][RNAP]K_{O1}K_{O3}K_{13}K_{O2}K_{P1}^{-1}+ [R'']^2K_{13}K_{O1}K_{O3}K_{O2_I}+$$
$$[R']^2[RNAP]K_{O1}K_{13}K_{O2}K_{O3}K_{P1}^{-1}+ [R']^2[RNAP]K_{O1}K_{P1}^{-1}K_{O3*}+$$
$$[R][R''][RNAP]K_{O1}K_{O2_I}K_{P1}^{-1}K_{O3*}K_{13*}+ [R'][R''][RNAP]K_{12}K_{O1}K_{O2}K_{O3}K_{P1}^{-1}+$$
$$[R'][R'']K_{O1}K_{O2}K_{O3*}K_{13*}+ [R'][R''][RNAP]K_{13}K_{O1}K_{O2}K_{O3}K_{P2}^{-1}+$$
$$[R][RNAP]K_{O2}K_{P1}^{-1}K_{O3*}K_{23*}+ [R'][R''][RNAP]K_{O1_I}K_{P2}^{-1}K_{O3*}+$$
$$[R][RNAP]K_{P1}^{-1}K_{O3*}+$$

$$[R'']^2[RNAP][CAP]K_{CAPO3}K_{O2}K_{O3}K_{O1_I}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{23c}+$$
$$[R][R'']^2[RNAP][CAP]K_{CAPO3}K_{O1}K_{O2}K_{12}K_{O3_I}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}+$$
$$[R']^2[CAP]K_{O1}K_{12}K_{O2}K_{CAP}^{-1}K_{O3*}+ [R'][R'']^2[RNAP]K_{O3}K_{13}K_{O1}K_{O2_I}K_{P1}^{-1}+$$
$$[R][R'']K_{O1}K_{O2_I}+ [R'][R'']^2[RNAP]K_{O2}K_{12}K_{O1}K_{O3_I}K_{P1}^{-1}+$$
$$[R][R'][CAP]K_{CAPO3}K_{O1}K_{O3}K_{O2}K_{CAP}^{-1}K_{13c}+ [R'']^2[CAP]K_{CAPO3}K_{O1_I}K_{O3_I}K_{CAP}^{-1}+$$
$$[R'']^2[R'][RNAP]K_{O1}K_{O2_I}K_{P2}^{-1}K_{O3*_I}+ [R']^2K_{O1}K_{12}K_{O2}K_{O3}+$$
$$[R'][R'']^2[RNAP]K_{O1}K_{O2}K_{P2}^{-1}K_{O3*}K_{23*}+ [R']^2[CAP]K_{CAPO3}K_{O1}K_{O2}K_{O3}K_{CAP}^{-1}K_{13c}+$$
$$[R'][R'']^2[RNAP]K_{23}K_{O1}K_{O2}K_{O3}K_{P2}^{-1}+ [R][R'][RNAP]K_{O2}K_{O1}K_{P2}^{-1}+$$
$$[R][R'][RNAP]K_{O1}K_{O2}K_{P2}^{-1}+ [R][R'][CAP]K_{O1}K_{O2}K_{12}K_{CAP}^{-1}K_{O3*}+$$
$$[R']^2[RNAP][CAP]K_{CAPO3}K_{O3}K_{O1}K_{O2}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{13c}+$$
$$[R'']^2[RNAP][CAP]K_{CAPO3}K_{12}K_{O1}K_{O2}K_{O3_I}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}+$$
$$[R']^2[CAP]K_{O1}K_{O2}K_{CAP}^{-1}K_{O3*}K_{23*c}+ [R'][R'']^2[RNAP]K_{13}K_{O1}K_{O2}K_{O3}K_{P1}^{-1}+$$
$$[R'][R'']^2[RNAP]K_{O1_I}K_{P1}^{-1}K_{O3*}+ [R][R']K_{O1}K_{O2}K_{12}K_{O3}+$$
$$[R]^2[RNAP][CAP]K_{O1}K_{O2}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*}K_{13*c}+ [R'][R'']K_{O2}K_{12}K_{O1}K_{O3*_I}+$$
$$[R][R'']^2[RNAP]K_{O1}K_{P2}^{-1}K_{O3*_I}+$$
$$[R']^2[RNAP][CAP]K_{O2}K_{12}K_{O1}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*}+$$
$$[R']^2[RNAP][CAP]K_{CAPO3}K_{O2}K_{O3}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}+$$
$$[R'][R'']^2[CAP]K_{CAPO3}K_{O3}K_{O1}K_{O2_I}K_{CAP}^{-1}K_{13c}+$$
$$[R][R'][CAP]K_{O2}K_{O1}K_{CAP}^{-1}K_{O3*}K_{23*c}+ [R'][R'']K_{O2_I}K_{O3*}+$$
$$[R']^2[R'']^2[CAP]K_{O1}K_{O2}K_{CAP}^{-1}K_{O3*_I}+ [R']K_{O2}K_{O3*}K_{23*}+$$
$$[R']^2[RNAP][CAP]K_{CAPO3}K_{O3}K_{O1}K_{O2}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{23c}+$$
$$[R][R']K_{O1}K_{O3}K_{13}K_{O2}+ [R'][R'']^2[CAP]K_{12}K_{O1}K_{O2}K_{CAP}^{-1}K_{O3*}+$$
$$[R'][R'']^2[CAP]K_{CAPO3}K_{O3}K_{O2_I}K_{CAP}^{-1}+ [R'']^2[R'][RNAP]K_{O1}K_{O2_I}K_{P1}^{-1}K_{O3*_I}+$$
$$[R'][R'']^2[RNAP]K_{O1}K_{O2}K_{P1}^{-1}K_{O3*}K_{23*}+ [R][R'']^2[CAP]K_{O2}K_{CAP}^{-1}K_{O3*_I}+$$
$$[R']^2[RNAP]K_{O2}K_{P2}^{-1}K_{O3*}+ [R'][R'']^2[RNAP]K_{23}K_{O1}K_{O2}K_{O3}K_{P1}^{-1}+$$
$$[R']^2K_{O1}K_{13}K_{O2}K_{O3}+ [R']^2K_{O1}K_{O3*}+ [R][R'][RNAP]K_{O2}K_{O1}K_{P1}^{-1}+$$

$$[R][R'][RNAP]K_{O1}K_{O2}K_{P1}^{-1}+ [RNAP]K_{O1_I}K_{O2_I}K_{P2}^{-1}K_{O3*_I}+$$

$$[R][R'']K_{O1}K_{O2_I}K_{O3*}K_{13*}+ [R'][R'']K_{12}K_{O1}K_{O2}K_{O3}+$$

$$[R'']^2[RNAP]K_{O2}K_{O1_I}K_{P2}^{-1}K_{O3*}K_{23*}+$$

$$[R'][R''][CAP]K_{CAPO3}K_{O1}K_{O2}K_{O3}K_{CAP}^{-1}K_{13c}+ [R'']^2[RNAP]K_{23}K_{O2}K_{O3}K_{O1_I}K_{P2}^{-1}+$$

$$[R'']^2[R'][RNAP]K_{O1}K_{O2_I}K_{O3_I}K_{P2}^{-1}+ [R]K_{O2}K_{O3*}K_{23*}+$$

$$[R][R''][RNAP]K_{O2}K_{O1_I}K_{P2}^{-1}+ [R']^2[CAP]K_{CAPO3}K_{O1}K_{O3}K_{CAP}^{-1}+$$

$$[R][R'][RNAP][CAP]K_{CAPO3}K_{O2}K_{O3}K_{O1}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{23c}+$$

$$[R]K_{O3*}+ [R][R'][RNAP]K_{O1}K_{O2}K_{P2}^{-1}K_{O3*}K_{13*}+$$

$$[R']^2[RNAP]K_{O2}K_{12}K_{O1}K_{O3}K_{P2}^{-1}+ [R'][R''][CAP]K_{O2}K_{O1_I}K_{CAP}^{-1}K_{O3*}K_{23*c}+$$

$$[R']^2[RNAP][CAP]K_{CAPO3}K_{O2}K_{O1}K_{O3}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{23c}+$$

$$[R][CAP]K_{CAPO3}K_{O3}K_{CAP}^{-1}+ [R'][R''][CAP]K_{CAPO3}K_{O2}K_{O3_I}K_{CAP}^{-1}+$$

$$[R][R''][RNAP]K_{O1}K_{P1}^{-1}K_{O3*_I}+$$

$$[R']^2[RNAP][CAP]K_{CAPO3}K_{O1}K_{12}K_{O2}K_{O3}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}+$$

$$[R'][R'']K_{O3}K_{13}K_{O1}K_{O2_I}+ [R'][R'']K_{O2}K_{12}K_{O1}K_{O3_I}+$$

$$[R][R''][RNAP]K_{O1}K_{O3_I}K_{P2}^{-1}+ [R'][R''][CAP]K_{CAPO3}K_{O1}K_{O2}K_{O3}K_{CAP}^{-1}K_{23c}+$$

$$[R][R''][CAP]K_{O2}K_{O1_I}K_{CAP}^{-1}K_{O3*}K_{23*c}+ [R']^2[RNAP]K_{O3}K_{13}K_{O1}K_{O2}K_{P2}^{-1}+$$

$$[R'']^2[R'][CAP]K_{O2}K_{O1_I}K_{CAP}^{-1}K_{O3*_I}+ [R']^2[RNAP]K_{O2}K_{P1}^{-1}K_{O3*}+$$

$$[R'][R''][RNAP][CAP]K_{CAPO3}K_{O3}K_{O2}K_{O1_I}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{23c}+$$

$$[R][R'][RNAP][CAP]K_{CAPO3}K_{O1}K_{O2}K_{12}K_{O3}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}+$$

$$[R']^2[RNAP]K_{O1}K_{12}K_{O2}K_{P2}^{-1}K_{O3*}+ [R'][R''][RNAP]K_{O1}K_{O2_I}K_{P2}^{-1}K_{O3*}K_{13*}+$$

$$[RNAP]K_{O1_I}K_{O2_I}K_{P1}^{-1}K_{O3*_I}+$$

$$[R][R'][RNAP][CAP]K_{O1}K_{O2}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*}K_{23*c}+$$

$$[R'']^2[RNAP]K_{O2}K_{O1_I}K_{P1}^{-1}K_{O3*}K_{23*}+ [R'']^2[RNAP]K_{23}K_{O2}K_{O3}K_{O1_I}K_{P1}^{-1}+$$

$$[R'][R'']K_{13}K_{O1}K_{O2}K_{O3}+ [R'']^2[R'][RNAP]K_{O1}K_{O2_I}K_{O3_I}K_{P1}^{-1}+$$

$$[R'][R'']K_{O1_I}K_{O3*}+ [R][R''][RNAP]K_{O2}K_{O1_I}K_{P1}^{-1}+$$

$$[RNAP]K_{O1_I}K_{O2_I}K_{O3_I}K_{P2}^{-1}+ [R][R'][RNAP]K_{O1}K_{O2}K_{P1}^{-1}K_{O3*}K_{13*}+$$

$$[R']^2[RNAP]K_{O2}K_{12}K_{O1}K_{O3}K_{P1}^{-1}+ [R'][R''][CAP]K_{CAPO3}K_{O3}K_{O1_I}K_{CAP}^{-1}+$$
$$[R][R''][RNAP][CAP]K_{CAPO3}K_{O2}K_{O3}K_{O1_I}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{23c}+$$
$$[R][R'][RNAP]K_{O1}K_{O2}K_{12}K_{P2}^{-1}K_{O3*}+ [R][R'']CAP]K_{O1}K_{O2}K_{CAP}^{-1}K_{O3*}K_{23*c}+$$
$$[R']^2[RNAP]K_{O3}K_{23}K_{O1}K_{O2}K_{P2}^{-1}+ [R'][R''][CAP]K_{CAPO3}K_{O1}K_{12}K_{O2}K_{O3_I}K_{CAP}^{-1}+$$
$$[R'][R''][RNAP][CAP]K_{CAPO3}K_{O2}K_{O3}K_{O1_I}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{23c}+$$
$$[R'][R''][RNAP][CAP]K_{CAPO3}K_{12}K_{O1}K_{O2}K_{O3}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}+$$
$$[R][R''][RNAP]K_{O1}K_{O3_I}K_{P1}^{-1}+ [R'']^2[R']K_{O1}K_{O2_I}K_{O3*_I}+$$
$$[R'][R'']K_{O1}K_{O2}K_{O3*}K_{23*}+ [R'][R'']K_{23}K_{O1}K_{O2}K_{O3}+$$
$$[R']^2[R''][RNAP][CAP]K_{O1}K_{O2_I}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*}+$$
$$[R']^2[RNAP]K_{O3}K_{13}K_{O1}K_{O2}K_{P1}^{-1}+ [R][R']K_{O2}K_{O1}+ [R][R']K_{O1}K_{O2}+$$
$$[R'']^2[CAP]K_{CAPO3}K_{O2}K_{O3}K_{O1_I}K_{CAP}^{-1}K_{23c}+ [R][R'][RNAP]K_{O2}K_{O3}K_{23}K_{O1}K_{P2}^{-1}+$$
$$[R][R''][CAP]K_{CAPO3}K_{O1}K_{O2}K_{12}K_{O3_I}K_{CAP}^{-1}+ [R']^2[RNAP]K_{O1}K_{12}K_{O2}K_{P1}^{-1}K_{O3*}+$$
$$[R']^2[R''][RNAP]K_{O1}K_{O2}K_{P2}^{-1}K_{O3*_I}+ [R']^2[RNAP]K_{O2}K_{O1}K_{P2}^{-1}K_{O3*}K_{23*}+$$
$$[R']^2[RNAP]K_{O2}K_{23}K_{O1}K_{O3}K_{P2}^{-1}+ [R'][R''][RNAP]K_{O1}K_{O2_I}K_{P1}^{-1}K_{O3*}K_{13*}+$$
$$[R][R''][RNAP][CAP]K_{O2_I}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*}+$$
$$[R'][R''][RNAP]K_{12}K_{O1}K_{O2}K_{P2}^{-1}K_{O3*}+ [R][R''][RNAP]K_{O2}K_{P2}^{-1}K_{O3*_I}+$$
$$[R'][R''][RNAP][CAP]K_{CAPO3}K_{O2}K_{12}K_{O1}K_{O3_I}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}+$$
$$[R][R'']K_{O1}K_{O3*_I}+ [RNAP]K_{O1_I}K_{O2_I}K_{O3_I}K_{P1}^{-1}+$$
$$[R][R'][RNAP][CAP]K_{O1}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*}+$$
$$[R][R'][RNAP][CAP]K_{O1}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*}+$$
$$[R][R'][RNAP]K_{O1}K_{O2}K_{12}K_{P1}^{-1}K_{O3*}+ [R']^2[RNAP]K_{O3}K_{23}K_{O1}K_{O2}K_{P1}^{-1}+$$
$$[R']^2[CAP]K_{CAPO3}K_{O3}K_{O1}K_{O2}K_{CAP}^{-1}K_{13c}+ [R'][R''][RNAP]K_{O3}K_{23}K_{O2}K_{O1_I}K_{P2}^{-1}+$$
$$[R'']^2[CAP]K_{CAPO3}K_{12}K_{O1}K_{O2}K_{O3_I}K_{CAP}^{-1}+ [R']^2[R''][RNAP]K_{O1}K_{O3}K_{O2_I}K_{P2}^{-1}+$$
$$[R']^2K_{O2}K_{O3*}+ [R]^2[CAP]K_{O1}K_{O2}K_{CAP}^{-1}K_{O3*}K_{13*c}+ K_{O1_I}K_{O2_I}K_{O3*_I}+$$
$$[R][R''][RNAP]K_{O3}K_{O2_I}K_{P2}^{-1}+ [R'']^2K_{O2}K_{O1_I}K_{O3*}K_{23*}+$$
$$[R][R'][RNAP]K_{O2}K_{O3}K_{23}K_{O1}K_{P1}^{-1}+$$

$$[R'']^2K_{23}K_{O2}K_{O3}K_{O1_I}+ [R']^2[CAP]K_{O2}K_{12}K_{O1}K_{CAP}^{-1}K_{O3*}+$$
$$[R'']^2[R'][RNAP][CAP]K_{O1_I}K_{O2_I}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*}+ [R'']^2[R']K_{O1}K_{O2_I}K_{O3_I}+$$
$$[R']^2[CAP]K_{CAPO3}K_{O2}K_{O3}K_{CAP}^{-1}+ [R']^2[R''][RNAP]K_{O1}K_{O2}K_{P1}^{-1}K_{O3*_I}+$$
$$[R][R'']K_{O2}K_{O1_I}+ [R']^2[RNAP]K_{O2}K_{O1}K_{P1}^{-1}K_{O3*}K_{23*}+$$
$$[R']^2[RNAP]K_{O2}K_{23}K_{O1}K_{O3}K_{P1}^{-1}+ [R][R''][RNAP]K_{O2}K_{O3}K_{23}K_{O1_I}K_{P2}^{-1}+$$
$$[R'][R''][RNAP]K_{12}K_{O1}K_{O2}K_{P1}^{-1}K_{O3*}+ [R'']^2[R'][RNAP]K_{O2}K_{O1_I}K_{P2}^{-1}K_{O3*_I}+$$
$$[R][R']K_{O1}K_{O2}K_{O3*}K_{13*}+ [R']^2K_{O2}K_{12}K_{O1}K_{O3}+$$
$$[R'][R'']K_{O2}K_{O1_I}K_{P2}^{-1}K_{O3*}K_{23*}+$$
$$[R][R'][RNAP][CAP]K_{O2}K_{O1}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*}K_{13*c}+$$
$$[R'][R'']K_{O2}K_{23}K_{O3}K_{O1_I}K_{P2}^{-1}+ [R][R'']K_{O2}K_{P1}^{-1}K_{O3*_I}+$$
$$[R][R'][RNAP]K_{O3}K_{O1}K_{P2}^{-1}+ [R][R'][RNAP]K_{O1}K_{O3}K_{P2}^{-1}+$$
$$[R']^2[R'']K_{O1}K_{O2}K_{O3_I}K_{P2}^{-1}+ [R']^2[CAP]K_{CAPO3}K_{O3}K_{O1}K_{O2}K_{CAP}^{-1}K_{23c}+$$
$$[R][R'']K_{O2}K_{O3_I}K_{P2}^{-1}+ [R][R'']K_{O1_I}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*}+$$
$$[R][R'']K_{O1}K_{O3_I}+ [R'][R'']K_{O3}K_{23}K_{O2}K_{O1_I}K_{P1}^{-1}+$$
$$[R']^2K_{O3}K_{13}K_{O1}K_{O2}+ [R']^2[R'']K_{O1}K_{O3}K_{O2_I}K_{P1}^{-1}+$$
$$[R][R'][CAP]K_{CAPO3}K_{O2}K_{O3}K_{O1}K_{CAP}^{-1}K_{23c}+ [R']^2K_{O1}K_{12}K_{O2}K_{O3*}+$$
$$[R'']^2[R'][RNAP]K_{O3}K_{O1_I}K_{O2_I}K_{P2}^{-1}+ [R']^2[CAP]K_{CAPO3}K_{O2}K_{O1}K_{O3}K_{CAP}^{-1}K_{23c}+$$
$$[R'][R'']K_{O1}K_{O2_I}K_{O3*}K_{13*}+ [R][R'']K_{O3}K_{O2_I}K_{P1}^{-1}+$$
$$[R']^2[CAP]K_{CAPO3}K_{O1}K_{12}K_{O2}K_{O3}K_{CAP}^{-1}+ [R][R'']K_{O2}K_{O3}K_{23}K_{O1_I}K_{P1}^{-1}+$$
$$[R'']^2[R][RNAP][CAP]K_{O1}K_{O2_I}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*_I}+$$
$$[R']^2[RNAP]K_{O1}K_{O2}K_{P2}^{-1}K_{O3*}K_{13*}+ K_{O1_I}K_{O2_I}K_{O3_I}+$$
$$[R'']^2[R'][RNAP]K_{O2}K_{O1_I}K_{P1}^{-1}K_{O3*_I}+ [R'][R''][RNAP]K_{O2}K_{O1_I}K_{P1}^{-1}K_{O3*}K_{23*}+$$
$$[R'][R''][RNAP]K_{O2}K_{23}K_{O3}K_{O1_I}K_{P1}^{-1}+ [R]^2[RNAP][CAP]K_{O1}K_{O2}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}+$$
$$[R][R'][RNAP]K_{O3}K_{O1}K_{P1}^{-1}+ [R][R'][RNAP]K_{O1}K_{O3}K_{P1}^{-1}+$$
$$[R']^2[R'']K_{O1}K_{O2}K_{O3_I}K_{P1}^{-1}+ [R][R']K_{O1}K_{O2}K_{12}K_{O3*}+$$
$$[R']^2K_{O3}K_{23}K_{O1}K_{O2}+ [R][R'']K_{O3}K_{O1_I}K_{P2}^{-1}+$$

$$[R'']^2[R'][RNAP]K_{O2}K_{O1_I}K_{O3_I}K_{P2}^{-1}+$$

$$[R'][R''][CAP]K_{CAPO3}K_{O3}K_{O2}K_{O1_I}K_{CAP}^{-1}K_{23c}+$$

$$[R][R'][CAP]K_{CAPO3}K_{O1}K_{O2}K_{12}K_{O3}K_{CAP}^{-1}+\ [R][R''][RNAP]K_{O2}K_{O3_I}K_{P1}^{-1}+$$

$$[R][R'][CAP]K_{O1}K_{O2}K_{CAP}^{-1}K_{O3*}K_{23*c}+$$

$$[R'']^2[R'][RNAP][CAP]K_{CAPO3}K_{O1}K_{O2_I}K_{O3_I}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}+$$

$$[R][R']K_{O2}K_{O3}K_{23}K_{O1}+$$

$$[R']^2[RNAP][CAP]K_{CAPO3}K_{O2}K_{12}K_{O1}K_{O3}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}+$$

$$[R']^2[R'']K_{O1}K_{O2}K_{O3*_I}+\ [R'']^2[R'][RNAP]K_{O3}K_{O1_I}K_{O2_I}K_{P1}^{-1}+$$

$$[R']^2K_{O2}K_{O1}K_{O3*}K_{23*}+\ [R']^2K_{O2}K_{23}K_{O1}K_{O3}+$$

$$[R][R''][CAP]K_{CAPO3}K_{O2}K_{O3}K_{O1_I}K_{CAP}^{-1}K_{23c}+$$

$$[RNAP][CAP]K_{O1}K_{O2}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*}+\ [R'][R'']K_{12}K_{O1}K_{O2}K_{O3*}+$$

$$[R'][R''][CAP]K_{CAPO3}K_{O2}K_{O3}K_{O1_I}K_{CAP}^{-1}K_{23c}+\ [R][R'']K_{O2}K_{O3*_I}+$$

$$[R']^2[RNAP]K_{O1}K_{O2}K_{P1}^{-1}K_{O3*}K_{13*}+$$

$$[R][R'][RNAP][CAP]K_{O2}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*}+$$

$$[R][R'][RNAP][CAP]K_{O2}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*}+$$

$$[R][R''][RNAP][CAP]K_{CAPO3}K_{O1}K_{O3_I}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}+$$

$$[R'][R''][CAP]K_{CAPO3}K_{12}K_{O1}K_{O2}K_{O3}K_{CAP}^{-1}+\ [R']^2[RNAP]K_{O2}K_{12}K_{O1}K_{P2}^{-1}K_{O3*}+$$

$$[R][R'][RNAP]K_{O1}K_{O2}K_{P2}^{-1}K_{O3*}K_{13*}+\ [R']^2[R''][CAP]K_{O1}K_{O2_I}K_{CAP}^{-1}K_{O3*}+$$

$$[R][R''][RNAP]K_{O3}K_{O1_I}K_{P1}^{-1}+\ [R'']^2[R'][RNAP]K_{O2}K_{O1_I}K_{O3_I}K_{P1}^{-1}+$$

$$[R'][R'']K_{O3}K_{23}K_{O2}K_{O1_I}+\ [R']^2[R'']K_{O1}K_{O3}K_{O2_I}+\ [R][R''][CAP]K_{O2_I}K_{CAP}^{-1}K_{O3*}+$$

$$[R'][R''][CAP]K_{CAPO3}K_{O2}K_{12}K_{O1}K_{O3_I}K_{CAP}^{-1}+\ [R][R'']K_{O3}K_{O2_I}+$$

$$[RNAP][CAP]K_{CAPO3}K_{O1_I}K_{O2_I}K_{O3_I}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}+$$

$$[RNAP]K_{O1}K_{O2}K_{O3}K_{P2}^{-1}+\ [R][R'][CAP]K_{O1}K_{CAP}^{-1}K_{O3*}+$$

$$[R][R'][CAP]K_{O1}K_{CAP}^{-1}K_{O3*}+\ [R][R'']K_{O2}K_{O3}K_{23}K_{O1_I}+$$

$$[R'']^2[R']K_{O2}K_{O1_I}K_{O3*_I}+\ [R][R'][RNAP]K_{O3}K_{O2}K_{P2}^{-1}+$$

$$[R][R'][RNAP]K_{O2}K_{O3}K_{P2}^{-1}+\ [R'][R'']K_{O2}K_{O1_I}K_{O3*}K_{23*}+$$

$$[R'][R'']K_{O2}K_{23}K_{O3}K_{O1_I}+ [R']^2[R''][RNAP][CAP]K_{O2}K_{O1_I}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*}+$$
$$[R][R']K_{O3}K_{O1}+ [R][R']K_{O1}K_{O3}+ [R']^2[R'']K_{O1}K_{O2}K_{O3_I}+$$
$$[R][R'][RNAP][CAP]K_{O1}K_{12}K_{O2}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*}+$$
$$[R']^2[RNAP]K_{O2}K_{12}K_{O1}K_{P1}^{-1}K_{O3*}+ [R][R'][RNAP]K_{O1}K_{O2}K_{P1}^{-1}K_{O3*}K_{13*}+$$
$$[R]^2[RNAP][CAP]K_{CAPO3}K_{O1}K_{O2}K_{O3}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{13c}+ [R][R'']K_{O2}K_{O3_I}+$$
$$[R][R''][RNAP]K_{O2}K_{O1}K_{P2}^{-1}K_{O3*}K_{13*}+$$
$$[R][R'][RNAP][CAP]K_{CAPO3}K_{O1}K_{O2}K_{O3}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{13c}+$$
$$[R'']^2[R'][CAP]K_{O1_I}K_{O2_I}K_{CAP}^{-1}K_{O3*}+$$
$$[R]^2[RNAP][CAP]K_{O1}K_{O2}K_{12}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*}+$$
$$[R][R'][CAP]K_{O2}K_{O1}K_{CAP}^{-1}K_{O3*}K_{13*c}+$$
$$[R][R'][RNAP][CAP]K_{O1}K_{O2}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*}K_{23*c}+$$
$$[R'']^2[R']K_{O3}K_{O1_I}K_{O2_I}+ [RNAP]K_{O1}K_{O2}K_{O3}K_{P1}^{-1}+$$
$$[R']^2[R''][RNAP]K_{O2}K_{O3}K_{O1_I}K_{P2}^{-1}+ [R']^2K_{O1}K_{O2}K_{O3*}K_{13*}+$$
$$[R][R''][CAP]K_{O1_I}K_{CAP}^{-1}K_{O3*}+ [R][R'][RNAP]K_{O3}K_{O2}K_{P1}^{-1}+$$
$$[R][R'][RNAP]K_{O2}K_{O3}K_{P1}^{-1}+ [R][R'][RNAP]K_{O1}K_{O3}K_{12}K_{O2}K_{P2}^{-1}+$$
$$[R]^2[RNAP][CAP]K_{O1}K_{O2}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*}K_{23*c}+$$
$$[R']^2[R''][RNAP][CAP]K_{CAPO3}K_{O1}K_{O3}K_{O2_I}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}+$$
$$[R][R'][R''][RNAP][CAP]K_{O2}K_{O1}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*_I}+$$
$$[R][R'][R''][RNAP][CAP]K_{O1}K_{O2}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*_I}+$$
$$[R][R'']K_{O3}K_{O1_I}+ [R'']^2[R']K_{O2}K_{O1_I}K_{O3_I}+$$
$$[R][R''][RNAP][CAP]K_{12}K_{O1}K_{O2}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*}+$$
$$[R][R''][RNAP][CAP]K_{CAPO3}K_{O3}K_{O2_I}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}+$$
$$[R][R''][RNAP]K_{O2}K_{O1}K_{P1}^{-1}K_{O3*}K_{13*}+ [R]^2[RNAP]K_{O1}K_{O2}K_{O3}K_{12}K_{P2}^{-1}+$$
$$[R'']^2[R][CAP]K_{O1}K_{O2_I}K_{CAP}^{-1}K_{O3*_I}+$$
$$[R][R''][RNAP][CAP]K_{CAPO3}K_{O2}K_{O1}K_{O3}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{13c}+$$
$$[R']^2[R''][RNAP]K_{O1}K_{O2_I}K_{P2}^{-1}K_{O3*}+ [R']^2[RNAP]K_{O1}K_{O2}K_{P2}^{-1}K_{O3*}K_{23*}+$$

$$[R]^2[CAP]K_{O1}K_{O2}K_{CAP}^{-1}+ [R][R'][RNAP][CAP]K_{CAPO3}K_{O3}K_{O1}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}+$$
$$[R][R'][RNAP][CAP]K_{CAPO3}K_{O1}K_{O3}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}+$$
$$[R']^2[R''][RNAP][CAP]K_{CAPO3}K_{O1}K_{O2}K_{O3_I}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}+$$
$$[R][R''][RNAP]K_{O2_I}K_{P2}^{-1}K_{O3*}+$$
$$[R][R''][RNAP][CAP]K_{CAPO3}K_{O2}K_{O3_I}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}+$$
$$[R']^2[R''][RNAP]K_{O2}K_{O3}K_{O1_I}K_{P1}^{-1}+ [R]^2[RNAP]K_{O1}K_{O2}K_{O3}K_{13}K_{P2}^{-1}+$$
$$[R][R'][RNAP]K_{O2}K_{O1}K_{P2}^{-1}K_{O3*}K_{23*}+ [R][R'][RNAP]K_{O1}K_{O3}K_{12}K_{O2}K_{P1}^{-1}+$$
$$[R']^2K_{O2}K_{12}K_{O1}K_{O3*}+ [R][R']K_{O1}K_{O2}K_{O3*}K_{13*}+$$
$$[R'']^2[R'][CAP]K_{CAPO3}K_{O1}K_{O2_I}K_{O3_I}K_{CAP}^{-1}+ [R][R'][RNAP]K_{O1}K_{O2}K_{13}K_{O3}K_{P2}^{-1}+$$
$$[R][R'][RNAP]K_{O1}K_{P2}^{-1}K_{O3*}+ [R][R'][RNAP]K_{O1}K_{P2}^{-1}K_{O3*}+$$
$$[R][R''][RNAP]K_{O3}K_{12}K_{O1}K_{O2}K_{P2}^{-1}+$$
$$[R][R''][RNAP][CAP]K_{CAPO3}K_{O1}K_{O2}K_{O3}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{23c}+$$
$$[R']^2[CAP]K_{CAPO3}K_{O2}K_{12}K_{O1}K_{O3}K_{CAP}^{-1}+$$
$$[R'']^2[R'][RNAP][CAP]K_{CAPO3}K_{O3}K_{O1_I}K_{O2_I}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}+$$
$$[R'']^2[R][RNAP][CAP]K_{O2}K_{O1_I}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*_I}+ [CAP]K_{O1}K_{O2}K_{CAP}^{-1}K_{O3*}+$$
$$[R]^2[RNAP]K_{O1}K_{O2}K_{O3}K_{12}K_{P1}^{-1}+ [R']^2[R''][RNAP]K_{O1}K_{O2_I}K_{P1}^{-1}K_{O3*}+$$
$$K_{O1}K_{O2}K_{O3}+ [R']^2[RNAP]K_{O1}K_{O2}K_{P1}^{-1}K_{O3*}K_{23*}+$$
$$[R][R'][CAP]K_{O2}K_{CAP}^{-1}K_{O3*}+ [R][R'][CAP]K_{O2}K_{CAP}^{-1}K_{O3*}+$$
$$[R][R''][CAP]K_{CAPO3}K_{O1}K_{O3_I}K_{CAP}^{-1}+ [R'']^2[R'][RNAP]K_{O1_I}K_{O2_I}K_{P2}^{-1}K_{O3*}+$$
$$[R'][R''][RNAP]K_{O2}K_{O1_I}K_{P2}^{-1}K_{O3*}K_{23*}+ [R][R''][RNAP]K_{O2_I}K_{P1}^{-1}K_{O3*}+$$
$$[R][R']K_{O3}K_{O2}+ [R][R']K_{O2}K_{O3}+$$
$$[R][R''][RNAP][CAP]K_{CAPO3}K_{O3}K_{O1_I}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}+$$
$$[R'']^2[R'][RNAP][CAP]K_{CAPO3}K_{O2}K_{O1_I}K_{O3_I}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}+$$
$$[R]^2[RNAP]K_{O1}K_{O2}K_{O3}K_{13}K_{P1}^{-1}+ [R][R'][RNAP]K_{O2}K_{O1}K_{P1}^{-1}K_{O3*}K_{23*}+$$
$$[R][R'][RNAP]K_{O1}K_{O2}K_{13}K_{O3}K_{P1}^{-1}+ [R][R'][RNAP]K_{O1}K_{P1}^{-1}K_{O3*}+$$
$$[R][R'][RNAP]K_{O1}K_{P1}^{-1}K_{O3*}+ [R][R''][RNAP]K_{O2}K_{O1_I}K_{P2}^{-1}K_{O3*}K_{23*}+$$

$$[R][R'']RNAP]K_{O3}K_{12}K_{O1}K_{O2}K_{P1}^{-1}+ [R][R'']K_{O2}K_{O1}K_{O3*}K_{13*}+$$
$$[CAP]K_{CAPO3}K_{O1_I}K_{O2_I}K_{O3_I}K_{CAP}^{-1}+ [R][R'']RNAP]K_{O2}K_{13}K_{O1}K_{O3}K_{P2}^{-1}+$$
$$[R][R'']RNAP]K_{O1_I}K_{P2}^{-1}K_{O3*}+ [R']^2[R'']CAP]K_{O2}K_{O1_I}K_{CAP}^{-1}K_{O3*}+$$
$$[R][R']CAP]K_{O1}K_{12}K_{O2}K_{CAP}^{-1}K_{O3*}+ [R'']^2[R']RNAP]K_{O1_I}K_{O2_I}K_{P1}^{-1}K_{O3*}+$$
$$[R]^2CAP]K_{CAPO3}K_{O1}K_{O2}K_{O3}K_{CAP}^{-1}K_{13c}+ [R']^2[R'']K_{O2}K_{O3}K_{O1_I}+$$
$$[R'][R'']RNAP]K_{O2}K_{O1_I}K_{P1}^{-1}K_{O3*}K_{23*}+ [R'']^2[R]RNAP]K_{O1}K_{O2_I}K_{P2}^{-1}K_{O3*_I}+$$
$$[R][R']K_{O1}K_{O3}K_{12}K_{O2}+ [R][R'']RNAP]K_{O1}K_{O2}K_{P2}^{-1}K_{O3*}K_{23*}+$$
$$[R][R']CAP]K_{CAPO3}K_{O1}K_{O2}K_{O3}K_{CAP}^{-1}K_{13c}+ [R][R'']RNAP]K_{O1}K_{23}K_{O2}K_{O3}K_{P2}^{-1}+$$
$$[R]^2RNAP]K_{O1}K_{O2}K_{P2}^{-1}+ [R]^2CAP]K_{O1}K_{O2}K_{12}K_{CAP}^{-1}K_{O3*}+$$
$$[R][R']RNAP]CAP]K_{CAPO3}K_{O2}K_{O3}K_{O1}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{13c}+$$
$$[R][R']CAP]K_{O1}K_{O2}K_{CAP}^{-1}K_{O3*}K_{23*c}+ [R][R'']RNAP]K_{O2}K_{O1_I}K_{P1}^{-1}K_{O3*}K_{23*}+$$
$$RNAP]CAP]K_{CAPO3}K_{O1}K_{O2}K_{O3}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}+$$
$$[R][R'']RNAP]K_{O2}K_{13}K_{O1}K_{O3}K_{P1}^{-1}+ [R][R'']RNAP]K_{O1_I}K_{P1}^{-1}K_{O3*}+$$
$$[R]^2K_{O1}K_{O2}K_{O3}K_{12}+ [R']^2[R'']K_{O1}K_{O2_I}K_{O3*}+$$
$$[R][R']RNAP]CAP]K_{O2}K_{12}K_{O1}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*}+ [R']^2K_{O1}K_{O2}K_{O3*}K_{23*}+$$
$$[R][R']RNAP]CAP]K_{CAPO3}K_{O3}K_{O2}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}+$$
$$[R][R']RNAP]CAP]K_{CAPO3}K_{O2}K_{O3}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}+$$
$$[R]^2[CAP]K_{O1}K_{O2}K_{CAP}^{-1}K_{O3*}K_{23*c}+ [R']^2[R'']CAP]K_{CAPO3}K_{O1}K_{O3}K_{O2_I}K_{CAP}^{-1}+$$
$$[R][R'']K_{O2_I}K_{O3*}+ [R][R'][R'']CAP]K_{O2}K_{O1}K_{CAP}^{-1}K_{O3*_I}+$$
$$[R][R'][R'']CAP]K_{O1}K_{O2}K_{CAP}^{-1}K_{O3*_I}+ [RNAP]K_{O1}K_{O2}K_{P2}^{-1}K_{O3*}+$$
$$[R][R']RNAP]CAP]K_{CAPO3}K_{O1}K_{O3}K_{O2}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{23c}+$$
$$[R]^2K_{O1}K_{O2}K_{O3}K_{13}+ [R][R'']CAP]K_{12}K_{O1}K_{O2}K_{CAP}^{-1}K_{O3*}+$$
$$[R][R'']CAP]K_{CAPO3}K_{O3}K_{O2_I}K_{CAP}^{-1}+ [R'']^2[R]RNAP]K_{O1}K_{O2_I}K_{P1}^{-1}K_{O3*_I}+$$
$$[R][R']K_{O2}K_{O1}K_{O3*}K_{23*}+ [R][R'']RNAP]K_{O1}K_{O2}K_{P1}^{-1}K_{O3*}K_{23*}+$$
$$[R][R']RNAP]K_{O2}K_{P2}^{-1}K_{O3*}+ [R][R']RNAP]K_{O2}K_{P2}^{-1}K_{O3*}+$$
$$[R][R'']RNAP]K_{O1}K_{23}K_{O2}K_{O3}K_{P1}^{-1}+ [R][R']K_{O1}K_{O2}K_{13}K_{O3}+ [R][R']K_{O1}K_{O3*}+$$

$$[R][R']K_{O1}K_{O3*}+ [R]^2[RNAP]K_{O1}K_{O2}K_{P1}^{-1}+$$
$$[R][R'']K_{O3}K_{12}K_{O1}K_{O2}+ [R][R''][CAP]K_{CAPO3}K_{O2}K_{O1}K_{O3}K_{CAP}^{-1}K_{13c}+$$
$$[R'']^2[R][RNAP]K_{O1}K_{O2_I}K_{O3_I}K_{P2}^{-1}+ [R][R'][CAP]K_{CAPO3}K_{O3}K_{O1}K_{CAP}^{-1}+$$
$$[R][R'][CAP]K_{CAPO3}K_{O1}K_{O3}K_{CAP}^{-1}+ [R']^2[R''][CAP]K_{CAPO3}K_{O1}K_{O2}K_{O3_I}K_{CAP}^{-1}+$$
$$[R]^2[RNAP][CAP]K_{CAPO3}K_{O1}K_{O2}K_{O3}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{23c}+$$
$$[R]^2[RNAP]K_{O1}K_{O2}K_{P2}^{-1}K_{O3*}K_{13*}+ [R][R'][RNAP]K_{O2}K_{O3}K_{12}K_{O1}K_{P2}^{-1}+$$
$$[R][R'][RNAP][CAP]K_{CAPO3}K_{O1}K_{O2}K_{O3}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{23c}+$$
$$[R']^2[R''][RNAP][CAP]K_{CAPO3}K_{O2}K_{O3}K_{O1_I}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}+$$
$$[R][R''][CAP]K_{CAPO3}K_{O2}K_{O3_I}K_{CAP}^{-1}+$$
$$[R][R'][RNAP][CAP]K_{CAPO3}K_{O1}K_{O3}K_{12}K_{O2}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}+$$
$$[R'']^2[R'][K_{O1_I}K_{O2_I}K_{O3*}+ [R'][R'']K_{O2}K_{O1_I}K_{O3*}K_{23*}+ [RNAP]K_{O1}K_{O2}K_{P1}^{-1}K_{O3*}+$$
$$[R][R''][CAP]K_{CAPO3}K_{O1}K_{O2}K_{O3}K_{CAP}^{-1}K_{23c}+$$
$$[R'']^2[R'][CAP]K_{CAPO3}K_{O3}K_{O1_I}K_{O2_I}K_{CAP}^{-1}+ [R][R'][RNAP]K_{O2}K_{O3}K_{13}K_{O1}K_{P2}^{-1}+$$
$$[R'']^2[R][CAP]K_{O2}K_{O1_I}K_{CAP}^{-1}K_{O3*_I}+ [R']^2[R''][RNAP]K_{O2}K_{O1_I}K_{P2}^{-1}K_{O3*}+$$
$$[R][R'][RNAP]K_{O2}K_{P1}^{-1}K_{O3*}+ [R][R'][RNAP]K_{O2}K_{P1}^{-1}K_{O3*}+$$
$$[R]^2[RNAP][CAP]K_{CAPO3}K_{O1}K_{O2}K_{O3}K_{12}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}+$$
$$[R][R'][RNAP]K_{O1}K_{12}K_{O2}K_{P2}^{-1}K_{O3*}+ [R][R'']K_{O2}K_{O1_I}K_{O3*}K_{23*}+$$
$$[R][R'']K_{O2}K_{13}K_{O1}K_{O3}+ [R'']^2[R][RNAP]K_{O1}K_{O2_I}K_{O3_I}K_{P1}^{-1}+$$
$$[R][R'']K_{O1_I}K_{O3*}+ [R]^2[RNAP]K_{O1}K_{O2}K_{P1}^{-1}K_{O3*}K_{13*}+$$
$$[R][R'][RNAP]K_{O2}K_{O3}K_{12}K_{O1}K_{P1}^{-1}+ [R][R'']CAP]K_{CAPO3}K_{O3}K_{O1_I}K_{CAP}^{-1}+$$
$$[R'']^2[R'][CAP]K_{CAPO3}K_{O2}K_{O1_I}K_{O3_I}K_{CAP}^{-1}+ [R]^2[RNAP]K_{O1}K_{O2}K_{12}K_{P2}^{-1}K_{O3*}+$$
$$[R][R'][RNAP]K_{O1}K_{O3}K_{23}K_{O2}K_{P2}^{-1}+$$
$$[R][R''][RNAP][CAP]K_{CAPO3}K_{O3}K_{12}K_{O1}K_{O2}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}+$$
$$[R'']^2[R]K_{O1}K_{O2_I}K_{O3*_I}+ [R][R'']K_{O1}K_{O2}K_{O3*}K_{23*}+$$
$$[R][R'']K_{O1}K_{23}K_{O2}K_{O3}+ [R][R'][R''][RNAP][CAP]K_{O1}K_{O2_I}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*}+$$
$$[R][R'][R''][RNAP][CAP]K_{O1}K_{O2_I}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*}+$$

$$[R][R'][RNAP]K_{O2}K_{O3}K_{13}K_{O1}K_{P1}^{-1}+ [R]^2K_{O1}K_{O2}+$$
$$[R']^2[R'']RNAP]K_{O2}K_{O1_I}K_{P1}^{-1}K_{O3*}+ [R]^2[RNAP]K_{O1}K_{O2}K_{O3}K_{23}K_{P2}^{-1}+$$
$$[R][R'][RNAP]K_{O1}K_{12}K_{O2}K_{P1}^{-1}K_{O3*}+ [R][R'][R'']RNAP]K_{O2}K_{O1}K_{P2}^{-1}K_{O3*_I}+$$
$$[R][R'][R'']RNAP]K_{O1}K_{O2}K_{P2}^{-1}K_{O3*_I}+ [R][R'][RNAP]K_{O1}K_{O2}K_{P2}^{-1}K_{O3*}K_{23*}+$$
$$[R][R'][RNAP]K_{O1}K_{O2}K_{23}K_{O3}K_{P2}^{-1}+ [R][R'']RNAP]K_{12}K_{O1}K_{O2}K_{P2}^{-1}K_{O3*}+$$
$$[R]^2[RNAP][CAP]K_{O1}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*}+ [R]^2[RNAP]K_{O1}K_{O2}K_{12}K_{P1}^{-1}K_{O3*}+$$
$$[R][R'][RNAP]K_{O1}K_{O3}K_{23}K_{O2}K_{P1}^{-1}+ K_{O1}K_{O2}K_{O3*}+$$
$$[R][R'][CAP]K_{CAPO3}K_{O2}K_{O3}K_{O1}K_{CAP}^{-1}K_{13c}+ [R][R'][R'']RNAP]K_{O3}K_{O1}K_{O2_I}K_{P2}^{-1}+$$
$$[R][R'][R'']RNAP]K_{O1}K_{O3}K_{O2_I}K_{P2}^{-1}+ [CAP]K_{CAPO3}K_{O1}K_{O2}K_{O3}K_{CAP}^{-1}+$$
$$[R][R']K_{O2}K_{O3*}+ [R][R']K_{O2}K_{O3*}+ [R]^2[RNAP]K_{O1}K_{O2}K_{O3}K_{23}K_{P1}^{-1}+$$
$$[R][R'][CAP]K_{O2}K_{12}K_{O1}K_{CAP}^{-1}K_{O3*}+$$
$$[R'']^2[R][RNAP][CAP]K_{O1_I}K_{O2_I}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*}+$$
$$[R'']^2[R]K_{O1}K_{O2_I}K_{O3_I}+ [R][R'][CAP]K_{CAPO3}K_{O3}K_{O2}K_{CAP}^{-1}+$$
$$[R][R'][CAP]K_{CAPO3}K_{O2}K_{O3}K_{CAP}^{-1}+ [R][R'][R'']RNAP]K_{O2}K_{O1}K_{P1}^{-1}K_{O3*_I}+$$
$$[R][R'][R'']RNAP]K_{O1}K_{O2}K_{P1}^{-1}K_{O3*_I}+ [R][R'][RNAP]K_{O1}K_{O2}K_{P1}^{-1}K_{O3*}K_{23*}+$$
$$[R][R'][RNAP]K_{O1}K_{O2}K_{23}K_{O3}K_{P1}^{-1}+ [R][R'']RNAP]K_{12}K_{O1}K_{O2}K_{P1}^{-1}K_{O3*}+$$
$$[R'']^2[R][RNAP]K_{O2}K_{O1_I}K_{P2}^{-1}K_{O3*_I}+ [R]^2K_{O1}K_{O2}K_{O3*}K_{13*}+$$
$$[R][R']K_{O2}K_{O3}K_{12}K_{O1}+ [R]^2[RNAP]K_{O1}K_{O3}K_{P2}^{-1}+$$
$$[R][R'][R'']RNAP]K_{O2}K_{O1}K_{O3_I}K_{P2}^{-1}+ [R][R'][R'']RNAP]K_{O1}K_{O2}K_{O3_I}K_{P2}^{-1}+$$
$$[R][R'][CAP]K_{CAPO3}K_{O1}K_{O3}K_{O2}K_{CAP}^{-1}K_{23c}+ [R][R']K_{O2}K_{O3}K_{13}K_{O1}+$$
$$[R][R'][R'']RNAP]K_{O3}K_{O1}K_{O2_I}K_{P1}^{-1}+ [R][R'][R'']RNAP]K_{O1}K_{O3}K_{O2_I}K_{P1}^{-1}+$$
$$[R']^2[R'']K_{O2}K_{O1_I}K_{O3*}+ [R]^2[CAP]K_{CAPO3}K_{O1}K_{O2}K_{O3}K_{CAP}^{-1}K_{23c}+$$
$$[R][R']K_{O1}K_{12}K_{O2}K_{O3*}+ [R'']^2[R][RNAP]K_{O3}K_{O1_I}K_{O2_I}K_{P2}^{-1}+$$
$$[R][R'][CAP]K_{CAPO3}K_{O1}K_{O2}K_{O3}K_{CAP}^{-1}K_{23c}+$$
$$[R']^2[R''][CAP]K_{CAPO3}K_{O2}K_{O3}K_{O1_I}K_{CAP}^{-1}+$$
$$[R][R'][CAP]K_{CAPO3}K_{O1}K_{O3}K_{12}K_{O2}K_{CAP}^{-1}+ [R][R'][RNAP]K_{O2}K_{O1}K_{P2}^{-1}K_{O3*}K_{13*}+$$

$$[R'']^2[R][RNAP]K_{O2}K_{O1_I}K_{P1}^{-1}K_{O3*_I}+ [R]^2[RNAP]K_{O1}K_{O3}K_{P1}^{-1}+$$
$$[R][R'][R''][RNAP]K_{O2}K_{O1}K_{O3_I}K_{P1}^{-1}+ [R][R'][R''][RNAP]K_{O1}K_{O2}K_{O3_I}K_{P1}^{-1}+$$
$$[R]^2K_{O1}K_{O2}K_{12}K_{O3*}+ [R][R']K_{O1}K_{O3}K_{23}K_{O2}+$$
$$[R'']^2[R][RNAP]K_{O2}K_{O1_I}K_{O3_I}K_{P2}^{-1}+ [R]^2[CAP]K_{CAPO3}K_{O1}K_{O2}K_{O3}K_{12}K_{CAP}^{-1}+$$
$$[R'']^2[R][RNAP][CAP]K_{CAPO3}K_{O1}K_{O2_I}K_{O3_I}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}+$$
$$[R]^2K_{O1}K_{O2}K_{O3}K_{23}+$$
$$[R][R'][RNAP][CAP]K_{CAPO3}K_{O2}K_{O3}K_{12}K_{O1}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}+$$
$$[R][R'][R'']K_{O2}K_{O1}K_{O3*_I}+ [R][R'][R'']K_{O1}K_{O2}K_{O3*_I}+$$
$$[R'']^2[R][RNAP]K_{O3}K_{O1_I}K_{O2_I}K_{P1}^{-1}+ [R][R']K_{O1}K_{O2}K_{O3*}K_{23*}+$$
$$[R][R']K_{O1}K_{O2}K_{23}K_{O3}+$$
$$[R']^2[R][RNAP][CAP]K_{O1}K_{O2}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*}+$$
$$[R']^2[R][RNAP][CAP]K_{O2}K_{O1}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*}+$$
$$[R']^2[R][RNAP][CAP]K_{O1}K_{O2}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*}+ [R][R'']K_{12}K_{O1}K_{O2}K_{O3*}+$$
$$[R][R'][RNAP]K_{O2}K_{O1}K_{P1}^{-1}K_{O3*}K_{13*}+$$
$$[R]^2[RNAP][CAP]K_{O2}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*}+$$
$$[R][R''][CAP]K_{CAPO3}K_{O3}K_{12}K_{O1}K_{O2}K_{CAP}^{-1}+$$
$$[R][R'][RNAP]K_{O2}K_{12}K_{O1}K_{P2}^{-1}K_{O3*}+$$
$$[R][R'][R''][CAP]K_{O1}K_{O2_I}K_{CAP}^{-1}K_{O3*}+ [R][R'][R''][CAP]K_{O1}K_{O2_I}K_{CAP}^{-1}K_{O3*}+$$
$$[R'']^2[R][RNAP]K_{O2}K_{O1_I}K_{O3_I}K_{P1}^{-1}+ [R][R'][R'']K_{O3}K_{O1}K_{O2_I}+$$
$$[R][R'][R'']K_{O1}K_{O3}K_{O2_I}+ [R']^2[R][RNAP]K_{O3}K_{O1}K_{O2}K_{P2}^{-1}+$$
$$[R']^2[R][RNAP]K_{O2}K_{O1}K_{O3}K_{P2}^{-1}+ [R']^2[R][RNAP]K_{O1}K_{O2}K_{O3}K_{P2}^{-1}+$$
$$[R]^2[CAP]K_{O1}K_{CAP}^{-1}K_{O3*}+ [R'']^2[R]K_{O2}K_{O1_I}K_{O3*_I}+$$
$$[R]^2[RNAP]K_{O2}K_{O3}K_{P2}^{-1}+$$
$$[R][R'][R''][RNAP][CAP]K_{O2}K_{O1_I}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*}+$$
$$[R][R'][R''][RNAP][CAP]K_{O2}K_{O1_I}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*}+ [R]^2K_{O1}K_{O3}+$$
$$[R][R'][R'']K_{O2}K_{O1}K_{O3_I}+ [R][R'][R'']K_{O1}K_{O2}K_{O3_I}+$$

$$[R][R'][RNAP]K_{O2}K_{12}K_{O1}K_{P1}^{-1}K_{O3*}+ [R'']^2[R][CAP]K_{O1_I}K_{O2_I}K_{CAP}^{-1}K_{O3*}+$$
$$[R'']^2[R]K_{O3}K_{O1_I}K_{O2_I}+ [R']^2[R][RNAP]K_{O3}K_{O1}K_{O2}K_{P1}^{-1}+$$
$$[R']^2[R][RNAP]K_{O2}K_{O1}K_{O3}K_{P1}^{-1}+ [R']^2[R][RNAP]K_{O1}K_{O2}K_{O3}K_{P1}^{-1}+$$
$$[R][R'][R''][RNAP]K_{O3}K_{O2}K_{O1_I}K_{P2}^{-1}+ [R][R'][R''][RNAP]K_{O2}K_{O3}K_{O1_I}K_{P2}^{-1}+$$
$$[R][R']K_{O2}K_{O1}K_{O3*}K_{13*}+ [R]^2[RNAP]K_{O2}K_{O3}K_{P1}^{-1}+$$
$$[R][R'][R''][RNAP][CAP]K_{CAPO3}K_{O3}K_{O1}K_{O2_I}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}+$$
$$[R][R'][R''][RNAP][CAP]K_{CAPO3}K_{O1}K_{O3}K_{O2_I}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}+$$
$$[R]^2[R''][RNAP][CAP]K_{O1}K_{O2}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*_I}+ [R'']^2[R]K_{O2}K_{O1_I}K_{O3_I}+$$
$$[R][R'][R''][RNAP]K_{O1}K_{O2_I}K_{P2}^{-1}K_{O3*}+ [R][R'][R''][RNAP]K_{O1}K_{O2_I}K_{P2}^{-1}K_{O3*}+$$
$$[R][R'][RNAP]K_{O1}K_{O2}K_{P2}^{-1}K_{O3*}K_{23*}+$$
$$[R]^2[RNAP][CAP]K_{CAPO3}K_{O1}K_{O3}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}+$$
$$[R][R'][R''][RNAP][CAP]K_{CAPO3}K_{O2}K_{O1}K_{O3_I}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}+$$
$$[R][R'][R''][RNAP][CAP]K_{CAPO3}K_{O1}K_{O2}K_{O3_I}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}+$$
$$[R][R'][R''][RNAP]K_{O3}K_{O2}K_{O1_I}K_{P1}^{-1}+ [R][R'][R''][RNAP]K_{O2}K_{O3}K_{O1_I}K_{P1}^{-1}+$$
$$[R]^2[RNAP]K_{O1}K_{O2}K_{P2}^{-1}K_{O3*}K_{23*}+ [R][R']K_{O2}K_{12}K_{O1}K_{O3*}+$$
$$[R'']^2[R][CAP]K_{CAPO3}K_{O1}K_{O2_I}K_{O3_I}K_{CAP}^{-1}+ [R]^2[RNAP]K_{O1}K_{P2}^{-1}K_{O3*}+$$
$$[R][R'][CAP]K_{CAPO3}K_{O2}K_{O3}K_{12}K_{O1}K_{CAP}^{-1}+$$
$$[R'']^2[R][RNAP][CAP]K_{CAPO3}K_{O3}K_{O1_I}K_{O2_I}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}+$$
$$[R']^2[R][CAP]K_{O1}K_{O2}K_{CAP}^{-1}K_{O3*}+ [R']^2[R][CAP]K_{O2}K_{O1}K_{CAP}^{-1}K_{O3*}+$$
$$[R']^2[R][CAP]K_{O1}K_{O2}K_{CAP}^{-1}K_{O3*}+ [R][R'][R''][RNAP]K_{O1}K_{O2_I}K_{P1}^{-1}K_{O3*}+$$
$$[R][R'][R''][RNAP]K_{O1}K_{O2_I}K_{P1}^{-1}K_{O3*}+ [R']^2[R]K_{O3}K_{O1}K_{O2}+$$
$$[R']^2[R]K_{O2}K_{O1}K_{O3}+ [R']^2[R]K_{O1}K_{O2}K_{O3}+$$
$$[R][R'][RNAP]K_{O1}K_{O2}K_{P1}^{-1}K_{O3*}K_{23*}+ [R]^2[CAP]K_{O2}K_{CAP}^{-1}K_{O3*}+$$
$$[R'']^2[R][RNAP]K_{O1_I}K_{O2_I}K_{P2}^{-1}K_{O3*}+ [R]^2K_{O2}K_{O3}+$$
$$[R'']^2[R][RNAP][CAP]K_{CAPO3}K_{O2}K_{O1_I}K_{O3_I}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}+$$
$$[R]^2[RNAP]K_{O1}K_{O2}K_{P1}^{-1}K_{O3*}K_{23*}+ [R]^2[RNAP]K_{O1}K_{P1}^{-1}K_{O3*}+$$

$$[R][R'][R''][CAP]K_{O2}K_{O1_I}K_{CAP}^{-1}K_{O3*}+ [R][R'][R''][CAP]K_{O2}K_{O1_I}K_{CAP}^{-1}K_{O3*}+$$
$$[R'']^2[R][RNAP]K_{O1_I}K_{O2_I}K_{P1}^{-1}K_{O3*}+ [R][R'][R'']K_{O3}K_{O2}K_{O1_I}+$$
$$[R][R'][R'']K_{O2}K_{O3}K_{O1_I}+$$
$$[R']^2[R][RNAP][CAP]K_{CAPO3}K_{O3}K_{O1}K_{O2}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}+$$
$$[R']^2[R][RNAP][CAP]K_{CAPO3}K_{O2}K_{O1}K_{O3}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}+$$
$$[R']^2[R][RNAP][CAP]K_{CAPO3}K_{O1}K_{O2}K_{O3}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}+$$
$$[R][R'][R'']K_{O1}K_{O2_I}K_{O3*}+ [R][R'][R'']K_{O1}K_{O2_I}K_{O3*}+ [R][R']K_{O1}K_{O2}K_{O3*}K_{23*}+$$
$$[R]^2[RNAP][CAP]K_{CAPO3}K_{O2}K_{O3}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}+$$
$$[R][R'][R''][CAP]K_{CAPO3}K_{O3}K_{O1}K_{O2_I}K_{CAP}^{-1}+$$
$$[R][R'][R''][CAP]K_{CAPO3}K_{O1}K_{O3}K_{O2_I}K_{CAP}^{-1}+ [R]^2[R''][CAP]K_{O1}K_{O2}K_{CAP}^{-1}K_{O3*_I}+$$
$$[R']^2[R][RNAP]K_{O1}K_{O2}K_{P2}^{-1}K_{O3*}+ [R']^2[R][RNAP]K_{O2}K_{O1}K_{P2}^{-1}K_{O3*}+$$
$$[R']^2[R][RNAP]K_{O1}K_{O2}K_{P2}^{-1}K_{O3*}+ [R]^2K_{O1}K_{O2}K_{O3*}K_{23*}+$$
$$[R]^2[RNAP]K_{O2}K_{P2}^{-1}K_{O3*}+ [R]^2K_{O1}K_{O3*}+ [R]^2[CAP]K_{CAPO3}K_{O1}K_{O3}K_{CAP}^{-1}+$$
$$[R][R'][R''][CAP]K_{CAPO3}K_{O2}K_{O1}K_{O3_I}K_{CAP}^{-1}+$$
$$[R][R'][R''][CAP]K_{CAPO3}K_{O1}K_{O2}K_{O3_I}K_{CAP}^{-1}+$$
$$[R][R'][R''][RNAP][CAP]K_{CAPO3}K_{O3}K_{O2}K_{O1_I}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}+$$
$$[R][R'][R''][RNAP][CAP]K_{CAPO3}K_{O2}K_{O3}K_{O1_I}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}+$$
$$[R'']^2[R]K_{O1_I}K_{O2_I}K_{O3*}+ [R']^2[R][RNAP]K_{O1}K_{O2}K_{P1}^{-1}K_{O3*}+$$
$$[R']^2[R][RNAP]K_{O2}K_{O1}K_{P1}^{-1}K_{O3*}+ [R']^2[R][RNAP]K_{O1}K_{O2}K_{P1}^{-1}K_{O3*}+$$
$$[R'']^2[R][CAP]K_{CAPO3}K_{O3}K_{O1_I}K_{O2_I}K_{CAP}^{-1}+ [R][R'][R''][RNAP]K_{O2}K_{O1_I}K_{P2}^{-1}K_{O3*}+$$
$$[R][R'][R''][RNAP]K_{O2}K_{O1_I}K_{P2}^{-1}K_{O3*}+ [R]^2[RNAP]K_{O2}K_{P1}^{-1}K_{O3*}+$$
$$[R'']^2[R][CAP]K_{CAPO3}K_{O2}K_{O1_I}K_{O3_I}K_{CAP}^{-1}+$$
$$[R]^2[R''][RNAP][CAP]K_{O1}K_{O2_I}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*}+$$
$$[R][R'][R''][RNAP]K_{O2}K_{O1_I}K_{P1}^{-1}K_{O3*}+ [R][R'][R''][RNAP]K_{O2}K_{O1_I}K_{P1}^{-1}K_{O3*}+$$
$$[R]^2[R''][RNAP]K_{O1}K_{O2}K_{P2}^{-1}K_{O3*_I}+ [R']^2[R]K_{O1}K_{O2}K_{O3*}+ [R']^2[R]K_{O2}K_{O1}K_{O3*}+$$
$$[R']^2[R]K_{O1}K_{O2}K_{O3*}+ [R]^2[R''][RNAP]K_{O1}K_{O3}K_{O2_I}K_{P2}^{-1}+$$

$$[R']^2[R][CAP]K_{CAPO3}K_{O3}K_{O1}K_{O2}K_{CAP}^{-1}+ [R']^2[R][CAP]K_{CAPO3}K_{O2}K_{O1}K_{O3}K_{CAP}^{-1}+$$
$$[R']^2[R][CAP]K_{CAPO3}K_{O1}K_{O2}K_{O3}K_{CAP}^{-1}+ [R]^2K_{O2}K_{O3*}+$$
$$[R]^2[CAP]K_{CAPO3}K_{O2}K_{O3}K_{CAP}^{-1}+ [R]^2[R'']RNAP]K_{O1}K_{O2}K_{P1}^{-1}K_{O3*_I}+$$
$$[R]^2[R'']RNAP]K_{O1}K_{O2}K_{O3_I}K_{P2}^{-1}+ [R]^2[R'']RNAP]K_{O1}K_{O3}K_{O2_I}K_{P1}^{-1}+$$
$$[R][R'][R'']K_{O2}K_{O1_I}K_{O3*}+ [R][R'][R'']K_{O2}K_{O1_I}K_{O3*}+$$
$$[R][R'][R'']CAP]K_{CAPO3}K_{O3}K_{O2}K_{O1_I}K_{CAP}^{-1}+$$
$$[R][R'][R'']CAP]K_{CAPO3}K_{O2}K_{O3}K_{O1_I}K_{CAP}^{-1}+$$
$$[R]^2[R'']RNAP]K_{O1}K_{O2}K_{O3_I}K_{P1}^{-1}+ [R]^2[R'']K_{O1}K_{O2}K_{O3*_I}+$$
$$[R]^2[R'][RNAP][CAP]K_{O2}K_{O1}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*}+$$
$$[R]^2[R'][RNAP][CAP]K_{O1}K_{O2}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*}+$$
$$[R]^2[R'][RNAP][CAP]K_{O1}K_{O2}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*}+$$
$$[R]^2[R'']CAP]K_{O1}K_{O2_I}K_{CAP}^{-1}K_{O3*}+ [R]^2[R'']K_{O1}K_{O3}K_{O2_I}+$$
$$[R]^2[R'][RNAP]K_{O2}K_{O3}K_{O1}K_{P2}^{-1}+ [R]^2[R'][RNAP]K_{O1}K_{O3}K_{O2}K_{P2}^{-1}+$$
$$[R]^2[R'][RNAP]K_{O1}K_{O2}K_{O3}K_{P2}^{-1}+$$
$$[R]^2[R'']RNAP][CAP]K_{O2}K_{O1_I}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*}+ [R]^2[R'']K_{O1}K_{O2}K_{O3_I}+$$
$$[R]^2[R'][RNAP]K_{O2}K_{O3}K_{O1}K_{P1}^{-1}+ [R]^2[R'][RNAP]K_{O1}K_{O3}K_{O2}K_{P1}^{-1}+$$
$$[R]^2[R'][RNAP]K_{O1}K_{O2}K_{O3}K_{P1}^{-1}+ [R]^2[R'']RNAP][CAP]K_{O2}K_{O1_I}K_{O3*}+$$
$$[R]^2[R'']RNAP][CAP]K_{CAPO3}K_{O1}K_{O3}K_{O2_I}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}+$$
$$[R]^2[R'']RNAP]K_{O1}K_{O2_I}K_{P2}^{-1}K_{O3*}+$$
$$[R]^2[R'']RNAP][CAP]K_{CAPO3}K_{O1}K_{O2}K_{O3_I}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}+$$
$$[R]^2[R'']RNAP]K_{O2}K_{O3}K_{O1_I}K_{P1}^{-1}+ [R]^2[R'][CAP]K_{O2}K_{O1}K_{CAP}^{-1}K_{O3*}+$$
$$[R]^2[R'][CAP]K_{O1}K_{O2}K_{CAP}^{-1}K_{O3*}+ [R]^2[R'][CAP]K_{O1}K_{O2}K_{CAP}^{-1}K_{O3*}+$$
$$[R]^2[R'']RNAP]K_{O1}K_{O2_I}K_{P1}^{-1}K_{O3*}+ [R]^2[R']K_{O2}K_{O3}K_{O1}+ [R]^2[R']K_{O1}K_{O3}K_{O2}+$$
$$[R]^2[R']K_{O1}K_{O2}K_{O3}+ [R]^2[R'']CAP]K_{O2}K_{O1_I}K_{CAP}^{-1}K_{O3*}+$$
$$[R]^2[R'']K_{O2}K_{O3}K_{O1_I}+ [R]^2[R'][RNAP][CAP]K_{CAPO3}K_{O2}K_{O3}K_{O1}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}+$$
$$[R]^2[R'][RNAP][CAP]K_{CAPO3}K_{O1}K_{O3}K_{O2}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}+$$

$$[R]^2[R'][RNAP][CAP]K_{CAPO3}K_{O1}K_{O2}K_{O3}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}+$$
$$[R]^2[R'']K_{O1}K_{O2_I}K_{O3*}+ [R]^2[R''][CAP]K_{CAPO3}K_{O1}K_{O3}K_{O2_I}K_{CAP}^{-1}+$$
$$[R]^2[R'][RNAP]K_{O2}K_{O1}K_{P2}^{-1}K_{O3*}+ [R]^2[R'][RNAP]K_{O1}K_{O2}K_{P2}^{-1}K_{O3*}+$$
$$[R]^2[R'][RNAP]K_{O1}K_{O2}K_{P2}^{-1}K_{O3*}+ [R]^2[R''][CAP]K_{CAPO3}K_{O1}K_{O2}K_{O3_I}K_{CAP}^{-1}+$$
$$[R]^2[R''][RNAP][CAP]K_{CAPO3}K_{O2}K_{O3}K_{O1_I}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}+$$
$$[R]^2[R'][RNAP]K_{O2}K_{O1}K_{P1}^{-1}K_{O3*}+ [R]^2[R'][RNAP]K_{O1}K_{O2}K_{P1}^{-1}K_{O3*}+$$
$$[R]^2[R'][RNAP]K_{O1}K_{O2}K_{P1}^{-1}K_{O3*}+ [R]^2[R''][RNAP]K_{O2}K_{O1_I}K_{P2}^{-1}K_{O3*}+$$
$$[R]^2[R''][RNAP]K_{O2}K_{O1_I}K_{P1}^{-1}K_{O3*}+ [R]^2[R']K_{O2}K_{O1}K_{O3*}+ [R]^2[R']K_{O1}K_{O2}K_{O3*}+$$
$$[R]^2[R']K_{O1}K_{O2}K_{O3*}+ [R]^2[R'][CAP]K_{CAPO3}K_{O2}K_{O3}K_{O1}K_{CAP}^{-1}+$$
$$[R]^2[R'][CAP]K_{CAPO3}K_{O1}K_{O3}K_{O2}K_{CAP}^{-1}+ [R]^2[R'][CAP]K_{CAPO3}K_{O1}K_{O2}K_{O3}K_{CAP}^{-1}+$$
$$[R]^2[R'']K_{O2}K_{O1_I}K_{O3*}+ [R]^2[R''][CAP]K_{CAPO3}K_{O2}K_{O3}K_{O1_I}K_{CAP}^{-1}+$$
$$[RNAP][CAP]K_{O1}K_{O2}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}K_{O3*}+ [RNAP]K_{O1}K_{O2}K_{O3}K_{P2}^{-1}+$$
$$[RNAP]K_{O1}K_{O2}K_{O3}K_{P1}^{-1}+ [CAP]K_{O1}K_{O2}K_{CAP}^{-1}K_{O3*}+ K_{O1}K_{O2}K_{O3}+$$
$$[RNAP][CAP]K_{CAPO3}K_{O1}K_{O2}K_{O3}K_{P1}^{-1}K_{CAP}^{-1}K_{coop}^{-1}+ [RNAP]K_{O1}K_{O2}K_{P2}^{-1}K_{O3*}+$$
$$[RNAP]K_{O1}K_{O2}K_{P1}^{-1}K_{O3*}+ K_{O1}K_{O2}K_{O3*}+ [CAP]K_{CAPO3}K_{O1}K_{O2}K_{O3}K_{CAP}^{-1}.$$