ABSTRACT

Title of Document: Differentiating a Landscape of Lac Repressor Mediated DNA Loops Using FRET

Aaron Raymond Haeusler, Doctor of Philosophy, 2011

Directed By: Associate Professor, Jason Kahn, Department of Chemistry and Biochemistry

The variety of DNA loops formed by a protein such as Lac Repressor (LacI), predicted to populate multiple loop topologies and geometries, increases the uncertainty of models that have not been experimentally verified. We created a method for evaluating the multitude of LacI•DNA looped states using FRET with fluorophore position variants (FPVs) on a landscape DNA constructs with dual operator phasing variants bracketing an intrinsically curved DNA sequence. The FRET signature for each FPV corresponds to a specific LacI•DNA loop topology, and when applied to the landscape of constructs we could systematically explore DNA sequence dependent LacI mediated DNA looping. The FRET efficiencies calculated for each FPV landscape revealed the first unambiguous detection of antiparallel (A1 and A2) and distributions among parallel loop (P1) and antiparallel loops. The FRET efficiency maximum of each FPV landscape was used to calculate the loop population distribution for each topology assuming a uniform population at the peaks. The differences in the observed peak FRET among FPV

landscapes led to an unavoidable estimate of loop populations with an extended LacI•DNA loop geometry, and also slightly extended LacI•DNA loop in the cases of antiparallel loops. The addition of saturating inducer to the FPV landscape demonstrated that IPTG•LacI•DNA loops have different properties than uninduced loops: they have increased instability, increased competition among previously energetically unfavorable states, and presumably have increased switching between specific and nonspecific LacI•DNA interactions. This multivariate data set confirms that rod mechanics models of DNA looping should also consider protein flexibility, and also demonstrates the importance of protein flexibility in modeling genetic regulatory circuits that are similar to the "hydrogen atom of gene regulation", the *lac* operon.

Differentiating a Landscape of Lac Repressor Mediated DNA Loops Using FRET

By

Aaron Raymond Haeusler

Dissertation submitted to the Faculty of the Graduate School of the

University of Maryland, College Park, in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy
2011

Advisory Committee:

Associate Professor Jason Kahn, Chair/Advisor
Professor Dorothy Beckett
Associate Professor Douglas Julin
Professor Steven Rokita
Professor Jonathan Dinman, Dean's Representative

Dedication

To my parents,

my siblings,

and my partner, Kim,

for all their generous support, patience, and encouragement throughout my career.

My success is as much theirs as it is mine.

Acknowledgements

I would first like to thank Dr. Jason Kahn. He allowed me to work independently and grow as an individual while providing me guidance whenever I requested it. Jason was always able to help me resolve any scientific question or experimental quandaries. He also provided me with many life-changing conversations that motivated me scientifically and challenged me conceptually. I will miss our coffee pot discussions.

I would like to thank the Department of Chemistry and Biochemistry for supporting me throughout Graduate School by allowing me to be a Teaching Assistant. They relieved me of a financial burden that would have prevented me from being able to fulfill a major life goal had it been carried alone. My teaching experiences at the university helped define me as a scientist and person, and liberated me from the monotony of experimentation.

It was a pleasure knowing the late Dr. Herman Kraybill; he was a guiding force and motivation for me. My experience with him will always remind me what is possible as a scientist and a person.

All of the members in Jason Kahn's lab have been integral in my development. When I originally joined the lab with previous experience in an organic lab, I had no idea even what 6X buffer meant. The long trip to my completion would not have been possible without the lab members that started graduate school with me: Kathy Goodson, and Sarah Sucayan, and especially Dan Gowetski who shared a similar love for science

and music. I would also like to thank the other Kahn lab members who assisted me as well: Jeffrey Leupp, Lucas Tricoli, and Xiaoyu Wang. Additionally, Dr. Patrick McTamney, a former member of Dr. Steven Rokita's lab with whom I began graduate school continues to help me develop experimentally and professionally. He is always a daily source of something humorous.

My Ph.D. committee members knew exactly what to say to motivate and challenge me. I still remember their remarks from Candidacy and my Independent Proposal, and their praises and criticisms were always appreciated. I think about those comments every time I prepare a presentation or paper. I could not have asked for a better experience.

My parents, Adam and Tammi, have supported me throughout life, and have continued to help me stay grounded to my core values and beliefs. Although I now live further away, they have continued to support me as much as before and have helped me keep my head held high in even the most trying of times.

Finally, I would like to thank my partner Kim. She has been a rock that I can always rely on and support myself with. I am lucky to have her in my life.

Table of Contents

List of Figures

Figure page

List of Abbreviations

A	adenine
A_E	acceptor fluorophore placed externally relative to the operator and the central bend
A_I	donor fluorophore placed internally relative to operator and the central bend
A	antiparallel loop (as in A1 and A2)
C	cytosine
CIP	calf-intestinal phosphatase
DB	doubly bound
DNA	deoxyribose nucleic acid
D_E	donor fluorophore placed externally relative to the operator and the central bend
D_I	donor fluorophore placed internally relative to operator and the central bend
E	extended loop
EDTA	ethylenediaminetetraacetic acid
EMSA	electrophoretic mobility shift assay
EtBr	ethidium bromide
EtOH	ethanol
FPV	fluorophore position variant
FRET	Förster resonance energy transfer
G	guanine

IDT	Integrated DNA Technologies
IPTG	isopropyl-β-D-thiogalactopyranoside
LacI	lac repressor
NEB	New England Biolabs
O_1	primary operator
O_2	secondary operator
O_3	secondary operator
O_{sym}	symmetrical operator
P	parallel loop (as in P1 and P2)
PAGE	polyacrylamide gel electrophoresis
PNK	polynucleotide kinase
PCR	polymerase chain reaction
SB	singly bound
s-ns	specific-nonspecific loop
T	thymine
TBE	buffer containing tris, borate, and EDTA
TE	buffer containing tris and EDTA
WA	wraps away from central complex
WT	wraps towards central complex

Chapter 1

Introduction

1.1 The Discovery of DNA Structure and the Genetic Code

Halfway through the 20th Century our world of genetics evolved from the macroscopic understanding of Mendelian and Darwinian models for genetic inheritance to the submicroscopic level of an underlying structurally based genetic code. In 1952 deoxyribose nucleic acids (DNA) and not amino acids were identified as containing the fundamental components necessary for transferring genetic information between organisms. The four major DNA bases that formed polymers, shown in Figure 1.1A, always paired up stoichiometrically with Adenine (A) = Thymine (T) and Guanine (G) = Cytosine (C), but the relative proportions of each pair varied between organisms.

Monumental X-Ray crystallography work by Rosalind Franklin resolved the structure of DNA polymers, and it was evident to Watson and Crick that the two DNA polymers formed a regular helical structure. The structure of the right-handed DNA helix has two polymers running antiparallel to each other with specific base pair complementarity forming the major groove and sugar-phosphate backbones forming minor grooves shown in Figure

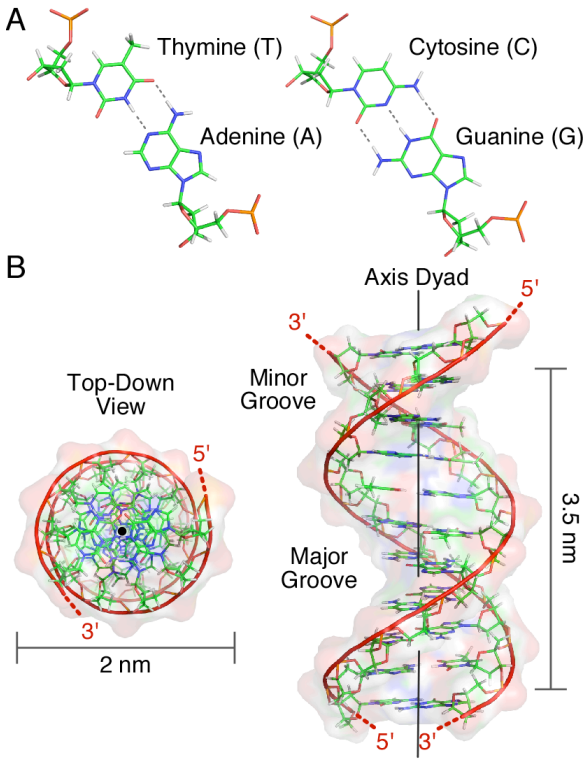


Figure 1.1: *The DNA double helix.* Modeled from ideal B-DNA. (A) Complementary Watson and Crick base pairing between G-C and A-T. (B) B-DNA shown from two angles to indicate the size of the right-handed helix.

1.1. The complementary base pairs stack in a plane that is 2 nm in diameter with a helical repeat around a central dyad axis every 10.45 base pairs or 3.5 nm.

The universal helical DNA structure laid the foundation for a definitive genetic code based on DNA sequences, and early in the 21st century a major milestone was reached with the sequencing of all the genes present in the human genome. This and other biochemical technological achievements have ushered in a new era of understanding genes and their influences on genetic traits. However, one of the greatest biochemical challenges remaining to be conquered this century will be the universal understanding of gene regulation, and therefore, controlling gene regulation. For what good is knowing a genetic code without understanding where, when and how it is accessed?

1.2 The *lac* Operon: Genetic Regulation by a Protein

Gene regulation at the DNA level requires intricate control of accessibility to genes that are generally tied up in DNA-protein structural elements and tightly compacted by supercoiling. For example, the contour length of DNA in an *E. coli* cell is 1600 μm, but remarkably occupies only 0.5 μm^3 in its nucleoid form (Bond, Peters et al. 2010). Therefore, a gene must be presented at an appropriate spatial location and time in the cell to promote RNA polymerase binding and transcription of essential genes to maintain cellular functions and respond to intracellular and extracellular environmental factors. Furthermore, many events that trigger a genetic signal must traverse through cascades of regulatory circuits before eliciting a genetic response. To effectively

understand and develop accurate models for global genetic responses, all the complexities of the local circuitry at each level of the regulatory cascade must be well defined.

The simplest genetic circuit, in its purest form, acts as a binary switch turning genes either off or on. Jacob and Monod first suggested a binary type of genetic switch for the Lactose (*lac*) operon in *E. coli* (Jacob and Monod 1961). They noticed that at high concentrations of lactose there was increased synthesis of the proteins ß-galactosidase, permease, and transacetylase that correspond to genes *LacZ*, *LacY*, and *LacA* respectively. They hypothesized that an agent acted to repress gene transcription (off), but the gene was induced in the presence of lactose (on). The agent behaving as a molecular switch was identified as a 152 kDA, homotetrameric protein called Lac Repressor (LacI) (Gilbert and Müller-Hill 1966; Müller-Hill, Crapo et al. 1968; Beyreuther, Adler et al. 1973; Farabaugh 1978). LacI negatively regulates the *lac* operon by binding a DNA operator site and restricting access of RNA polymerase to the promoter site as shown in Figure 1.2A (Gilbert and Müller-HillMüller-Hill 1967; Maizels 1973; Schlax, Capp et al. 1995). When *E. coli* is grown with lactose

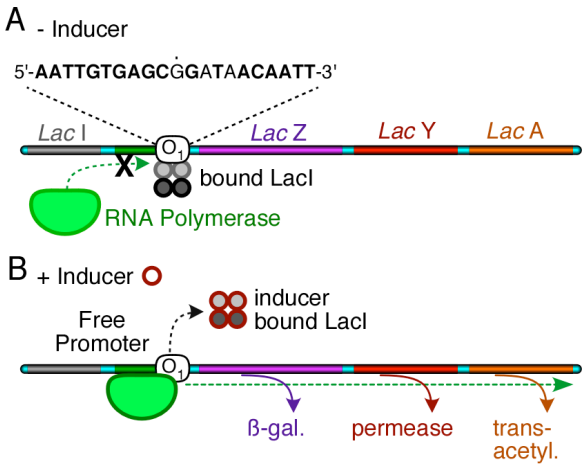


Figure 1.2: *The* lac *operon*. (A) In the absence of inducer, Lac Repressor, (LacI) is bound to the Operator (O_1) and overlaps the downstream polymerase promoter region. RNA is unable to bind which prevents the transcription of the three genes downstream of the promoter, *Lac* Z, Y, and A. The operator DNA sequence is shown with the symmetrical sequence shown in bold, and the dot above G represents the symmetry axis. (B) In the presence of an inducer, such as allolactose, LacI is bound with the inducer and allosterically releases from the O_1. RNA polymerase can bind the promoter region and transcribe the genes that make proteins essential for metabolizing lactose.

as the primary carbon source *ß*-galactosidase converts lactose → glucose + galactose + allolactose, and as allolactose concentrations increase in the cell LacI binds with this natural inducer and is allosterically affected (Monod, Wyman et al. 1965; Jobe and Bourgeois 1972). The conformational change induced by allolactose reduces the affinity of LacI for its operator, and because the LacI binding site overlaps the promoter region, the promoter becomes more accessible for RNA polymerase binding and therefore gene transcription increases as depicted in Figure 1.2B.

The natural *lac* operator consists of 21 base pairs that are pseudo-symmetric shown in Figure 1.2A, but only 17 of the base pairs are necessary for full binding affinity (Gilbert and Maxam 1973; Bahl, Wu et al. 1977). The affinity of LacI for the pseudo-symmetric operator sequence is incredibly strong ($K_d \sim 10^{-13}$ M), and it has a binding rate to non-specific DNA that is 10^2 to 10^3 orders of magnitude faster than the diffusion-controlled limit (Bourgeois and Riggs 1970). Any reduction to the symmetry of the DNA operator sequence also reduces the affinity of LacI since specificity is being compromised (Gilbert and Maxam 1973). In contrast, increasing the symmetry of the operator to an ideal inverted repeat with a central dyad axis, O_{sym}, (5'-AATTGTGAGC:GCTCACAATT-3', with the colon indicating the center of symmetry) increased the affinity of LacI approximately 10-fold relative to the natural operator sequence (Sadler, Sasmor et al. 1983; Simons, Tils et al. 1984).

LacI has been proposed to bind nonspecifically and search DNA through one-dimensional diffusion or hopping until locking onto the operator DNA (Berg and Blomberg 1976; Loverdo, Benichou et al. 2009; Furini, Domene et al. 2010). To put this

into perspective, there are typically less than 10 molecules of LacI in an *E. coli* cell (Gilbert and Müller-Hill 1966), and the operator accounts for ~ 4.3 x 10^{-4}% of the total DNA in the *E. coli* genome (Blattner, Plunkett et al. 1997). It is estimated that these 10 LacI molecules spend 98% of their life bound to DNA, and a majority of that time LacI is nonspecifically interacting with DNA (Lin and Riggs 1975; Elf, Li et al. 2007), yet remarkably LacI has 95% repression efficiency *in vivo* (Oehler, Eismann et al. 1990). This led some to believe that there might be another mechanism underlying the high efficiency of repression, which could possibly be due to secondary tetrameric LacI•DNA interactions (Kania and Müller-Hil 1977).

1.3 Enhanced Repression Through the Formation of Protein-DNA Loops

The two pseudo-operators were discovered, O_2 and O_3, which are located proximal to the pseudo-symmetrical operator (O_1) shown in Figure 1.3 (Reznikoff, Winter et al. 1974; Gilbert 1975). Initially, O_2 and O_3 were considered to be functionless because DNA fragments that contained all three operators had similar LacI binding affinities to a fragment that contained only O_1 (Pfahl, Gulde et al. 1979). However, *in vitro* experiments indicated that LacI could bind two separate DNA operator sequences at once and form a sandwich complex (O'Gorman, Dunaway et al. 1980; Culard and Maurizot 1981), and when two operators were placed on the same DNA strand LacI could bind both operators simultaneously and form a LacI•DNA looped complex (Kramer, Niemoller et al. 1987; Kramer, Amouyal et al. 1988).

The specific regulatory role for the pseudo operators *in vivo* was less certain until Müller-Hill measured the repression efficiency of the lac operon by destroying combinations of the three operators represented in Figure 1.3A (Oehler, Eismann et al. 1990). Each destruction of the pseudo operator(s) decreased repression and their roles were rewritten as "auxiliary operators" that are necessary for full repression of the *lac* operon. Furthermore, when both pseudo operators were destroyed repression was drastically decreased, which suggested the formation of LacI•DNA loops. Additionally, the classic signal for looping, observed as the periodicity of repression efficiency, is observed when the interoperator distances are varied (Mossing and Record 1986; Oehler, Amouyal et al. 1994; Becker, Kahn et al. 2005; Bond, Peters et al. 2010). These results established that LacI•DNA looping is an essential mechanism in the repression of the *lac* operon, and based on affinities and occupancy at each operator site, looping primarily occurs between O1 – O2 and O1 – O3 as shown in Figure 1.3B (Borowiec, Zhang et al. 1987; Flashner and Gralla 1988; Sasse-Dwight and Gralla 1988; Oehler, Eismann et al. 1990).

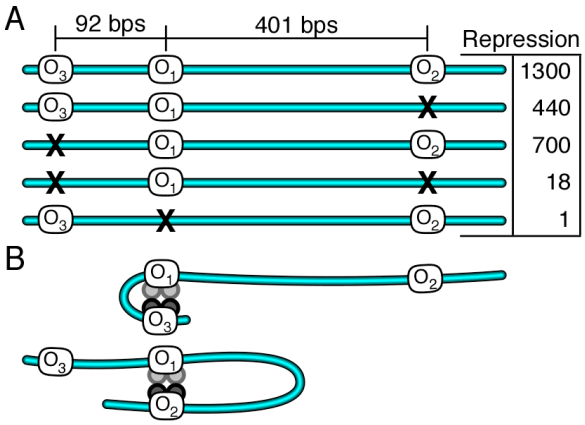


Figure 1.3: *The cooperativity of repression for the three operators of the* lac *operon* (Oehler, Eismann et al. 1990). Modified from the original paper. (A) Destruction of the primary operators (O_1) and/or secondary operators (O_2 and O_3) decreases repression. Destroying the primary operator decreases repression more than 1000 fold, but the secondary operators are necessary for full repression. (B) LacI binding to the secondary operators enhances repression by increasing the relative local [O_1] and looping intervening DNA to form a LacI•DNA loop.

DNA looping was being described in other regulatory systems, concurrently with the lac operon, as a biological means to link transcriptional regulatory events that are often spatially separated along the DNA, such as the L-arabinose operon in E. coli, the Xenopus rRNA gene promoter in *X. laevis*, and NTrC transcriptional activation in enteric bacteria (Hahn, Dunn et al. 1984; Dunaway and Droge 1989; Wedel, Weiss et al. 1990). The enhancement in regulatory circuits via DNA looping was ascribed to local concentration effects that increase the probability of molecular collisions by reducing the conformational space available (Shore and Baldwin 1983). For example, if a single tetrameric LacI were to bind two separate operators in a sphere the size of an E. coli cell (~2.0 x 10-12 cm3) compared to the sphere with a diameter corresponding to the distance between O1 – O2 (1 x 10-14 cm3) the latter would increase the concentration ~ 200 fold compared to the former (Müller-Hill 1998). Of course this neglects the contributions of LacI flexibility and DNA flexibility, which can potentiate or prevent looping.

In prokaryotes the majority of chromosomal DNA is supercoiled and tightly compacted, and eukaryotes experience a similar level of compaction albeit by complex chromatin structures. In both domains of life the compaction often results in enhancing the formation of protein-DNA loops by bringing distant sites closer together spatially, which increases the local concentration. “Action at a distance” is being more and more recognized for its biological importance and therapeutic potentials in such applications as gene and cancer research (Stenger, Tegtmeyer et al. 1994; Rippe, von Hippel et al. 1995; Yasmin, Yeung et al. 2004; Galande, Purbey et al. 2007). The *lac* operon provides a

scaffold to build protein facilitated DNA looping models for the advancement of these and many other disciplines.

1.4 Structure Suggests Lac Repressor Conformational Adaptability in Loop Formation

The mechanism(s) for LacI mediated DNA looping became more evident upon the discovery of the LacI structure. The LacI monomers are 360 amino acids and form a homotetramer that exists as a dimer of dimers under biologically relevant concentrations in solution (Fickert and Müller-Hill 1992; Levandoski, Tsodikov et al. 1996). Each dimer binds one DNA operator site through N-terminal headpieces (Kaptein, Zuiderweg et al. 1985), binds one or two inducers in a central core domain (Friedman, Fischmann et al. 1995; Daber, Stayrook et al. 2007), and binds the other dimer through a shared C-terminal four-helix bundle shown in Figure 1.4A (Chen and Matthews 1994; Friedman, Fischmann et al. 1995; Lewis, Chang et al. 1996). The landmark crystal structure of LacI bound to two operators is a V-shaped protein with DNA operator segments (O_{sym}) held closely together, and a C2 symmetry axis running through the center of the V-shaped cleft of the LacI tetramer.

A few flexible features noted in the crystal structure of LacI have been suggested to enhance looping and stabilize the LacI•DNA looped complex. Short peptide segments that lack secondary structure and have varied conformations make the connections between the core domain and the tetramerization domain. This arrangement has been hypothesized to permit opening of the dimer-dimer interface into an extended linear form,

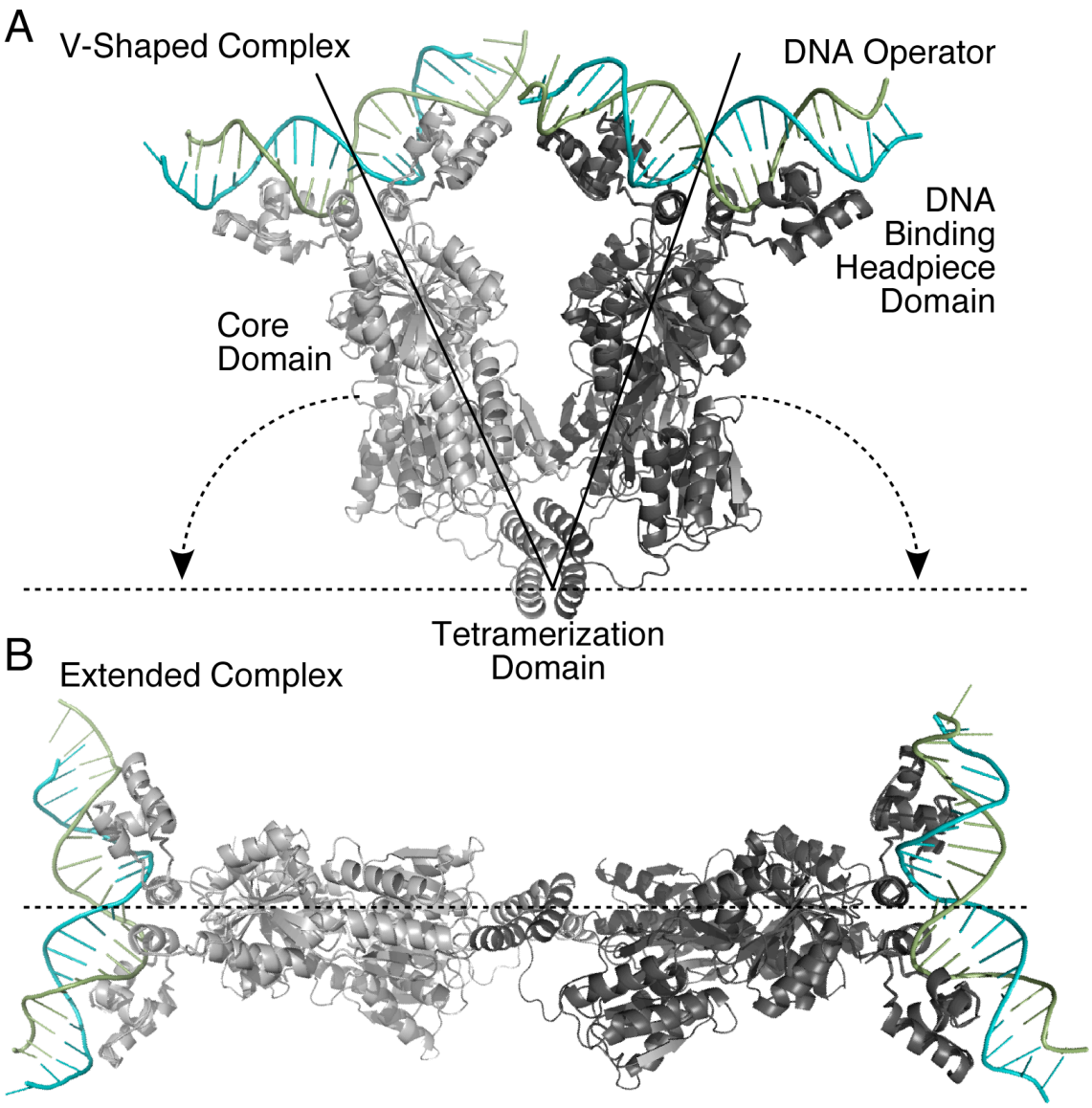


Figure 1.4: *The co-crystal structure of the V-shaped LacI bound with two DNA operator fragments*. The coccrystal, (Lewis, Chang et al. 1996) was pair-fitted with a core and tetramerization domain (Friedman, Fischmann et al. 1995) in Pymol (Delano) to represent α-helix and β-sheet structures not present in c-α Lewis co-crystal structure. (A) The homotetramer exists as a dimer of dimers with the N-terminal DNA binding headpiece, a helix-turn-helix motif. The DNA binding headpiece sits on top of the N-terminal core domain and is connected to the domain through a flexible hinge-helix. At the vertex of the V is the tetramerization domain, a four-helix bundle that stabilizes the dimer of dimers. (B) The flexible linkers that connect the C-terminal core domain to the tetramerization domain may permit opening of the LacI tetramer into and extended linear form, while maintaining the tetrameric state of the protein. The flexibility of the headpiece and tetramerization domain have been proposed to stabilize many loop topologies and loop shapes.

while maintaining the tetrameric oligomeric state depicted in Figure 1.4B (Friedman, Fischmann et al. 1995; Ruben and Roos 1997; Rutkauskas, Zhan et al. 2009). The interface between the DNA binding headpiece and the core domain, appropriately called the hinge-helix, has also been proposed to be flexible (Villa, Balaeff et al. 2005) and the headpiece itself has an inherent plasticity that allows stable non-specific binding to non-operator DNA (Kalodimos, Biris et al. 2004; Romanuka, Folkers et al. 2009). The nonspecific interactions may accelerate loop formation by allowing one-dimensional diffusion or hopping along DNA (Berg and Blomberg 1976; Loverdo, Benichou et al. 2009; Furini, Domene et al. 2010). Still, the extent to which all of the LacI features described above contribute to the free energy when forming a looped complex is unknown.

1.5 Multiplicity of Loop Topologies Predicted from the V-shaped Lac Repressor

The formation of any observed LacI•DNA loop includes free energy contributions from specific LacI-operator binding, the cooperativity between headpiece-DNA binding interactions (Levandoski, Tsodikov et al. 1996), and DNA bending and twisting in the loop. Specifically, loops that have intervening DNA shorter than the persistence length, ~150 base pairs have a high enthalpic penalty for bending the DNA to juxtapose the operators in a V-shaped looped complex. Conversely, at distances much larger than the persistence length there are entropic penalties for the juxtaposition of operators that are now part of a much larger conformational space.

The balance between enthalpy and entropy associated with DNA looping has been evaluated using the ratio of equilibrium constants for forming a cyclized DNA product over the bimolecular product, called the J-factor for a given DNA length (Jacobson and Stockmayer 1950). The J-factor provides insight into the intrinsic DNA sequence-dependent stiffness or persistence length as a function of length. By varying the DNA length in cyclization experiments two other DNA sequence dependent parameters can be extracted: the DNA helical repeat seen as the substructure periodicity in the function of J, and the free energy of DNA twisting, the torsional modulus, corresponding to the amplitude of the periodicity.

The DNA persistence length and torsional rigidity calculated have been applied to protein-mediated DNA loops to predict the free energy associated with the formation and shape of a loop. In particular, the symmetry of LacI and the DNA operator binding sites offer a unique system that is suggested to occupy many energetically degenerate loop states. DNA loops anchored by a V-shaped protein can adopt different DNA loop topologies referred to as parallel (P) and antiparallel (A) according to the relative directions of the DNA strands in the two DNA operators shown in Figure 1.5 (Semsey, Virnik et al. 2005). Furthermore, the parallel and antiparallel DNA trajectories can be subdivided into P1, P2, A1, and A2 according to whether the upstream DNA operator is directed toward or away from the downstream dimer-DNA unit (Swigon, Coleman et al. 2006). In addition to the loops described solely on DNA trajectories, the positioning of LacI inside or outside of a parallel DNA loop can be defined as wrapping towards (WT) and wrapping away (WA) respectively as shown in Figure 1.5 (Levandoski, Tsodikov et

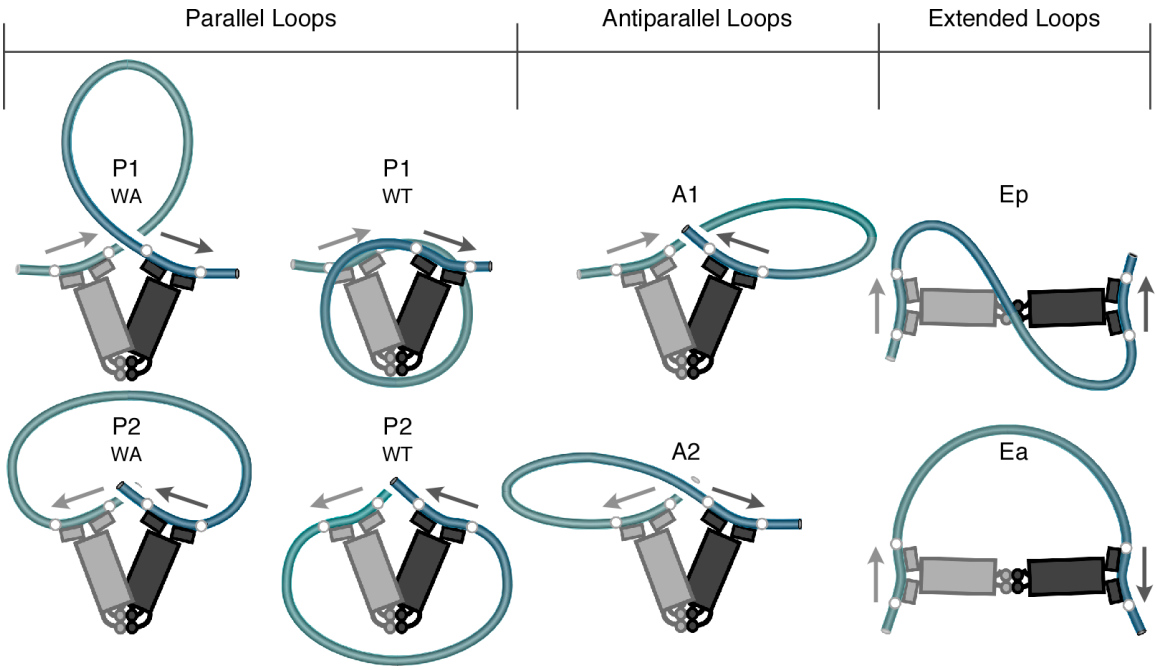


Figure 1.5: *The multiplicity of LacI•DNA loops*. The two V-shaped loop topologies for LacI•DNA complexes, parallel (P) and antiparallel (A) are defined by the direction of the DNA strands in a loop (arrows). The two topologies can be subdivided further into P1 and P2, and A1 and A2, based on the direction of the DNA operators relative to the LacI complex. Additionally, the P1 and P2 can be further subdivided into wrapping away (WA) and wrapping towards (WT) depending on the position of LacI relative to the intervening DNA loop. Finally, if the extended LacI•DNA looped complex exists it should have two significantly different topologies, parallel and antiparallel. The length of the DNA in the loop, the relative operator phasings, and flexible contributions of LacI control the preferred looped complex.

al. 1996; Mehta and Kahn 1999; Wong, Guthold et al. 2008). Finally, if the V-shaped LacI opens up to a more linear form, then an extended (E) loop needs to be considered (Villa, Balaeff et al. 2005; Swigon, Coleman et al. 2006).

1.6 The Remaining Paradigms for Gene Regulation by Lac Repressor

Describing the genetic circuit for the *lac* operon has evolved from a simple molecular switch to a cooperative, multivalent circuit that requires knowledge of the physical properties of DNA and protein to accurately model. Additionally, environmental

conditions, especially the ionic strength (Record, Lohman et al. 1976) and the presence or absence of other DNA binding proteins (Becker, Kahn et al. 2005; Becker, Kahn et al. 2007; Bond, Peters et al. 2010), affect the free energy of protein-DNA binding and DNA deformation. Sequence-directed bending or enhanced flexibility in the DNA can potentiate or prevent looping (Lobell and Schleif 1991; Mehta and Kahn 1999).

The physical chemistry of looping, and especially the connection between *in vitro* properties of DNA and *in vivo* loop stability, is not completely understood. *In vivo* loops are often observed to be surprisingly stable, and while this has been rationalized by recent controversial suggestions that DNA is substantially more flexible than had previously been thought (Cloutier and Widom 2005; Wiggins, van der Heijden et al. 2006), protein flexibility and multivalency must also be considered. A complete thermodynamic description of looping is needed for an accurate quantitative description of the regulation of the *lac* operon (Zhang, McEwen et al. 2006; Han, Garcia et al. 2009). Our experimental efforts address the multiplicity of LacI•DNA looping (Chapter 3), inducer effects on the loop states, (Chapter 4), and the kinetics of the looped complexes (Chapter 4 and 5) to improve upon the thermodynamic description of the *lac* operon and guide similar efforts at describing more complex gene regulatory circuits. The implications and significance of these results are described along with possible future directions (Chapter 6).

Chapter 2

Experimental Design Derived from Previous Work and Collaborations from the Kahn Lab

Overview

There are two straightforward methods to distinguish between protein and DNA energetic contributions in the formation of a looped complex: a rigid protein with a flexible DNA, or a rigid DNA and a flexible protein. This chapter details the key work by Jason Kahn's lab and his collaborators, which uses intrinsically curved DNA constructs to compare the flexibility of LacI to the flexibility of DNA. Our immediate experimental goals were to distinguish the contributions of protein flexibility and DNA flexibility in looping, and disentangle a multitude of possibilities predicted by models for LacI•DNA looping (Semsey, Virnik et al. 2005; Swigon, Coleman et al. 2006; Levandoski, Tsodikov et al. 1996; Mehta and Kahn 1999; Wong, Guthold et al. 2008). To accomplish these goals previous DNA constructs were expanded into a landscape of intrinsically bent DNA constructs with dual rotation of operators flanking a central bend. Furthermore, previous FRET locations were generalized to both sides of the two operator sequences to provide unique FRET signatures for different LacI•DNA loop topologies. Coupling multiple fluorophore positions with a DNA construct landscape produces a global analysis of LacI mediated looping.

2.1 Intrinsically Curved DNA

The ideal model of B-DNA, shown in Figure 1.1, has the same tilt and roll for each base pair step, lacks any dependence on the DNA sequence, and is linear at a thermodynamic minimum. However, DNA can have a natural sequence-dependent DNA curvature, and this dependence was first demonstrated in the form of repetitive sequences of A-T base pairs with minicircles from the kinetoplast DNA of *L. tarentolae* (Marini, Levene et al. 1982). Since then, sequence-directed bends have been reviewed in eukaryotes for their enhancement in nucleosome formation (Crothers, Haran et al. 1990), and its important in nucleoid compaction in prokaryotes has also been demonstrated (Mrazek 2010). Additionally, DNA regions with enhanced flexibility or sequence-directed bends have been detected near transcription initiation sites and are suggested to promote transcriptional events (Gimenes, Takeda et al. 2008). Although it is often easier to consider DNA as a linear polymer, the evolution of DNA sequences would disagree with this oversimplification.

When A-T base pairs are sequentially placed together in runs of 4-6, called A-tracts, these regions have a tendency to deviate from the ideal B-DNA and bend slightly in the direction of tilt at the edges of the A-tracts (Koo, Wu et al. 1986). This compresses the DNA minor groove in an effort to maintain continuos DNA stacking interactions, which results in intrinsically curving the DNA by 16-18° relative to the central axis dyad shown in Figure 2.1 (Levene, Wu et al. 1986). Placing these bends in phase with the double stranded helix repeat (repeats every 10-11 base pairs) additively curves DNA in one direction.

DNA cyclization experiments on DNA fragment indicated the presence of phased A-tract bends lowers the free energy of DNA deformation and also increases the local concentration of the two DNA ends relative to a linear DNA fragment (Koo, Drak et al. 1990; Crothers, Drak et al. 1992). DNA cyclization is conceptually similar to looping DNA fragments, since its occurrence depends on the probability of bringing the DNA ends together in a specific spatial arrangement (Hagerman and Ramadevi 1990; Rippe, von Hippel et al. 1995; Goyal, Lillian et al. 2007). Therefore, intrinsically bent DNA should have the same energetic benefits in looping as in cyclization, and the loop stability should depend on the operator phasings and the protein flexibility in compensation to the DNA bending and torsional requirements to form a loop.

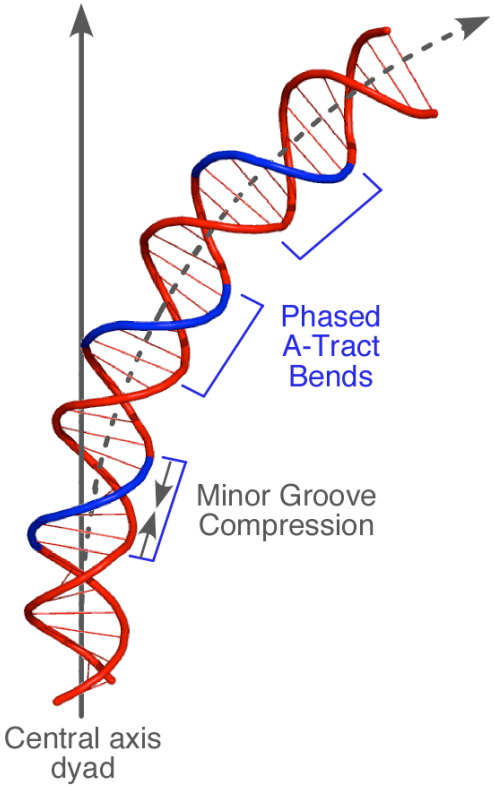


Figure 2.1: *Intrinsically bent DNA sequences*. DNA assembled using a junction model (Koo, Drak et al. 1990). A-tract regions, sequential runs of 4-6 A-T base pairs, compress the minor groove of the DNA in order to stabilize base pair interactions. This causes the DNA to bend by 16-18°, and can additively curve the initial DNA trajectory when placed in phase with the helical repeat.

2.2 LacI Operator Phasing Modulates the Preferred Loop Topologies

The DNA curvature as a result of intrinsically bent DNA sequences have an energetic minimum that is biased in a direction, and this loss of spatial freedom can be

used to probe LacI flexibility. To this end, two DNA looping constructs were synthesized by bracketing a central DNA bend of eight helically phased A-tracts with two phased operators as shown in Figure 2.3A (Mehta and Kahn 1999). The nomenclature used for these constructs, 9C14 and 11C12*, specify the DNA adaptor lengths surrounding a central bend (C) in the operator-adaptor-bend-adaptor-operator design. The differences in DNA adaptor lengths confer rotation of the operating phasings relative to the central bend, and the different operator phasings of the constructs favor different LacI•DNA looped complexes and LacI geometries shown in Figure 2.3A. The construct 11C12* is missing one base pair in the A-tract region relative to 9C14, and differs from the construct 11C12 that is used throughout my research.

The DNA constructs 9C14 and 11C12* form stable LacI•DNA looped complexes at 1:1 stoichiometry of DNA:LacI, and based on competition studies both form hyperstable loops with lifetimes of days. Cyclization studies, shown in Figure 2.3B, indicated the construct 9C14 predominantly formed a P1-WA loop with a V-shaped LacI geometry (50%) and two other populations, possibly an antiparallel loop (40%) and a P1-WT (10%) (Mehta and Kahn 1999). On the other hand, 11C12* had a similar topological behavior ± LacI for the cyclized product, which suggested 11C12* must be forming an open/extended geometry since the complex has a lifetime of days in the presence of excess DNA competitor.

DNA cyclization studies were effective at determining LacI•DNA loop populations, but potential salt-bridge interactions between the extended DNA tails and the LacI cores in a looped complex might have influenced the cyclization results (Saecker

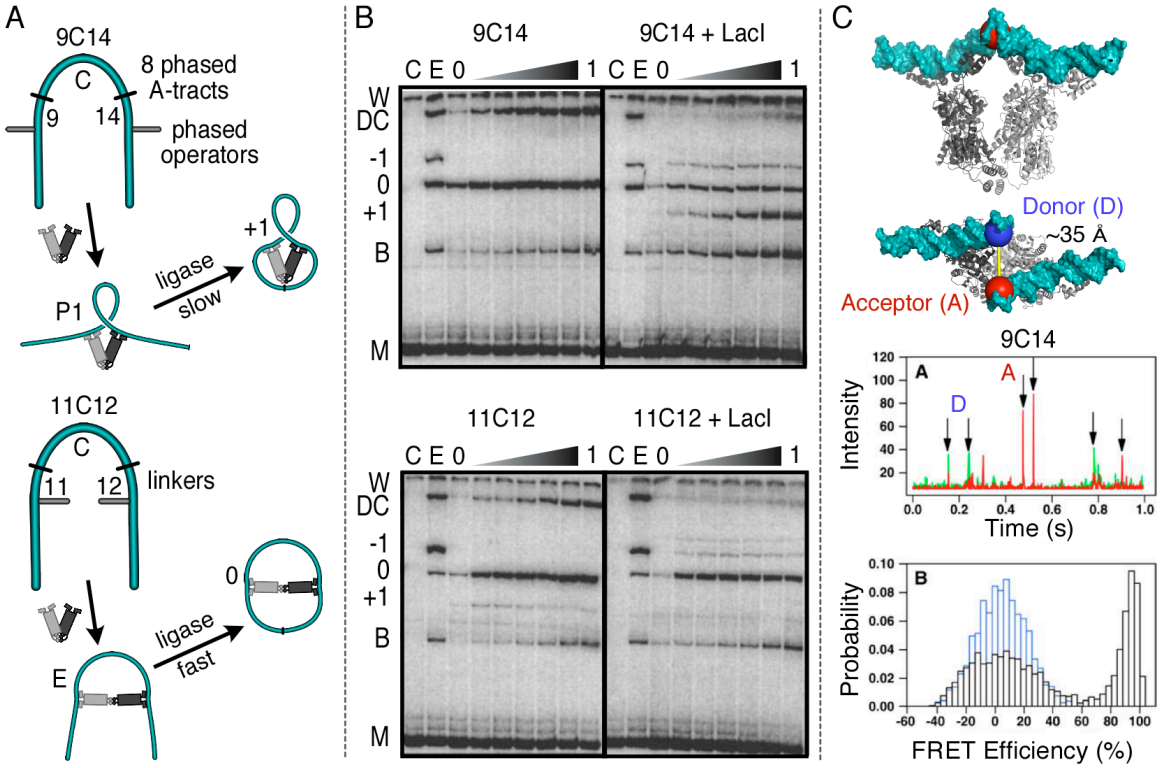


Figure 2.2: *Operator phasings, relative to an intrinsically bent DNA sequence, modulate the preferred LacI•DNA loop topology and geometry*. (A) The two constructs designed to form to different loops: 9C14 will preferentially form a V-shaped P1 loop with slow cyclization kinetics, and 11C12*, with operators facing the central bend has preference for an extended (E) LacI loop and fast cyclization kinetics. The nomenclatures of the constructs are derived from the linker/ adaptor regions between the central 8-phased A-tract bends and the bracketing operators. (B) Cyclization studies in the presence of LacI indicated 9C14 could form multiple populations having varying kinetic rates of formation (Mehta and Kahn 1999). The slower formation of a cyclized product suggests a V-shaped P1 looped complex. 11C12* is predominately one population that had a fast rate of circle formation suggesting an open or E LacI•DNA looped complex. (C) Single-molecule studies on 9C14 LacI•DNA complexes (Morgan, Okamoto et al. 2005). Donor and acceptor fluorophores were positioned two base pairs outside of each operator, between the operator and intrinsic bend, modeled using the V-shape co-crystal structure and ideal B-DNA to extend the operator fragments (Lewis, Chang et al. 1996). Single molecule diffusions were monitored over time with acceptor photons shown as in red and the donor photons shown as green line in the presence of LacI. These single molecule events are summarized in the histogram below where DNA only is shown in blue and the black histograms indicates the addition of LacI.

and Record 2002; Morgan, Okamoto et al. 2005). Therefore, a more direct method,

Förster resonance energy transfer (FRET), was used to elucidate the relative geometries

in a LacI•DNA looped complex. In the FRET process a donor fluorophore is excited from

the ground state by a photon. That energy is transferred to an acceptor fluorophore through dipole-dipole interactions, and a red-shifted photon is released as the acceptor relaxes back down to the ground state (Förster 1948). The efficiency of energy transfer between the donor and acceptor fluorophores is a sensitive function of the distance between the fluorophores, their relative orientations, and the quantum yields of the fluorophores.

Acceptor and donor fluorophores were positioned two base pairs away from each operator on constructs 9C14 and 11C12*, located between the operator and central bend. This allowed the LacI geometry in a DNA looped complex to be directly measured. High bulk FRET efficiencies were observed for construct 9C14 (70%) while 11C12* showed low efficiency (10%) (Edelman, Cheong et al. 2003). These measurements were in agreement with cyclization experiments for the relative geometries of the two constructs, a V-shaped geometry for 9C14 and extended geometry for 11C12*, but the ambiguity of bulk FRET methods made it difficult to confidently know both population distributions and geometries simultaneously. For example, is 9C14 forming a looped complex where 100% of the molecules have 70% FRET efficiency, or is it 70% of the molecules with 100% efficiency, or any combination in between?

Single-molecule (s-m) FRET diffusion studies were performed to resolve the ambiguities of bulk FRET measurements for 9C14. Since this method only looks at individual molecules diffusing through a beam, it is possible to see subpopulations that make up the observed bulk population distributions. The s-m FRET results shown in Figure 2.2C suggested that 100% of the molecules had ~90% FRET efficiency, and

suggested that LacI is in a V-shaped geometry and forms a P1-WA loop with construct 9C14 (Morgan, Okamoto et al. 2005). The remaining zero energy transfer peak population appeared to be donor only molecules.

2.3 Rod Mechanical Models for Looping Intrinsically Bent DNA

Modeling LacI mediated DNA looping can provide tremendous insight into converging experimental observations with the underlying energetic parameters, as well as motivate future experiments. One of the most promising methods for modeling DNA behavior considers DNA as a rod (Manning, Maddocks et al. 1996) and defines the intrinsic properties using only two parameters: the stiffness of bending the DNA and stiffness of the torsional modulus as shown in Figure 2.3A (Lillian, Goyal et al. 2008). Specifically, this DNA rod mechanical model has been applied to the results for the looping constructs 9C14 and 11C12* to examine the properties of LacI•DNA. The DNA was modeled using a straight-helix-straight (SHS) simplification shown in Figure 2.3A and suggested that the experimental observations can be explained by a rigid protein coupled with careful consideration of all the possible loop topologies as well as twist variations within the loop (Goyal, Lillian et al. 2007). The calculations generate predictions for loop energies, topologies, and geometries that can be directly compared to experiments. To test whether the existing descriptions for DNA and protein flexibility can quantitatively explain all the free energy contributions underlying protein-mediated DNA looping, the computational analysis was then extended to a complete landscape of possible operator phasing isomers for intrinsically curved DNA (Lillian, Goyal et al.

2008). The hyperstable looping constructs, due to their intrinsic curvature, are particularly suitable for studying loops with a wide range of free energies and geometries, in contrast to loops formed on intrinsically straight DNA. The latter are many times less stable with respect to other complexes, such as DNA bound with two LacI.

The calculated energy landscape for LacI•DNA looped complexes using the SHS rod mechanics suggested that antiparallel (A1 and A2) loops would be more prevalent than parallel (P1 and P2) shown in Figure 2.3B. LacI•DNA complexes that form parallel loops have a larger DNA deformation penalty that must be overcome to form a looped complex; the DNA must bend 360° in parallel

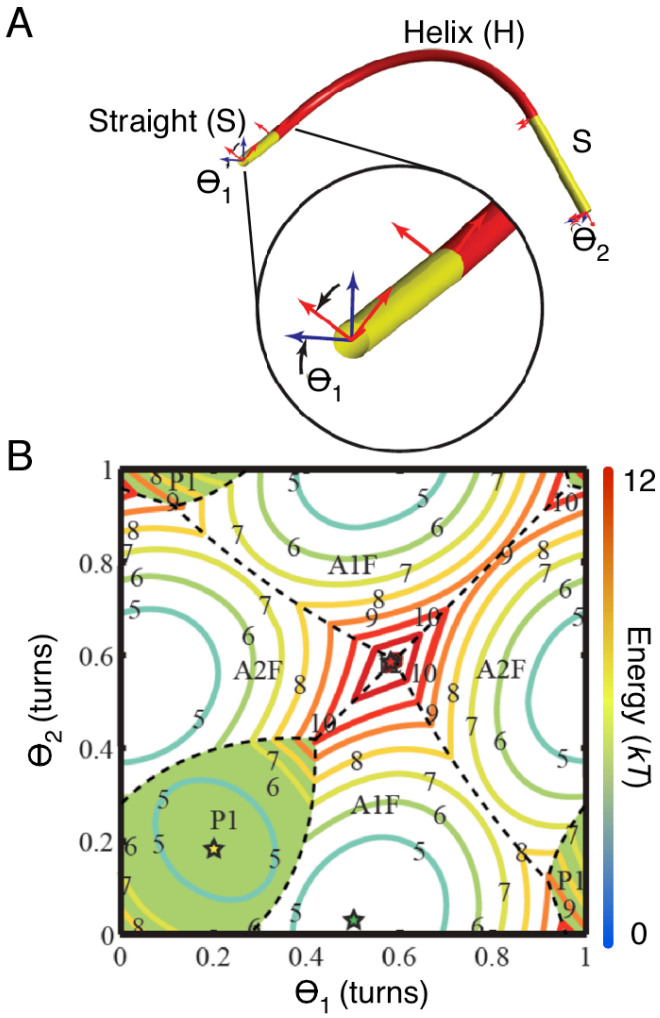


Figure 2.3: *Statistical rod mechanics modeling of LacI mediated DNA looping on a family of intrinsically curved DNA*. (A) To calculate the underlying energetics for LacI facilitated looping a straight-helix-straight interpretation was used with different DNA sequnece phasings (Θ) to model DNA torsional and bending stiffness when confined to the avaiable trajectories defined by the DNA co-crystal structure. (B) An energetic LacI•DNA loop landscape calculated for a family of intrinsically curved DNA (Lillian, Goyal et al. 2008). The yellow, green, and red starts denote locations for low energy P1 loops, low energy A1 loops, and a high energy antiarallel loop respectively.

compared to 180º degrees for antiparallel loops. The free energy calculations for this DNA sequence space also predicted that P2 loops would be energetically costly because of the extremely high local DNA deformation energy needed to bend the intervening DNA in a looped complex, which was defined by the initial operator trajectories seen in the V-shape LacI co-crystal structure.

Earlier protein flexibility models and DNA rod mechanics models that focus on DNA flexibility make some clear predictions that can be distinguished experimentally. In particular, the protein flexibility model predicts that some loops, specifically 11C12* looped complexes, have headpieces that are too far apart to support FRET. The rod mechanics models using a rigid V-shaped LacI suggested that the initial positioning of the fluorophores for 9C14 and 11C12* (internally located between the operator and central bend) preferentially detects P1 loop topologies and not antiparallel loops. Furthermore, they predict that changing the fluorophore positions relative to the operators should bring the fluorophores into close contact in the context of an antiparallel, A1 or A2 loop.

2.4 A DNA Construct Landscape

The initial DNA construct designs confirmed that operator phasings relative to an intrinsic bend could modulate the preferred LacI geometry in a LacI•DNA looped complex. The constructs 9C14 and 11C12* provide two different extreme perspectives with operators phasing towards or away from the central bend respectively. The ambiguous experimental population results of the two constructs and the disagreement between theory and experimentation suggested that a more in depth analysis could help

theory and experiment coalesce. Achieving these goals requires a less coarse analysis of looping that can put these two extremes into perspective. Therefore, the two constructs should be generalized to examine operator phasings relative to intrinsically curved DNA bend.

The constructs 9C14 and 11C12* can be related by describing 9C14 as gaining two base pairs on one adaptor side and losing two base pairs on the other adaptor to become 11C12*. Performing only one of these two base pair additions or subtractions on either adaptor would create 11C14 and 9C12 respectively. However, this would only examine a small LacI•DNA looping sequence space. Since the primary goal is to examine LacI and DNA flexible contributions in looping as well as the multiplicity of the loop states, the experimental results must broaden our understanding of the underlying energetics of both partners involved in the looping process. A family of related DNA looping substrates should provide examples of all of the accessible loop topologies, and systematic study of the landscape of loops as a function of DNA shape and sequence can help disentangle all of the loop free energy contributions discussed earlier. Furthermore, an entire DNA period could allow direct comparison to *in vivo* periodicity studies, and our results could be directly compared to theoretical values calculated for the torsional modulus for a similar DNA sequence space. Therefore, analyzing LacI mediated looping on an entire DNA landscape of dual operator phasing variants relative to a central bend would provide a means to these ends.

The most straightforward and cost effective method to control the insertion or deletion of base pairs into a DNA construct was by site-directed mutagenesis. The

plasmid pRM9C14 was used as the mutagenic template for designing an entire construct landscape centered on construct 9C14 shown in Figure 2.4A. The mutagenic primers alter the helical phasing adaptors, rotating the operators on both sides by 0, ± 2, or ± 4 base pairs relative to the operator-bend phasing in 9C14, care was taken to avoid including any pseudo LacI recognition sites. Successful mutagenesis on one side would generate a new plasmid that can serve as a template for adaptor modifications on the other DNA end. For example, the initial single-round of mutagenesis that modified the "downstream" end of 9C14 creates a cross section through the landscape composed of the plasmids pAH9C10 (-4), pAH9C12(-2), pAH9C16(+2), and pAH9C18(+4). The first number indexes the bend-operator separation on the donor side, the second indexes the acceptor side, and the number in parenthesis is the base pair change relative to construct 9C14. Then the other adaptor could be modified in the second round of mutagenesis, i.e. pAH9C10 could be subsequently modified using site-directed mutagenesis on the other DNA end to create pAH5C10(-4,-4), pAH7C10(-2,-4), pAH11C10(+2,-4), and pAH13C10(+4,-4). Completing this systematic expansion of the construct landscape using the four mutagenic primers generates a 5 × 5 matrix of DNA constructs shown in Figure 2.4B.

It is worth noting that the landscape of constructs prepared does not cover an entire period. However, to increase the landscape into a 6 × 6 matrix of constructs would require 11 more constructs, would increase the work 1½ fold, and would still not cover an entire period. Since each of the constructs varies by two base pairs relative to its neighbor, we considered the extra efforts to stretch a model an additional ½ base pair at the edges of the construct landscape is easier than the experimental work. Additionally,

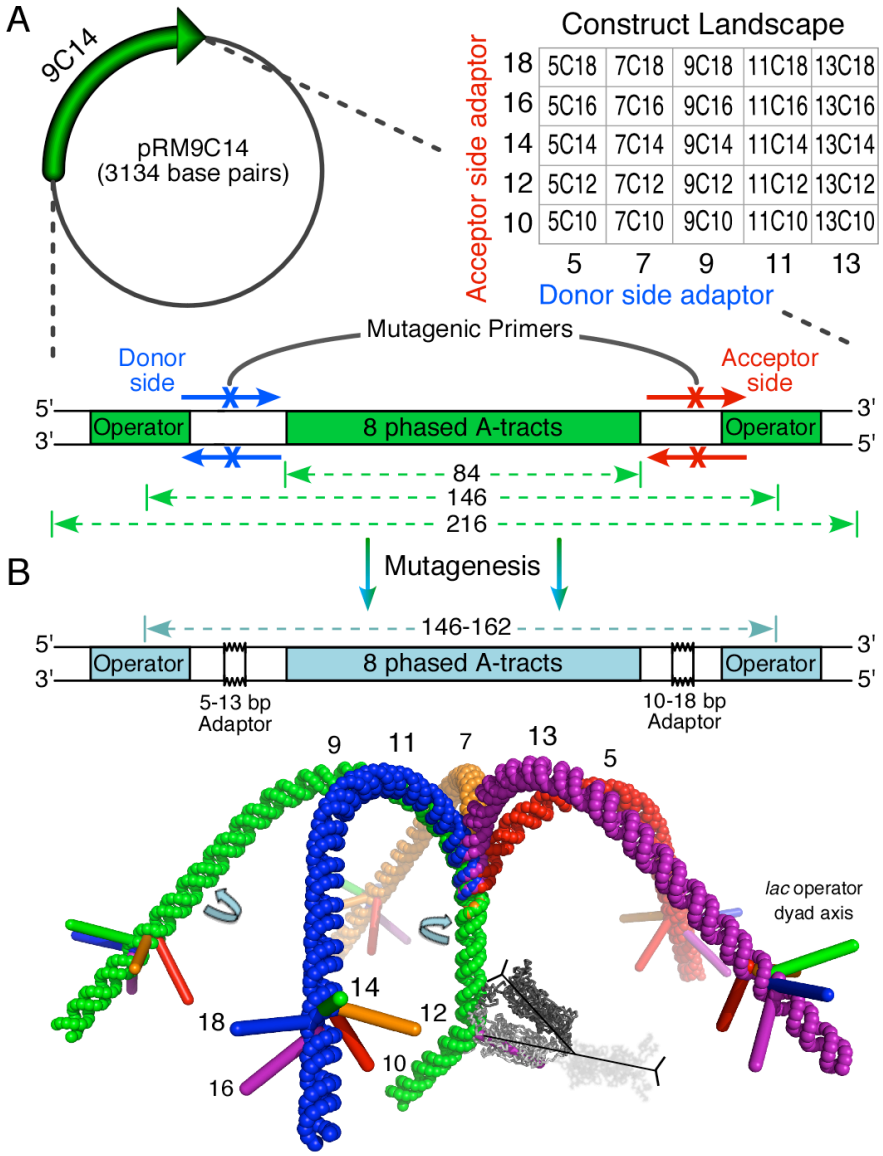


Figure 2.4: *Design for synthesizing a DNA construct landscape to evaluate LacI mediated DNA looping*. (A) The plasmid, pRM9C14, containing the looping construct 9C14 would be used in combination with 4 sets of site-directed mutagenic primers to modify adaptor regions 0, ±2, and ±4 base pairs between the operators and A-tract bends. Then each modified plasmid can be used as a template to generate 25 possible combinations that create a 5 x 5 matrix of DNA constructs. The nomenclature is derived from the previous adaptor-bend adaptor nomenclature: the first number indexes the bend-operator separation on the donor side and the second indexes the acceptor side. (B) The family of constructs is modeled using the junction model for A-tract DNA (Koo, Drak et al. 1990) and the LacI•DNA co-crystal structure (Lewis, Chang et al. 1996). The donor sides of the DNA constructs are superimposed. LacI is shown bound in one of two possible orientations relative to the operator. The extended LacI is shown as a ghost image.

the construct 11C12 that resulted from the 9C14 family is not the same as 11C12* described earlier. This 11C12 varies only from its neighbor by adaptor lengths and does not have a missing base pair in the A-tract that was present in the original 11C12* DNA construct.

2.5 Fluorophore Position Variants

A DNA construct landscape makes it possible to generate all loop topologies for LacI•DNA looped complexes. However, previous methods showed their limitations in determining LacI geometries, loop topologies, and loop populations simultaneously. Additionally, the multiplicity of LacI•DNA loops needs to be addressed in a systematic method that can be directly compared to LacI geometries and also result in loop population distributions. FRET showed the greatest promise as a method in addressing both loop geometries and loop population distributions simultaneously, but has a limited range over which the FRET efficiency can be related to distance.

Inspired by the DNA rod mechanics models, we propose that varying the fluorophore positioning attachment sites may allow FRET sensitivity to alternative loop topologies and geometries. Specifically, moving a single fluorophore from its original attachment site two base pairs “outside” of the operator to the analogous location on the “inside” of the operator would increase the sensitivity for antiparallel loop topologies, which have different trajectories than P1 loops. Further generalizing these predictions, if we had four different fluorophore position variants (FPV), two bracketing each operator as modeled in Figure 2.5A, we could detect any V-shape LacI•DNA looped complex. The

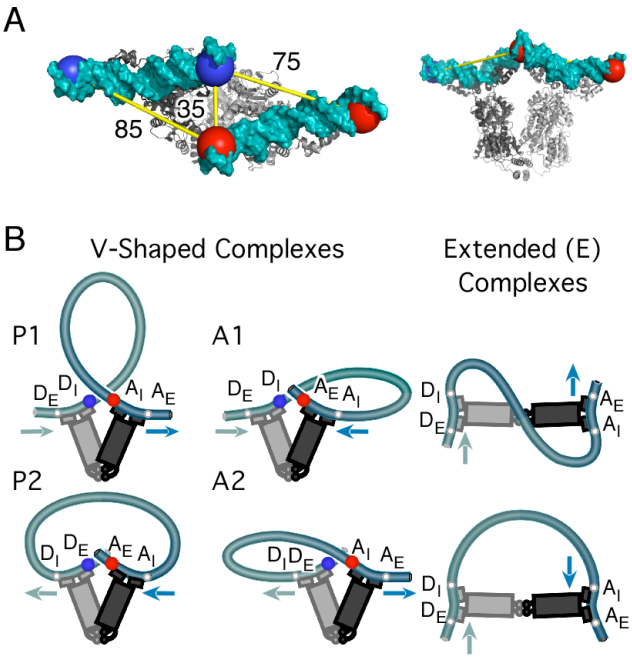


Figure 2.5: *The FRET signatures of Fluorophore Position Variants (FPVs) can differentiate among loop topologies*. (A) The operator DNA in the V-shaped LacI•DNA co-crystal (Lewis, Chang et al. 1996) structure was extended with ideal B-DNA to show the four fluorophore attachment points two base pairs outside of each operator. The distances shown are calculated between the C5 carbons of the T residues to which the fluorophores are attached. (B) The FPVs can be mapped onto loop topologies anticipated for this DNA sequence space. Assuming that the LacI V-shape does not change, FRET for each FPV should be maximal for one of the four loop topologies: the P1 loop should give maximal FRET with the D_IA_I FPV, P2 with D_EA_E, A1 with D_IA_E, and A2 with D_EA_I. FRET should be decreased or absent in extended complexes.

FPVs for the donor (D) and acceptor (A) combinations are denoted D_IA_I D_IA_E, D_EA_I, and D_EA_E, with "I" indicating internal fluorophore attachment, between the operator and the A-tract bend as in previous FRET studies for the constructs, and "E" indicating external attachment. Each construct in the looping landscape could be prepared with each of the FPVs using PCR amplification and four different primer combinations. For example primer containing the D_I FPV can be mixed with FPVs A_I or A_E, to result in D_IA_I and D_IA_E and similar for primers with D_E mixed with A_I or A_E.

The different DNA trajectories in the V-shaped LacI•DNA complexes would provide unique FPV signatures that are sensitive for the different V-shaped loop topologies. In Figure 2.5B the FPVs are mapped onto the four loop topologies. Assuming

that the LacI V-shape does not change, the FRET for each FPV combination should then be maximal for one of the four loop topologies; the P1 loop should give maximal FRET with the D_IA_I FPV, P2 with D_EA_E, A1 with D_IA_E, and A2 with D_EA_I. If LacI exists in an extended LacI•DNA loop then no FRET would be observed, regardless of the FPV.

Chapter 3

A Landscape of LacI Mediated DNA Loops

Overview

The preparations of the DNA construct landscape and the FPV sub-landscapes are discussed in detail in this Chapter, followed by experimental results and discussion. Electrophoretic mobility shift assays were performed on the construct landscape to evaluate if each construct could form a LacI•DNA looped complex. Then bulk FRET with the different FVPs on each construct was used to explore the preferred loop topology and loop geometry for each DNA sequence. The results of the bulk FRET were represented as four 5 × 5 matrices that correspond to specific loop topologies defined by the FPVs. We find that the FRET efficiency maximum for each landscape is a FRET minima on the other FPV landscapes, population distributions were calculated from the peak FRET amplitudes of each FPV landscape, assuming a uniform population at the peaks. Because of the observation of lower FRET amplitude in comparison to parallel loops, our population calculations resulted in an unavoidable interpretation of slightly extended LacI•DNA looped complexes for antiparallel loop populations. Additionally, since FRET is not uniform across the landscape, but all constructs appear to form stable looped complexes, we also have population distributions for a "dark" population that we ascribe to an extended LacI•DNA looped complexes that is incapable of FRET.

This chapter is summarized in the submission of the manuscript to NAR (Haeusler, A., Goodson, K., Lillian, T., Goyal, S., Perkins, N., and Kahn, J. FRET Studies of Landscape of Lac Repressor Mediated DNA Loops. 2011).

3.1 Materials and Methods

All chemicals were purchased from Fisher. All restriction enzymes, kinases, and DNA polymerases came from New England Biolabs (NEB). The DNA oligonucleotides were purchased from Integrated DNA Technologies (IDT).

3.1.1 Synthesis of a DNA Construct Landscape of Dual Operator Phasing Variants

The DNA constructs for the looping landscape originated from the template plasmid pRM9C14 (Mehta and Kahn 1999). The plasmid at subpicomolar concentrations was transformed into XL1-blue electro-competent cells, expressed, and purified following manufacturer's protocol for a Maxiprep (Stratagene). All unlabeled DNA concentrations were determined from UV scans from 220-340 nm (Cary Varian, Bio 100). A DNA construct matrix was prepared from pRM9C14 using the primers shown below (complements not shown) to insert and/or remove two and four base pairs in the adaptor region on either side of the central A-tract bend with a QuikChange kit (Stratagene). This conferred rotation of the lac operators relative to the central bend, which created the DNA construct looping landscape. Primer design considerations included the avoiding the introduction of pseudo-operator sites into the original template plasmid, as well as addition/removal of a unique restriction recognition site.

E1(+4)top 5′-AGATCTCA**G**AT**C**TCGT**CG**ACGGATCCGG-3′

E1(+2)top 5′-AGATCTCA**G**AT**C**TCGT--ACGGATCCGG-3′

Template1 5′-AGATCTCA-AT-TCGT--ACGGATCCGG-3′

E1(-2)top 5′-AGATCTC----TTCGT--ACGGATCCGG-3′

E1(-4)top 5′-AGATCT------TCGT--ACGGATCCGG-3′

E2(+4)top 5′-GCGCTGAACGCGTCCTAGA**CGCT**ATCGAAGC-3′

E2(+2)top 5′-GCGCTGAACGCGTCCTAGA**CG**--ATCGAAGC-3′

Template2 5′-GCGCTGAACGCGTCCTAGA----ATCGAAGC-3′

E2(-2)top 5′-GCGCTGAACGCGTCCTA--ATCGAAGCTAGC-3′

E2(-4)top 5′-GCGCTGAACGCGTCC----ATCGAAGCTAGC-3′

The E1 primers modified the donor side of the 9C14 template (template 1) of the constructs and the newly developed plasmids, pAH5C14, pAH7C14, pAH11C14, and pAH13C14 names were derived from the adaptor lengths. The DNA mutagenic sequence modifications were amplified with a thermocycler (Eppendorf MasterGradient) following Stratagene's recommended PCR conditions. The subsequent *Dpn I* digest was extended for a total of two hours to guarantee digestion of the parent plasmid DNA before transformation by electroporation (Eppendorf, Electroporator 2510). The plasmids were then transformed into XL1-Blue electro-competent cells following suggested protocols for cell growth (Stratagene). Minipreps (Strataprep) were performed on all plasmids were grown from individual colonies and the constructs verified through DNA sequencing. Similarly, the acceptor side adaptor template (template 2) for the pRM9C14 construct plasmid was modified using the E2 primers to create pAH9C10, pAH9C12, pAH9C16,

and pAH9C18. The mutagenic process was then repeated using the eight newly created pAHplasmids as templates along with the appropriate primers to generate a 5 × 5 DNA construct matrix that in which each construct varies by two base pairs relative to its neighbors.

The 25 purified pAHplasmids served as PCR templates. Generally 2 µg of each plasmid was digested with *BstN I* as recommended (NEB), but incubations were performed overnight. The desired ~420 base pair DNA fragment was purified away from the rest of the plasmid fragments. A 7.5% (40:1, acrylamide:*bis*-acrylamide) PAGE for 2 hrs at 20 V/cm and 20º C was pre-electrophoresed for at least 30 min in 1X TBE buffer (40 mM tris, 20 mM borate pH 8.3, 10 mM EDTA) before separation was carried out. The gel was stained with Ethidium Bromide (EthBr), visualized through UV transillumination, and then the appropriate band excised. Each product had appropriate mobility for the ~420 base pair intrinsically curved DNA fragment and for the other *BstN I* fragments. The DNA was eluted by a crush, freeze, and soak method using 400 µL of gel elution buffer (50 mM sodium acetate, 1 mM EDTA, pH 7.0) at 37º C. After soaking overnight, the elution buffer containing the DNA was removed, reduced in volume but not completed dehydrated in a speed vacuum., and meanwhile another 200 µL of gel elution buffer was added to the gel slices for an additional two-hour incubation at 37º C. The second elution was combined with the first and the DNA was phenol-chloroform extracted followed by an ethanol precipitation. The purified DNA was resuspended in TE buffer (10 mM Tris, 1 mM EDTA, pH 7.0) and its concentration was determined through UV absorbance.

3.1.2 Fluorescently Labeling DNA Constructs

The amino modified PCR primers below (IDT) were labeled in house with Alexa555 (LEFL56 and D_E Top) and Alexa647 (KTTM56 and A_E Bot), shown below. The location of the amino modified thymine is underlined and capitalized two base pairs outside of the symmetrical Lac operator (in capital letters) in the primer. The amino modified primers were labeled using purchased dye labeling kits (Invitrogen)following the manufacturer's recommendations, except 5 nmol of primer and the dye were allowed to react for three hours instead of the recommended one hour (a highly detailed description is of this process is given in Appendix 1). Additionally, the dye-primer reaction was mixed at least every 15 minutes in a sonicating bath for 30 seconds. Unreacted label was removed using a P6 spin column (BioRad) and primers were subsequently ethanol precipitated to remove salts and then resuspended in 10 µL of TE buffer containing 8 M urea. The primers were heated at 90^o C for 10 minutes and purified on a pre-electrophoresed (at least 30 minutes) 12% (40:1 acrylamide:*bis*-acrylamide) 8 M urea denaturing PAGE gel for two hours at 55 watts using 1X TBE. The appropriate band, corresponding to the labeled primer, could be visualized by eye (either a red or blue band) and was eluted as previously described for PCR templates. The labeling efficiencies were determined by scanning the UV-Vis spectrum from 220-800 nm and were calculated using the manufacturer's provided extinction coefficients for the donor and acceptor (for Alexa 555 ε max = 150000 M^{-1} cm^{-1} and for Alexa 647 ε max = 239000 M^{-1} cm^{-1})and corresponding DNA absorbance contributions calculated from the nearest neighbor rules absorbances and contributions from the dyes. This typically resulted in the

labeling efficiencies of the primers being 100%. Considerations for possible hypochromacity effects via secondary structure formation of the O_{sym} in the DNA primers still lead to values 100%.

LEFL56 5′-ctgcaggtcagtctaggtAATTGTGAGCGCTCACAATTa**T**atctcaattcgtacgg-3′

KTTM56 5′-caagctttaccatcaacgAATTGTGAGCGCTCACAATTa**T**ctagcttcgattctag-3′

D_E Top 5′-ctgcaggtcagtctag**T**tAATTGTGAGCGC-3′

A_E Bot 5′-caagctttaccatcaa**T**gAATTGTGAGCGC -3′

3.1.3 LacI Expression and Purification

The protein expression and purification was adapted from unpublished work by Larry Edelman, and published protocols (Brenowitz, Mandal et al. 1991; Wilson, Das et al. 2005) with modifications mentioned below. LacI was expressed in E. coli BL21(DE3) pLysS previously transformed with pETLE1a, a plasmid that contained the LacI gene and ampicillin resistance provided by Brenowtiz. Cells were grown from a single colony in a 1 mL LB starter culture overnight, then added to 1 L of LB media containing 50 μg/mL of both ampicillin and chloramphenicol for a 2 hr growth at 37°C. The cells were then induced with 0.1 mM IPTG and an additional 50 μg/mL of ampicillin. Growth was continued for 3 hrs and then harvested by cell pelleting at 8,000 g and 4°C for 10 min in a JLA-10.5 rotor (Beckman, Avanti J-25I). The cell pellet was resuspended to 0.25 g/mL in LacI Resuspension Buffer (200 mM Tris-HCl, 200 mM KCl, 10 mM MgCl2, 5% (v/v) glycerol, 0.3 mM DTT, 1 mM NaN3, 1 mM PMSF, pH 7.2) and lysed by four passes through a French Press (American Instrument Company). The lysed cells were mixed with 0.4 mg/mL of DNase I for 30 minutes on ice with constant rocking. The debris was

pelleted from the lysate at 39,000 g and 4°C for 20 minutes using a JA 25.5 rotor. The crude LacI was precipitated from the supernatant with ~35% ammonium sulfate saturation (confirmed by a 1% SDS-PAGE) at 4°C, and then spun again for 20 min at 39,000 g. The LacI pellet was resuspended in 20 mL of dialysis buffer (50 mM KPO4, 0.1 mM EDTA, 5% (w/v) glucose, 0.3 mM DTT, 1 mM NaN3, pH 7.2) and placed in the same dialysis buffer overnight at 4°C in a Slide-A-Lyzer 7K dialysis cassette. The dialyzed sample was spun at 39,000 g for 20 min at 4°C and the supernatant filtered through a 0.2 μm syringe filter (Millipore).

LacI was purified on a Mono S HR 5/5 Column using a GE AKTA FPLC in a (0.1 mM EDTA, 5% (w/v) glucose, 0.3 mM DTT, 1 mM NaN3, pH 7.2) 50 mM to 500 mM KH_2PO4/K_2HPO4 , 40-column volume linear gradient and a 1.0 mL/min flow rate. The fractions containing LacI (usually eluting between 250 and 300 mM KCl) were pooled and concentrated (if required) with a Centricon (Amicon, Millipore). Fractions that contain 95% purity (determined from a 1% SDS-PAGE silver staining analysis) were further purified using a Sephadex 300, 2 x 15 cm gel filtration column at 1.0 mL/min. flow rate using LacI Storage Buffer (200 mM KPO4, 2 mM EDTA, 400 mM NaCl, 1 mM DTT, 1 mM NaN3, pH 7.2). The concentrated samples were dialyzed in LacI Storage Buffer G (LacI Storage buffer that includes 50% glycerol) with multiple changes at 4°C overnight.

The concentration of LacI tetramer was determined from the A_{280} with the extinction coefficient of 22,500 for each monomer in the tetramer with a typical yield of 5 nmol (2 mL of 25 μM) LacI tetramer per liter of LB. The active protein concentration

was determined by electrophoretic mobility shift assay (EMSA), using a 7.5% gel (75:1 acrylamide:bis-acrylamide) at 20 V/cm for 2 hours at 20°C in TBE buffer, with constant ^{32}P labeled DNA construct 9C14 (14.4 nM) and LacI tetramer ranging in concentration from 0–32 nm in LacI buffer (25 mM Tris, 100 mM KCl, 5 mM MgCl2, 2 mM DTT, 50 µg/µL, 0.02% Nonidet P40 (NP40) detergent, pH 7.8). The fraction of DNA shifting was used to calculate the active LacI concentration relative to the protein A_{280} and these results are available in Appendix 1. Radiolabeled constructs were prepared by labeling 100 µM of the D_E Top (unmodified) primer with 70 µCi γ-32P-ATP (Perkin-Elmer) following manufacturer's protocol for T4 polynucleotide kinase PNK (NEB). The primer was separated from unincorporated ATP using a P6 column (BioRad). The labeled primer was mixed with the A_E Bot (unmodified) primer and radiolabeled 9C14 construct was produced following the PCR protocol for externally labeled fluorophores described above. The labeled product was visualized on a STORM, and gel purified as discussed. The final DNA concentration was determined by scintillation counting (Packard 1600TR) with the specific activity being known from that of the labeled primer.

Body labeling the construct during PCR was used later, described in chapter 4, and provided more accurate quantification of radiolabeled DNA concentrations.

3.1.4 LacI Electrophoretic Mobility Shift Assays on the Construct Landscape

The radiolabeled landscape DNA constructs were prepared by digesting either the plasmid or a PCR product with *EcoRV* following manufacturer's suggestions (NEB), but allowing the digest to occur overnight. The PCR products were derived from the

procedures described by Mehta using the extended reverse primer (ERP) and T7 promoter primer. The digested products were phenol-chloroform extracted and EtOH precipitated followed by dephosphorylation of the 5' ends using calf-intestinal phosphatase (CIP) following the manufacturer's protocol (NEB). The dephosphorylated product was gel purified as described earlier and was again phenol chloroform extracted and EtOH precipitated. The 5' ends were radiolabeled using 50 μCi of γ-P^{32}-ATP per DNA construct and following the manufacturer's conditions for PNK, but reacting was maintained for 4 hours. Unreacted label was separated using a P6 or P30 biospin column (BioRad), and the final DNA concentration was determined based on the cpm's and assuming 100% labeling efficiencies for both ends. As stated earlier, body-labeling PCR products, discussed in chapter 4, would have provided more accurate determinations of DNA concentrations.

DNA constructs, 2 nM, were mixed with LacI (0.5 , 2, 4, and 8 nM) at room temperature for 15 minutes. Samples were mixed with loading dye and loaded onto a pre-electrophoresed 6% PAGE gel (75:1, acrylamide:*bis*-acrylamide) at 20 V/cm and 16° C for 2 hrs. Each series, for example 5C10-18, was processed together and loaded onto gels derived from the same stock for relative comparison. After running, the gel was dried on a filter paper and exposed overnight to a phosphor which was scanned on a STORM. This was repeated for all landscape constructs.

3.1.5 Bulk FRET Studies on the LacI•DNA FPV Landscape

All bulk FRET experiments were performed on a Varian Cary Eclipse Fluorescent Spectrophotometer courtesy of the Herman Sintim Lab, and excitation of donor (Alexa 555) was performed at 514 nm and the emission spectrum was scanned from 550-750 nm with a 10 nm slit width. The acceptor (Alexa 647) was directly excited at 600 nm and emission was scanned from 650-750 nm for fluorescence intensity comparisons to other FPV combinations. To verify that donor quenching and acceptor fluorescence was a result of FRET and not some other artifact, excitation spectra were also obtained. In these cases the donor excitation spectra was scanned from 400 to 550 nm and monitored at 570 nm, and acceptor excitation spectra was scanned from 400 to 650 nm and monitored at 670 nm.

For each FPV combination for each construct on the landscape, donor and acceptor emission spectra scanned for a 2 nM DNA sample. Because the on-rates for loop formation are very fast, the DNA was mixed with LacI for 1 min before measuring emission spectra. Possible re-equilibration will be discussed in Chapter 5. After scanning the LacI•DNA complexes, 5 mM IPTG was added and allowed to incubate for another minute before rescanning. For one set of experiments the FPVs that still showed FRET after IPTG saturation had an additional 3 nM of LacI added and were rescanned after another minute of incubation.

Competition studies were performed with the FPV that showed the highest FRET for each of the 24 constructs. As before, 2 nM of DNA was incubated with LacI and then emission spectra were scanned. Then 20 nM of unlabeled DNA competitor was added

and the sample was again scanned. Each construct was competed with the corresponding unlabeled construct. Since the lifetimes of previous constructs were on the order of days, the samples were measured after 2, 24, 48, 72, and 96 hours. However, the buffer signal and donor enhancement began to overwhelm the acceptor signal after 24 hours and the information from the donor and eventually the acceptor became compromised. This was overcome and the method and results are described in Chapter 5.

3.1.6 Calculations for Bulk FRET Efficiencies

The FRET efficiencies reported in the FPV landscapes were obtained from the emission spectra. All spectra were buffer corrected and spectrally decomposed into reference spectra components using Matlab 2009b as described later in this section. The calculated components are used to calculate the energy transfer. The final equation for calculating the energy transfer for both donor quenching, ET_{donor}, and acceptor enhancement $ET_{acceptor}$, is shown below. The ET_{donor} is the fraction of the donor fluorescence that is lost relative to the coefficient for the no transfer state, $C_{donor}(\text{DNA only})$, due to the energy transfer from the donor, $C_{donor}(\text{+xfer})$ to the acceptor when energy transfer is possible. This is expressed in terms of spectral components in the equation below. The ET_{donor} is dependent on $1/f_A$ because energy transfer is only possible when the acceptor is present and ET_{donor} does not depend on the quantum yield of the acceptor, Θ_A.

$$ET_{donor} = \left[\frac{C_{donor}(\text{DNA only}) - C_{donor}(+\text{xfer})}{C_{donor}(\text{DNA only})}\right]\frac{1}{f_A} = \left[1 - \frac{C_{donor}(+\text{xfer})}{C_{donor}(\text{DNA only})}\right]\frac{1}{f_A}$$

$$ET_{acceptor} = \left[\frac{C_{acceptor}(+\text{xfer}) - C_{acceptor}(\text{DNA only})}{C_{donor}(\text{DNA only})}\right]\frac{\Theta_D/\Theta_A}{f_A}$$

The $ET_{acceptor}$ is calculated from its fluorescence enhancement. In this case the increase in the acceptor coefficient, in a state that can do energy transfer, $C_{acceptor}(+\text{xfer})$, relative to a state that undergoes no energy transfer, $C_{acceptor}(\text{DNA only})$, is normalized $C_{donor}(\text{DNA only})$. However, acceptor energy transfer depends on the quantum yield of both fluorophores, Θ_D / Θ_A, and on the labeling efficiency of the donor, f_D, but not the f_A. Therefore,this was scaled to the ratio of the quantum yield of the donor, Θ_D, relative to the quantum yield for the acceptor, Θ_A, and divided by f_A. The relationship of Θ_D / Θ_A is described later in this section and outlined in the Matlab Script in the Appendix for calculating the energy transfer efficiency.

The spectral decomposition is worth describing in more detail because we put significant effort into making comparisons among molecules as robust as possible. First of all for convenience we wanted to express the coefficients in units of DNA concentration. To get the best possible estimate of the true concentration of the singly labeled reference sample, we normalized the integrated emission for all reference samples to the average integrated emissions for the all doubly labeled samples for the corresponding FPVs.

The spectral decomposition began by calculating the coefficients $C_{FPV}^{D}(\text{cond})$ and $C_{FPV}^{A}(\text{cond})$, where "cond" represents the conditions of the experiment as described below, for the emission spectra for doubly labeled DNA samples, according to the representative equation below.

$$\left[\begin{array}{cc} C_{FPV}^{D}(\text{cond}) & C_{FPV}^{A}(\text{cond}) \end{array}\right] = \left[\begin{array}{cc} D_{FPV}^{reference}(\text{cond}) & A_{FPV}^{reference}(\text{cond}) \end{array}\right] \backslash \left[FPV(\text{cond})\right]$$

Where the (1 × 2) coefficient matrix $\left[\begin{array}{cc} C_{FPV}^{D}(\text{cond}) & C_{FPV}^{A}(\text{cond}) \end{array}\right]$ is obtained by solving a linear system of equations using the backslash operator in Matlab. This equation represents the decomposition of the experimental emission spectra (excited at 514 nm) of the FPVs under each condition (DNA only, LacI + DNA, and DNA + LacI + IPTG), $\left[FPV(\text{cond})\right]$, using the reference emission spectra for singly labeled donor and acceptor, $\left[\begin{array}{cc} D_{FPV}^{reference}(\text{cond}) & A_{FPV}^{reference}(\text{cond}) \end{array}\right]$ (excited at 514 and 620 nm respectively). Each spectra are 201 element column vectors representing the respective emission spectra as a function of wavelength.

It was necessary to use the corresponding conditions for reference and experiment because of the apparent quantum yield effects with the addition of LacI and/or IPTG observed as donor fluorophore enhancement. Therefore, each experimental condition required a corresponding singly labeled construct reference under identical conditions to decompose the experimental data. However, this could be a source of error because these changes may be sequence dependent and/or dependent on the environmental conditions of each LacI-DNA loop formed. Simply stated, a more quantitative analysis requires a landscape of singly labeled reference samples for each FPV and condition.

Before the coefficients could be directly related to FRET efficiency they had to be scaled to the concentrations of the fluorophores. However, because of the low yield for the purified doubly labeled constructs, typically trace acrylamide and phenol contaminated the UV spectra. The errors in determining an absolute DNA concentration was accepted. Therefore, even though all DNA was adjusted to 2 nM based on the A_{260}, the observed fluorescent intensity of the excited fluorophores across the construct

landscape varied by a factor of two most likely as a result of these impurities. However, all of the constructs were prepared in parallel from the same fluorescently labeled primer stock, hence the uncertainty in concentrations could be reduced by using the fluorescence intensity to determine the relative concentrations of reference samples. To do this, the integrated emission spectra for direct excitation of each dye was collected for all the FPV landscape constructs and the average was considered to represent a ~ 2 nM DNA sample. For example, averaging the integrated internal donor emission spectra for all constructs, *n*, in the FPV combination involving D_I solved for the internal donor fluorophore FPV,

$$\int_{550nm}^{640nm} \overline{D_I}.$$

$$\int_{550nm}^{640nm} \overline{D_I} = \frac{1}{2n}\sum_{i=1}^{n}\left(i\int_{550nm}^{640nm} D_I A_I + i\int_{550nm}^{640nm} D_I A_E \right) \simeq 2\ nM$$

This generated a single value that was used with the FPV reference spectra to create the generalized normalized reference spectra, $\left[D_{FPV}^{reference}\right]$ and $\left[A_{FPV}^{reference}\right]$, which compares the integrated reference DNA spectra with the average integrated fluorescent intensity for all FPVs that share that specific combination.

$$\left[D_{FPV}^{reference}\right] = \frac{\int_{550nm}^{640nm} D_{FPV}}{\int_{550nm}^{640nm} \overline{D}_{FPV}} \quad \text{and} \quad \left[A_{FPV}^{reference}\right] = \frac{\int_{641nm}^{750nm} A_{FPV}}{\int_{641nm}^{750nm} \overline{A}_{FPV}}$$

The final normalized coefficients $C_{FPV}^{D}(norm)$ and $C_{FPV}^{A}(norm)$ are simply given by scaling the $C_{FPV}^{D}(\text{cond})$ and $C_{FPV}^{A}(\text{cond})$ to the relative concentration of the corresponding reference, $\left[D_{FPV}^{reference}\right]$ or $\left[A_{FPV}^{reference}\right]$.

$$C_{FPV}^{D}(norm) = C_{FPV}^{D} * \left[D_{FPV}^{reference}\right] \quad \text{and} \quad C_{FPV}^{A}(norm) = C_{FPV}^{A} * \left[A_{FPV}^{reference}\right]$$

The equation for energy transfer include f_A and f_D, which we measured from the absorbances of the purified primer. However, our calculations indicated that f_A and f_D

were > 100 % for the purified primers.Therefore, we believe the labeling efficiencies may not accurately reflect the doubly labeled DNA construct used in the experimental setting. Furthermore, after PCR, gel purification, storage, and use of the double-labeled dsDNA molecules, absorbance measurements on the dsDNA constructs indicate that the labeling efficiencies could be as low as 50%. We believe this is the lower limit based on comparison of the fluorescence intensity of the UV absorbance of the double stranded DNA versus the labeled primers, which suggests they may have decreased ~ 50 % of the initial value for double labeled constructs. Low PCR yields (due to primers that run through symmetrical operators) make absolute labeling efficiency impossible to measure precisely, and since we also do not know the environmental influences on the fluorophore properties (Delon, Wang et al. 2010), we chose not to rescale the apparent energy transfer efficiency. However, all constructs originated from the same primer stock and were systematically processed in parallel. Since each FPV shared one fluorophore with another FPV, each FPV landscape could be normalized to the other FPV landscapes by using the integrated emission spectra of the donor and acceptor in doubly labeled DNA-only samples. For example, the relative integrated emission area ratio of external-to-internal acceptor f_{A_E}' / f_{A_I}', was calculated with respect to their shared partner for the internal donor, as shown in the equation below. $\int_{641nm}^{750nm} \overline{A_E}$ is the average integrated fluorescent intensity for all internal and external acceptor emission spectra respectively described earlier. Similarly, the relative integrated emission area ratio of externally labeled donor, f_{D_E}', to internally labeled donor, f_{D_I}', was calculated with respect to their shared FPV.

$$\frac{f_{A_E}'}{f_{A_I}'} = \frac{\int_{641nm}^{750nm} \overline{A}_E}{\int_{550nm}^{640nm} \overline{D}_I} \Bigg/ \frac{\int_{641nm}^{750nm} \overline{A}_I}{\int_{550nm}^{640nm} \overline{D}_I}$$

The labeling efficiency of the external acceptor, f_{A_E}, with respect to internal acceptor, f_{A_I}, was calculated from the equation shown below.

$$f_{A_E} = f_{A_I} * \left(\frac{f_{A_E}'}{f_{A_I}'} \right)$$

The value for the internal acceptor and donor labeling efficiency, f_{A_I} and f_{D_I} respectively, had to be assumed to determine the relative labeling efficiencies for the other FPVs; the values were considered to be 100% for our calculations. The relationship of f_{D_E} to f_{D_I} was calculated similarly to f_{A_I}.

The final consideration for calculating FRET efficiency in our experiment involved solving for the quantum yield of acceptor and donor. First, Invitrogen's values for the Alexa555 and Alexa647 fluorophores are for free dye in solution and may not be applicable to our dyes conjugated to an oligonucleotide. Because of this possibility we began by calculating our own extinction coefficient values based on the UV-Vis absorbance spectra of the primers. In the bulk FRET experiments the donor, Alexa555, was excited at 514 nm and its emission maxima appeared at 555 nm, with an absorbance ratio of 514/555 = 0.7 (Invitrogen reported 0.4). The primer labeled with acceptor, Alexa647, had a ratio of 0.4 (Invitrogen reported 0.3) based on the excitation at 600 and emission maxima at 647. However, we were unable to measure our own extinction coefficients, ε, Invitrogen's reported values, 150,000 and 239,000 $cm^{-1}M^{-1}$ for donor and acceptor respectively, were used directly with confidence in the manufacturers values.

Finally, this required relating the fluorescent intensity relative to their concentrations, extinction coefficients, and labeling efficiencies as shown in the equation below.

$$\frac{\Theta_D}{\Theta_A} = \left(\frac{\int_{550nm}^{640nm} rD}{[rD]} \Big/ \frac{\int_{641nm}^{750nm} rA}{[rA]} \right) * \left(\frac{\varepsilon_A}{\varepsilon_D} \right) * \left(\frac{fA}{fD} \right)$$

Note that as described earlier this expression means that the calculated energy transfer for acceptor enhancement does not depend on the f_A .

The backslash operator used in matlab provided a a best fit solution of the reference spectra to the experimental spectra. We then calculated a fit from this solution, as shown in the equation below, and then determined residuals $\left[\; res\_D_{FPV}(\text{cond}) \quad res\_A_{FPV}(\text{cond}) \; \right]$ by subtracting the fit from the experimental spectra. Each matrix in the equation is a 201 element column vector except for the coefficient matrix, which has a singular value (1 × 2 matrix).

$$\left[\; res\_D_{FPV}(\text{cond}) \quad res\_A_{FPV}(\text{cond}) \; \right] = \left(\left[\; D_{FPV} \quad A_{FPV} \; \right] - \left[\; D_{FPV}^{reference} \quad A_{FPV}^{reference} \; \right] * \left[\; C_{FPV}^{D} \quad C_{FPV}^{A} \; \right] \right)(\text{cond})$$

3.1.7 Extracting Loop Population Distributions from Bulk FRET Efficiencies

The loop population distributions were derived from the energy transfer values for each construct using assumptions that were based on observed trends in the data. (1) Each uniform loop type (l = P1, A1, A2, or E) is assumed to have a uniform FRET efficiency (ET_l) within each FPV landscape: for example, all D_lA_l constructs have the same ET when they are in the P1 loop conformation. (2) Based on the observation that the peak ET in each FPV landscape corresponds to minima in the other two landscapes, we assume that the ET for the dominant loop topology for each landscape is simply the maximum

observed ET: for example, ET for A1 loops is taken to be 18% based on the maximum ET observed in the D_IA_E landscape. Furthermore, we assume no "crosstalk" between landscapes, so ET = 0 for all loops other than the dominant loop. Quantitative calculation of crosstalk (non-zero ETs for each loops in each FPV) would require more FPV landscapes to allow for simultaneous calculation of the ETs as well as populations. (3) The ET for extended (E) loops is taken to be uniformly zero, which is required in order to calculate populations for three loops (P1, A1, and A2) from three FPV landscapes (D_IA_I, D_IA_E, and D_EA_I). Hence, the calculated amounts of E loops are minimum estimates; this is also true because the total observed energy transfer is < 100% for all constructs. (4) The populations fl for each loop type sum to 1 based on the observed hyperstability of all the loops. Any "missing FRET" corresponds to. extended loops. For IPTG saturated LacI•DNA looped complexes used the same ET_l values, but "E" now represents a dark state that includes contributions of E loops and other complexes that result in decreased FRET efficiency.

In general, the f_l are given by the solution to the set of three equations below, solved individually for each construct, subject to the constraints that each f_l is between 0 and 1 and the f_l sum to 1:

$$ET_{FPV} = \sum_l f_l \cdot ET_l, \text{ for } FPV = D_IA_I, D_IA_E, \text{or } D_EA_I$$

where the ET_{FPV} are the observed ET for each FPV. Given the assumptions above, the populations fl for P1, A1, and A2 should reduce to the ratios ET_{FPV}/ET_l and the fraction f_E found in extended looped states should be simply $f_E = 1 - f_{P1} - f_{A1} - f_{A2}$. In practice, because some of the calculated ET values are negative (although small) and some of the

constructs exhibit crosstalk, the ET_{FPV} equations do not give exact solutions for the f_l. Therefore, we used the constrained minimization tools in Matlab to give best-fit values.

The ET_l input values were based on the FRET peak amplitudes of each FPV or l, which we consider to generally represent a uniform population. Therefore, the ET_{P1} representing the maximum observed energy transfer for a P1 topology landscape (D_IA_{Imax}) had the value D_IA_{Imax}, 0, 0, and 0% energy transfer for the FPV's D_IA_I, D_IA_E, D_EA_I, and E loop respectively. This was applied similarly for ET_{A1} (0, D_IA_{Emax}, 0, and 0%) and ET_{A2} (0, 0, D_EA_{Imax}, and 0%) with 0% energy transfer for all FPV's in ET_E. Our analysis of the IPTG saturated LacI•DNA looped complexes used the same input values, but the ET_E (which is always represents a 0 % FRET efficiency) now represents a dark state that includes contributions of E loops and any other complexes that contribute to a decrease in the observable FRET.

The residuals were obtained by subtracting the calculated f_l multiplied by our ET_l values used to solve for the populations from the observed energy transfer ET_{obs}.

$$residuals = ET_{obs} - (ET_l * f_l).$$

The result of the residuals for population modeling indicated there is more FRET than can be explained by our model observed in a subset of constructs. This suggests the values for each ET_l may be set too low, i.e. there is enhanced fluorescence in this subset of LacI•DNA looped complexes or there is crosstalk between the FPV landscapes. Since fluorescent enhancement and FPV crosstalk could not be directly addressed, modifying the ET_l constraints allowed further manipulations to ET in order provided alternative solutions to our population modeling to make sure our qualitative conclusions were

appropriate and did not depend on the assumed values for ET_l. The Matlab scripts and functions used for all of the calculations are included in Appendix 3.

To test the contributions of crosstalk between FPV landscapes, two additional populations were included that had non zero D_IA_I—D_IA_E and D_IA_I—D_EA_I ET_l values. Different combinations of ET_l values were used to minimize the residuals of the fit, while using the initial values for the four populations that provided the minimal estimate of the E loop complex described earlier. (the maximum of each FPV landscape was used for the ET_l and is described further in the results in this chapter).

3.2 Results

3.2.1 EMSAs demonstrate all constructs can form LacI•DNA Loops

The EMSAs performed on the construct landscape indicated all constructs could form a LacI•DNA looped complex. In Figure 3.2.1 three gels from the landscape are shown for constructs 7C12, 9C14, and 11C10. The LacI•DNA looped complex for all constructs appeared to shift at approximately the same location and be composed of one band, not multiple. The difference in gel conditions makes it difficult to quantify any mobility changes as a result of differences in loop shape or affinity, and requires a looping ladder standard, such as construct 9C14, for each gel.

In an attempt to obtain some qualitative information from the gels, each EMSA was examined for tightness of bands for LacI•DNA looped complexes relative to DNA only, which we inferred to represent stability of a complex or the possibility of multiple loop topologies or a single topology. Additionally, the stoichiometry at which LacI shifted

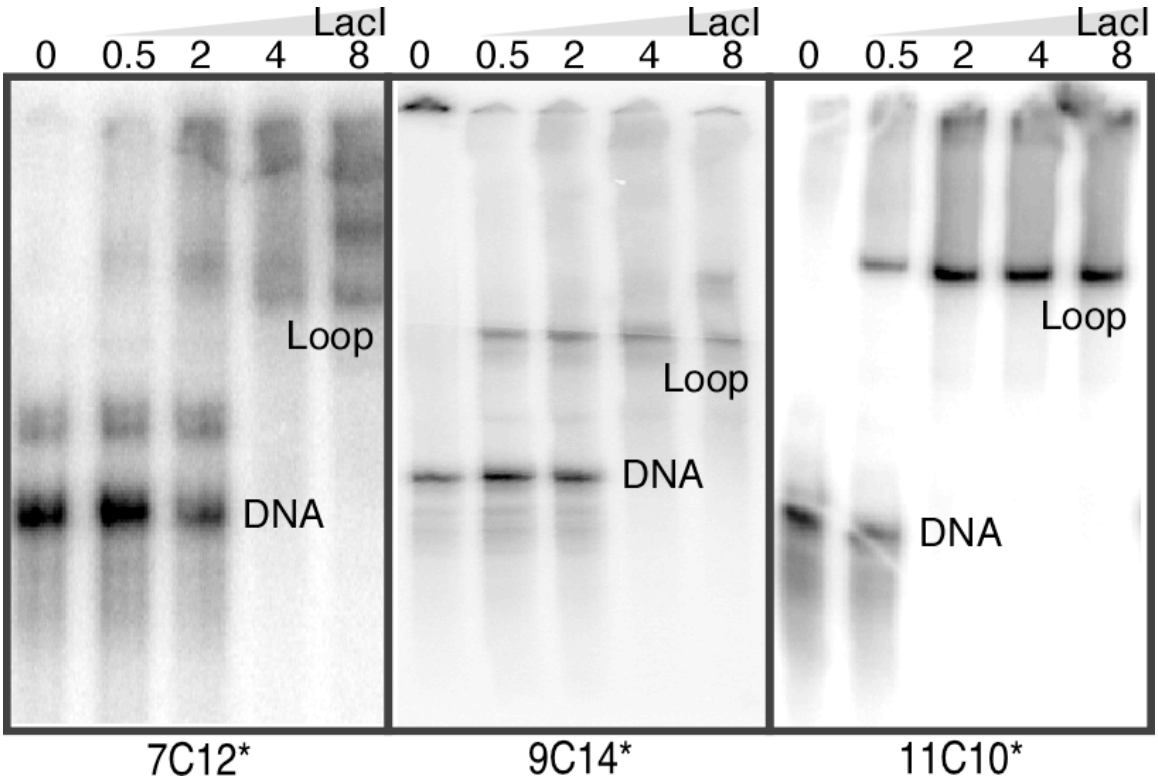


Figure 3.2.1: *All constructs in the landscape form LacI•DNA looped complexes in an Electrophoretic Mobility Shift Assay (EMSA)*. LacI (0, 0.5, 2, 4, and 8 nM) was incubated with 2 nM ^{32}P end labeled DNA for 15 minutes. Samples were loaded onto a 6% PAGE (75:1, acrylamide:bis-acrylamide) at 20 V/cm and 16o C and were ran for 2 hours in TBE buffer. Constructs 7C12, 9C14 and 11C10 all formed LacI•DNA looped complexes (Loop) but differed on the stoichiometric requirements before complete shift of free DNA (DNA). Variations in gel conditions, radiolabeling, and overall experimental quality prevented quantification of these experiments as described in the Materials and Methods in Chapter 3. Qualitative assessment of the entire construct landscape is shown in Appendix 1.

all DNA was qualitatively handled, and because of the nature of these data both of these qualitative results are include in the Appendix 1.

3.2.2 All Constructs Show FRET in at Least One FPV When Looped by LacI

FRET efficiency was measured as described in Materials and Methods in Chapter 3 for each of the 96 DNA constructs in the presence of saturating LacI and also LacI+IPTG (not 100: the sequence of the 5C12 amplicons were found be incorrect with follow up

sequencing after the experiments were completed). Fluorescence spectra for the LacI-mediated looping constructs with the most efficient energy transfer drawn from each of the D_IA_I, D_IA_E, and D_EA_I FPV landscapes are shown in Figure 3.2.2 (all spectra taken for one set of experiments are shown in Appendix 1). Each fluorescence emission spectrum was decomposed into a linear combination of donor and acceptor emission components using the reference spectra obtained. FRET efficiency was calculated separately from both donor quenching and enhanced acceptor emission data. Enhanced donor quenching and acceptor emission are both qualitatively obvious in the spectra of Figure 3.2.2.

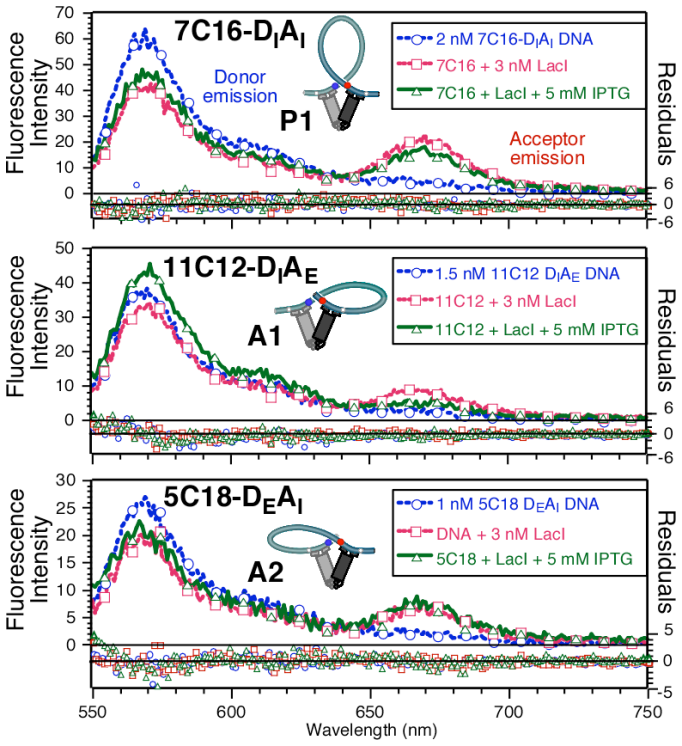


Figure 3.2.2: *Different FPV/construct combinations identify the three FRET-observable loop types.* Emission spectra are shown for DNA alone (thin dashed line), DNA + 3 nM LacI (solid line) and DNA + LacI + 5 mM IPTG (dashed line), for the three constructs exhibiting the highest energy transfer for each FPV landscape. Enhanced donor quenching and acceptor emission are both qualitatively obvious in the spectra. As described in Materials and Methods, the spectra were decomposed into a linear combination of donor and acceptor emission components using reference spectra obtained for singly labeled construct 9C14 ± LacI or LacI+IPTG, with the residuals plotted below the spectra as dots. Using the emission components, FRET efficiency was calculated separately for both donor quenching and enhanced acceptor emission. LacI binding increases the fluorescence of singly-labeled donor (D_I and D_E) 9C14 by about 5% and donor fluorescence increases a further 10% with the addition of IPTG, but direct acceptor emission was not affected by addition of LacI and/or IPTG. Because of this enhanced donor quantum yield, the decrease in donor emission due to energy transfer is less than would be anticipated in comparison to the dramatic increase in the acceptor emission.

However, LacI binding increases the fluorescence of singly-labeled (DI) 9C14 by about 5% and donor fluorescence increases a further 10% with the addition of IPTG, but direct acceptor emission was not affected by addition of LacI and/or IPTG. Therefore, quantitatively the decrease in the donor emission is less than would be expected in comparison to the dramatic increase in the acceptor emission. We cannot quantify the effect that enhanced donor fluorescence has on the observed acceptor emission in the context of a doubly-labeled sample undergoing energy transfer, but believe both values should be considered in unbiased FRET efficiency interpretation. Therefore, all of the quantitative results described henceforth are reported as averages of the donor and acceptor efficiency values, but all of the qualitative features of the landscapes are apparent in the individual data sets as well as the average as described in Section 3.2.3.

Energy transfer efficiencies for the constructs giving maximal FRET for each FPV (Figure 3.2.2) are 30% for 7C16-D_IA_I (expected to provide maximal FRET in the P1 topology), 18% for 11C12-D_IA_E (A1 topology), and 18% for 5C18-D_EA_I (A2 topology). No FRET was observed for any D_EA_E construct, suggesting that P2 loops do not exist, and "sandwich complexes" do not contribute substantial FRET for any of these molecules under our conditions. We would expect that any sandwich complex formation should show up equally on each landscape, and the flat FRET efficiency surface for the D_EA_E landscape implies that any FRET contributions of the sandwich complex are negligible or non-existent. The overall lower than anticipated FRET could indicate some overestimation of the true fluorophore labeling efficiencies, which we take as 100% based on the absorbance spectra of labeled primers. In fact, after PCR, gel purification,

storage, and use of the double-labeled dsDNA molecules, absorbance measurements on the dsDNA constructs show that the labeling efficiencies are at least 50%, but because low PCR yields make absolute labeling efficiency impossible to measure precisely, we have not rescaled the apparent energy transfer efficiency. The quoted energy transfer efficiencies are, however, adjusted for the small (~10%) differences in the relative labeling efficiencies of D_I versus D_E and of A_I versus A_E primer sets described in the Materials and Methods in Chapter 3.

3.2.3 The FPV Landscapes for LacI Mediated DNA Looping

Figure 3.2.3 presents the four FPV FRET efficiency landscape as 5 x 5 matrices with each column sharing a donor-side phasing adapter and each row an acceptor-side adapter. The landscapes were obtained on sets of molecules that were prepared and analyzed in parallel from the same batches of fluorescent PCR primers. The same qualitative features were observed for landscapes obtained from different primer batches with varying labeling efficiencies, as well as individual FRET measurements from donor and acceptor in Figure 3.2.3A and 3.2.3B respectively. Therefore, comparisons among molecules within a landscape and among separate landscapes are robust; in particular, the quantitative differences in maximum ET among the different landscapes are greater than any estimated experimental errors and are reproducible. The standard deviations for multiple trials on the landscapes as well as the FRET efficiencies are presented in Appendix 1.

The distinct peaks in the separate FPV landscapes allow us to identify the regions where P1, A1 and A2 loops dominate. Initially we assume a rigid LacI protein as in

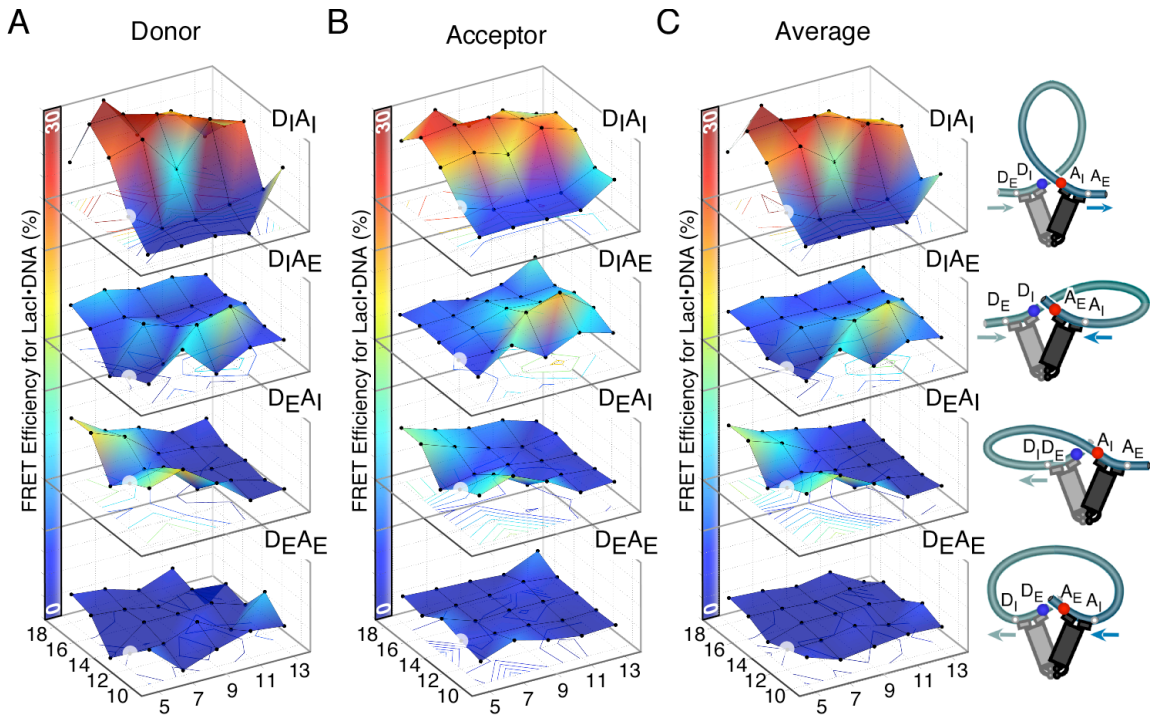


Figure 3.2.3: *FRET efficiencies for LacI•DNA landscape presented as individual FPV landscapes*. The four FPVs for each construct in the looping landscape were prepared in parallel PCR reactions using one set of four primers. This yielded a total of 96 double-labeled molecules; the white spherical blurs for 5C12 are averaged from adjacent values on the map for visual continuity. The FRET efficiency calculated for donor quenching (A) and acceptor enhancement (B) are described in the Materials Methods. (C) The average between the acceptor and donor FRET efficiencies show all the same qualitative features, but vary slightly depending on the FPV landscape.

previous theoretical work (Lillian and Goyal et al. 2008), and then we ask to what extent this constraint must be modified in response to experimental data. The constructs giving maximum energy transfer in one FPV landscape generally show little or no transfer in the others, confirming the implications in Figure 2.5 (the Experimental Design in Chapter 2) that a loop of a particular topology shows strong energy transfer for only one arrangement of fluorophores. For example, 9C16 shows 28% transfer in the D_IA_I (P1-sensitive) landscape but no more that 3-4% in either of the others, and 11C12 shows 18% transfer in the D_IA_E (A1-sensitive) landscape and no transfer in either of the others).

The peaks in the D_IA_E (A1) and D_EA_I (A2) landscapes do not show significant FRET in the other FPV landscape, suggesting that the A1 and A2 peaks represent populations generally inclusive to A1 and A2 antiparallel loops. The fluorophores are all in analogous positions relative to the lac operators, thus we expected that the maximal level of energy transfer would be the same for each of the loop topologies on its corresponding FPV. In contrast, Figure 3.2.3 shows that the maximum FRET efficiency is much lower (18%) on the antiparallel loop landscapes than on the P1 landscape (30%). Additionally a subset of constructs 5,7(C)16,18 show efficient transfer in both the P1 and A2 landscapes. The simplest explanation for the energy transfer differences among loop topologies is that LacI is deformed in such a way that fluorophores are farther apart in the A1 and A2 loops compared to P1, and LacI is deformed or compressed to have FRET crosstalk between FPV landscapes. Models for LacI flexibility and the implications of enhanced/diminished FRET are proposed in the Discussion of Chapter 3.

3.2.4 LacI•DNA Loops Appear Resistant to DNA Competitor Challenge

The total FRET for all FPVs is not a flat surface in Figure 3.2.4A and varies substantially across the landscape. This suggests that either some loops are unstable, which would lower FRET for that complex, or LacI is deformed enough in some looped complexes to reduce the overall FRET for some constructs. Competition experiments using unlabeled DNA competitors show in Figure 3.2.4B that the observed acceptor FRET efficiency for every construct, when measured for its highest-efficiency FPV, is resistant to competitor DNA challenge for many hours. Unlooped singly bound or doubly bound species would be unstable toward competitor DNA via intermediate sandwich

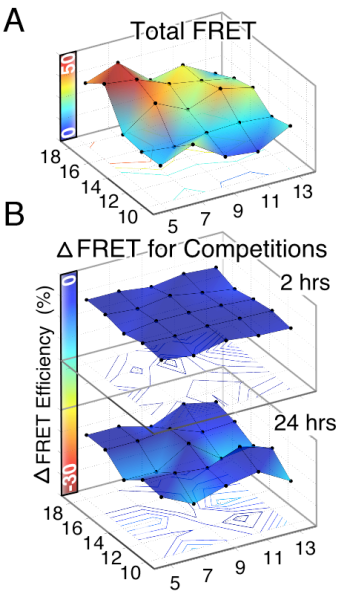


Figure 3.2.4: *High-FRET LacI•DNA complexes resist DNA competitions for hours*. (A) The total FRET, summed for all FPVs in Figure 3.2.3, is not a flat surface. This could be due to unstable LacI•DNA loops or extended LacI•DNA loops that contribute no FRET. (B) The LacI•DNA complexes for each construct's FPV that had the highest looping FRET efficiencies (DNA:LacI at 2:3 nM ratios) were preincubated before the addition of 20 nM unlabeled competitor DNA. Samples were scanned for days, but dramatic changes to the donor signal after one day convoluted data after 24 hours. The change in acceptor enhancement and not donor quenching are reported here following 2 and 24 hours of incubation with competitor DNA. Since, FRET did not drastically change for the acceptor in the presence of competitor over 24 hours, the loops must be stable. The missing FRET in (A) must be the result of dark loops that do no FRET.

complex formation. This suggests, perhaps surprisingly, that all of the constructs form stable loops and therefore that the total of the loop topology populations should be 100% for each construct, especially in the absence of DNA competitor. However, initial analysis showed that this is not the case. Four constructs, 5C16, 5C18, 7C16, and 7C18 show an apparent population greater than 100%; as mentioned above, we ascribe this to crosstalk between FPV landscapes. Other constructs **(**5C14, 7C12, 9C10, 11C10, 13C10, 13C12**)** show a total population less than 90% based on FRET**.** Moreover, construct 11C10 has 7% FRET for the FPV D_IA_E and no other FRET contributions in the other FPV landscapes, yet in EMSA studies shown in Figure 3.2.1 formed LacI•DNA complexes at near stoichiometric conditions. These observations could be due to substantial variation in transfer efficiency among constructs within a given topology, or there may be looped states that exhibit much lower energy transfer efficiency than P1, A1, and A2 (D_IA_I,

D_IA_E, and D_EA_I respectively). Either possibility requires deformation away from V-shaped LacI, but our bulk FRET data cannot distinguish between them.

3.2.5 Loop Topology Population Distributions Calculated from FRET Studies of the FPV Landscapes

The energy transfer landscapes of Figure 3.2.3 each provide a different window on the underlying distribution among loop topologies. To extract these populations rigorously would require knowledge of the energy transfer efficiency for each topology for each individual construct, which is not available from our data. Instead we simply assume that the energy transfer efficiency for a given topology is constant, i.e. that the deformation of LacI is similar for each construct within a given loop topology. This allows the population of each topology to be calculated by dividing the energy transfer in the corresponding FPV landscape by the maximum amount of transfer in that landscape.

The simplest way to model the population distributions is to assume a fourth looped state that does no energy transfer, such that the population of this state quantifies the total "missing energy transfer" across the landscapes. We refer to this as the Extended or E loop state, as in previous models (Mehta Kahn, 1999; Swigon et al. 2006a). Inclusion of this state allows us to compute final loop topology populations as described in Materials and Methods in Chapter 3, using the assumption that energy transfer is uniform for each topology. The calculated landscapes for P1, A1, and A2 (D_IA_I, D_IA_E, and D_EA_I respectively) loops for all constructs shown in Figure 3.2.5 are similar in shape to the observed FRET efficiency shown in Figure 3.2.3.

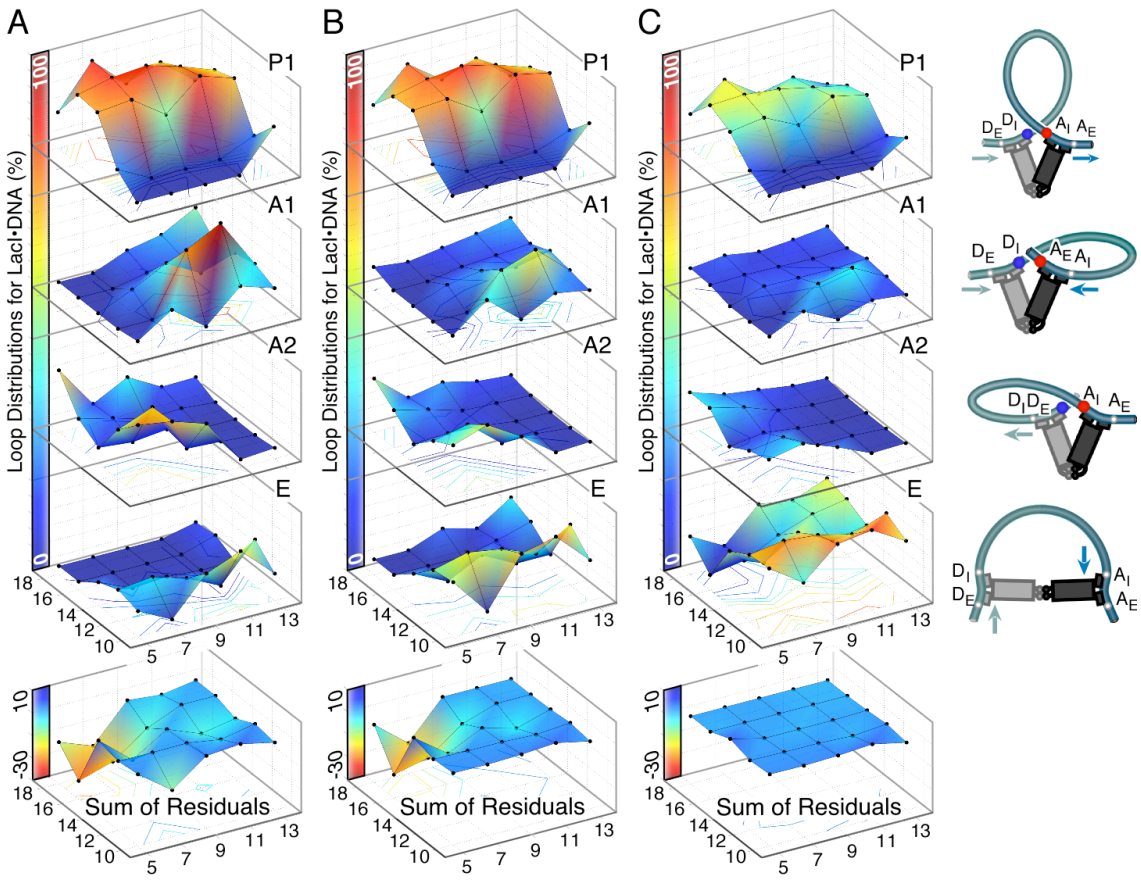


Figure 3.2.5: *Loop population distributions landscapes calculated from the FRET efficiencies of the FPV landscapes*. (A) The four populations were calculated with two assumptions as described in Materials and Methods in Chapter 3: the four loop populations must add up to 100% for an individual construct, and the FRET efficiency maxima for each FPV represents a uniform population that corresponds to a specific loop topology. The E landscape accounts for the difference between the observed total FRET and the expected total FRET for a population of V-shaped LacI. It represents a putative 0% FRET complex that, based on kinetic stability results (Figure 3.2.4), is still looped. Averaged over the entire landscape, the P1, A1, and A2 topologies contribute up to 43%, 22%, and 22% of the total population respectively, with the E loops contributing at least 13%. The residuals of the population fit are shown at the bottom, and the negative values for constructs (5,7)C(16,18) suggest possible LacI loop geometries with enhanced FRET or some crosstalk between FPV landscapes. (B) Increasing the maxima for each uniform population, to the same FRET maxima observed for all FPV landscapes (~ 30%), decreases the antiparallel loop contributions and increases E loop contributions. This analysis would suggest that antiparallel loops are mixtures of V-shaped loops and E loops, but has little effect on the populations calculated for V-shaped P1 loops and the residuals (C) Assuming a uniform V-shaped loop topology is equal to FRET maxima observed for the sum over all FPV landscapes (~ 50%) reduces all of the V-shaped looped populations and improves the sum of the residuals.

The residuals for the fit between the FPV landscapes and the population modeling, shown at the bottom of Figure 3.2.5, indicated that there is more FRET observed in a subset of constructs mentioned earlier. This suggests the constraints for each ET_l may be set too low, there is enhanced fluorescence in this subset of LacI•DNA looped complexes, or there is FRET crosstalk between the FPV landscapes. Since fluorescent enhancement could not be addressed, and FPV crosstalk between landscapes requires a deeper understanding of LacI headpiece flexibility in a looped state, further manipulations to ET_l in the constrained minimization function were performed. All rational adjustments still maintained the general shapes shown in Figure 3.2.5 and in the cases where ET_l was increased for all loops, as in Figure 3.2.5B-C the allocation of the E loop populations were equally amplified. More specifically, if we assume that the maximum FRET efficiency of all FPV landscapes should be equal to the observed FRET maxima across all landscapes (~ 30%), this results in decreased population distributions for the closed-geometry antiparallel loops and thereby increases the calculated E loop populations shown in Figure 3.2.5B. Although, this may more accurately reflect the maximum allowed FRET representing a uniform LacI•DNA looped complex observed for P1 loops, it also imposes expectations that each loop topology should share the same FRET maximum, which may not necessarily be true for this DNA sequence space. Finally, to minimize the residual differences between the FPV landscapes and the calculated population distributions a final calculation was made. In Figure 3.2.5C the sum of all FRET measured for all landscapes (~ 50%) was applied to all ET_l, excluding E loops, which still contributed no FRET on any surface. This drastically increased the

distributions calculated for the E loop, diminished all V-shaped loops, but minimized the residuals to almost zero. Previous calculations did not effect the P1 loops as seen in Figure 3.2.5A-B, and all three calculations for the E loop landscape indicate that all of the constructs that have substantial E contributions are otherwise forming antiparallel loops, and this significance will be described in the Discussion in Chapter 3.

3.3 Discussion

3.3.1 Direct Detection of Loop Topology Populations Using FRET FPV Landscapes

The peaks and valleys in the FRET landscapes for different FPVs reflect the free energy landscape of different LacI•DNA loop topologies, as in Figures 3 and 4. Comparisons between FPV landscapes demonstrate unambiguous FRET detection of antiparallel loops for the first time. The FRET maxima for each FPV landscape appear for constructs that show near-zero FRET in all the other landscapes, so the maxima represent constructs that adopt a uniform population of a single loop type. The boundary constructs that show significant FRET on multiple FPV landscapes adopt more than one loop topology, presumably because the sequence is not optimal for any single topology. Surprisingly, all of the bend-operator phasing variants form hyperstable loops, and because even the loops with low energy transfer efficiency appear resistant to competitor challenge, we infer that there is a substantial population of "dark" looped states that do not give FRET for any FPV.

3.3.2 FPV Landscapes Suggest LacI Flexibility

The maximum energy transfer efficiency for antiparallel loops (A1 and A2) is lower than for parallel (P1) loops for this set of constructs. This was initially puzzling, since a V-shaped LacI should enforce the same fluorophore distances for each topology. The lower FRET suggests a more open form of the loop, which would require LacI deformation. Figure 3.3.2A shows a plausible model for the driving force for this conformational change: an increase in the DNA radius of curvature, a response to DNA bending strain, tends to contract the V-shaped LacI in a P1 loop and result in higher

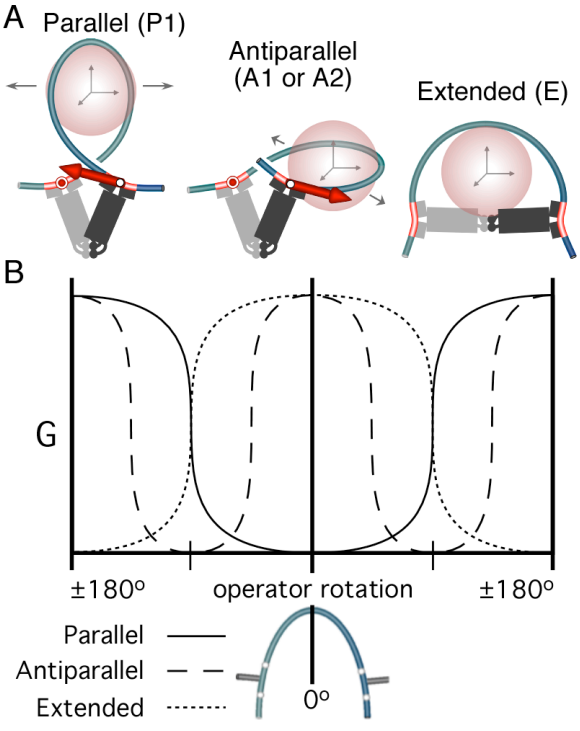


Figure 3.3.2: *Relief of DNA bending and torsional strain alters LacI•DNA complex geometries*. (A) The membrane represents the favorable free energy of DNA expansion, within the tightly bent DNA, enforced upon looping. Parallel loops contract LacI in response to the DNA bending strain. Antiparallel loops decrease DNA bend strain by opening LacI. Extended LacI minimizes DNA strain at the maximum energetic cost for LacI deformation. (B) Energetic representation of the preferred landscapes when both operators are rotated towards each other in either direction, but in unison. This representation assumes DNA bending energy, which is ~ 1/L2, energetically partitions preferred topologies less than the DNA torsion, ~ 1/L. As shown, V-shaped parallel loops should be lowest in energy when operators are facing away from the central DNA bend (180°) and are unfavorable when both operators are facing towards the bend. However, the 180° rotation of both operators favors the formation of the extended LacI•DNA loop. The energies of antiparallel loops would lie between these two extremes. The energy diagram suggests specific operator phasings could compete between V-shaped parallel and antiparallel loops, but parallel loops do not compete with extended loops.

FRET, as well as potentially inducing crosstalk between FPV landscapes as indicated in Figure 3.2.5 and is discussed later in Section 3.3.4. In contrast, DNA expansion in A1 and A2 loops would open up the LacI dimer-dimer interface and decrease FRET. With R_0 = 51 Å for this fluorophore pair, small changes in LacI geometry could result in large changes in FRET: a change from 30% to 18% efficiency requires only an 8 Å increase in distance.

As seen in the results in Figure 3.2.5 the E loop landscape contributes to the antiparallel loops and to a lesser extent the parallel loops. We believe this observation indicates, as shown in Figure 3.3.2B, the effects of operator phasings relative to an intrinsic bend. For example, the operator phasings that prefer P1 loops (operators oriented outward relative to the intrinsic DNA bend) would convert to extended looped complexes with the largest DNA twist energies, and they are therefore inaccessible. The E loop dark state free energy should be minimal when the operators are both oriented toward the intrinsic bend, which does not support the most stable P1, A1, or A2 loops.

3.3.3 The Extended LacI•DNA Looped Complex Exists, but the Extent is Unknown.

Given the assumption of uniform loops at the FRET maxima, the analysis reported in Figures 3.2.5 quantitates the population of dark looped states, which could result from further expansion of LacI to an open form that we refer to as the extended (E) form loop. We feel confident that with our FPV system we should be able to detect any loop that forms a V-shaped loop complex. However, with bulk data we cannot separate possible contributions from low-FRET versus dark states, so lower total FRET, observed

for antiparallel loops relative to parallel loops, could be due to uniformly expanded loops or to mixtures of V-shape and extended LacI loop geometries. We would expect that if LacI does open to an extended state there should be a continuum of LacI loop geometries that appear dark as suggested in Figure 3.3.3. The R_0 for our dye pair should allow us to measure 1/3rd of the potential space for LacI•DNA looped geometries.

Since the calculated total FRET, summed over all the FPVs, never reaches 90% (the FRET efficiency limit for this fluorophore pair determined from interfluorophore distances measured in Figure 2.5) for any construct it is difficult to know the extent to which the E loop population contributes. In the results in Figure 3.2.5 we reported three different possibilities of an extended LacI•DNA looped complex assuming different

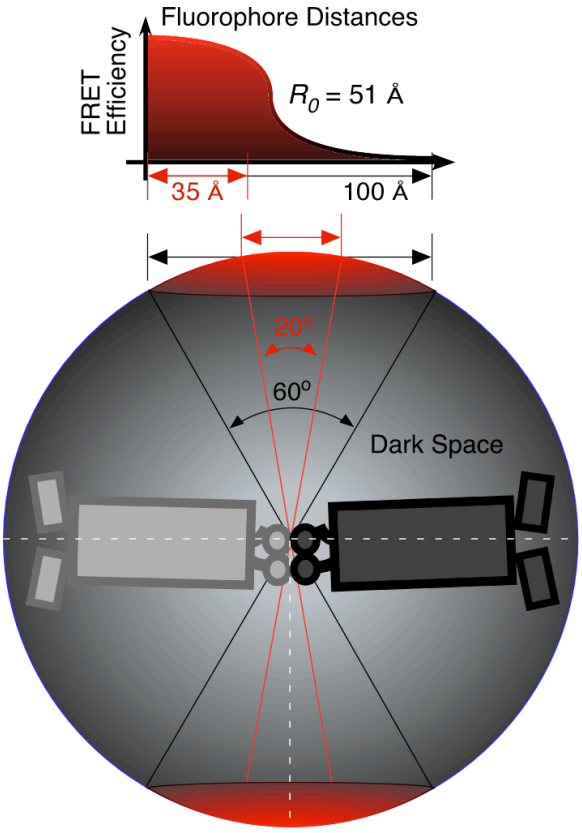


Figure 3.3.3: *FRET detection of the accessible space for LacI•DNA loop geometries.* The V-shaped LacI co crystal structure has an inscribed angle of ~ 20°, that separates the fluorophores in a looped DNA complex by ~ 35 Å corresponding to 90% FRET efficiency for the Alexa555–Alexa647 pair (R0 = 51). Our FPV detection can efficiently detect any of these geometries, independent of the initial DNA trajectories, but FRET detection is limited to LacI geometries that do not exceed an ~ 60° V-shaped looped complex (100 Å interfluorophore distance). Since LacI is symmetrical, this FRET detectable surface area accounts for 33% of the total surface, assuming LacI can form the depicted extended geometry there could be a continuum of LacI geometries in DNA looping

FRET maxima for uniform V-shaped topologies. The total E loop population contributions relative to all calculated loops range from 13%, the minimal estimate in Figure 3.2.5A, to 49% shown in Figure 3.2.5C. Because of the uncertainties in the absolute labeling efficiencies, the changes in fluorophore environment, and the contributions from extended loop states, we do not attempt to convert FRET efficiencies to interfluorophore distances, but rather fit the qualitative nature of the data with loop population interpretations. Single-molecule FRET could address all of these deficiencies, and techniques like FIONA and SHRIMP (Balci, Ha et al. 2005; Yildiz and Selvin 2005) could provide direct measurements of our "dark" state geometries.

3.3.4 Including Crosstalk Between FPV Landscapes in the Loop Population Analysis

Our population modeling in Figure 3.2.5 does not consider LacI headpiece movement, which LacI crosslinking has shown to be important for loop formation (Rutkauskas, Zhan et al. 2009). However, our population does allow for interpretations of crosstalk between FPV landscapes, but defining the extent to which a loop appears on another FPV landscape is difficult in the absence of any high-resolution LacI•DNA loop structures. However, in Figure 3.3.4 we considered the possibility of two additional deformed LacI•DNA loop topologies populations that crosstalk into two FPV landscapes. Inclusion of these two crosstalking populations demonstrated the constructs previously mentioned (5C16, 5C18, 7C16, and 7C18) as well as a few others could represent a deformed headpiece or a compressed LacI•DNA looped complex as depicted in Figure 3.3.2 and in Figure 3.3.4A. However, the boundaries between uniform loop topologies,

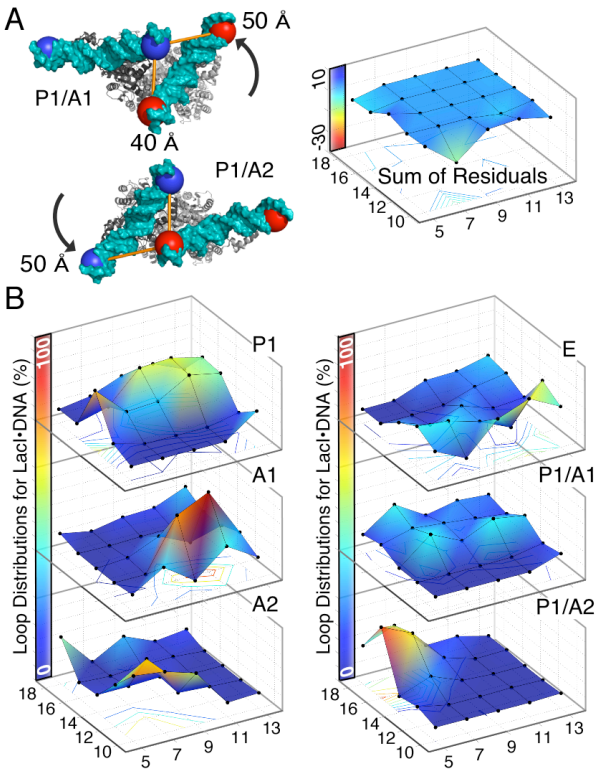


Figure 3.3.4: *Crosstalk between FPV landscapes and the calculated loop population distributions with two additional compressed V-shaped loops*. (A) Crosstalk between FPV landscapes could be caused by deformation of the LacI headpiece in a looped state away from the V-shape crystal structure (Lewis, Chang et al. 1996). For example, rotation of the acceptor side headpiece in a looped complex results in crosstalk between P1/AI FPV landscapes, and rotation of the donor side results in crosstalk between the P1/A2 FPV landscapes. Minimizing the residual fit, for calculating the loop population distributions with the addition of these two populations, required that the maximum FRET efficiencies for the D_IA_I and D_IA_E FPV landscapes be used for the ET_I constraint for a P1/A1 loop, and similarly for the D_IA_I and D_EA_I landscape and a P1/A2 landscape. The other four ET_I's were set to the original settings. (B) The six populations landscapes calculated from the FPV landscapes. The inclusion of two populations that crosstalk between FPV landscapes (P1/A1 and P1/A2) drastically reduces the P1 loop distributions and the boundaries between the loop states. Previously these boundaries were considered constructs that formed mixtures of loop populations and have FRET appearing in two FPV landscapes, but due to the settings and inclusion of crosstalk many of those loop mixtures now appear as populations that crosstalk.

which presumably correspond to constructs that form mixtures of high-energy V-shaped loop topologies appearing on multiple FPV landscapes, are now also calculated to represent loops that crosstalk between FPVs shown in Figure 3.3.4B and not only loops that can occupy more than one loop topology.

It is clear from this calculation that one can solve for loop population distributions or crosstalk between the FPV landscapes, but it is unlikely both can be solved simultaneously using our population modeling approach. Moreover, if the headpiece

deforms, the extent of deformation would vary with the DNA torsion and bending enforced in a LacI•DNA loop. Therefore, using one or two uniform deformed population distributions that crosstalk to describe a continuum of deformed crosstalking state is a serious oversimplification. It is likely that a small subset may show crosstalk between FPVs in our landscapes, but the crosstalk must be partiality confined to a specific area on the landscape and requires structural knowledge of LacI•DNA looped complexes beyond the scope of this data and other current data. Varying fluorophore position base pair by base pair could aid in the identification of the local headpiece movements, or the use of immobilized single molecule techniques could distinguish among multiple loop populations and crosstalk between FPV landscapes for a given construct.

3.3.5 Comparison With Previous Work and Theoretical Models

Previous work with hyperstable looping constructs indicated that 9C14 clearly formed a P1 loop topology, but the cyclization data and the single-molecule FRET conflicted on the extent of the population distribution. In these results construct 9C14-D_IA_I transferred energy with intermediate efficiency, at 15% (can be compared with other values in Appendix 1). Based on previous FRET results showing much higher-efficiency transfer in analogous 9C14 fluorescein-TAMRA (R_0= 55 Å) (Edelman, Cheong et al. 2003) and Cy3-Cy5 (R_0= 60 Å) (Morgan, Okamoto et al. 2005) constructs, we had anticipated more efficient energy transfer for the Alexa555 – Alexa647 fluorophore pair used here (R_0 = 51 Å). However, we did not scale the calculated FRET efficiencies to the labeling efficiencies, which could be as low as 50%, partially because we do not know the effects, if any, for the quantum yield for our fluorophore pair, and as described in the

Materials and Methods we lack confidence in the accuracy of these values. Therefore, the FRET results may not be directly comparable to previous FRET work, but our population models should have been able to address the discrepancies in labeling efficiencies.

In this study 9C14 demonstrated substantial energy transfer in two other FPV landscapes. The calculated population distributions for 9C14, 48% for P1, 43% for A1, and 9% for A2 (for the minimal estimate of an extended complex), are strikingly similar to the cyclization studies observation of multiple loop topology populations that generated different ΔLks populations, 50% with +1, 40% with 0, and 10% with -1 (Mehta and Kahn 1999). Only the 3rd estimate for loop population distributions in Figure 3.2.5C does not sensibly match previous cyclization data, but neither population interpretation corroborates previous FRET interpretations. However, it is difficult to know with certainty if 9C14 occupies one or multiple loop populations from our results because it is feasible that LacI deformation in the headpiece could contribute to crosstalk between FPV landscapes (appearing as A1 and A2 loops described earlier). This could show a lower overall FRET efficiency for 9C14, in the context of a landscape of FRET data when compared with individual measurements. However, without an immobilized single-molecule FRET experiment on all the FPVs for 9C14 it is difficult to disambiguate why the FPV landscape observations do not match previous FRET results and instead tend to agree with 9C14 cyclization experiments that observed multiple loop populations for 9C14.

The results in this Chapter agree with the existence of constructs with multiple high-energy loops separating constructs with uniform loop topologies, as predicted by rod

mechanics models for this landscape (Lillian et al. 2008), although we cannot quantify the free energy differences among the constructs. The alignment of the experimental in Figure 3.2.3 and theoretical maps shown in Figure 2.3 is uncertain, but even so, the shapes of the calculated boundaries have no obvious correspondence with our results, and the diagonal symmetry built in to the computational representation was also not observed, especially for P1 loops (the observed experimental asymmetry will be discussed in Chapter 4). These discrepancies may reflect an inadequate underlying model for sequence-dependent DNA shape, the imposition of a central bend position in the modeling, and/or the omission of protein flexibility from these initial stages of modeling. Given the small size of our constructs and the phasing of the operators relative to the A-tracts, we do not believe that a wrapping-toward loop can explain the results (La Penna Perico, 2010), although these structures may be stable for other DNA sequences (Tsodikov et al. 1999). The existence of wrapping-toward versus wrapping-away parallel loops could be evaluated using AFM (Wong et al. 2008).

In spite of the quantitative uncertainties, the qualitative trends among the FPV landscapes strongly support the existence of extended or open form LacI loops. The inclusion of protein flexibility in rod mechanics models will eventually allow us to determine whether existing models for DNA shape and flexibility can explain the landscape data.

Chapter 4

The LacI•DNA Looping Landscape with Saturating IPTG and Subsequent Addition of Excess LacI

Overview

The inducer effects determined from previous work on LacI are discussed to present the context of our study described in this Chapter. Then we discuss the use of the landscape constructs, designed to form hyperstable LacI•DNA loops, to investigate the effects of inducer on looped states. To emphasize the point that inducer bound LacI is still a DNA looping protein, all experiments were performed with saturating synthetic inducer, IPTG. Each FPV had a differential response to inducer, and the results suggest that LacI•DNA loops in the presence of inducer have different properties than in the absence of inducer. Since saturating inducer generally decreases FRET for all FPVs there was an increase in unknown distributions of "dark" states. In order to partition the different possible species underlying the dark components of inducer complexes, excess LacI was added. The differential response to this treatment demonstrated that saturating inducer does not completely release LacI from DNA. Additional LacI following inducer saturation also suggested that in some cases the LacI•DNA complex ± inducer is similar.

This chapter is summarized in the submission of the manuscript to NAR (Haeusler, A., Goodson, K., Lillian, T., Goyal, S., Perkins, N., and Kahn, J. FRET Studies of Landscape of Lac Repressor Mediated DNA Loops. 2011).

4.1 Introduction

4.1.1 Allosteric Binding Reorients the LacI Core Domain and Destabilizes DNA Binding to the Headpiece

A LacI•DNA complex occupies the promoter region 95% of the time and efficiently represses the *lac* operon to the same extent. (Reznikoff, Winter et al. 1974; Oehler, Eismann et al. 1990). In the presence of synthetic inducer shown in Figure 4.1.1, isopropyl-β,D-thiogalactoside (IPTG) a homologue of the natural inducer allolactose, there is almost a 1000-fold decrease in affinity of LacI for the operator (Barkley, Riggs et al. 1975). One IPTG inducer molecule can bind one monomer of LacI in its core domain (Ohshima, Mizokoshi et al. 1974; Schmitz, Schmeissner et al. 1976; Friedman, Fischmann et al. 1995; Lewis, Chang et al. 1996; Daber, Stayrook et al. 2007) with a K_D = 5.0 x 10^{-6} M (O'Gorman, Dunaway et al. 1980; Donner, Caruthers et al. 1982).

Early *in vivo* studies of LacI's interaction with the lac operators as a function of IPTG concentration demonstrate a sigmoidal curve. This led to a belief in either a cooperative or a two-step mechanism requiring two molecules of inducer to dissociate LacI from the operator (Boezi and Cowie 1961; Clark and Marr 1964; Yagil and Yagil 1971) but did not consider LacI binding two operators (Kania and Müller-Hill 1977; O'Gorman, Dunaway et al. 1980) or the presence of four inducer sites as opposed to two sites (Ohshima, Mizokoshi et al. 1974; Schmitz, Schmeissner et al. 1976). Oehler, building from the mechanism proposed by Yagil, suggested that the sigmoidal curve seen *in vivo* is a result of binding stabilization of an auxiliary operator through DNA loop

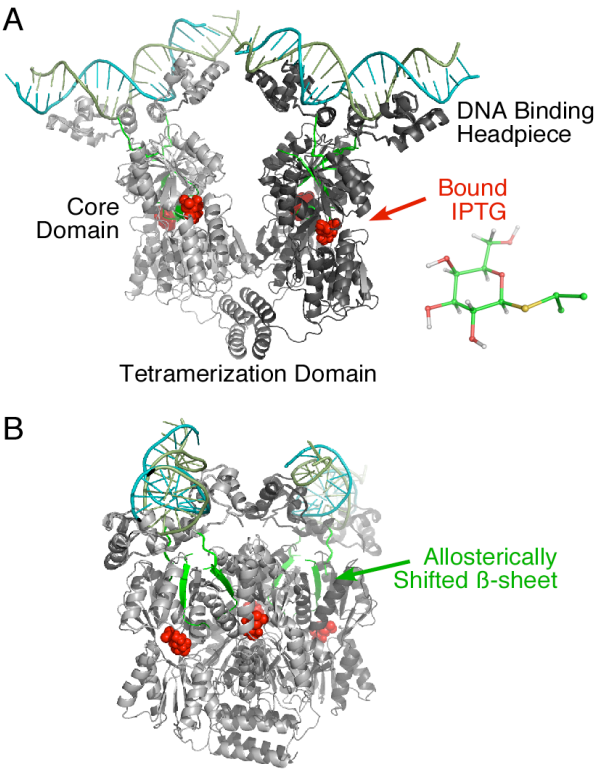


Figure 4.1.1: *IPTG destabilizes LacI•DNA interactions by reorienting the headpiece and core domain.* The DNA binding headpiece bound to an operator fragment was pair fitted (Lewis, Chang et al. 1996) onto the IPTG bound LacI core domains (Friedman, Fischmann et al. 1995) for visual continuity. (A) One IPTG molecule binds one monomer in the sugar pocket between the N-terminal and C-terminal core domains (IPTG is shown in red in when bound in the LacI core). (B) Binding of IPTG shifts an internal ß-sheet (green) down one hydrogen bond network towards sugar pocket. The allosteric transition places a tension on the DNA binding headpiece that is connected to the core via the ß-sheet. IPTG also decreases flexibility between the N-terminal and C-terminal core domains.

formation, assuming only one IPTG molecule is necessary for derepression (Oehler, Alberti et al. 2006).

The recent model described by Lewis, using a heterodimeric LacI coupled with IPTG binding pocket disruption studies, demonstrated that a single IPTG molecule is "not enough" to dissociate a dimer from DNA (Daber, Sharp et al. 2009). It is the reorientation of the core domain during inducer binding that increases the distances between headpieces in a dimer, thereby destabilizing the strong interactions between the headpiece and operator (Lewis 2005). The conformational change, upon IPTG binding in the core domain, affects the headpiece by inducing an allosteric transition of a central ß-sheet that links the two domains together shown in Figure 4.1.1. When LacI is induced this internal ß-sheet slides towards the binding pocket by one hydrogen bond network to

bind the inducer. This places a tension on the DNA binding headpiece, while decreasing protein flexibility in the sugar-binding pocket (Daber, Stayrook et al. 2007). An alteration to this specific ß-sheet, which also forms part of the monomer-monomer interface between dimers, diminishes the allosteric communication between the core domain and DNA binding headpiece domain (Markiewicz, Kleina et al. 1994; Zhan, Camargo et al. 2010).

4.1.2 The Nonspecifically Bound LacI•DNA Complex

IPTG reduces the affinity of LacI for DNA to nonspecific levels (Barkley, Riggs et al. 1975; Kalodimos, Biris et al. 2004), but the mechanism by which IPTG allosterically alters the interactions between LacI and DNA is not clearly understood. The binding of inducer has been proposed to alter the LacI conformation in the flexible hinge-helix located between the DNA binding headpiece and the N-terminal core domain shown in Figure 4.1.2 (Spronk, Slijper et al. 1996; Falcon and Matthews 2001). The hinge helix is connected to the core via the ß-sheet emphasized in Figure 4.1.1, and is disordered until specifically bound to DNA as shown in figure 4.1.2. NMR studies on the DNA binding headpiece substantiate that the plasticity of the hinge helix region can sample the backbone in a nonspecific DNA sequence, probe the DNA for specific binding sites, and then lock onto specific recognition sites by reorienting into the α-helical structure (Kalodimos, Biris et al. 2004). Therefore, a tensor on the headpiece that withdraws the hinge helix from DNA towards the core when induced, may limit the freedom of the hinge helix to sample the DNA surface or limit the specific insertion into minor groove of

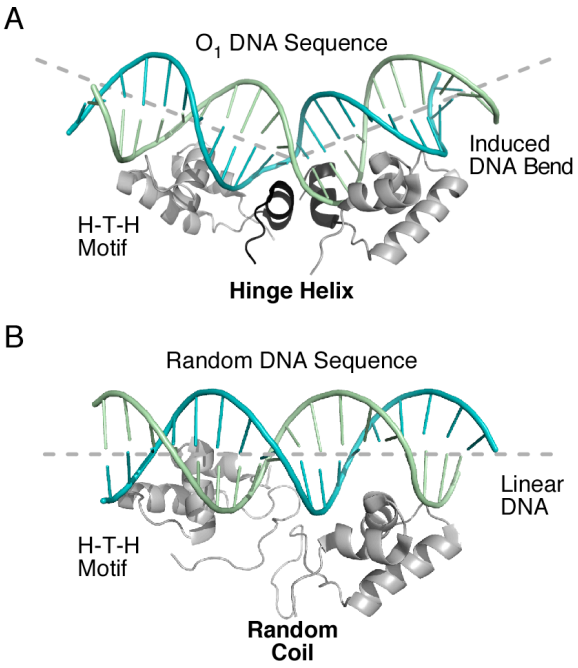


Figure 4.1.2: *The specific and nonspecific interactions of the DNA binding headpiece suggest a mechanism for inducer effects on LacI•DNA interactions* (Kalodimos, Biris et al. 2004). (A) When DNA is specifically bound (O1 in this example) each monomer forms the classical helix-turn-helix (H-T-H) motif and the hinge helix (black) takes on an ordered structure that unwinds the minor groove , as shown in NMR studies. This induces a bend in the DNA around the central symmetry axis of the DNA. (B) When LacI is nonspecifically interacting with DNA the H-T-H motifs still binds the major groove, but the hinge helix is now disordered in the DNA binding headpiece. The plasticity of the hinge helix may allow LacI to sample the DNA shape until recognizing a specific DNA sequence and locking into the site while forming an ordered α-helical structure. The hinge helix is also associated with the allosteric effects of the inducer, presumably by placing a tensor on the hinge helix that withdraws minor groove intercalators and thereby reduces the affinity of the DNA binding headpiece to nonspecific levels.

DNA. Either case would diminish the affinity of LacI for DNA. These observations suggest that the negative cooperativity of IPTG and DNA allows LacI to specifically interact with one or the other, but not both simultaneously. More specifically, the energy for a given LacI•DNA interaction must be higher than the inducer bound LacI for any allosteric transition to occur, and vice versa. Therefore, does IPTG bind when LacI is nonspecifically bound to DNA and prevent the transition to a specific LacI•DNA complex, or can IPTG bind a specific LacI•DNA complex and induce a transition to a nonspecific LacI•DNA interaction, or both?

4.1.3 Unknown Effects of Inducer on Lac Repressor-DNA Looped Complexes

Most studies of the inducer effects are performed on dimers of LacI that are unable to tetramerize. As described above, this does not take into consideration the cooperativity of binding multiple operator sites and LacI•DNA looping. *In vivo* evidence for LacI•DNA looping shows a classical periodicity in repression efficiency, and this periodicity is observed even in the presence of saturating inducer (Becker, Kahn et al. 2005; Bond, Peters et al. 2010). Additionally, the LacI•DNA loops are surprisingly more stable *in vivo* than can be predicted by inducer studies on LacI dimers that cannot form tetramers (Daber, Sharp et al. 2009), which could be partly due to other environmental proteins that reduce the energetic cost to form loops (Becker, Kahn et al. 2005; Becker, Kahn et al. 2007). The slight offset in the periodicity of repression ± IPTG suggests there is conformational change induced by IPTG that either requires an altered writhe and/or DNA unwinding (Becker, Kahn et al. 2005). Furthermore, IPTG saturated LacI•DNA looped complexes have a higher sensitivity to the torsional phasing of the operators. All of this biologically relevant insight into the effects of induction on repression cannot be provided by LacI dimer studies.

Single-molecule *in vivo* studies of *lac* permease expression demonstrated that short stochastic bursts of gene expression occurred as a result of partial dissociation of LacI•DNA looped complex (Choi, Cai et al. 2008). In the presence of inducer the bursts were larger, and they described this event as the complete release of LacI that when bound with inducer has a slower DNA rebinding rate. However, in the absence of direct measurements of the effects of inducer on tetrameric LacI•DNA looped complexes, we

decided to test the response of our construct landscape to saturating inducer. Since each construct has at least one FPV that showed substantial FRET, we could evaluate all operator-phasing variants relative to central bend to determine the stability of specific torsional phasings in a loop, the possible effects on LacI•DNA loop geometries, and the preferred loop topology in the presence of saturating inducer. Furthermore, to distinguish no FRET LacI•DNA complexes from free DNA, we added excess LacI following IPTG saturation to partition among potential species.

4.2 Materials and Methods

4.2.1 EMSA Studies Comparing Looping in the Absence and Presence of Inducer

Improving upon previous EMSA studies described in Chapter 3, one primer was radiolabeled for PCR amplifications to provide a more accurate quantification for radiolabeled samples. All primers were purchased from IDT and purified if required using a denaturing PAGE 12% (40:1) 8M urea, at 55 W for 2 hr. The gel was stained with ethidium bromide (EthBr) and the appropriate band extracted and gel purified. Then 50 μM of the D_E. primer was radio-labeled in 1X PNK Buffer with PNK (NEB) and ~1 uM of 500 μCi γ-^{32}P-ATP (Perkin-Elmer) in 10 μL, for one hour at 37^{o}C. The reaction was purified through a P-6 column (Bio-Rad) and then heat denatured at 80^{o}C for 20 minutes. The A_{260} was used to determine the final DNA concentration and was adjusted to 50 μM. PCR amplification was performed as previously described for the protocols involving D_E primer and A_E bot primers and PCR product purified following typical gel purification methods described in the Materials and Methods in Chapter 3. The concentration of the

radiolabeled product was determined by comparing the stock radiolabeled primer cpm to the cpm of the purified PCR product.

Labeled DNA was held at 1 nM concentration with titrating LacI, the LacI concentrations varied depending the experimental settings (typical conditions are shown in the Figure 4.3.1 in the Results). The DNA and LacI were incubated in absence or presence of 5 mM IPTG. Competitor DNA was then added to some samples, which also varied in concentration between experimental settings, and was incubated for 30 minutes before loading onto a gel. All gel conditions were as previously described for EMSAs, but 5 mM IPTG (Fisher) was added to the gel before polymerization.

A more reliable method that allowed direct quantification of the radiolabeled DNA came from body labeling the the PCR product. This was performed in 50 μL and included: 1 μM of each primer shown below, 0.04 ng/μL DNA template, 250 μM of each dNTP (USB), 50 μCi of ^{32}P labeled α-ATP (Perkin Elmer), and 2 units of Phusion polymerase (NEB) in HF buffer to body label the PCR product. This allowed the most effective method for quantifying the DNA concentrations based on: (construct A-T base pairs – primer AT-base pairs), and the relative concentration of P^{32} labeled α-ATP to unlabeled ATP. EMSAs ±IPTG were performed as described above.

4.2.2 Bulk FRET Studies on the FPV Landscapes with Saturating IPTG and Excess LacI

Bulk FRET showed negligible changes in acceptor and donor fluorescent intensity when the volume was changed by ~ 2%. Therefore, LacI•DNA looped complexes from Chapter 3 (with 50 μL total volume) had 1 μL of stock IPTG added to provide a final

concentration of 5 mM IPTG. Samples were scanned as before. Similarly, for a single trial another 1 μL addition of stock LacI was added following IPTG saturation and spectra obtained.

The small changes in volume could be included in our analysis of FRET efficiency, but the changes, if any, are within the deviation of each fluorescent intensity measurement. Therefore, the FRET efficiencies were calculated as described in the Materials and Methods in Chapter 3. The only exceptions were that the reference spectra used to decompose the experimental data for IPTG-LacI•DNA complexes and IPTG-LacI•DNA complexes in the presence of excess LacI inducer were decomposed from donor only and acceptor only labeled 9C14 for IPTG-LacI•DNA complexes.

All population calculations were performed as described in the Materials and Methods in Chapter 3 using the minimal estimate for extended populations.

4.3 Results

4.3.1 EMSA Demonstrates Lac Repressor-DNA Looping in the Presence of Inducer

EMSAs were performed in the absence and presence of saturating inducer shown in Figure 4.3.1. The EMSA results are highly sensitive to the pretreatment of the complexes before loading onto the gel. For example, in Figure 4.3.1A LacI•DNA complexes were vortexed more liberally, where as in Figure 4.3.1B the complexes were gently flicked. The variation between the gels is quite dramatic, and furthermore the presence of glucose (5 mM) or glycerol (0.1-1%) in the gels broke up the multiplexes in Figure 4.3.1B and gave similar results as in Figure 4.3.1A without vortexing. However, in

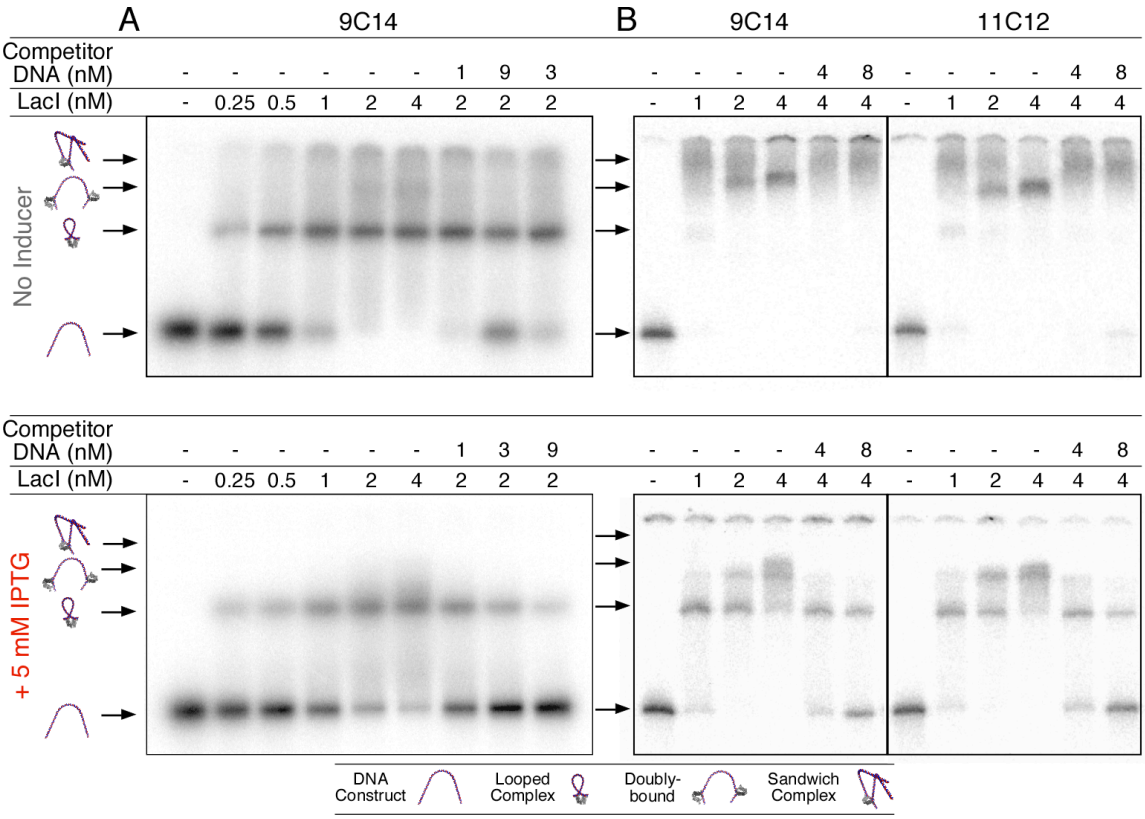


Figure 4.3.1: *EMSA on LacI•DNA complexes ± IPTG show clear differences between the induced and uninduced loop states*. The top panel indicates the absence of inducer and the bottom panel is in the presence of inducer. The gels (PAGE, 75:1 acrylaimde:bis-acrylamide) were prepared in parallel, except that 5 mM IPTG was added to one gel before polymerization. The ^{32}Plabeled DNA was at 1 nM concentrations and was premixed with 5 mM IPTG before the addition of titrating LacI as shown above each lane (order of addition experiments demonstrated that equilibria are very fast). Unlabeled DNA competitor, the same DNA sequence as the labeled construct, was added after 30 minutes of incubation for LacI•DNA complexes, and was allowed to equilibrate another 30 minutes. (A) Body labeled 9C14 comparisons ±IPTG demonstrated that LacI can still bind DNA in the presence of IPTG, but with less affinity than in its absence. The addition of competitor shows a greater increase in free DNA +IPTG. (B) 9C14 and 11C12 singly end labeled DNA indicated +IPTG reduces the formation of sandwich complexes compared to its absence. From both 9C14 and 11C12 it is apparent that ±IPTG there is still a formation of doubly bound LacI•DNA complexes (LacI2-DNA). In the presence of DNA competitor the sandwich complex does not form as readily and increases the free DNA +IPTG relative to –IPTG, but most likely occurs through a sandwich complex intermediate. The results in (B) were more easily reproduced; the observed results in (A) occurred from a more liberal vortexing and most likely physically dislodged nonspecific binding interactions that lead to LacI•DNA multiplexes. Apparently, in (B) this is not an issue when inducer is present or when glycerol/glucose is used in the gel (shown in Appendix 1).

the presence of glycerol the inducer had little to no effects on LacI•DNA complexes, presumably because glycerol and sugar may occupy the LacI sugar-binding pocket and thus require a higher [IPTG] to displace the "anti-inducers". Despite these discrepancies there are clear patterns for the general differences ±IPTG in all cases. In Figure 4.3.1A the addition of 5 mM IPTG to the gel demonstrates that the formation of a LacI•DNA looped complex is favored in the presence of saturating IPTG. Titrating in more LacI in the presence of IPTG stabilizes looping and requires a higher LacI:DNA stoichiometry to observe complete shifting of free DNA. In general, the presence of inducer decreases the appearance of doubly bound DNA (smearing from the loop upward), and the addition of competitor never results in DNA sandwich complexes, but both species appear in the absence of inducer. These multiplexes may be more stable in solution with inducer, but breakdown more readily due to the stresses of gel migration and the weakened stability of the induced LacI. The IPTG saturated LacI•DNA loops do have diminished stability relative to the non-induced counterparts as demonstrated by the total shift for free DNA in the LacI titrations, by the smearing seen from the loop states down to free DNA, and by the increased susceptibility of IPTG-LacI•DNA complexes to DNA competition.

In Figure 4.3.1B the same general trends are observed as in Figure 4.3.1A, but LacI•DNA multiplexes are present ± IPTG. The difference in handling for this set of EMSA may not have physically dislodged (flicking versus vortexing) nonspecific interactions or possibly broken up LacI aggregation. These could lead to LacI•DNA loops being held together by additional nonspecific LacI DNA interactions as shown in the absence of IPTG. However, in the presence of IPTG this effect goes away, and it is

unclear exactly why, but could be caused by IPTG stabilizing the LacI protein and preventing aggregation (sugars, as stated above stabilize the LacI protein) or diminished affinity of LacI for DNA complexes. Nonetheless, in Figure 4.3.1B there are two surprising observations for both constructs: the appearance of loops in the presence of IPTG, and the strong formation of doubly bound LacI•DNA looped complexes ± IPTG. Presumably these loops have an additional LacI bound nonspecifically to the loop, since with excess LacI a faint band begins to appear above a doubly bound band, suggesting a population of triply bound DNA in the presence of IPTG. The change in smearing patterns still demonstrates the instability of the IPTG complexes relative to the uninduced counterparts in the gel. Furthermore, the addition of competitor in the presence of IPTG has increased free DNA, which also supports IPTG reduces the stability of LacI•DNA complexes.

4.3.2 The FPV Landscapes Have Sequence Dependent Responses to Inducer

We expect that even for the hyperstable loops used here, the loop free energies vary widely across the landscape, but we cannot detect these variations at LacI concentrations (~nM) well above K_D. However, since all of our constructs exhibit FRET in at least one FPV landscape, at least we have experimental access to loops that presumably span this wide range of free energies. Previous work has shown that IPTG•LacI•DNA loops are detectable with FRET and sufficiently stable to repress transcription, but the induced loops must be much higher in free energy than uninduced loops (Edelman, Cheong et al. 2003). We expect the destabilization due to IPTG binding

should be more marked at the boundaries between topologies, which are not optimal for either topology and thus are higher in energy.

LacI•DNA loops were prepared and FRET efficiency was determined from the samples in Chapter 3, then saturating IPTG (5 mM) was added to the mixture and was allowed to equilibrate for one minute or more before collecting emission spectra. (The spectra did not change with time, and on the benchmark 9C14 construct pre-incubation of DNA with IPTG gave identical results, suggesting that equilibrium was reached.) The general shapes of the resulting FPV landscapes (FRET in Figure 4.3.2A, with calculated population distributions in Figure 4.3.2B) are similar with and without IPTG except for the A1 landscape. This confirms that LacI can still effectively loop DNA in the presence of IPTG as seen with the EMSA studies, *in vivo* studies, and previous bulk FRET studies (Edelman, Cheong et al. 2003; Becker, Kahn et al. 2005; Bond, Peters et al. 2010). However, the A1 landscape appears particularly sensitive to IPTG, again indicating that the A1 and A2 loop topologies have different stabilities and possibly preferred LacI•DNA loop geometries.

IPTG decreased but usually did not eliminate FRET for all constructs, suggesting that IPTG either partially breaks the loops or induces an extended (E) conformation of LacI in the loop. Due to differential loop stability, the decreases in FRET are not uniform across the landscapes. IPTG never markedly increases FRET, showing that induction does not cause substantial redistribution among the P1, A1, and A2 loops, which all presumably have V-shaped LacI. The overall loss of FRET (Figure 4.3.2C) upon IPTG saturation does increase the calculated population of dark states, which could be free

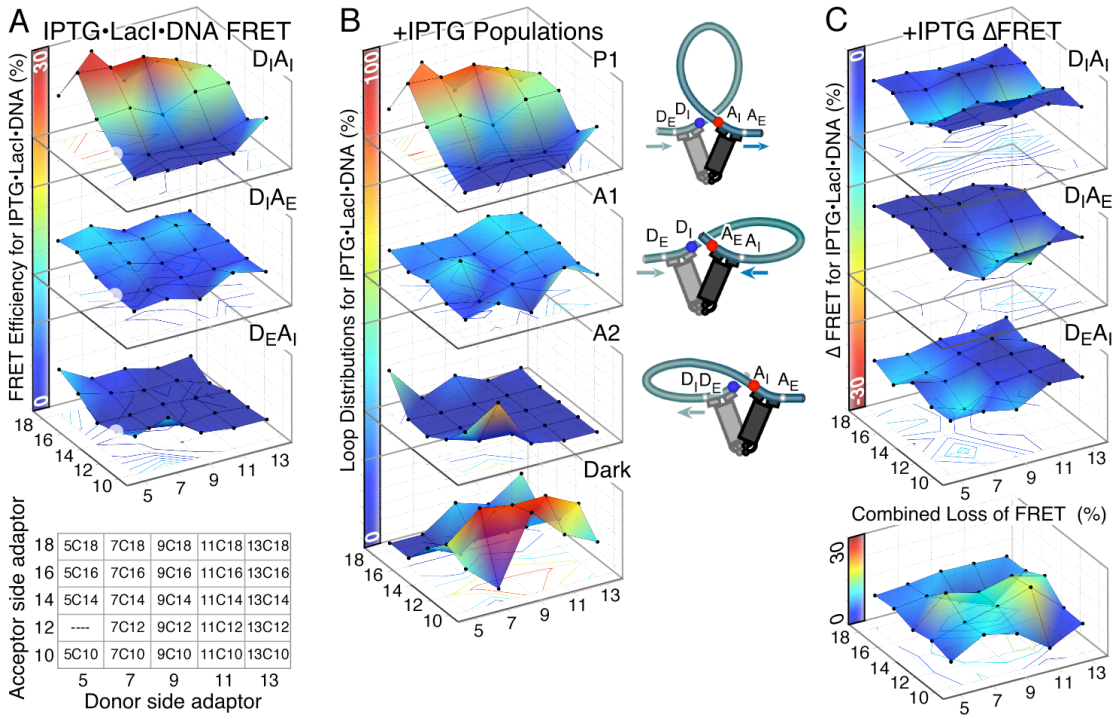


Figure 4.3.2: *High energy LacI•DNA looped complexes are more sensitive to IPTG.* (A) Pre-formed LacI•DNA looped complexes were equilibrated with saturating (5 mM) IPTG, and FRET efficiencies were measured for all constructs as in Figure 3.2.2. IPTG flattened each FPV surface to a different extent, and it never significantly increased FRET for any construct. (B) Loop population distributions calculated for the IPTG•LacI•DNA looped complexes using the minimum estimate parameters for E loop populations. IPTG increased the population of dark states, the nature of which is described in the text. (C) The FRET efficiency decrease upon IPTG saturation is most marked for the A1 loop topology or else at the boundaries between the most uniform loops on each FPV landscape. These boundary constructs should correspond to the least stable loops, because they do not have optimal shapes for any topology. The loss of FRET summed over all three FPVs in the bottom landscape emphasizes the loops that are most sensitive to IPTG.

DNA, singly bound, extended loops, or other low FRET loop states described in the discussion.

The landscape regions that are most dramatically affected by IPTG are identified by the differential FRET landscape of Figure 4.3.2C. These regions represent the high

energy loop boundaries that separate uniform loop topologies, but it surprising that the A1 loops are particularly sensitive to IPTG for our DNA sequence space.

4.3.3 Excess LacI Illuminates Differential Stability of IPTG-LacI•DNA Complexes

To understand the effects of IPTG on the landscape, especially the loss of FRET, it is necessary to establish the nature of the induced complexes. Data obtained at a single LacI concentration cannot identify whether decreased FRET reflects the appearance of free DNA, unlooped single-bound DNA, alternative loop shapes, or doubly-bound DNA. The addition of excess LacI can distinguish among these possibilities. If there is free DNA present, excess LacI should repopulate the loop state and increase FRET (and since the complexes are initially formed at nearly stoichiometric LacI:DNA ratios, free DNA must accompany the production of any doubly-bound DNA upon IPTG addition). On the other hand, excess LacI should partition unlooped single-bound complexes into double-bound unlooped complexes, as suggested by EMSAs, reducing the observed FRET further. Excess LacI can convert any loop to a double-bound complex, and the extent of conversion for a given [LacI] depends on the loop free energy. If induction yields alternative loop conformations, they should be differentially susceptible to excess LacI. Thus, if LacI addition after IPTG addition decreases FRET further, IPTG must have destabilized the loop enough to make it unstable relative to double-bound, whereas if IPTG addition causes a decrease in FRET but subsequent LacI addition causes no further change, we conclude that the IPTG-bound LacI still forms stable loops but with altered conformation(s).

FRET was measured upon increasing LacI from 3 nM to 6 nM after IPTG saturation for all FPVs that had initially shown FRET in the presence of LacI without IPTG. The results in Figure 4.3.3 show that IPTG with excess LacI decreases FRET with a different pattern than IPTG, suggesting that addition of LacI to an IPTG-bound complex causes new changes in loop stability and conformation. For the P1 loops, excess

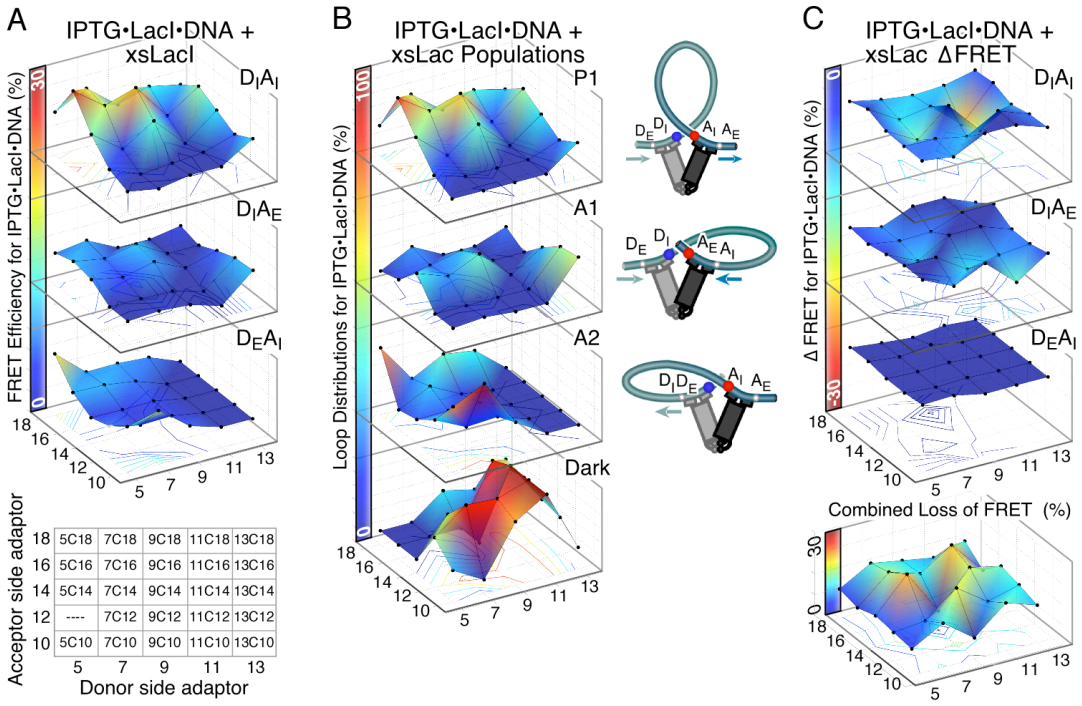


Figure 4.3.3:*The addition of excess LacI to IPTG saturated LacI•DNA complexes partitions amoung unstable loops*. (A) Following incubation of LacI•DNA complexes with IPTG excess LacI was added and fluorescence spectra measured (3 nM). Constructs that showed no FRET before the addition of IPTG were not measured, but zeros were placed in their location for visual continuity across the landscape (actual FRET efficiency measurements are shown in Appendix 1). Excess LacI decreased FRET further for A1 loops and some P1 constructs centered on 7C14 and 11C14, but A2 loops were unaffected. FRET did not increase for any of the constructs, confirming that the initial loss of FRET upon IPTG saturation was not due to the release of free DNA. (B) Population distributions for IPTG-LacI•DNA in the presence of excess LacI indicate a further increase in dark states described in the text. Distributions were calculated using the parameters that provide the minimum estimate for E loop populations. (C) Difference landscapes were measured for the addition of excess (3 nM) LacI to those IPTG•LacI•DNA looped complexes. The loss of FRET summed over all three FPVs in the bottom landscape showed the constructs that are most sensitive to excess LacI.

LacI decreases but still does not eliminate FRET, and the constructs that were not previously affected by IPTG are the most obviously affected, suggesting that they are readily converted to double-bound forms or that the loop conformation has changed. The decrease is most marked for the constructs centered around 11C14 in the D_IA_{II} (P1) landscape, and in this region of the D_IA_E FPV (A1) landscape we also observe slightly increased FRET, suggesting that some P1 loops are converted to A1 or else that an altered P1 conformation exhibits increased crosstalk into the D_IA_E channel. For the constructs that were previously A1 loops, excess LacI further decreased the remaining FRET for IPTG•LacI•DNA. Finally, the D_EA_I landscape (A2 loops) seems unaffected by excess LacI. All of the excess LacI results are consistent with the initial loss of FRET with IPTG being due to alternative induced loop conformations, or single-bound complexes but not free DNA.

4.4 Discussion

4.4.1 The Different Properties of IPTG bound LacI•DNA complexes

All of the loops studied here appear hyperstable in the absence of inducer, but addition of saturating IPTG reveals the free energy variations among them as well as changes in the relative stabilities of different loops formed for each construct. Some of the constructs change very little upon induction, and we suggest that the original specific-specific V-shape (LacI is bound specifically to both DNA operators) loop is stable for these constructs. Usually, however, FRET is decreased (Figure 4.3.2) most markedly for the boundary constructs and also for the A1 loops (D_IA_E). This was particularly surprising

since 11C12, primarily an A1 loop, still formed a looped complex in IPTG-EMSA studies in Figure 4.2.1B, but FRET diminished completely upon induction. Therefore, in Figure 4.4.1 we illustrate some of the possible lower-FRET states: single-bound unlooped LacI•DNA, or expanded or E loops, or loops formed between specific operators and non-specific DNA (s-ns loops). In the presence of inducer, we propose the s-ns loops are stabilized because conversion of an interface from specific to nonspecific is accompanied by DNA reorganization that minimizes the free energy of DNA deformation. The IPTG effects observed by FRET depend on the free energy of each specific-specific V-shape loop relative to all other possibilities in Figure 4.4.1. Bulk FRET cannot distinguish between alternative induced no FRET complexes for this fluorophore pair, and evidently EMSA studies cannot distinguish between them either. Because the EMSAs indicated that the looped complexes are still formed in the presence of inducer for construct 9C14 and 11C12, but we see decreased FRET for these constructs, especially 11C12 in solution. It is possible there could be a transiently bound IPTG-LacI•DNA looped state or a conformational change to an extended

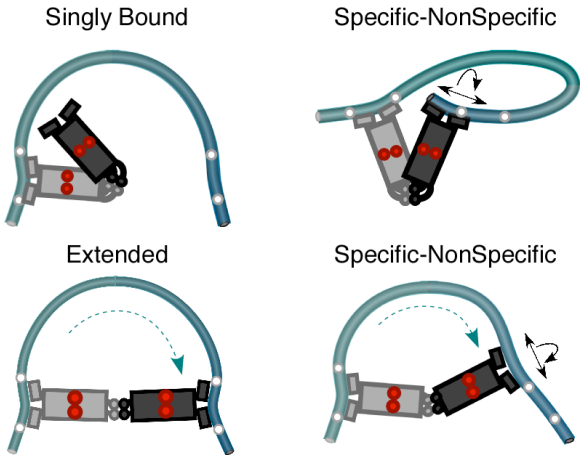


Figure 4.4.1: *Alternative IPTG-LacI•DNA complexes that may contribute to the observed increase in the dark state populations*. IPTG binding (shown in red) decreases specificity for the lac operator, thereby stabilizing proposed specific-nonspecific loops relative to the specific-specific loops shown previously, and IPTG could also alter the relative stability of open forms of LacI relative to the V-shape. Any looped complex will generally resist disruption by excess LacI or DNA. In contrast, LacI singly bound states will convert to doubly bound or sandwich complexes under excess conditions.

LacI•DNA looped state, both would contribute to a dark population, yet be difficult to compete.

4.4.2 Excess LacI Partitions Among the IPTG-LacI•DNA Complexes

To better understand the dark population, we added additional LacI to all the IPTG-saturated complexes that still showed substantial FRET. There was never an increase in FRET, showing that the initial FRET decrease had not been caused by the complete release of DNA (any free DNA would have been re-bound to give a state with at least some FRET). On the other hand, in many cases additional LacI also did not decrease FRET further, suggesting that the dark states are still stable loops (singly bound DNA and therefore any unstable loops would be converted to double-bound $LacI_2$-DNA that would not show FRET). In EMSAs we saw the presence of doubly bound DNA, but we could not distinguish if two LacI's are both specifically bound to an operator, or if a loop is formed with a nonspecifically bound LacI in these studies. Additionally, these results suggest that the free DNA observed in EMSAs must have transitioned via a sandwich complex intermediate, and not the complete release of repressor. Our results confirm and extend previous demonstrations (Edelman, Cheong et al. 2003; Becker, Kahn et al. 2005; Bond, Peters et al. 2010) that IPTG-LacI•DNA loops can be stable but they do not share the same properties as the uninduced loops.

4.4.3 Asymmetries Revealed in the LacI•DNA Looping Landscape with Inducer

Differences in the responses to IPTG and IPTG+LacI reflect the free energy landscape of the induced loops and the relationship of this landscape to the uninduced

landscape. Some of the constructs behave markedly differently in the presence of IPTG even though they are very similar in its absence. For example, the A1 and A2 antiparallel loops have similar FRET maxima, and they emerged as equivalent solutions from rod mechanics that was based on a symmetrical linker-bend-linker representation. However, the disappearance of FRET upon IPTG addition for A1 loops but not A2 loops in Figure 4.3.2 suggests that the asymmetric positioning of the A tracts relative to the operator has a much stronger influence than expected. Moreover, the offset between operators in the V-shaped LacI (the C2 symmetry seen in the co-crystal structure) apparently favors antiparallel V-shaped loops with different linker lengths.

There are also several P1 loops that are relatively insensitive to saturating IPTG but highly sensitive to additional LacI, seen as two valleys (centered on 11C16 and 7C16) on either side of the more stable constructs 9C14 and 9C16 in the D_IA_I landscape in Figure 4.3.3. IPTG destabilized the original loop but did not allow conversion to other loop forms. The two valleys may represent populations with similar FRET but different global structure, for example, reflecting either undertwisting or overtwisting in the intervening DNA loop. Torsional strain might be less easily tolerated by the less stable IPTG-bound loop as seen *in vivo* (Becker, Kahn et al. 2005), especially if s-ns or extended structures of lower free energy are accessible. The 11C16 valley appears to be more sensitive to added LacI than the 7C16 valley; we do not have a clear explanation for the difference. However, that whole side of the landscape had lower FRET throughout the study, and the families of constructs that share the same donor side adaptor (11C0-18 and even 13C10-18 to a lesser extent) may have extended looped complexes that are

energetically comparable to the V-shaped looped complexes. The families that share the same donor side adaptor on the other side of the landscape (5C10-18 and 7C10-18) tend to have high FRET, with some having higher than anticipated FRET. In contrast, these families may have compressed or deformed LacI•DNA loops that crosstalk between FPV landscapes and are energetically comparable to V-shaped loops.

4.4.4 Comparing the Free Energy LacI•DNA Looping Landscape ± Inducer

In the free energy landscapes of Figure 4.4.4 we summarize the uninduced LacI•DNA loop populations and speculate on the energetics underlying the differential responses of each landscape to treatment with inducer and additional LacI. Our observations all depend on the relative free energies of the most stable FRET-active states versus the most stable dark states. We consider singly bound, doubly bound, sandwich (DNA•LacI•DNA), and looped complexes, with the free energies of the first three being independent of DNA sequence. The free energy change for different species depends on the number of specific LacI•DNA interfaces present (Daber, Sharp et al. 2009; Daber, Stayrook et al. 2007).

The specific V-shaped loops, as well as doubly-bound and sandwich, all increase in free energy relative to singly bound because two dimer interfaces are disrupted. Subtle changes in loop shape due to conversion from the "R" form of LacI to the IPTG-bound "R*" are not considered (Daber, Sharp et al. 2009). The landscapes for the specific V-shaped loops are offset by a constant free energy because their relative free energies depend on the cost of DNA deformation in the loop, not the LacI•DNA interface.

The crux of the model is that the dark looped state landscape is sequence dependent, and that because the dark states are a mixture of E and s-ns loops, the sequence dependence is not the same for induced versus uninduced conditions. The V-shaped, E, and s-ns loops differ in the number of specific interfaces and in loop shape, so the tradeoffs among LacI deformation, DNA deformation, and specific versus nonspecific interfaces change upon induction. For example, the operator phasings that prefer P1 loops (operators oriented outward relative to the intrinsic DNA bend) would convert to

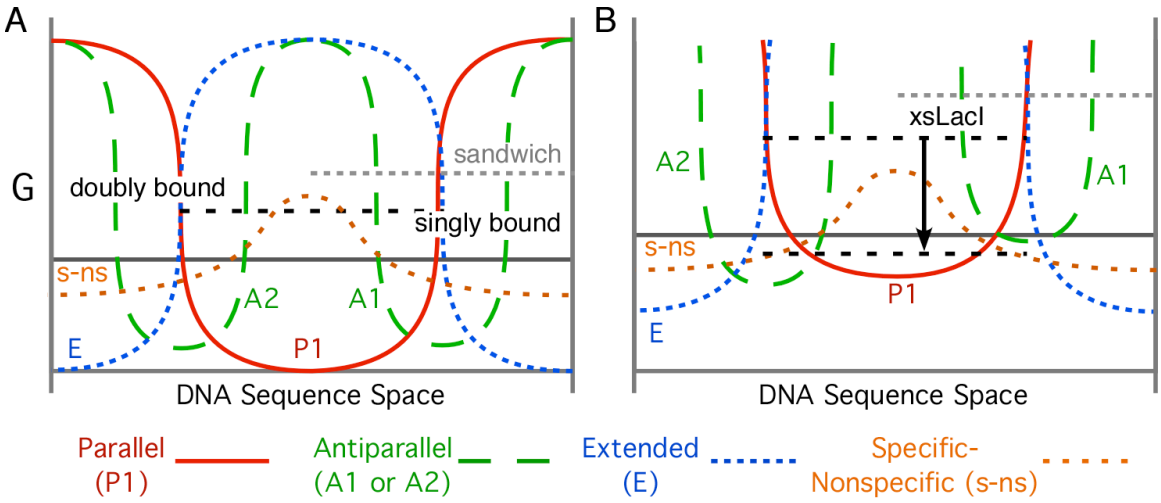


Figure 4.4.4: *The variation of the free energy of induced and uninduced complexes over the DNA sequence space*. (A) The relative energies extracted from the experiments are illustrated for the V-shaped loops (P1, A1, A2), extended (E) loops, specific-nonspecific loops (s-ns), and three species that should not compete at stoichiometric conditions: singly bound, doubly bound, and sandwich complexes. Only V-shaped loops and possibly some s-ns loops can be detected by FRET, and the other illustrated features, referred to as dark states, are described in the Discussions in Chapters 3 and 4. (B) The presence of saturating IPTG raises the free energy proportionally to the loss of specific LacI•DNA interfaces relative to the uninduced state. The compression of free energy differences increases competition between previously inaccessible states that are now energetically comparable in the presence of inducer. The high-energy V-shaped A1 loops could be driven to form either s-ns or E loops in the presence of IPTG and not singly bound based on EMSAs and DNA competitions. The addition of excess (xs) LacI lowered the energy for doubly bound complexes and partitioned singly bound states and presumably high-energy s-ns states into a doubly bound state that further reduced the observable FRET.

extended looped complexes or s-ns loops with the largest DNA twist energies in addition to the largest DNA bend energies among the loops. Therefore, the s-ns loops are inaccessible, consistent with the observation that IPTG does not decrease FRET for P1 constructs, but can be partitioned by excess LacI to IPTG•LacI•DNA complexes. In contrast, the s-ns loops will be stabilized if at least one operator is directed inward relative to the bend, because DNA twist and DNA bending strain should both be reduced relative to P1 loops. This operator phasing would favor the formation of either an A1, or A2 loop, and when both operators are phased towards the central bend favor the formation of an Extended loop. .In general we suggest that the energy minima in the dark state surface correspond to unstable V-shaped LacI loops. Therefore, from our results the boundaries between parallel and antiparallel loops are observed to be mixtures of V-shaped loop topologies. The boundaries between antiparallel loops and extended loops appear to be mixtures of FRET-active and dark geometries (refer to Figures 3.2.5 and 4.3.2).

The different responses of inducer-saturated complexes to excess LacI support the sequence-dependence of the dark state loop free energies. Excess LacI decreases the free energy of the doubly bound state, and we assume that all doubly bound constructs have the same free energy. The key observation is that the FRET in the D_EA_I landscape (A2 loops) decreases upon induction, consistent with conversion to a dark state, but the remaining FRET is insensitive to excess LacI, showing that the dark loop state has a lower free energy than the doubly-bound state. In contrast, the FRET for the D_IA_I landscape (P1 loops) is resistant to inducer because the dark states are high in energy for

these phasing variants, but among these P1 loops the less stable ones are sensitive to excess LacI, suggesting that the doubly-bound state is now lower in free energy than these P1 loops.

Chapter 5

Kinetic Studies on the LacI•DNA FPV Landscapes

Overview

The goals of this chapter were to seek closure on observations from a few individual experiments, such as DNA competition studies and other preliminary experimental work not reported here, as well as perform preliminary experiments to guide single-molecule immobilization studies. Some immediate experimental goals were to measure relative loop lifetimes, the stability of the loops in competition experiments, the interconversion between loop topologies, and distinguishing overtwisted and undertwisted DNA in LacI•DNA looped complexes for all FPV landscapes. However, a thorough investigation such as this on a landscape requires numerous controls that must be monitored simultaneously in parallel with the variables. Therefore, bulk FRET for kinetics on all FPV landscapes was performed using a microplate on a typhoon imager for high-throughput analysis. The results verified previous equilibrium bulk FRET measurements reported in Chapters 3 and 4, but also indicated that loop interconversion could not be detected on the time scales measured. The response of a LacI•DNA looped complex to excess DNA competitor indicated that all the constructs form hyperstable LacI•DNA loops, which frequently have lifetimes of days.

5.1 Introduction

5.1.1 Kinetic Loops Could Compete with Thermodynamic Loops

LacI can bind a single symmetrical operator DNA in two ways: forward and reverse as shown in Figure 5.1.1. If we assume the reorientation rate of LacI from forward to reverse states is slower than the rate of loop formation (LacI is a V-shape and is more rigid than DNA), then the initial binding of LacI could promote the formation of a kinetically stable complex, and not the thermodynamically optimum state represented in Figure 5.1.1. The interconversion between the kinetic loop and thermodynamic loop would require LacI to lose a specific LacI•DNA interface to rebind the DNA operator to form the energetically minimized loop. Depending on the free energy differences and the barrier height between the kinetic and thermodynamic loops the interconversion rate may be faster when the kinetic loop is much higher in energy than the thermodynamic and be extremely slow when kinetic is approximately equal to the thermodynamic.

If LacI were considered more flexible than the DNA, the free energy differences between the kinetic and thermodynamic looped states would be compressed. Therefore, the equilibria between different loops could be established immediately since the energetic barriers separating all loop states would be lower for a flexible LacI and depend only on the elastic properties of the DNA sequence. Furthermore, if the free energy barriers separating the looped states are low enough, LacI could freely interconvert among loops. It has been suggested that LacI can interconvert directly between two different loops and bypass the unlooped LacI•DNA state (Rutkauskas, Zhan et al. 2009),

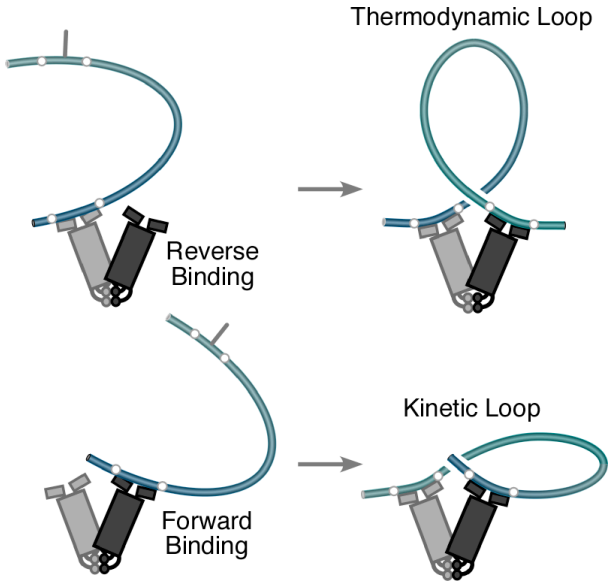


Figure 5.1.1 *Kinetic and thermodynamic LacI•DNA loops could compete and interconvert*. The symmetrical V-shaped LacI can bind a symmetrical DNA operator in two orientations equally: forward and reverse. Assuming LacI maintains the V-shaped form, the initial DNA trajectory and LacI binding orientation could result in a kinetically driven looped complex and not the thermodynamic loop. Interconversion between the two loops would require the loss of one specific LacI•DNA interface, reorientation of protein and/or DNA, and reformation of a specific interface. If the free energy of kinetic loops is much greater than thermodynamic loops then interconversion would be almost spontaneous and difficult to measure. As the free energy of kinetic loops approaches the thermodynamic loops, kinetic loops may take longer to interconvert to thermodynamic loops. Therefore, it may be possible to measure the interconversion between the two looped states. However, if LacI were flexible, especially in the DNA binding headpiece and the tetramerization domain, then interconversion would be extremely fast since the free energy barriers between the different loop states would be lower and only depend on the elastic energy of the DNA.

but the current data has not detailed the mechanism behind this prediction. However, our FPV system should be able to detect any interconversion between loops given two conditions: the transition occurs in the time scale of our experiment, and the loop interconversion does not exceed the detectable FRET limits of our FPVs.

The addition of IPTG described in Chapter 4 provided some insight into the energies for the loops formed for each FPV landscapes. The results also indicated that our loop population analyses, where we assumed each peak maximum represented a uniform looped complex geometry and would form more unstable loops at the boundaries, was a good first approximation. However, the addition of excess LacI to IPTG saturated LacI•DNA complexes indicated that our P1 loops could be comprised of two peaks,

possibly representing overtwisted and undertwisted loop populations. This suggested that the energies and/or loop shapes for P1 loops might vary drastically, even though they share similar FRET efficiencies. The ultimate goal of this project is to understand LacI•DNA loop geometries, loop shapes and the corresponding loop free energies for each DNA sequence. Therefore we considered it essential to study the relative lifetimes of each loop in the presence of DNA competitor to determine the relative stabilities of each loop detected in the looping landscape.

5.2 Materials and Methods

5.2.1 Bulk FRET FPV Landscape Studies on a Microplate

Bulk FRET measurements and experiments on a microplate followed the general methodology previously described for analysis on a Typhoon Trio Imager (GE) (Hieb, Halsey et al. 2007; Kugel 2008). This technique is a way to measure FRET on a 384 well bora-silicate microplate (Greiner, VWR) by producing an image from the measured fluorescence intensity for the samples (an example image is shown in Figure 5.3.2).

The DNA construct landscape was prepared separately for the four FPVs with DNA at 4 nM in 10 µL volumes in 7 different wells. Then 10 µL of either LacI buffer, 6 nM LacI in LacI buffer, 750 nM chloroquine in LacI buffer, or 30 nM DNA competitor in LacI buffer was added to the appropriate wells for a total volume of 20 µL each. Preliminary FRET measurements indicated that volumes less than 15 µL showed a meniscus in the well that interfered with a uniform fluorescent intensity reading. The plasmid pAH5C16, prepared in Chapter 3, served as the DNA competitor for all

landscape constructs, and was prepared using a maxiprep kit following manufacturer's protocol (Qiagen). The construct 9C14-D_IA_I as well as all FPVs for construct 5C12 were not used in this study. The Typhoon was adjusted to a normal sensitivity (600 V), the platen set to +3 mm with 200 μm scan sizes. Three different fluorescent measurements were performed: the direct excitation of donor (532 nm) and measurement of its emission (580), the direct excitation of acceptor (633 nm) and measurement of its emission (670 nm), and the direct excitation of donor (532 nm) while monitoring acceptor emission (670 nm). The three measurements on 4 microplates took ~ 15 minutes, but all three measurements were necessary to adjust for any fluorescence variations over time, which was previously observed during DNA competition studies (Chapter 3).

After two hours an additional 10 μL of the solutions of either buffer, LacI, competitor DNA, or chloriquine as described above were added to the appropriate wells and fluorescent measurements were continued over a 18 hours. Finally, after this time, 5 μL was removed from each sample, and 5 μL of 150 mM IPTG was added, and FRET was measured for another ~ 4 hours.

5.2.2 FRET Efficiency Calculations and Data Fitting

FRET efficiencies were calculated as shown below, similarly to Chapter 3.

$$ET_{donor} = \left[1 - \frac{C_{donor}(\text{+xfer})}{C_{donor}(\text{DNA only})}\right]\frac{1}{f_A}$$

$$ET_{acceptor} = \left[\frac{C_{acceptor}(\text{+xfer}) - C_{acceptor}(\text{DNA only})}{C_{donor}(\text{DNA only})}\right]\frac{\Theta_D/\Theta_A}{f_A}$$

However, because the doubly labeled DNA constructs samples provided similar spectra to the singly labeled DNA samples (can be seen in the residuals in Figure 3.3.2) the

doubly labeled samples were used directly to compare energy transfer samples to no transfer for each time point. Since the concentration can vary between construct samples on the microplates, each measurement was normalized to emission for the direct excitation of the acceptor for the DNA only sample, $C^{exc}_{acceptor}(\text{DNA only})$, relative to the direct excitation of the acceptor for a sample mixed with a ligand that can possibly cause energy transfer, $C^{exc}_{acceptor}(+\text{xfer})$.

$$C_{acceptor}(+\text{xfer})' * C^{exc}_{acceptor}(+\text{xfer}) = C_{acceptor}(+\text{xfer}) * \left(\frac{C^{exc}_{acceptor}(+\text{xfer})}{C^{exc}_{acceptor}(\text{DNA only})} \right)$$

$$C_{donor}(+\text{xfer})' * C^{exc}_{acceptor}(+\text{xfer}) = C_{donor}(+\text{xfer}) * \left(\frac{C^{exc}_{acceptor}(+\text{xfer})}{C^{exc}_{acceptor}(\text{DNA only})} \right)$$

Therefore, the donor FRET efficiency could be calculated from donor quenching by measuring the direct excitation of the donor while monitoring its emission. Acceptor was calculated from fluorescence enhancement measured by exciting the donor and monitoring acceptor emission.

The overall fluorescence intensity was calculated by integrating the volume from the Typhoon image for each microplate well using a circle of equal diameter to the well in ImageQuant. However, due to imperfections in the signal intensities, caused by defects on the plate surface, the typhoon surface, scratches, or dust, etc., the median value was used to represent the fluorescence intensity calculated from the integrated area from each well (a calculated output generated by ImageQuant).

The decay of the FRET efficiencies for LacI•DNA complexes with excess DNA competitor was fit to a single exponential. However, first the values for LacI•DNA + DNA competitor samples, ET_t^{DLC}, needed to be normalized to the decay of FRET

efficiency over time observed with LacI•DNA only without DNA competitor, ET_t^{DL0}, shown below. These calculations were all handled as matrices of each FPV landscape ($t \times n$ constructs per FPV). The FRET efficiency measured for LacI•DNA only at the time competitor DNA was added, ET_i^{DL0}, was subtracted from each LacI•DNA only time point, ET_t^{DL0}. Then this was subtracted from ET_t^{DLC} to correct possible contributions of ET_t^{DL0} decay from the decay of ET_t^{DLC} to give $ET_t^{DLC}{}'$ data.

$$ET_t^{DLC}{}' = ET_t^{DLC} - (ET_t^{DL0} - ET_i^{DL0})$$

The ET_t^{DLC} was then fit to a single exponential by setting the amplitude to the maximum value of ET_t^{DLC}, $\max\_ET_t^{DLC}$, minus the endpoint, $\min\_ET_t^{DCL}$, which was a matched sample of DNA + competitor DNA premixed before the addition of LacI. The $\min\_ET_t^{DCL}$ was also used as the plateau value for the exponential fit. The decay rate constant, *k*, was calculated as shown below by minimizing the sum of squares of the residuals of the fit, fit, minus the experimental data in Matlab with the *fmincon* function. All of the Matlab scripts and function are included in Appendix 3.

$$fit = (\max\_ET_t^{DLC}{}' - \min\_ET_t^{DCL}) * e^{-k*t} + \min\_ET_t^{DCL}$$

The half-lives were calculated using the relationship $t^{1/2} = \ln 2 / k$. The half-lives were calculated using the $\min\_ET_t^{DCL}$ as the plateau constraint for a single exponential fit in Prism 4 for comparison to Matlab calculations and also visual representation of the fit to the data shown in the Results.

5.3 Results

5.3.1 LacI•DNA Loops Equilibrate in Less Than One Minute

The FRET efficiency for LacI mixed with DNA was monitored for almost two hours for all FPV landscapes. LacI•DNA loops apparently equilibrate before the first FRET measurement. These results corroborate that previous FRET measurements were performed at equilibrium. FRET did not increase over time for any construct; instead, Figure 5.3.1 shows FRET efficiency tended to have decrease over the two-hour period. Initially, this was puzzling until we discovered that FRET returned to its original levels upon remixing of the microplates after two hours. Therefore, we believe the decrease in FRET is caused by protein binding to the sides of the microplate, and not related to the fluorescent properties of the dyes. However, this important observation was crucial in analyzing the half-lives for loops described later in the Results and in the Materials and Methods as the correction to the calculation of loop half-lives.

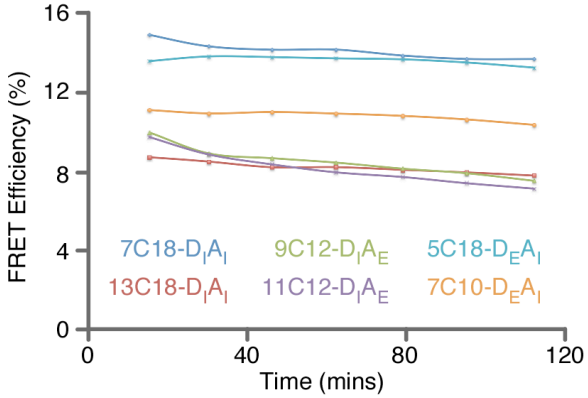


Figure 5.3.1: *LacI•DNA complexes form rapidly and do not interconvert on the hour time scale*. FRET efficiencies were measured at 15, 30, 46, 62, 79, 95, and 112 minutes on a Typhoon Imager with 2 nM fluorescently labeled DNA mixed with 2 nM LacI in LacI buffer. Image scanning and conversion to FRET efficiencies are described in the Materials and Methods in Chapter 5. Two representative examples that had significant FRET are shown for each FPV landscape. FRET decreased slightly over time for each construct and never increased. Since FRET can be fully recovered by mixing, the slow decay of FRET is attributed to protein-DNA aggregation and/or protein binding to the microplate walls.

Interconversion between V-shaped loops was not observed, since FRET never increased proportionally to a decrease in FRET for any single construct. Therefore, the V-shaped loops formed must be exchanged rapidly. Possible reasons for these observations are presented in the Discussion.

5.3.2 The Construct Landscape Forms Hyperstable LacI•DNA Loops

Previously, DNA competition studies on the high FRET FPVs in the construct landscape indicated these constructs were stable for at least two hours. However, the measurements at twenty-four hours had fluorescent interference from the buffer and/or donor that prevented us from evaluating the loops past the two-hour period (refer to Figure 3.2.4). Performing bulk FRET on a microplate allowed us to evaluate the looped complexes for all FPVs continuously for many hours, and the controls could easily be evaluated in parallel as shown in Figure 5.3.2A. Therefore, any unusual buffer behaviors of the experimental system for each FPV could be addressed in this analysis, as well as changes in fluorescence. In Figure 5.3.2B the three FPV landscape D_IA_I, D_IA_E, and D_EA_I (P1, A1, and A2 respectively) are shown with an exponential fit to each data set. The slow decay of FRET efficiencies over 20 hours in the presence of a 5-fold excess of unlabeled DNA demonstrated the loops formed are hyperstable. The half-lives, calculated for each FPV with greater than 3% FRET efficiency, range between ~ 10–100 hours for loops as shown in Figure 5.3.2C. The 3% FRET cutoff was necessary to address FRET signals that did not change, and thus their calculated lifetimes are infinite. However, this threshold may not have been set high enough since in the A1 landscape a half life was

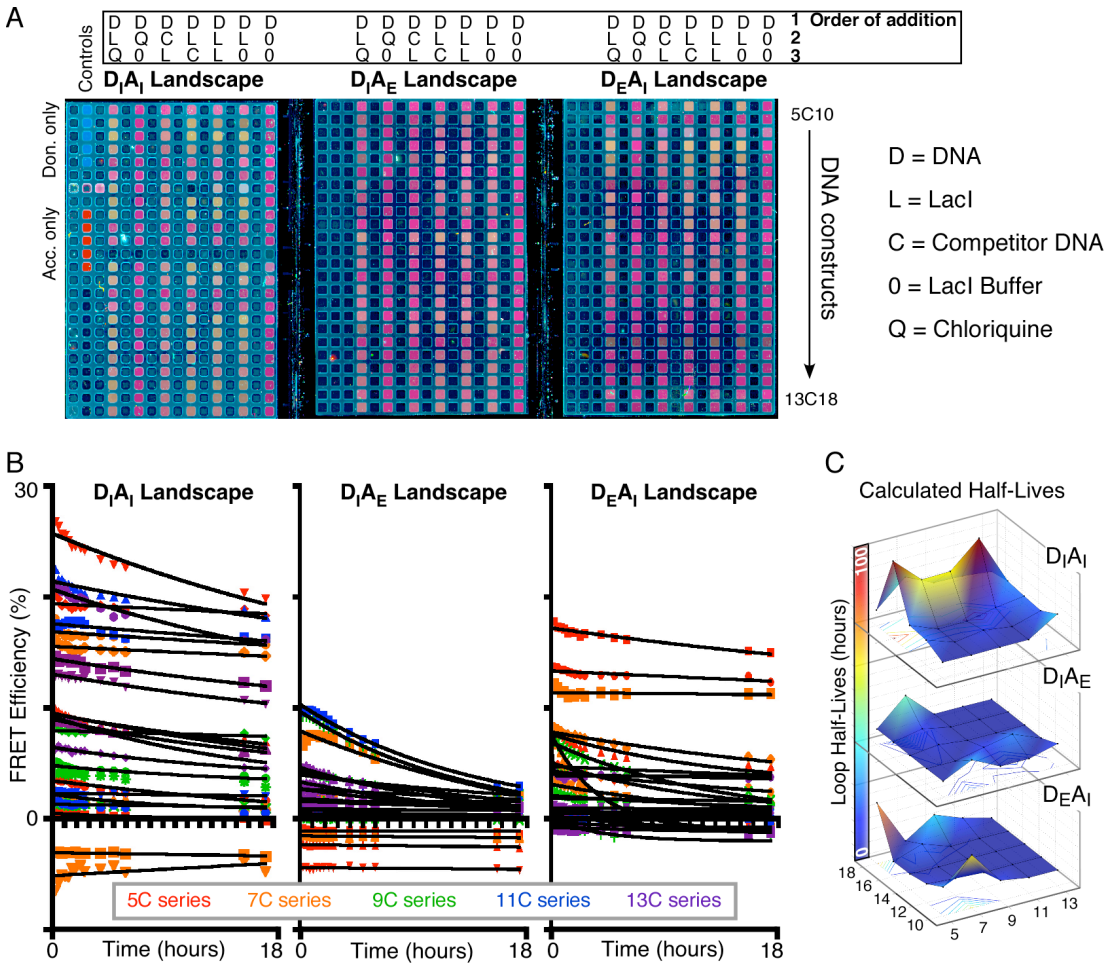


Figure 5.3.2: *The designed construct landscape forms hyperstable LacI•DNA loops.* (A) An example Typhoon overlayed image of the three fluorescence intensity measurements imediately after the second addition as described in the Materials and Methods. (B) FRET was measured for all FPV landscapes 15 minutes after the addition of 5-fold excess DNA competitor (10 nM of plasmid, pAH5C16) to all LacI•DNA complexes that were preincubated for two hours. FRET was calculated as described in the Materials and Methods, but only the FRET for the acceptor is reported. The D_EA_E landscape was a flat landscape and is not shown. 5C12 and 9C14-D_IA_I were not available for FRET measurements. The FRET data was corrected for FRET decay and each FPV was fit to a single exponential decay in Matlab and Prism (shown) as described in the Materials and Methods (C) The half-lives calculated for LacI•DNA looped complexes in the presence of excess DNA competitor. The decay rate from the single exponential was used to determine the half-lives for each FPV landscape. A2 (D_EA_I) and P1 loops (D_IA_I) have the longest half-lives, while A1 V-shaped loops (D_IA_E) are more susceptible to DNA competition. The lifetimes correlate well with the boundaries separating uniform loop populations being higher energy loops. However, the P1 loops also appear to have two discrete half-life maxima, although previously P1 appeared as a high-FRET ridge. This suggests that even though some loops share similar FRET maxima, the underlying energies for the loops are different and presumably so are the loop shapes.

calculated even though their appears to be no significant FRET present for construct 7C18-D_IA_E.

FRET efficiency decayed quickly for A1 loops and also for the boundaries separating uniform loop topologies, similar to the effects of saturating IPTG in Chapter 4. The P1 (D_IA_I) landscape separated into two discrete half-life maxima that in our previous results could be inferred, especially with the addition of excess LacI to IPTG saturated complexes, but we still considered to be slightly speculative. The absence of 9C14-D_IA_I (averaged from neighboring points for visual continuity) in the experiment may have contributed to the discontinuous half-lives for this surface. However, this reoccurring theme may be significant and is detailed in the Discussion in this Chapter.

5.3.3 IPTG•LacI•DNA Loops Equilibrate Rapidly and the Relative Stability Among Constructs Can Be Differentiated

Figure 5.3.3 shows the order of addition controls for kinetic experiments that were performed on the construct 5C18-D_EA_I. In this particular competition study, 20 nM of unlabeled construct 5C14 was used as the DNA competitor. When LacI was added to the mixture of 2 nM labeled and 20 nM unlabeled DNA little to no FRET was observed. Upon the addition of saturating (5 mM) IPTG FRET appeared after less than minute, although not to the same extent observed for 5C18-LacI complexes in the absence of IPTG. The appearance of FRET after the addition of IPTG in the presence of competitor, but not before, can only be observed if an IPTG-LacI•DNA looped complex is formed for the doubly labeled construct. Furthermore, since 5C14 competitor was at 10-fold excess

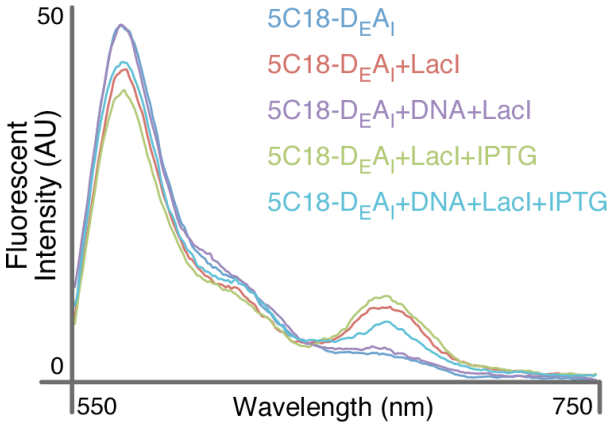


Figure 5.3.3: *LacI•DNA complexes equilibrate rapidly in the presence of IPTG.* The order of addition experiments with 2 nM fluorescently labeled DNA construct 5C18-D_EA_I, the unlabeled 20 nM DNA competitor 5C14, and 5 mM IPTG in LacI buffer. The spectra obtained from exciting the donor (514 nm) and scanning the emission spectra from 550-750 nm for each condition. The spectra were smoothed in Matlab (a 10 span sampling) for a clearer visual separation between the scans. The presentation of each variable represents the order in which it was added; 5C18-D_EA_I + DNA + LacI, competitor DNA 5C14 added before the addition of LacI. The addition of unlabeled DNA competitor before LacI shows no significant FRET for 5C18-D_EA_I. However, the addition of IPTG to 5C18-D_EA_I + DNA + LacI, shows substantial FRET, and FRET is not affected for 5C18-DNA looped complexes in the presence of saturating IPTG. This indicates that even though 5C14 is in 10-fold excess relative to 5C18- D_EA_I its stability in the presence of inducer is lower than the loops formed by 5C18- D_EA_I. The equilibration between loops in the presence of inducer occurred in less than one minute, and suggests IPTG bound LacI is more transiently associating and sampling the DNA surface.

relative to 5C18-D_EA_I the population had drastically shifted in less than one minute as observed with FRET. This must be due to the loop(s) formed by 5C14 being much lower in stability than the loop formed by 5C18-D_EA_I in the presence of inducer. This indicates that LacI•DNA equilibrium in the presence of inducer is very fast.

Bulk FRET studies on a microplate with saturating IPTG reached equilibrium before the first measurement at 15 minutes. However, DNA competition studies with 5-fold excess plasmid 5C16 DNA diminished FRET universally for all FPV landscapes in the presence of saturating IPTG, and the flat FPV landscapes are not shown. Because the goal of the DNA competition studies was to determine the overall stability of each looped complex, we considered it more important to use a plasmid that not only contains a

looping complex but also has ~ 14-fold more nonspecific DNA sites and two additional secondary operators (O_1 and O_2). Therefore, it was not surprising that FRET was lost for all FPV landscapes. However, doing this allowed us to determine the half-lives more effectively, but did not allow the energy differences between two different looping constructs in the presence of saturating inducer to be determined as suggested by the results in Figure 5.3.3. Not surprisingly, this informs us that in order to study IPTG effects on the partitioning among preferred LacI•DNA loops for a given set of operator phasings the DNA should be comparable in both length and sequence, or using less DNA competitor to bring the equilibria closer together.

5.4 Discussion

5.4.1 Interconversion Among LacI•DNA Loops Depends on the Energy Differences Among Kinetic and Thermodynamic Products

Interconversion between LacI•DNA loop states was briefly described in the Introduction. We believed that if FRET was monitored for an extended period of time it might be possible to see an individual construct decrease FRET for one FPV and have a corresponding increase for another FPV. This would only be possible if the free energy differences between one loop state and the alternative were very small but the barrier for interconversion substantial enough to prevent direct interconversion. Interconversion between loops would most likely occur at the boundaries that separate uniform loop populations, where presumably the looped states are higher in energy and the energy barriers are lower relative to these loop states. Therefore, the initial LacI•DNA complex

orientation, whether it is a singly bound, unlooped complex or a specific nonspecific looped complex should partition the final loop state.

If we consider initial mixing of LacI DNA, with no predisposition for a specific orientation when binding DNA, there should be a 50/50 chance of LacI binding a symmetrical operator in the forward or reverse orientation. Of course LacI probably initially binds nonspecifically with one dimer and begins the operator search, but we will assume one dimer is fixed to the operator with one free dimer for simplicity. The binding of the second dimer would most probably result in the formation of a specific-nonspecific loop until locking onto the other operator. The final equilibrium should represent the energy differences between the two specific-specific looped states. However, we never saw any substantial evidence over a two-hour time period for the increase in FRET efficiency in one FPV while decreasing simultaneously in another. This suggests we cannot measure loop interconversions on this time scale. It could be argued that the loops at the boundaries could be interconverting, but their interconversion could be changing FRET so little that it is within the noise of the experiment. More specifically, the boundary constructs typically have lower FRET and therefore a smaller signal to noise ratio, and therefore a population change from 50% – 60% or 50% – 40%, would be more difficult to quantify in these situations.

An alternative explanation why interconversion was not experimentally observed can be ascribed to protein flexibility or to rapid interconversion among bound forms of the LacI DNA binding headpiece in a DNA complex. For example, interconversion of loops via specific to nonspecific transitions without passing through a singly bound state.

In general, LacI flexibility would lower the energy barrier between LacI•DNA loop states and allow LacI to sample more states in a short time period. The order of addition study presented in Figure 5.3.2 demonstrated that flexibility, at least in the presence of IPTG was essential in operator searching and loop formation. No FRET was observed when LacI was added to a mixture of labeled 5C18-D_EA_I and 10-fold excess unlabeled 5C14 competitor DNA. Upon the addition of saturating IPTG, substantial FRET was observed for 5C18-D_EA_I in less than one minute. Remarkably, this suggests that in the presence of IPTG, LacI sampled through many DNA molecules to form the most stable loop on the minority of fluorescently labeled 5C18 DNA. The half-lives calculated for the two constructs suggest the stability of each loop is different (5C14-D_IA_I = 19 hours and 5C18-D_EA_I = 75 hours), so it is not surprising that this occurred, but the rate at which it occurred is surprising. The fast re-association rates suggest that LacI does actively sample many different loop states and is transiently associating with the DNA in the presence of IPTG. LacI presumably has a similar underlying behavior in the absence of inducer, but with 1000-fold higher DNA association rates (Barkley, Riggs et al. 1975). Therefore, LacI flexibility could still allow reconfiguration of a looped complex over longer time scales in the absence of inducer, but still equilibrate before our first experimental measurement. Single-molecule FRET on an immobilized LacI•DNA complex would answer the possibility of interconversions of loops, measure loop lifetimes, and loop state fluctuations in the microsecond timescale.

5.4.2 The P1 Landscape Has Two Distinct Populations

The addition of excess LacI to LacI•DNA looped complexes (Chapter 4) in the presence of inducer indicated that although FRET was still high for some loops, their responses to inducer were remarkably different. We previously suggested that P1 loops, and more generally all loops, should have high FRET constructs on each FPV landscapes where the relative operator phasings result in DNA overtwisted or undertwisted in a LacI•DNA looped complex. Distinguishing the two higher-energy twists is important for modeling the torsional limits of DNA twist in a LacI•DNA looped complex and also understanding the contributions of protein flexibility in a loop.

In Figure 5.3.2 the half-lives for each FPV landscape show that two distinct peaks occur when the LacI•DNA looped complexes were challenged with excess DNA competitor. These results support previous observations of two unique maxima for P1 (D_IA_I) loop populations appearing when challenged with excess LacI experiments in the presence of inducer. Initially we expected that these two peaks should overlap each other, with each peak centered on the same construct, since both should be providing insight into the underlying stability of the loop. However, one peak does not align with the other as re-shown in Figure 5.4.2, and there are no clear explanations for these results and the differences between

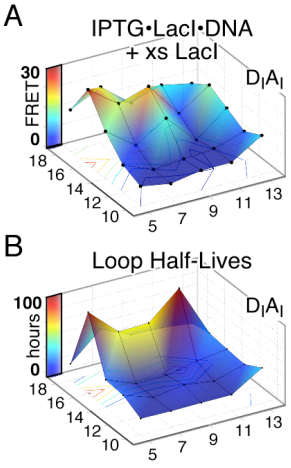


Figure 5.4.2: *Competition studies with excess LacI or DNA both indicate two different P1 loops populations*. (A) The addition of excess LacI to IPTG-LacI•DNA complexes for the D_IA_I (P1) landscape from Figure 4.3.3. (B) The half-lives calculated from the addition of excess unlabeled DNA to LacI•DNA complexes for the D_IA_I landscape from Figure 5.3.2.

them. It is possible that these two landscapes in Figure 5.4.2 should not align, since we are comparing the looping free energies ± inducer, and therefore may not share identical loop properties as seen with *in vivo* studies (Becker, Kahn et al. 2005; Bond, Peters et al. 2010). Moreover, the *in vivo* studies suggest that inducer saturated loops are more sensitive to the torsional phasing of the operators and the inducer saturated loops also had altered writhe and/or local DNA unwinding in comparison to uninduced loops. Therefore, it is possible that the results for these two different competition experiments agree perfectly with each other, and the differences between the two landscapes elaborate the subtle variations in DNA shape and torsion in a P1 V-shaped loop. However, it is difficult to address subtle shape and torsional variations among loops with our bulk data, but DNA modeling that considers protein flexibility and inducer effects on the looped complex might better address these considerations for two distinct P1 loop populations and their context in the presence of inducer.

Chapter 6

Significance of Results and Future Prospects

6.1 Physical Models for Predicting DNA and Protein Flexibility

The results described in this work will be important in developing physical models for gene regulation that include the contributions of DNA, protein, and allosteric effectors. We have thoroughly investigated a large DNA sequence space, and these results will generate a deeper understanding of the underlying flexibility of LacI and DNA that can improve current models. Optimistically, this will include expanding the current models for DNA sequence dependence effects and the inherent flexibility for a given DNA sequence. Additionally, our simple population models determined from FRET efficiencies can be further developed to include or be compared to models that include LacI headpiece movement in addition to the opening of the tetramer leading to the extended state, as our current understanding of protein flexibility becomes more applicable to detailing these potential contributions.

In Figure 6.1 the calculated population distributions are displayed next to the operator axis dyads for each construct. This sequence dependent illustration demonstrates some clear patterns for the looping preferences of LacI•DNA loops based on both operator phasings relative to the central bend. Foremost, any time an operator is pointed outward from the central bend it will have a large contribution of P1 loops regardless of the orientation of the other operator. Although this observation depends on the underlying model used to generate the Pymol rendering, it does suggest that simple observations such as this may aid in the development of a more detailed understanding of sequence dependent DNA flexibility and how DNA binding proteins accommodate to the sequence dependent requirements. Further conclusions can be made from the operator

arrangements and the corresponding populations as shown in Figure 6.1. However, we do not want to emphasize this pymol model, but more importantly demonstrate the value of this work in detailing the fundamental sequence dependent properties of LacI•DNA loops.

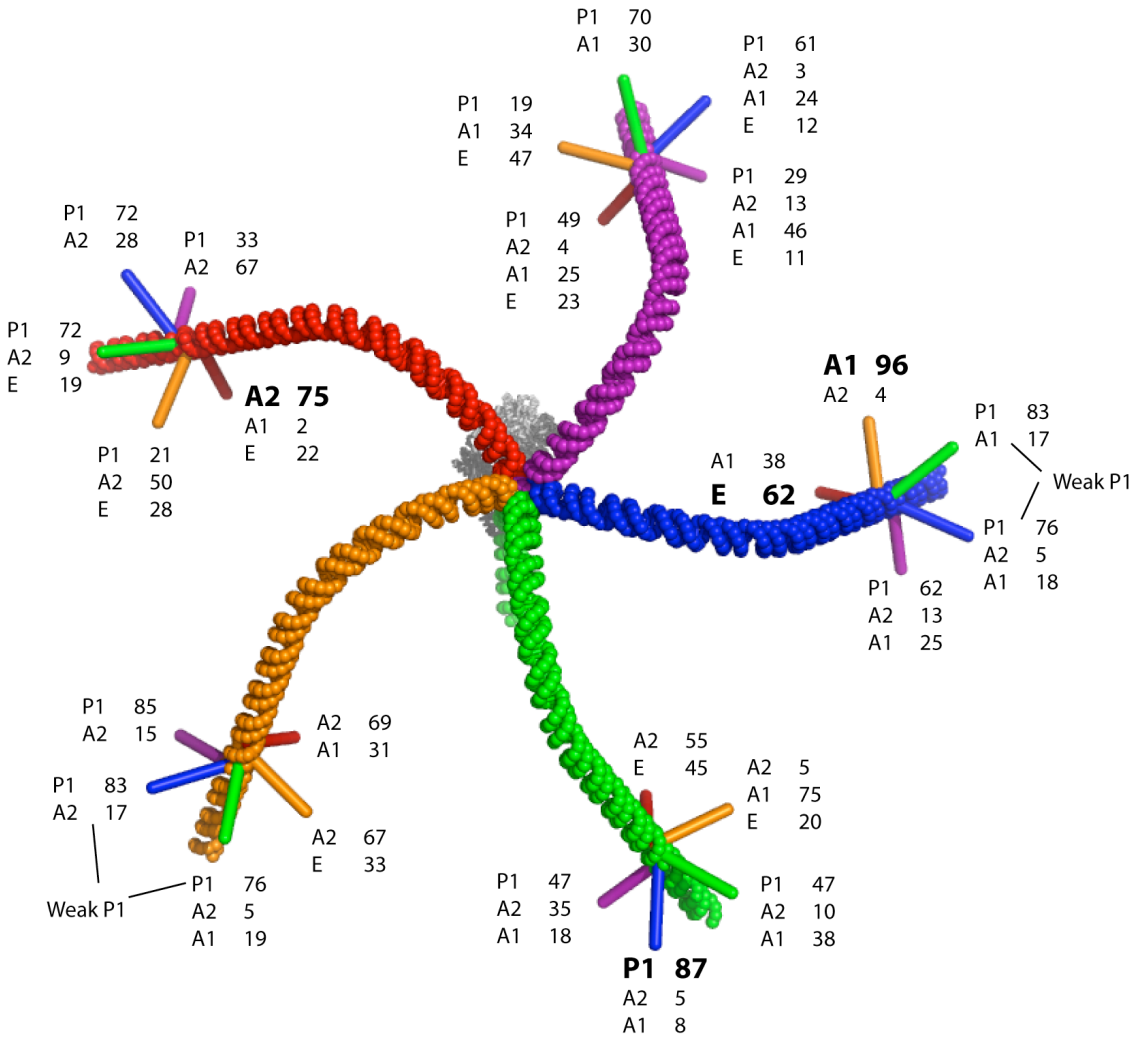


Figure 6.1: *Calculated loop population distributions set to the corresponding operator phasings for the DNA construct landscape*. The landscape was rendered in Pymol as described in Figure 2.4 with LacI shown bound to the donor side operator (gray). The populations calculated in Figure 3.2.5A are shown next to the corresponding operator axis dyad. Each loop type, P1, A1, and A2 are highlighted on the sequence that had the highest calculated population distribution. The weak P1 loops indicate the constructs that decreased FRET valleys (in Figure 4.3.3) in response to excess LacI to IPTG saturated LacI•DNA complexes.

Our work should be applicable to all models for similar protein-DNA systems. The FPVs can provide detailed information on the DNA trajectories for any protein DNA complex, especially since the fluorophore positions can be generalized to many locations on the DNA or even on the protein. These features will be of greater importance as the R_0's for fluorophore pairs increase, and allow us to shed light on previously dark states.

The relative ease at which many different variations of FPVs can be made for a single DNA sequence suggests that this approach could be scaled to micro-array style proportions to quickly examine a plethora of DNA sequences. Furthermore, FPVs can be applied to anything that brings two DNA polymers in close proximity and is not limited to protein-DNA interactions.

Eventually, the FPVs, may earn the name "nano positioning system" (NPS) (Brunger, Strop, et al. 2011) and provided quantifiable triangulation measurements similar to the fundamental ideas behind GPS.

6.2 Nanotechnology Applications

The ability of DNA to self-assemble and its stiffness make it an excellent building-block for nanotechnology applications. Additionally, the sequence can be easily controlled and manipulated to include regions of sequence dependence stiffness, flexibility, or intrinsic bends. Our results also suggest that in addition to the intrinsic properties of DNA ,the presence of a protein, such as LacI, can manipulate the DNA shape and create complexes that can have long lifetimes. Our results can be applied to many DNA nanotechnology applications. For example, inclusion of two appropriately

phased operators can lead to controlling a loop with a large or small diameter. The long lifetimes of these loops would allow additional structural layers to be built upon the scaffolding of the original shape still maintained by LacI•DNA complex.

Another application of our results is in DNA nano-circuitry or nano-structure. For example, fixing two DNA strands running parallel to each other upon a surface while being held in position by multiple DNA-LacI•DNA sandwiches. After fixing the DNA to the surface, the LacI could be washed away with an IPTG solution to leave only the DNA behind. This could be easily elaborated upon and could have many possibilities as DNA circuits, attaching markers to the DNA to act as molecular probes, or simply architectural scaffolding.

6.3 Biological Implications

As the complexities of cellular gene regulatory networks become more realized, we rely more upon our simplest models to disentangle the fundamental components of each circuit. The *lac* operon has served our basic understanding of gene regulation well for over 50 years, but this simple system must be fully understood to continue being a paradigm for describing similar systems. We took on LacI in a systematic evaluation to discover that the key to LacI's repression efficiency is its flexibility and symmetry. The flexibility allows LacI to occupy a continuum of geometries while the plasticity of the DNA binding headpiece samples and adjust to the surface of the DNA. The symmetry of LacI can bind DNA in multiple orientations which increases the likelihood that it remains bound to the operator, even if the operator is asymmetrical as in the natural occurring O_1,

O_2, and O_3 *lac* Operon. Although we considered including an asymmetrical operator in our system to mimic the natural system, we decided to focus on exploring all possible geometries of LacI repressor to better understand its contribution in gene regulation.

The presence of inducer allosterically effects LacI as described in Chapter 4, but not as dramatically as we initially imagined. LacI is still a DNA looping protein in the presence of saturating inducer, but appears to increase switching between specific and nonspecific DNA interactions. Surprisingly, we never saw LacI completely release from the DNA, which indicates that it should still be effective in gene repression. Therefore, it is possible that inducer-bound loops, or less stable loops in general, can be controlled by mechanisms such as DNA supercoiling, or mechanical forces resulting from translocation by the transcriptional machinery (Finkelstein, Visnapuu et al. 2010). The role of some allosteric effectors may be to reduce the affinity of the repressor for the gene so it can be more easily displaced and not necessarily to allosterically release the repressor.

Appendix 1

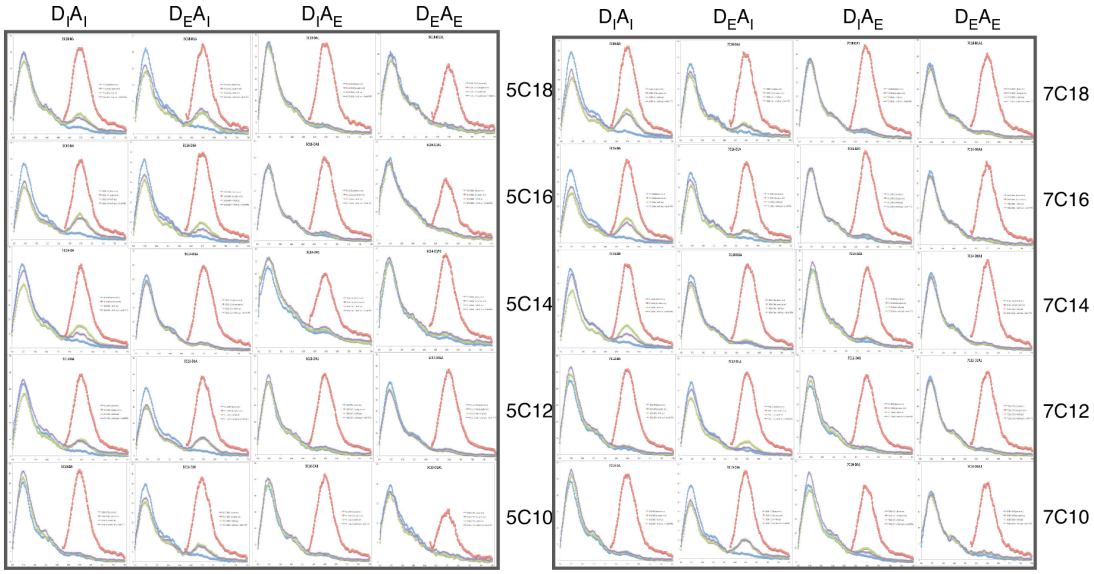
Data, Analysis, Results, and Additional Figures

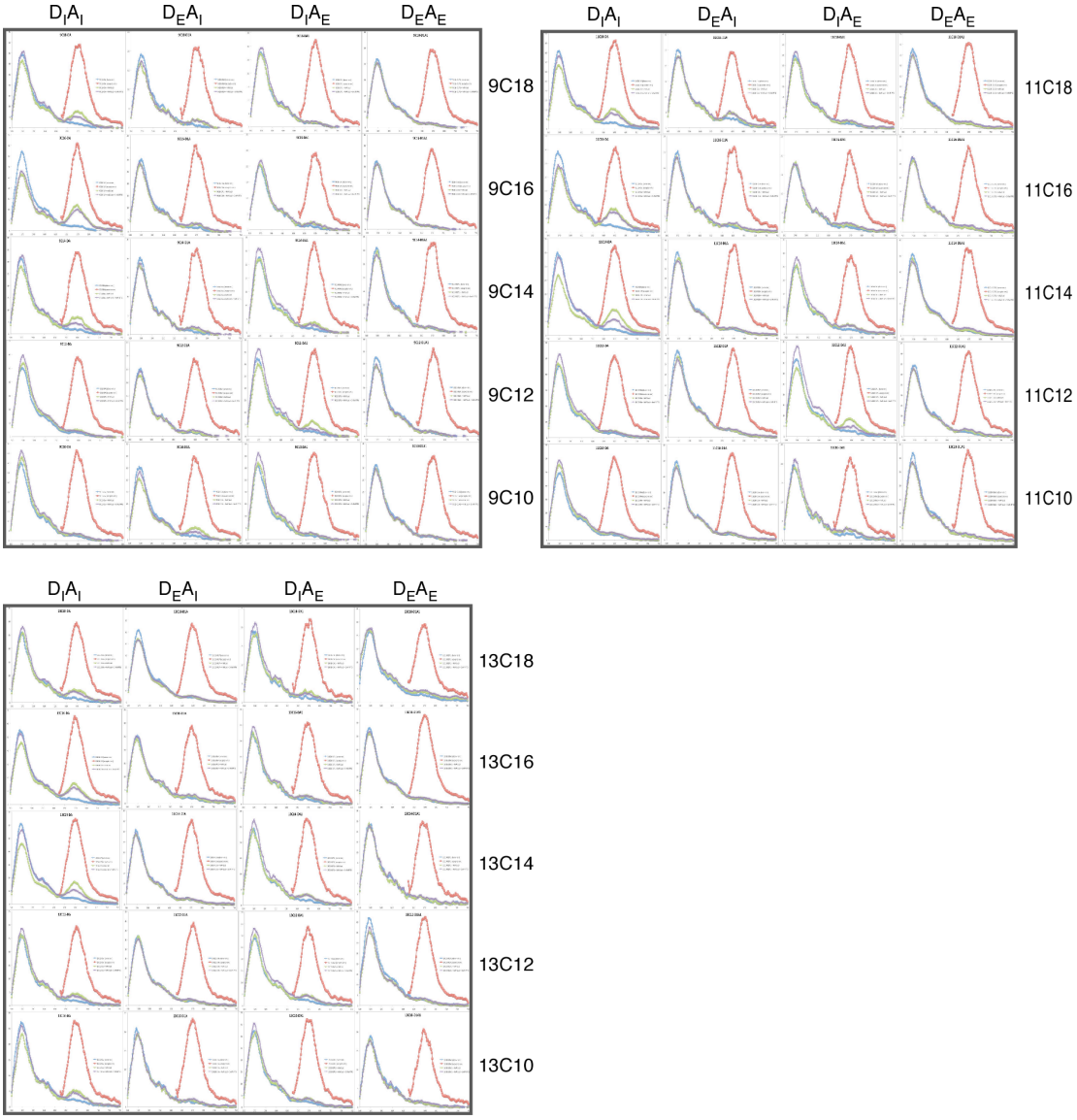
1.1 Qualitative Assessment of EMSAs on the LacI•DNA looping landscape

EMSA		Donor Side				
		5C__	7C__	9C__	11C__	13C__
Acceptor Side	_C10	5C10	7C10	9C10	11C10	13C10
	_C12	5C12	7C12	9C12	11C12	13C12
	_C14	5C14	7C14	9C14	11C14	13C14
	_C16	5C16	7C16	9C16	11C16	13C16
	_C18	5C18	7C18	9C18	11C18	13C18
green-high affinity and tight bands						
yellow-high affinity and smearing						
blue-low affinity and tight bands						
red-low affinity and smearing						
white-not sufficient information						

Appendix 1.1: Qualitative interpretation of LacI•DNA EMSAs. Two criteria were considered in this analysis, stoichiometry to completely shift DNA to a LacI•DNA looped complex and tightness of band relative to DNA only. These results were inconclusive due to the high variability between gel shift conditions and DNA labeled concentrations. All gels were not made from the same gel stock, and did not have internal looping controls, which would make a quantitative analysis more reliable. The variation in DNA* concentrations derived from the assumption that both DNA ends were 100% ^{32}P labeled, when actual labeling efficiency varied between samples. Insufficient information, meant that visualization of bands was difficult, or gel was damaged during handling and prevented any interpretation.

1.2 Individual FPV Bulk Spectra for the Construct Landscape





Appendix 1.2: The individual Spectra for the FPVs for each construct grouped in the Donor side adaptor families. All spectra share the same x-axis (550-750 nm), but the y-axis varies with the fluorescence intensity for each FPV. The blue line indicates direct excitation of the donor (514 nm) and the emission scanned (550-750 nm). Red corresponds to direct excitation of the Acceptor (600 nm) and the emission scanning (640-750). The green line is the excitation of the donor in the presence of LacI and purple shows the addition of IPTG to LacI•DNA complexes while scanning the emission spectrum. The construct "5C12" DNA sequence is actually the construct 5C18, due to sequencing errors discovered after the experiments.

1.3 Calculated Bulk FRET Efficiencies and Their Statistical Significance

1.3.1 Bulk FRET Efficiencies, Averages, and Standard Deviations for LacI•DNA Loops

LacI-DNA Looped Complexes

FPV	Donor ET					Acceptor ET					
DIAI	-5.7	-5.8	-5.3	-8.1	17.6	2.1	2.4	2.2	2.2	12.0	_10
	4.5	-4.8	-4.9	-2.1	0.8	8.4	3.6	3.5	3.1	10.5	_12
	24.1	28.3	11.2	29.6	24.5	19.6	22.3	17.5	25.5	20.6	_14
	32.3	32.7	28.9	23.8	15.7	26.0	27.8	27.1	22.4	21.2	_16
	8.2	32.1	12.6	19.3	3.9	17.7	27.7	16.5	19.5	13.9	_18
DIAE	-1.1	10.4	-4.4	3.5	2.9	2.3	12.5	3.2	10.2	6.1	_10
	-5.1	-3.0	10.6	15.0	2.0	3.0	2.7	16.3	21.4	10.5	_12
	-11.4	8.2	2.8	3.9	5.4	3.9	6.5	11.2	9.9	10.3	_14
	3.0	4.3	0.5	1.2	2.2	4.5	4.9	7.5	5.4	6.5	_16
	6.7	9.6	6.0	7.3	4.4	2.5	4.7	1.5	4.1	12.5	_18
DEAI	15.5	20.0	9.8	-1.9	-0.8	12.0	16.8	10.4	1.4	2.2	_10
	10.7	14.0	0.2	0.1	-3.6	7.7	10.5	1.8	2.4	1.1	_12
	2.6	6.2	-0.1	-1.7	-1.8	0.7	3.6	3.9	1.7	0.8	_14
	19.6	13.9	2.3	1.0	-2.0	14.4	9.1	4.5	1.0	3.1	_16
	19.3	9.7	7.1	0.9	3.2	14.7	9.9	6.5	6.2	1.7	_18
DEAE	5.5	-7.8	-2.3	2.0	0.2	0.3	5.8	1.9	-0.1	0.7	_10
	-3.1	-2.8	3.1	-4.3	8.7	2.6	1.9	1.6	-0.3	0.5	_12
	-12.0	0.0	-0.2	0.7	-4.0	5.6	0.6	1.4	1.3	1.2	_14
	0.2	-1.1	-3.9	-0.2	-1.5	0.1	0.7	1.2	3.3	1.5	_16
	2.8	1.1	-3.3	2.5	-18.5	1.9	1.2	0.6	-0.4	6.3	_18
	5_	7_	9_	11_	13_	5_	7_	9_	11_	13_	Adaptor

LacI-DNA Looped Complexes Averaged

FPV	Average ET for All Trials					ET STDEV					
DIAI	0.9	0.3	0.7	-0.6	14.5	4.7	4.1	4.2	5.0	4.5	_10
	7.6	0.9	2.1	1.2	8.0	3.1	3.9	5.0	2.3	6.1	_12
	21.0	24.5	15.0	21.6	20.4	6.5	4.4	2.9	8.4	6.6	_14
	27.6	27.2	28.2	20.5	16.4	3.7	4.3	2.3	5.2	4.8	_16
	15.1	25.6	11.7	18.1	10.8	5.0	6.0	3.6	4.5	5.6	_18
DIAE	1.7	11.6	2.0	7.0	6.6	2.1	2.1	5.3	2.9	5.1	_10
	1.8	1.5	14.6	14.9	10.3	4.8	3.5	3.9	4.7	8.8	_12
	2.3	9.0	8.2	7.6	9.6	9.4	5.0	3.8	6.1	3.4	_14
	4.5	5.2	6.2	6.3	4.7	1.4	1.4	5.0	6.9	2.2	_16
	4.3	7.2	4.2	5.1	9.3	2.1	3.5	2.7	2.3	4.2	_18
DEAI	13.5	16.7	11.2	0.0	1.3	2.6	3.1	3.2	1.5	1.5	_10
	8.7	10.8	2.4	1.7	0.4	2.0	2.7	2.8	1.6	3.0	_12
	1.8	5.0	2.5	1.5	1.1	0.9	1.6	1.8	3.1	3.0	_14
	14.5	10.5	3.4	0.9	1.7	3.8	4.2	1.4	0.9	2.8	_16
	16.7	8.7	7.4	6.6	2.2	3.3	2.1	2.9	4.9	1.0	_18
DEAE	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	_10
	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	_12
	NaN	NaN	1.9	NaN	NaN	NaN	NaN	2.5	NaN	NaN	_14
	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	_16
	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	_18
	5_	7_	9_	11_	13_	5_	7_	9_	11_	13_	Adaptor

Appendix 1.3.1: The FRET Efficiencies calculated for LacI•DNA looped complexes. The individual measurements for Donor quenching and acceptor enhancement are as described in the methods (Left). Large differences between donor and acceptor are ascribed to the sensitivity of the Donor properties that are environmentally specific. The Acceptor properties appeared less effected by environmental changes. Average FRET between multiple LacI•DNA complexes and the standard deviation between measurements (Right) demonstrate that larger variations occur when FRET efficiencies are low.

1.3.2 Bulk FRET Efficiencies, Averages, and Standard Deviations for IPTG saturated LacI•DNA Loops

IPTG-LacI-DNA Looped Complexes

FPV	Donor ET					Acceptor ET					
DIAI	-2.4	-2.4	-3.2	-3.3	15.8	2.5	2.3	2.0	2.6	8.5	_10
	4.3	-2.2	-2.3	-3.1	2.2	5.1	2.7	1.8	2.0	7.0	_12
	17.5	17.9	6.8	16.1	17.3	10.0	12.0	9.0	12.7	11.7	_14
	32.6	31.1	31.2	25.0	13.5	23.5	22.2	23.6	19.7	16.9	_16
	7.7	30.6	9.6	15.5	3.3	13.6	23.8	9.8	14.1	10.4	_18
DIAE	1.7	1.6	-4.6	5.0	4.1	2.8	6.3	3.9	11.5	5.4	_10
	-0.5	-2.6	1.9	-3.5	-2.1	5.2	4.2	4.3	8.4	9.2	_12
	-0.7	8.7	-1.4	0.5	0.5	8.6	5.5	6.4	8.7	9.2	_14
	12.1	13.1	6.5	10.3	4.1	6.6	5.9	8.0	8.8	7.0	_16
	10.8	15.9	7.1	6.9	2.0	4.5	5.6	1.5	6.0	7.8	_18
DEAI	7.3	10.7	-1.6	-2.7	-3.6	9.0	15.6	4.6	1.7	2.7	_10
	1.5	-1.5	-6.4	-4.2	-4.8	4.3	2.8	2.5	3.4	1.3	_12
	-1.2	-2.4	-3.2	-2.6	-4.1	1.2	1.5	4.8	1.8	1.1	_14
	6.5	-1.9	-4.1	-0.1	-12.6	7.2	8.5	4.7	1.4	3.8	_16
	6.8	-1.5	1.1	-6.3	0.6	16.8	7.6	5.5	5.8	1.4	_18
DEAE	-2.8	-12.6	-3.8	-3.7	-3.6	3.5	6.7	1.5	2.0	2.0	_10
	-11.3	-8.6	-0.4	-12.2	4.5	4.6	3.1	1.7	2.4	1.2	_12
	-22.5	-4.0	-2.8	-3.8	-10.2	7.2	1.6	1.6	3.6	2.8	_14
	-7.3	-6.4	-6.3	-5.0	-6.0	2.0	2.6	1.1	4.6	1.8	_16
	-5.9	-6.4	-4.8	-2.1	-25.2	3.6	3.0	1.1	2.0	8.2	_18
	5_	7_	9_	11_	13_	5_	7_	9_	11_	13_	Adaptor

IPTG-LacI-DNA Looped Complexes Averaged

FPV	Average ET for All Trials					ET STDEV					
DIAI	2.7	2.3	NaN	NaN	13.4	5.4	4.8	NaN	NaN	6.9	_10
	6.8	1.7	3.4	NaN	7.4	5.7	3.2	6.7	NaN	6.9	_12
	16.1	15.4	10.3	14.1	15.9	10.3	6.5	5.1	6.5	9.2	_14
	28.5	26.8	28.3	23.3	16.8	7.4	6.3	7.1	8.1	7.5	_16
	12.8	26.4	12.7	17.2	10.0	5.3	6.9	6.6	7.5	6.6	_18
DIAE	NaN	6.0	NaN	8.4	8.2	NaN	5.1	NaN	4.9	9.3	_10
	NaN	NaN	6.2	2.5	8.4	NaN	NaN	5.6	5.0	11.1	_12
	5.9	9.6	4.8	NaN	7.2	5.1	8.7	4.7	NaN	5.9	_14
	10.5	10.4	10.6	NaN	NaN	6.6	6.7	9.5	NaN	NaN	_16
	7.7	12.1	NaN	NaN	6.8	5.4	8.9	NaN	NaN	5.7	_18
DEAI	8.9	13.1	3.1	NaN	NaN	2.5	2.7	3.6	NaN	NaN	_10
	3.6	1.3	-0.1	NaN	NaN	1.9	2.0	4.4	NaN	NaN	_12
	0.7	0.7	1.2	NaN	NaN	1.5	2.2	3.3	NaN	NaN	_14
	6.1	5.0	1.5	NaN	NaN	1.1	5.6	3.9	NaN	NaN	_16
	14.5	4.6	5.5	2.3	NaN	6.2	4.6	4.7	6.2	NaN	_18
DEAE	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	_10
	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	_12
	NaN	NaN	1.1	NaN	NaN	NaN	NaN	3.1	NaN	NaN	_14
	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	_16
	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	_18
	5_	7_	9_	11_	13_	5_	7_	9_	11_	13_	Adaptor

Appendix 1.3.2: FRET Efficiencies calculated for IPTG-LacI•DNA looped complexes. The environmental influences of saturating IPTG affected the FRET efficiency calculated for Donor to a greater extent than Acceptor. This is apparent in the standard deviation values ±IPTG. Some samples were only measured once with IPTG because FRET was not observed for LacI•DNA, and therefore when averaged appear as NaN (Not a Number nomenclature in Matlab).

1.3.3 Bulk FRET Efficiencies for IPTG saturated LacI•DNA Loops in the Presence of Excess LacI

IPTG-LacI-DNA Looped Complexes + xs LacI

FPV	Donor ET					Acceptor ET					
DIAI	-0.6	-6.0	NaN	NaN	12.3	0.6	0.7	NaN	NaN	7.0	_10
	5.4	-5.6	-3.5	NaN	6.0	2.1	0.7	1.0	NaN	3.5	_12
	22.6	0.5	17.2	-6.5	14.8	5.1	5.0	7.6	1.4	4.2	_14
	34.4	17.0	23.6	3.1	18.7	21.3	16.6	19.5	4.6	9.7	_16
	12.6	20.0	5.9	14.4	9.0	8.8	16.5	6.8	7.9	7.1	_18
DIAE	NaN	13.0	NaN	13.6	22.0	NaN	2.4	NaN	1.2	1.5	_10
	NaN	NaN	14.5	0.8	23.5	NaN	NaN	3.7	3.3	3.0	_12
	11.0	22.0	9.9	NaN	14.3	4.3	1.4	5.7	NaN	3.9	_14
	19.1	18.7	24.5	NaN	NaN	4.7	4.1	4.2	NaN	NaN	_16
	13.6	22.9	NaN	NaN	14.4	2.8	4.6	NaN	NaN	3.5	_18
DEAI	16.3	20.5	6.7	NaN	NaN	8.0	11.6	2.3	NaN	NaN	_10
	7.9	2.4	4.7	NaN	NaN	3.3	1.6	0.2	NaN	NaN	_12
	5.0	4.3	3.3	NaN	NaN	0.4	0.6	2.0	NaN	NaN	_14
	10.1	12.1	7.8	NaN	NaN	5.7	3.0	2.8	NaN	NaN	_16
	25.4	10.0	11.2	9.6	NaN	13.9	3.9	2.9	2.4	NaN	_18
DEAE	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	_10
	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	_12
	NaN	NaN	8.0	NaN	NaN	NaN	NaN	1.1	NaN	NaN	_14
	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	_16
	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	_18
	5_	7_	9_	11_	13_	5_	7_	9_	11_	13_	Adaptor

Appendix 1.3.3: FRET Efficiencies calculated for IPTG-LacI•DNA looped complexes with excess LacI. Only one set was attempted, so an average is not available.

1.4 The DNA Sequences for the Construct Landscape Shown Between the EcoRV Sites that Flank the Operators

5C10

ATCTGCAGGTCAGTCTAGGTAATTGTGAGCGCTCACAATTAGATCTTCGT
ACGGATCCGGTTTTTTGCCCGTTTTTTGCCGTTTTTTGCCCGTTTTTTGC
CGTTTTTTGCCCGTTTTTTGCCGTTTTTTGCCCGTTTTTTGCGCTGAACG
CGTCCATCGAAGCTAGCTAATTGTGAGCGCTCACAATTCGTTGTGGTAAA
GCTTTGAT

5C12

ATCTGCAGGTCAGTCTAGGTAATTGTGAGCGCTCACAATTAGATCTTCGT
ACGGATCCGGTTTTTTGCCCGTTTTTTGCCGTTTTTTGCCCGTTTTTTGC
CGTTTTTTGCCCGTTTTTTGCCGTTTTTTGCCCGTTTTTTGCGCTGAACG
CGTCCTAATCGAAGCTAGCTAATTGTGAGCGCTCACAATTCGTTGTGGTA
AAGCTTTGAT

5C14

ATCTGCAGGTCAGTCTAGGTAATTGTGAGCGCTCACAATTAGATCTTCGT
ACGGATCCGGTTTTTTGCCCGTTTTTTGCCGTTTTTTGCCCGTTTTTTGC

CGTTTTTTGCCCGTTTTTTGCCGTTTTTTGCCCGTTTTTTGCGCTGAACG
CGTCCTAGAATCGAAGCTAGCTAATTGTGAGCGCTCACAATTCGTTGTGG
TAAAGCTTTGAT

5C16

ATCTGCAGGTCAGTCTAGGTAATTGTGAGCGCTCACAATTAGATCTTCGT
ACGGATCCGGTTTTTTGCCCGTTTTTTGCCGTTTTTTGCCCGTTTTTTGC
CGTTTTTTGCCCGTTTTTTGCCGTTTTTTGCCCGTTTTTTGCGCTGAACG
CGTCCTAGACGATCGAAGCTAGCTAATTGTGAGCGCTCACAATTCGTTGT
GGTAAAGCTTTGAT

5C18

ATCTGCAGGTCAGTCTAGGTAATTGTGAGCGCTCACAATTAGATCTTCGT
ACGGATCCGGTTTTTTGCCCGTTTTTTGCCGTTTTTTGCCCGTTTTTTGC
CGTTTTTTGCCCGTTTTTTGCCGTTTTTTGCCCGTTTTTTGCGCTGAACG
CGTCCTAGACGCTATCGAAGCTAGCTAATTGTGAGCGCTCACAATTCGTT
GTGGTAAAGCTTTGAT

7C10

ATCTGCAGGTCAGTCTAGGTAATTGTGAGCGCTCACAATTAGATCTCTTC
GTACGGATCCGGTTTTTTGCCCGTTTTTTGCCGTTTTTTGCCCGTTTTTT
GCCGTTTTTTGCCCGTTTTTTGCCGTTTTTTGCCCGTTTTTTGCGCTGAA
CGCGTCCATCGAAGCTAGCTAATTGTGAGCGCTCACAATTCGTTGTGGTA
AAGCTTTGAT

7C12

ATCTGCAGGTCAGTCTAGGTAATTGTGAGCGCTCACAATTAGATCTCTTC
GTACGGATCCGGTTTTTTGCCCGTTTTTTGCCGTTTTTTGCCCGTTTTTT
GCCGTTTTTTGCCCGTTTTTTGCCGTTTTTTGCCCGTTTTTTGCGCTGAA
CGCGTCCTAATCGAAGCTAGCTAATTGTGAGCGCTCACAATTCGTTGTGG
TAAAGCTTTGAT

7C14

ATCTGCAGGTCAGTCTAGGTAATTGTGAGCGCTCACAATTAGATCTCTTC
GTACGGATCCGGTTTTTTGCCCGTTTTTTGCCGTTTTTTGCCCGTTTTTT
GCCGTTTTTTGCCCGTTTTTTGCCGTTTTTTGCCCGTTTTTTGCGCTGAA
CGCGTCCTAGAATCGAAGCTAGCTAATTGTGAGCGCTCACAATTCGTTGT
GGTAAAGCTTTGAT

7C16

ATCTGCAGGTCAGTCTAGGTAATTGTGAGCGCTCACAATTAGATCTCTTC
GTACGGATCCGGTTTTTTGCCCGTTTTTTGCCGTTTTTTGCCCGTTTTTT
GCCGTTTTTTGCCCGTTTTTTGCCGTTTTTTGCCCGTTTTTTGCGCTGAA
CGCGTCCTAGACGATCGAAGCTAGCTAATTGTGAGCGCTCACAATTCGTT
GTGGTAAAGCTTTGAT

7C18

ATCTGCAGGTCAGTCTAGGTAATTGTGAGCGCTCACAATTAGATCTCTTC
GTACGGATCCGGTTTTTTGCCCGTTTTTTGCCGTTTTTTGCCCGTTTTTT
GCCGTTTTTTGCCCGTTTTTTGCCGTTTTTTGCCCGTTTTTTGCGCTGAA
CGCGTCCTAGACGCTATCGAAGCTAGCTAATTGTGAGCGCTCACAATTCG
TTGTGGTAAAGCTTTGAT

9C10

ATCTGCAGGTCAGTCTAGGTAATTGTGAGCGCTCACAATTAGATCTCAAT
TCGTACGGATCCGGTTTTTTGCCCGTTTTTTGCCGTTTTTTGCCCGTTTT
TTGCCGTTTTTTGCCCGTTTTTTGCCGTTTTTTGCCCGTTTTTTGCGCTG
AACGCGTCCATCGAAGCTAGCTAATTGTGAGCGCTCACAATTCGTTGTGG
TAAAGCTTTGAT

9C12

ATCTGCAGGTCAGTCTAGGTAATTGTGAGCGCTCACAATTAGATCTCAAT
TCGTACGGATCCGGTTTTTTGCCCGTTTTTTGCCGTTTTTTGCCCGTTTT
TTGCCGTTTTTTGCCCGTTTTTTGCCGTTTTTTGCCCGTTTTTTGCGCTG
AACGCGTCCTAATCGAAGCTAGCTAATTGTGAGCGCTCACAATTCGTTGT
GGTAAAGCTTTGAT

9C14

ATCTGCAGGTCAGTCTAGGTAATTGTGAGCGCTCACAATTAGATCTCAAT
TCGTACGGATCCGGTTTTTTGCCCGTTTTTTGCCGTTTTTTGCCCGTTTT
TTGCCGTTTTTTGCCCGTTTTTTGCCGTTTTTTGCCCGTTTTTTGCGCTG
AACGCGTCCTAGAATCGAAGCTAGCTAATTGTGAGCGCTCACAATTCGTT
GTGGTAAAGCTTTGAT

9C16

ATCTGCAGGTCAGTCTAGGTAATTGTGAGCGCTCACAATTAGATCTCAAT
TCGTACGGATCCGGTTTTTTGCCCGTTTTTTGCCGTTTTTTGCCCGTTTT
TTGCCGTTTTTTGCCCGTTTTTTGCCGTTTTTTGCCCGTTTTTTGCGCTG
AACGCGTCCTAGACGATCGAAGCTAGCTAATTGTGAGCGCTCACAATTCG
TTGTGGTAAAGCTTTGAT

9C18

ATCTGCAGGTCAGTCTAGGTAATTGTGAGCGCTCACAATTAGATCTCAAT
TCGTACGGATCCGGTTTTTTGCCCGTTTTTTGCCGTTTTTTGCCCGTTTT
TTGCCGTTTTTTGCCCGTTTTTTGCCGTTTTTTGCCCGTTTTTTGCGCTG
AACGCGTCCTAGACGCTATCGAAGCTAGCTAATTGTGAGCGCTCACAATT
CGTTGTGGTAAAGCTTTGAT

11C10

ATCTGCAGGTCAGTCTAGGTAATTGTGAGCGCTCACAATTAGATCTCAGA
TCTCGTACGGATCCGGTTTTTTGCCCGTTTTTTGCCGTTTTTTGCCCGTT
TTTTGCCGTTTTTTGCCCGTTTTTTGCCGTTTTTTGCCCGTTTTTTGCGC
TGAACGCGTCCATCGAAGCTAGCTAATTGTGAGCGCTCACAATTCGTTGT
GGTAAAGCTTTGAT

11C12

ATCTGCAGGTCAGTCTAGGTAATTGTGAGCGCTCACAATTAGATCTCAGA
TCTCGTACGGATCCGGTTTTTTGCCCGTTTTTTGCCGTTTTTTGCCCGTT
TTTTGCCGTTTTTTGCCCGTTTTTTGCCGTTTTTTGCCCGTTTTTTGCGC
TGAACGCGTCCTAATCGAAGCTAGCTAATTGTGAGCGCTCACAATTCGTT
GTGGTAAAGCTTTGAT

11C14

ATCTGCAGGTCAGTCTAGGTAATTGTGAGCGCTCACAATTAGATCTCAGA
TCTCGTACGGATCCGGTTTTTTGCCCGTTTTTTGCCGTTTTTTGCCCGTT
TTTTGCCGTTTTTTGCCCGTTTTTTGCCGTTTTTTGCCCGTTTTTTGCGC
TGAACGCGTCCTAGAATCGAAGCTAGCTAATTGTGAGCGCTCACAATTCG
TTGTGGTAAAGCTTTGAT

11C16

ATCTGCAGGTCAGTCTAGGTAATTGTGAGCGCTCACAATTAGATCTCAGA
TCTCGTACGGATCCGGTTTTTTGCCCGTTTTTTGCCGTTTTTTGCCCGTT
TTTTGCCGTTTTTTGCCCGTTTTTTGCCGTTTTTTGCCCGTTTTTTGCGC
TGAACGCGTCCTAGACGATCGAAGCTAGCTAATTGTGAGCGCTCACAATT
CGTTGTGGTAAAGCTTTGAT

11C18

ATCTGCAGGTCAGTCTAGGTAATTGTGAGCGCTCACAATTAGATCTCAGA
TCTCGTACGGATCCGGTTTTTTGCCCGTTTTTTGCCGTTTTTTGCCCGTT
TTTTGCCGTTTTTTGCCCGTTTTTTGCCGTTTTTTGCCCGTTTTTTGCGC

TGAACGCGTCCTAGACGCTATCGAAGCTAGCTAATTGTGAGCGCTCACAA
TTCGTTGTGGTAAAGCTTTGAT

13C10

ATCTGCAGGTCAGTCTAGGTAATTGTGAGCGCTCACAATTAGATCTCAGA
TCTCGTCGACGGATCCGGTTTTTTGCCCGTTTTTTGCCGTTTTTTGCCCG
TTTTTTGCCGTTTTTTGCCCGTTTTTTGCCGTTTTTTGCCCGTTTTTTGC
GCTGAACGCGTCCATCGAAGCTAGCTAATTGTGAGCGCTCACAATTCGTT
GTGGTAAAGCTTTGAT

13C12

ATCTGCAGGTCAGTCTAGGTAATTGTGAGCGCTCACAATTAGATCTCAGA
TCTCGTCGACGGATCCGGTTTTTTGCCCGTTTTTTGCCGTTTTTTGCCCG
TTTTTTGCCGTTTTTTGCCCGTTTTTTGCCGTTTTTTGCCCGTTTTTTGC
GCTGAACGCGTCCTAATCGAAGCTAGCTAATTGTGAGCGCTCACAATTCG
TTGTGGTAAAGCTTTGAT

13C14

ATCTGCAGGTCAGTCTAGGTAATTGTGAGCGCTCACAATTAGATCTCAGA
TCTCGTCGACGGATCCGGTTTTTTGCCCGTTTTTTGCCGTTTTTTGCCCG
TTTTTTGCCGTTTTTTGCCCGTTTTTTGCCGTTTTTTGCCCGTTTTTTGC
GCTGAACGCGTCCTAGAATCGAAGCTAGCTAATTGTGAGCGCTCACAATT
CGTTGTGGTAAAGCTTTGAT

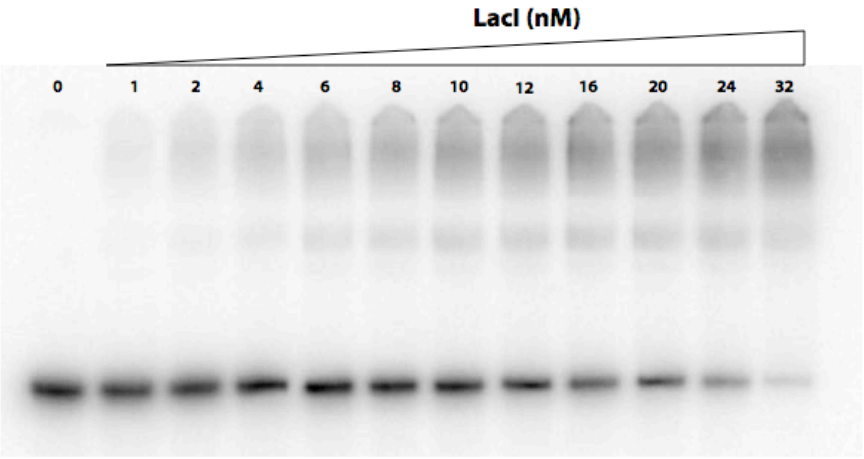
13C16

ATCTGCAGGTCAGTCTAGGTAATTGTGAGCGCTCACAATTAGATCTCAGA
TCTCGTCGACGGATCCGGTTTTTTGCCCGTTTTTTGCCGTTTTTTGCCCG
TTTTTTGCCGTTTTTTGCCCGTTTTTTGCCGTTTTTTGCCCGTTTTTTGC
GCTGAACGCGTCCTAGACGATCGAAGCTAGCTAATTGTGAGCGCTCACAA
TTCGTTGTGGTAAAGCTTTGAT

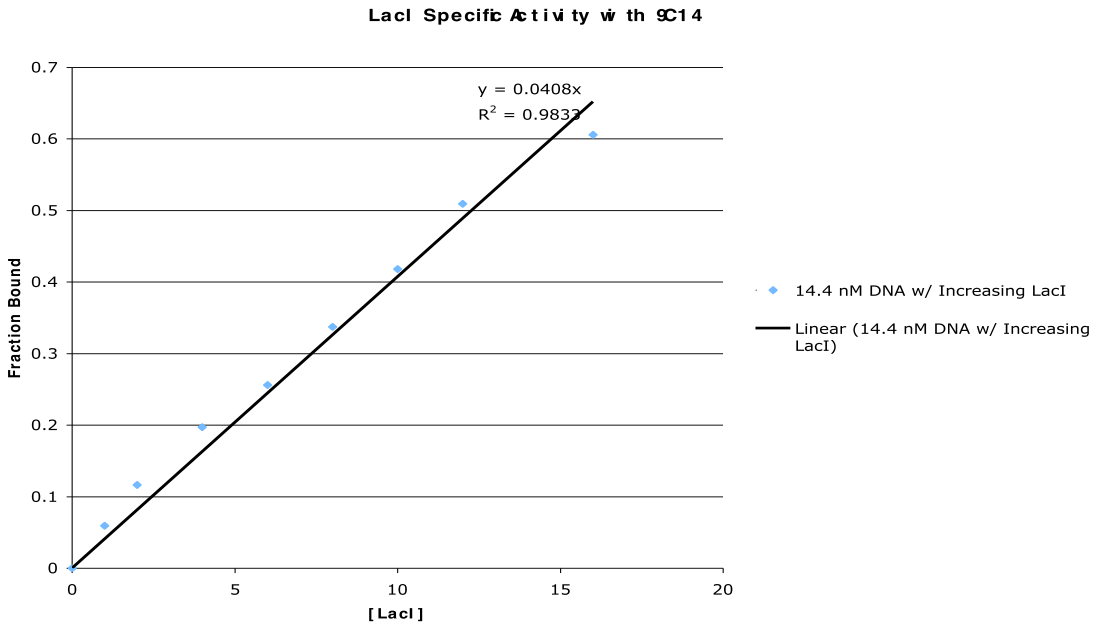
13C18

ATCTGCAGGTCAGTCTAGGTAATTGTGAGCGCTCACAATTAGATCTCAGA
TCTCGTCGACGGATCCGGTTTTTTGCCCGTTTTTTGCCGTTTTTTGCCCG
TTTTTTGCCGTTTTTTGCCCGTTTTTTGCCGTTTTTTGCCCGTTTTTTGC
GCTGAACGCGTCCTAGACGCTATCGAAGCTAGCTAATTGTGAGCGCTCAC
AATTCGTTGTGGTAAAGCTTTGAT

1.5 Determining the Specific Activity of LacI Using an EMSA with Construct 9C14



Appendix 1.5.1: Titrating levels of LacI were added to a constant radiolabeled DNA concentration of 14.4 nM ([DNA] based on the A_{260} and verified through specific activity comparison for cpm). LacI and 9C14 were incubated together for 15 min. at room temp. in LacI buffer and 8 μL loaded onto a 7.5% (75:1 acrylamide:bis-acrylamide) poly-acrylamide gel in TBE buffer and ran for 2 hours at 400 V and 16°C. The gel was then dried and image transferred to a phosphor screen overnight.



Appendix 1.5.2: the specific activity was calculated from the gel shown in figure 1 and graphed based on the DNA fraction bound relative to [LacI]. Only the linear portion of the graph is shown. From the slope of the line the% of active LacI of the 25 µM (A_{280}) was found to be around 65% or 15 µM of active LacI tetraner.

Appendix 2 Methods Further Detailed.

2.1 Fluorescently Labeling Amino-Modified DNA and PCR Amplification of a Labeled Construct

Day One

Preparation of DNA for Labeling *(early morning)*

- Spin down the purchased DNA and resuspend in dH_2O to a volume appropriate for a final concentration of 5 mM
  - It may be necessary to EtOH precipitate the purchased oligos immediately and resuspend in TE buffer (50 mM Tris, 1 mM EDTA, pH 8)
    - This would remove the high salts usually present after a purification, however, this could also chemically effect the amino groups during the precipitation stage in the presence of sodium acetate
  - A typical yield for purchased primers (IDT at least) allows us to perform almost 10 labeling experiments per individual primer order, based on the

concentration that has given the best yield below

Fluorescently Labeling an Amino Modified DNA *(morning)*

- Prepare 5 nmol (1 uL of 5 mM) of stock primer for labeling in 7 µL (resuspend in 8 µL if the DNA is totally dehrydrated) of labeling buffer [0.3 M sodium bicarbonate (250 mg in 10 mL), made fresh for each use and filtered through a 0.2 µm syringe filter]
    - It is important not to dilute the sodium bicarbonate, if you have a large DNA volume you may want to dry your DNA down and resuspend in the labeling buffer
        - Be mindful of the salts and buffers present in the dehydrated DNA sample and react accordingly
- Fully resuspend the dehydrated Fluorescent Dye in 2 µL DMSO (High Quality) then add the 8 µL of 5 nmol primers and mix thoroughly and spin down
- Incubate the samples (protected from light) at 37^{o} C for 3 hrs with mixing every 45 min through use of a sonicating bath for 10 sec intervals.
    - The sonicating bath offers the convenience of vortexing and spinning samples back down, but all in one step
- After the reaction, add 52 µL of labeling buffer (from above) and 3 µL of 4M Hydroxylamine to the labeling mixture and let sit at room temperature for 15 min in the dark with occasional mixing
    - The hydroxylamine quenches the reaction
- Separate the dye from the primers through a BioRad P6 microspin column (or a G-25) column for 4 min at 1.0 rcf (which is really 1000 rcf when scaled)
    - Prep the column beforehand following the manufacturer's protocol
        - Don't over dry the column by spinning to fast or too long, but also make sure that there is not too much buffer that it dilutes your sample
    - a subsequent 50 µL wash with dH_2O on the column increases the primer yield, but also slightly increases the amount of free dye eluted off, but this will be separated in the subsequent denaturing gel purification

Purifying the fluorescently labeled primer *(early afternoon)*

- Then add 20 µL 1M NaCl, 2 µL 1M MgCl2, and 108 uL ddH2O to sample (all filtered as well), mix and then add 400 µl of 100% EtOH and perform an EtOH precipitation
    - The addition of NaCl and $MgCl_2$ in the bicarbonate solution is the moral equivalent of the first step of an ethanol precipitation
    - If there is a requirement to take a break at this point, then allow the DNA to sit overnight in the first stage of the ethanol precipitation (frozen at -80^{o}C).
        - However, it is recommended that this freezing step last no longer than

an hour before continuing onto the rest of the precipitation since degradation always increases with time.

- Resuspend dried samples in 12 µL denaturing sample buffer (1XTE, 8 M Urea, ~15% glycerol) and heat at 90° C for 10 minutes
    - This denatures the primers before being loaded onto a gel for cleaner separation and to reduce any dsDNA to ssDNA
- Prepare a denaturing gel {12% PAGE (40:1 acrylamide:bis-acrylamide), 8 M Urea with wide spacers and large wells for best resolution) and prerun the gel for 30 minutes at 55 watts prior to loading the warmed sample and then run loaded for a minimum of 2 hrs.
    - This allows the gel to equilibrate to a higher temp required for denaturing gels
- Excise the most intense band and gel extract following a modified freeze/crush/ soak method
    - The band is usually visible by eye in the gel as red or blue, but should be verified by looking at the appropriate wavelength on the STORM depending on the fluorophore used

Day Two *(early morning)*

- Finish the gel extraction by performing a phenol/chloroform extraction and EtOH precipitation
- Resuspend the primers in at most 21 µL of TE (or in dH_2O if high salts are present) and use a 1 µL aliquot diluted appropriately for determining yield and labeling efficiencies on a UV-Vis scanning from 200-800 nm
    - A 100 fold dilution is usually appropriate base on previous yields and the blank should contain the same ingredients as the sample (minus the DNA)
- Adjust the [labeled-primer] to 50 µM for ease of use for PCR
    - Being mindful of the change in TE concentration if dehydrating and possibly performing an ethanol precipitation as an alternative.

PCR Amplification with fluorescently labeled primers *(late morning-early afternoon)*

- Mix together 40 pg of template (1 µL of 2 ng/µL), 0.5-1.0 µM of each labeled primer (1 µL of 50 µM but primers should be stoicheometrically equivalent), 250 µM of dNTP's (1 µL of 12.5 mM), and 1X HF buffer (NEB, 10 µL of 5X HF), and 35 µL of dH_2O. Lastly, add 2 units of Phusion Polymerase (NEB, 1 µL of 100 units/µL) to bring the total volume to 50 µL.
- There are two different methods for the thermocycler depending on the primers used for the constructs and the fluorophore position variants
    - For anything with an internally labeled fluorophore use the amplification settings described by Mehta and Edelmen
    - The double externally labeled (D_E and A_E primers) require a thermocycler

setting of: 95°C for 1 min., 63°C for 30 sec., 72°C for 1 min, and repeated 35 times, with an initial 95°C melt for 3 minutes.

Purification of Fluorescently labeled PCR Products *(mid afternoon)*

- During PCR prepare a 7.5% PAGE (40:1 acrylamide:bis-acrylamide) in a temperature controlled box with wide spacers and wide wells for best resolution.
    - Prerun the gel for at least 30 minutes at 400 V in TBE to equilibrate the salts or until a constant current is reached
- Prepare PCR product for gel by diluting to 1X loading dye (10 µL of 6X to 50 µL), then load gel and allow to run for 2.5 hrs at 400 V
    - You can reduce the volume before loading to increase the resolution, but I find that you are also increasing the amount of DNA stuck in the well without first phenol/chloroform extracting and ethanol precipitating
        - I usually just load all 60 µL and have had successful separation of single bands, but you cannot shorten less than a two hour run time
- Visualize the products on a STORM exciting at 635 and reading at 650
    - Making sure the glass plate is facing up when set down on the STORM
    - Print the image as an actual size and remember that this image is the reverse image of the gel when placed underneath the glass and needs to be placed over an illuminator to determine where bands are located through the paper
        - It may be necessary to place a pipette tip into the stock acceptor labeled primer and place into the gel for visual orientation
        - Alternatively you can set up the lanes in the gel so that it is obvious which band is which, for example 4 samples can be loaded with three consecutively, a space, and load the last lane
- Excise the bands and gel purify the PCR product as previously described.

Day Three *(morning)*

- Finish the gel extraction by performing a phenol/chloroform extraction and EtOH precipitation
- After dehydrating the sample, resuspend in 100 µL of TE buffer and scan entire sample from 220 to 340 nm to determine the DNA concentration at 260 nm.
    - Make sure that you are using the same TE as a blank
- Place sample in -20°C for storage and label concentration and construct name on the tube along with the final date of processing
    - It may be easier to adjust the final product to a concentration conducive for future experiments
        - Using a smaller amount of stocks and resuspending in the experimental buffer allows for almost negligible contributions of stock solution upon the final experimental buffer

General Notes:

- It is often better to do many samples at once so that they are handled identically and systematically
    - This allows us to immediately identify any problems with the purchased DNA/Fluorescent dyes and have other results to compare/contrast the experimental results to
        - If all fail, it could be due to technique, if only one fails then it is most likely not technique but rather a problem with the DNA or dyes
        - Be observant of the entire process and make notes of any peculiarities during the process that may show up in the final results
- Timeliness is key in having a successful labeling reaction and subsequent amplification of product through PCR
    - Degradation increases with time, especially when exposed to non-ideal buffers for an extended period
    - Avoid taking breaks when at all possible! Just commit a couple days of hard focused work and then enjoy your weekend.
    - Like most things the dye packs and the amino-modified DNA start going bad the day they are made, so use them at their peak and be prepared to work experimentally hard thereafter
    - The many breaks during long procedures allows ample amount of time to prepare the next stage before the samples are finished
- Buffers need to be filtered for all uses
    - If a buffer is required to be stored it can only go through one freeze thaw cycle and should never be used again at any point
    - This remove contaminates and limits the amount of premature precipitation of product that may occur
    - They need to be made fresh more often than not
        - Make sure you are measuring amounts confidently, for example 25 mg in 1 mL would be less wasteful than 250 mg in 10 mL, but the error will also be much larger
- When in doubt about your buffer and its compatibility with the next step in the process, EtOH precipitate then resuspend in the appropriate buffer to move forward

Appendix 3 Matlab Scripts and Functions.

```
STEP ONE
  Setting up your workspace variables
The FRET spectra saved as csv files must first be loaded into your matlab workspace using the import
function in matlab
  make sure to check the box that allows you to load the file with column
  names the file for reference emission spectra should also be loaded at this
  point
  the naming scheme for files should follow those specified by the
```

```
   etmatrix caller function

STEP TWO
   Normalization of dyes relative to each other to solve for fA and fD

Load the dye_exc_area.m function into your matlab path along with the
scripts: dye_pairs_normalization.1.0.m and dye_ratios.m
   allows you to evaluate your entire workspace using the defined nomenclature for variables.
   The function sums the area for the direct excitation of Donor and/or Acceptor
   for each construct's four FPV combinations separately
   After running dye_pairs_normalization.1.0, run the script dye_ratios.m
   this creates the workspace variables to be subsequently used in the following
   functions
run dye_pairs_normalization;

STEP THREE
   Calculating the energy transfer for each experimental condition

Load the ETCalc.m and the ETMatrix_universal.m into your path in matlab
   ETMatrix allows you to evaluate your workspace but inputs those
   variables into the ETCalc function as specified by the order in which variables
   are listed
   the function ETCalc determines the energy transfer for each
   experimental setting
   ET_quick_Calc can be used in place of ETCalc if only variables are
   needed and not the printed files
run ETMatrix_Universal;

STEP FOUR
   reshaping and FPV landscape calculations

load the script called ET_FPV_landscape_calculations.m
   this combines the ET data based on construct and FPV into four
   landscape, and further separates them into Donor and Acceptor ET.
   this averages values for ET and shows the diffence landscapes for pre and
   post IPTG
   it prepares the FPV landscape and the maximum ET for each landscape to be used in the
   subsequent population calculations
   at the end it plots 3 sets of 12 landscapes that compare acceptor,
   donor and their average ET for LacI and LacI + IPTG, and the final plot
   shows the averages and differences
run ET_FPV_landscape_calculations

STEP FIVE
   Looping distributions based on ET

load the function etmatrixcalc.m, which is a constrained minimization
function, as well as the script called Population_plot.m, and etm_resid.m
   this function allows us to use the maximum values of ET for each landscape to
   determine the limit of FRET for each loop type in our system
   the script reshapes the outputs of etmatrixcalc and plots them as
   landscapes based on the loop topologies P1, A1, A2, and E
run etmatrixcalc;
run Population_plot;
```

```
function sums = dye_exc_area(donor_exc,acceptor_exc);
%
% This sums up only the excitiation spectra for the donor and the
acceptor in the appropriate window of a 1:201 column corresponding to
550-750 nm (wavelength)
Donor_area = sum(donor_exc((1:90))); % sums the values from 550-640
Acceptor_area = sum(acceptor_exc(91:180)); %sums the values from
641-750
sums = [Donor_area Acceptor_area];
return


% Script to automate running the normalization of dyes relative to one
dye. In this case relative to DI, dye_pairs_normalization
% Run from the command line
% There should be an array called "con_names" defined, with variable
names like 11C10 padded to five characters (i.e. define 9C14 as 09C14).
con_names = ['05C10'; '05C12'; '05C14'; '05C16'; '05C18'; '07C10';
'07C12'; '07C14'; '07C16'; '07C18';'09C10'; '09C12'; '09C14'; '09C16';
'09C18'; '11C10'; '11C12'; '11C14'; '11C16'; '11C18'; '13C10'; '13C12';
'13C14'; '13C16'; '13C18';];
%Prepares the character array for all construct as arguments to place
each set of fluorophores (dyes) together for all constructs in the FPV
family.
numcons = size(con_names,1);
for j = 1:numcons;
construct = con_names(j,:);
evalin('base',['dyesum_DA(j,:) = dye_exc_area(x' construct '_DA_D, x'
construct '_DA_A)']);
evalin('base',['dyesum_DA1(j,:) = dye_exc_area(x' construct '_DA1_D, x'
construct '_D1A_A)']);
evalin('base',['dyesum_D1A(j,:) = dye_exc_area(x' construct '_D1A_D, x'
construct '_DA1_A)']);
evalin('base',['dyesum_D1A1(j,:) = dye_exc_area(x' construct '_D1A1_D,
x' construct '_D1A1_A)']);
end
% This begins the golden ratio of the same fluorescent pairs for each
FPV. That is, each donor is scaled to the acceptor and vice versa.
Since the same primer stock was used for all FPVS, this allows a
normalization of one dye to another for a relative labeling efficiency.
% This sums the columns from above,  providing one value for all 8 FPV
combination such as DI with respect to AI (DI__AI) emission areas.
DI__AI = sum(dyesum_DA);
DI__AE = sum(dyesum_DA1);
DE__AI = sum(dyesum_D1A);
DE__AE = sum(dyesum_D1A1);
%the sum of 50 total dye areas for one FPV regardless of its partner.
DI = DI__AI(:,1) + DI__AE(:,1);
DE = DE__AI(:,1) + DE__AE(:,1);
```

```
AI = DI__AI(:,2) + DE__AI(:,2);
AE = DI__AE(:,2) + DE__AE(:,2);
%This scales the relative concentration (rel_conc) of the 9C14 donor or
acceptor only reference spectra based on fluorescent intensity for the
direct excitation of donor or acceptor. Where AI = A0, and AE = A1 and
similarly for the donor. To run this part of the script requires that
there be reference spectra for each donor and acceptor only
combination, i.e. ref_A0 si the reference spectra for a singly labeled
DNA only sample in the AI position.
relconc_A0_ref = sum(ref_A0(91:180))/(AI/(2*(size(con_names,1))));
relconc_A1_ref = sum(ref_A1(91:180))/(AE/(2*(size(con_names,1))));
relconc_D0_ref = sum(ref_D0(1:90))/(DI/(2*(size(con_names,1))));
relconc_D1_ref = sum(ref_D1(1:90))/(DE/(2*(size(con_names,1))));
%this creates the ratio of DI with respect to DE and AI over AE for
each of the different FPV perspectives (which should agree and be two
duplicate values with different names).
DE_over_DI__AI = (DE/AI)/(DI/AI);
DE_over_DI__AE = (DE/AE)/(DI/AE);
AE_over_AI__DI = (AE/DI)/(AI/DI);
AE_over_AI__DE = (AE/DE)/(AI/DE);
%Now the relative labeling efficiencies for each of the 4 FPVs are
calculated called fAI, fAE, fDI, and fDE.  This sets the relationship
of the known "golden ratio" primers from above and sets the other donor
and acceptor relative to this original ratio.  This could solved
possibly better with a solver, but I looked at the outputs and set AI
as 1 since it had the higher emission relative to AE.
fAI = 1;
fDI = 1;
fAE = fAI*AE_over_AI__DI;
fDE = fDI*DE_over_DI__AI;
```

3.2 Bulk FRET Spectral Decomposition

```
function Energy_Transfer = ETCalc
(ET_spec,DNA_only,fA,fD,Donor_ref,relconc_Donor_ref,Acc_ref,relconc_Acc
_ref,lambdas)
%This function calculates energy transfer efficiency based for FRET,
returning both the ET calculated from donor quenching and ET calculated
from acceptor enhancement. The input variables are as follows:
% ET_spec is the emission spectrum for the experimental mix for which
we are calculating ET
% fA and  fD are the labeling efficiencies for donor and acceptor
% DNA_only is the control non-transfer spectrum for a double-labeled
sample
% Donor_ref is a set of two reference spectra for a known concentration
donor-only sample. To consider the possibility that LacI or changes in
conditions cause some quenching, we use separate donor reference
spectra for each condition. The first Donor_ref column is DNA-only, the
second column is DNA+LacI (more generally, DNA under transfer expt
conditions), [implement later: 3rd is DNA + IPTG, 4th is DNA+IPTG+LacI]
```

```
% The routine will check the dimensionality and if there is only one
column it will use that spectrum for all calculations.
% Acc_ref is a reference spectrum for a known concentration acceptor-
only sample. The known concentration aspect allows us to calculate
relative quantum yields from the data.
% All spectra are assumed to be background-corrected (and all values
non-negative??). Furthermore we assume "corrected spectra," i.e.
corrected for variations in source intensity and PMT sensitivity. (ie.
we can compare areas).
% All spectra are assumed to refer to the same set of wavelengths
lambdas is the vector of wavelengths used.
% conc_D_et, conc_D_nx, conc_A_et, and conc_A_nx are the estimated true
% concentrations of the reference donor-only and acceptor-only
molecules, probably traceable to a UV absorbance measurement for each
sample, with the concentrations of individual samples being estimated
from the observed fluorescence intensity relative to the average
fluorescence intensity over the whole set of samples, where the average
is assumed to be at ~ 2 nM.
% Get reference spectra, just duplicate if there's only one.
[nrows_D, ncols_D] = size(Donor_ref)
[nrows_A, ncols_A] = size(Acc_ref)
%If there's only one column, use it as the reference spectrum for both
conditions
if (ncols_D == 1)
    Donor_ref = [Donor_ref Donor_ref];
end
if (ncols_A == 1)
    Acc_ref = [Acc_ref Acc_ref];
end
% Reassort the reference spectra to make references for no-xfer and
+ET.
refspec_nx = [ Donor_ref(:,1) Acc_ref(:,1) ];
refspec_et = [ Donor_ref(:,2) Acc_ref(:,2) ];
% coeffs_nx and coeffs_et give the best-fit linear combination of donor
and acceptor reference spectra that make up the experimental spectra.
The backslash operator is the meat of the routine.
coeffs_nx = refspec_nx\DNA_only;
coeffs_et = refspec_et\ET_spec;
% Plot the best-fit linear combinations (fitspec_x) and the residuals
(resid_x)
fitspec_nx = refspec_nx*coeffs_nx;
fitspec_et = refspec_et*coeffs_et;
resid_nx = DNA_only - fitspec_nx;
resid_et = ET_spec - fitspec_et;
% Use the reference spectra at known concentrations to estimate
relative quantum yields under our conditions and also to normalize the
coefficients. Normalize everything to 2 nM just for concreteness; in
other words, we assume that the reference spectra were all obtained at
2 nM. The concentrations of the single-labeled reference samples is
specified relative to the average of all of the double-labeled test
samples,looking only at DNA-only samples. The lines below correct for
```

```
the possibility that the reference sample was at a slightly different
concentration than the test sample.
coeffs_nx(1) = coeffs_nx(1) * relconc_Donor_ref;
coeffs_nx(2) = coeffs_nx(2) * relconc_Acc_ref;
coeffs_et(1) = coeffs_et(1) * relconc_Donor_ref;
coeffs_et(2) = coeffs_et(2) * relconc_Acc_ref;
% Relative quantum yields?
% Use the second column of the reference spectra arrays because we want
the relative quantum yields in the presence of LacI
% Assume corrected spectra so we can directly compare areas under the
curves. Assume extinction coeff at lambdamax is correct from
Invitrogen.
% Calculate ext. coeff at our excitation lambda based on our own UV
spectra. For Alexa 555 - Alexa 647 excited at 514 nm and 600 nm
respectively: For Alexa 555 epsil max = 150000 published Epsil 514/
Epsil 555 = 0.680 from our measured absorbance spectrum of labeled
primer (vs. about 0.4 from published spec?). For Alexa 647 epsil max =
239000 publsished Epsil 600/ Epsil 647 = 0.437 from our measured
absorbance spectrum of labeled primer (vs. about 0.3 from published
spec?)
Epsil_A = 0.437*239000;   % Ext. coeff at 600 nm
Epsil_D = 0.680*150000;   % Ext. coeff at 514 nm
total_intensity_D = sum(Donor_ref(:,2))/relconc_Donor_ref;
total_intensity_A = sum(Acc_ref(:,2))/relconc_Acc_ref;
QD_over_QA = (0.1/0.3)*(fA/fD);
% Calculate energy transfer efficiency
et_donor = (1-coeffs_et(1)/coeffs_nx(1))/fA;
et_acc = (coeffs_et(2)-coeffs_nx(2))/coeffs_nx(1) * QD_over_QA/fA;
%
Energy_Transfer = [et_donor et_acc];
Return


% Script to automate running ET_quick_Calc
% Run from the command line
% This goes through the workspace looking for spectral data that
correspond to the spectra for each construct and FPV.
% There should be an array called "con_names" defined, with variable
names like 11C10 padded to five characters (i.e. define 9C14 as 09C14)
con_names = ['05C10'; '05C12'; '05C14'; '05C16'; '05C18'; '07C10';
'07C12'; '07C14'; '07C16'; '07C18';'09C10'; '09C12'; '09C14'; '09C16';
'09C18'; '11C10'; '11C12'; '11C14'; '11C16'; '11C18'; '13C10'; '13C12';
'13C14'; '13C16'; '13C18';];
%Prepares the character array for all construct as arguments
ETMnumcons = size(con_names,1);
for j = 1:ETMnumcons;
ETMcon = con_names(j,:);
%
if ~exist(['x' ETMcon '_D1A1_D'],'var');%this searches the workspace
for the variable (i e x05C10_DA_LACI) and if it does not exist then it
assigns NaN as a 201 x 1 array, and similar for each following if–else
logic.
```

```
else
    assignin('base',['x' ETMcon '_D1A1_D'],NaN(201,1));
end
if -exist(['x' ETMcon '_DA_LACI'],'var');
else
    assignin('base',['x' ETMcon '_DA_LACI'],NaN(201,1));
end
evalin('base',['ET_' ETMcon '_DA_LACI = ET_quick_Calc(x' ETMcon
'_DA_LACI, x' ETMcon '_DA_D,' num2str(fAI) ',' num2str(fDI) ',[ref_D0
ref_D0_lac],' num2str(relconc_D0_ref) ',[ref_A0 ref_A0_lac],' num2str
(relconc_A0_ref) ',Wavelength)']);
if -exist(['x' ETMcon '_DA_LACI_IPTG'],'var');
    else
    assignin('base',['x' ETMcon '_DA_LACI_IPTG'],NaN(201,1));
end
evalin('base',['ET_' ETMcon '_DA_LACI_IPTG = ET_quick_Calc(x', ETMcon,
'_DA_LACI_IPTG, x', ETMcon,'_DA_D,' num2str(fAI) ',' num2str(fDI) ',
[ref_D0 ref_D0_lac_iptg],',num2str(relconc_D0_ref),',[ref_A0
ref_A0_lac_iptg],',num2str(relconc_A0_ref),',Wavelength)'])
if -exist(['x' ETMcon '_DA_LACI_IPTG_LACI'],'var');
    else
    assignin('base',['x' ETMcon '_DA_LACI_IPTG_LACI'],NaN(201,1));
end
evalin('base',['ET_' ETMcon '_DA_LACI_IPTG_LACI = ET_quick_Calc(x',
ETMcon, '_DA_LACI_IPTG_LACI, x', ETMcon,'_DA_D,' num2str(fAI) ','
num2str(fDI) ',[ref_D0 ref_D0_lac_iptg],',num2str(relconc_D0_ref),',
[ref_A0 ref_A0_lac_iptg],',num2str(relconc_A0_ref),',Wavelength)'])
if -exist(['x' ETMcon '_DA1_LACI'],'var');
    assignin('base',['x' ETMcon '_DA1_LACI'],NaN(201,1));
end
evalin('base',['ET_' ETMcon '_DA1_LACI = ET_quick_Calc(x', ETMcon,
'_DA1_LACI, x', ETMcon,'_DA1_D,' num2str(fAE) ',' num2str(fDI) ',
[ref_D0 ref_D0_lac],',num2str(relconc_D0_ref),',[ref_A1
ref_A1_lac],',num2str(relconc_A1_ref),',Wavelength)'])
if -exist(['x' ETMcon '_DA1_LACI_IPTG'],'var');
    else
    assignin('base',['x' ETMcon '_DA1_LACI_IPTG'],NaN(201,1));
end
evalin('base',['ET_' ETMcon '_DA1_LACI_IPTG = ET_quick_Calc(x', ETMcon,
'_DA1_LACI_IPTG, x', ETMcon,'_DA1_D,' num2str(fAE) ',' num2str(fDI) ',
[ref_D0 ref_D0_lac_iptg],',num2str(relconc_D0_ref),',[ref_A1
ref_A1_lac_iptg],',num2str(relconc_A1_ref),',Wavelength)'])
if -exist(['x' ETMcon '_DA1_LACI_IPTG_LACI'],'var');
    else
    assignin('base',['x' ETMcon '_DA1_LACI_IPTG_LACI'],NaN(201,1));
end
evalin('base',['ET_' ETMcon '_DA1_LACI_IPTG_LACI = ET_quick_Calc(x',
ETMcon, '_DA1_LACI_IPTG_LACI, x', ETMcon,'_DA1_D,' num2str(fAE) ','
num2str(fDI) ',[ref_D0 ref_D0_lac_iptg],',num2str(relconc_D0_ref),',
[ref_A1 ref_A1_lac_iptg],',num2str(relconc_A1_ref),',Wavelength)'])
if -exist(['x' ETMcon '_D1A_LACI'],'var');
    else
```

```
    assignin('base',['x' ETMcon '_D1A_LACI'],NaN(201,1));
end
evalin('base',['ET_' ETMcon '_D1A_LACI = ET_quick_Calc(x', ETMcon,
'_D1A_LACI, x', ETMcon,'_D1A_D,' num2str(fAI) ',' num2str(fDE) ',
[ref_D1 ref_D1_lac],',num2str(relconc_D1_ref),',[ref_A0
ref_A0_lac],',num2str(relconc_A0_ref),',Wavelength)'])
if -exist(['x' ETMcon '_D1A_LACI_IPTG'],'var');
    assignin('base',['x' ETMcon '_D1A_LACI_IPTG'],NaN(201,1));
end
evalin('base',['ET_' ETMcon '_D1A_LACI_IPTG = ET_quick_Calc(x', ETMcon,
'_D1A_LACI_IPTG, x', ETMcon,'_D1A_D,' num2str(fAI) ',' num2str(fDE) ',
[ref_D1 ref_D1_lac_iptg],',num2str(relconc_D1_ref),',[ref_A0
ref_A0_lac_iptg],',num2str(relconc_A0_ref),',Wavelength)'])
if -exist(['x' ETMcon '_D1A_LACI_IPTG_LACI'],'var');
    else
    assignin('base',['x' ETMcon '_D1A_LACI_IPTG_LACI'],NaN(201,1));
end
evalin('base',['ET_' ETMcon '_D1A_LACI_IPTG_LACI = ET_quick_Calc(x',
ETMcon, '_D1A_LACI_IPTG_LACI, x', ETMcon,'_D1A_D,' num2str(fAI) ','
num2str(fDE) ',[ref_D1 ref_D1_lac_iptg],',num2str(relconc_D1_ref),',
[ref_A0 ref_A0_lac_iptg],',num2str(relconc_A0_ref),',Wavelength)'])
if -exist(['x' ETMcon '_D1A1_LACI'],'var');
    else
    assignin('base',['x' ETMcon '_D1A1_LACI'],NaN(201,1));
end
evalin('base',['ET_' ETMcon '_D1A1_LACI = ET_quick_Calc(x', ETMcon,
'_D1A1_LACI, x', ETMcon,'_D1A1_D,' num2str(fAE) ',' num2str(fDE) ',
[ref_D1 ref_D1_lac],',num2str(relconc_D1_ref),',[ref_A1
ref_A1_lac],',num2str(relconc_A1_ref),',Wavelength)'])
if -exist(['x' ETMcon '_D1A1_LACI_IPTG'],'var');
    else
    assignin('base',['x' ETMcon '_D1A1_LACI_IPTG'],NaN(201,1));
end
evalin('base',['ET_' ETMcon '_D1A1_LACI_IPTG = ET_quick_Calc(x',
ETMcon, '_D1A1_LACI_IPTG, x', ETMcon,'_D1A1_D,' num2str(fAE) ','
num2str(fDE) ',[ref_D1 ref_D1_lac_iptg],',num2str(relconc_D1_ref),',
[ref_A1 ref_A1_lac_iptg],',num2str(relconc_A1_ref),',Wavelength)'])
if -exist(['x' ETMcon '_D1A1_LACI_IPTG_LACI'],'var');
    else
    assignin('base',['x' ETMcon '_D1A1_LACI_IPTG_LACI'],NaN(201,1));
end
evalin('base',['ET_' ETMcon '_D1A1_LACI_IPTG_LACI = ET_quick_Calc(x',
ETMcon, '_D1A1_LACI_IPTG_LACI, x', ETMcon,'_D1A1_D,' num2str(fAE) ','
num2str(fDE) ',[ref_D1 ref_D1_lac_iptg],',num2str(relconc_D1_ref),',
[ref_A1 ref_A1_lac_iptg],',num2str(relconc_A1_ref),',Wavelength)'])
end
```

3.3 Loop Population Distribution Calculations

```
% solve for fraction of each loop type subject to constraints
```

```
m = 4 ; % number of loop types to consider, this number may change
depending on the interpretation. For the inclusion of two FPV
corsstalking landscape theis was adjusted to 6.
k = 3 ;% number of landscapes
j = 25 ;% number of constructs
% ET matrix is a given, the efficiency for each loop type in each
landscape. Each row is ET for each type given a particular label
position. D0A0 = first row, D0A1 = second row , D1A0 = third row. etmat
is k x m row x column. For example the 3 x 4 below could also be a 3 x
6 for an additional two populations:
etmat = [ D0A0_max_ET 0 0 0 ; 0 D0A1_max_ET 0 0 ; 0 0 D1A0_max_ET 0 ];
% need to copy it into block diagonal form
bd_etmat = zeros(j*k,j*m);
norm_mat = bd_etmat;
% Also need to create the normalization matrix
for i = 0:j-1
    bd_etmat(i*k+1:i*k+k , i*m+1:i*m+m) = etmat;
    norm_mat(i*k+1:i*k+k , i*m+1:i*m+m) = ones(size(etmat));
end
% The fvec is the vector of fractions of each type. It should be a
column vector with m*j elements. Initial estimate is a uniform
distribution.
fvec = ones(m*j,1)./m;
% The et_obs is the ET's for the j*k experiments, a column vector
grouped by construct (expvec(1) = ET of construct 1, loop type 1,
expvec(3) = ET of construct 1, loop type 3)
norm_vec = ones(size(et_obs));
% Define the function that calculates the difference between the
predicted ET and the observed, in terms of separately provided
et_resid = @(fvec)etm_resid(fvec,bd_etmat,et_obs);
%etm_resid (separate file to define the function) is just
%resid_vec = bd_etmat*fvec - expvec
%et_resid = sum(sum(resid_vec.^2))
options = optimset('MaxFunEvals',100000);
[fans,residual] = fmincon(et_resid,fvec,[],[],norm_mat,norm_vec,zeros
(j*m,1),ones(j*m,1),[],options)


function etm_resid = etm_resid( fvec, bd_etmat, expvec)
%et_resid Calculates residual for A*x - b for fmincon
%   fvec is the design variable, a flist of the fraction of each loop
type
%   bd_etmat is the energy transfer matrix copied into block diagonal
form
%   expvec is the observed energy transfer efficiency
resid_vec = bd_etmat*fvec - expvec
etm_resid = sum(sum(resid_vec.^2));
end
```

```
function Energy_Transfer = ET_calc_db_samp
(D_con_ecx_nx,A_con_exc_nx,con_exc_nx,D_con_exc_et,A_con_exc_et,con_exc
_et,fA,fD)
%This function works similar to ET_calc function used to determine the
energy transfer from acceptor and donor spectra. However, this applied
to microplates with a singular value for the median intensity of each
well, (average values and integrated areas had a tendency to
overcompensate for any imperfections in the wells).
% the calculated quantum yields of the donor and acceptor were used as
described in the function ET_calc.
QD_over_QA = (0.1/0.3)*(fA/fD);
% Since energy transfer (et) must be calculated relative to the no
transfer (nx) controls, in this case the doubly labeled DNA only
samples, the concentrations need to be scaled to each other to
normalize the coefficients. Therefore, the direct excitation of
acceptor and the corresponding emission were used to adjust the
experimental value realtive to the nx control. A_con_exc refers to the
direct excitation of the acceptor (633 nm) and monitoring the emission
(670 nm) for either nx or et. D_con_exc is the direct excitation of the
donor (532 nm) and monitoring the donor emission (580). Lastly,
con_exc_ is the excitation of the donor (532 nm) while monitoring the
emission of the acceptor (670 nm).
coeff_nx_D = D_con_ecx_nx.*(A_con_exc_et./A_con_exc_nx);
coeff_nx_A = con_exc_nx.*(A_con_exc_et./A_con_exc_nx);
coeff_et_D = D_con_exc_et.*(A_con_exc_et./A_con_exc_nx);
coeff_et_A = con_exc_et.*(A_con_exc_et./A_con_exc_nx);
% this is the final calculation for energy transfer for donor quenching
and
% acceptor enhancement
et_donor = (1-coeff_et_D./coeff_nx_D)./fA;
et_acc = (coeff_et_A-coeff_nx_A)./coeff_nx_D.*QD_over_QA./fA;
Energy_Transfer = [et_donor et_acc];


% This script uses the ET_calc_db_samp function and evaluates the work
space for any variables that could be applied to this function.
however, this requires you to organize the variable names in your
workspace in a way that reflects how you call for all the variables
below.
% This is a list of experimental variables (var_names) ran padded to 2
character spaces. When placed together, such as "LC" which means that
LacI was added 2 hrs after the competitor DNA was added. "00" would be
only the labeled DNA and two additions of buffer. "Q" corresponds to
the addition of chloriquine
var_names = ['00'; 'L0'; 'LL'; 'LC'; 'CL'; 'Q0'; 'LQ';];
% This is a list of the named constructs (con_names) used in the
microplate settings padded to 5 characters.
con_names = ['05C10'; '05C14'; '05C16'; '05C18'; '07C10'; '07C12';
'07C14'; '07C16'; '07C18';'09C10'; '09C12'; '09C14'; '09C16'; '09C18';
```

```
'11C10'; '11C12'; '11C14'; '11C16'; '11C18'; '13C10'; '13C12'; '13C14';
'13C16'; '13C18';];
% Here is the list of fluorophore postition variants (FPV_names) padded
to 4 character spaces.
FPV_names = ['DIAI'; 'DIAE'; 'DEAI';];
% LacI was shown to enhance the donor signal upon edition for donor-
only 9C14 by 1.05, the acceptor-only sample showed no effects and is
equal to 1.  However, this produces a slight bias to the loop shape of
9C14, therefore this generically applied to represent all FPVs.
%This Prepares the character array for all construct as arguements
ETMnumvars = size(var_names,1);
ETMnumcons = size(con_names,1);
ETMnumFPVs = size(FPV_names,1);
for i = 1:ETMnumvars;
    ETMvar = var_names(i,:);
for j = 1:ETMnumcons;
ETMcon = con_names(j,:);
for k = 1:ETMnumFPVs;
ETMFPV = FPV_names(k,:);
if -exist(['x' ETMcon '_' ETMFPV '_' ETMvar ''],'var');
else
    assignin('base',['x' ETMcon '_' ETMFPV '_' ETMvar ''],NaN(size
(ALEXA_555)));
end
if -exist(['x' ETMcon '_' ETMFPV '_00_D'],'var');
else
    assignin('base',['x' ETMcon '_' ETMFPV '_00_D'],NaN(size
(ALEXA_555)));
end
if -exist(['x' ETMcon '_' ETMFPV '_00_A'],'var');
else
    assignin('base',['x' ETMcon '_' ETMFPV '_00_A'],NaN(size
(ALEXA_555)));
end
if -exist(['x' ETMcon '_' ETMFPV '_' ETMvar ''],'var');
else
    assignin('base',['x' ETMcon '_' ETMFPV ''],NaN(size(ALEXA_555)));
end
evalin('base',['ET_' ETMcon '_' ETMFPV '_' ETMvar ' = ET_calc_db_samp
(x' ETMcon '_' ETMFPV '_00_D,x' ETMcon '_' ETMFPV '_00_A,x' ETMcon '_'
ETMFPV '_00,x' ETMcon '_' ETMFPV '_' ETMvar '_D,x' ETMcon '_' ETMFPV
'_' ETMvar '_A,x' ETMcon '_' ETMFPV '_' ETMvar ',1,1)']);
% This uses the DNA only sample with buffer and chloriquine to
determine the FRET efficiency and not the DNA only sample in buffer as
the script above.
evalin('base',['ET_' ETMcon '_' ETMFPV '_Q0 = ET_calc_db_samp(x' ETMcon
'_' ETMFPV '_Q0_D,x' ETMcon '_' ETMFPV '_Q0_A,x' ETMcon '_' ETMFPV
'_Q0,x' ETMcon '_' ETMFPV '_Q0_D,x' ETMcon '_' ETMFPV '_Q0_A,x' ETMcon
'_' ETMFPV '_Q0,1,1)']);
evalin('base',['ET_' ETMcon '_' ETMFPV '_LQ = ET_calc_db_samp(x' ETMcon
'_' ETMFPV '_Q0_D,x' ETMcon '_' ETMFPV '_Q0_A,x' ETMcon '_' ETMFPV
```

```
'_Q0,x' ETMcon '_' ETMFPV '_LQ_D,x' ETMcon '_' ETMFPV '_LQ_A,x' ETMcon
'_' ETMFPV '_LQ,1,1)']);
end
end
end


%This script turns individual ET values for each construct-FPV-variable
into a landscape that corresponds to each.
var_names = ['00'; 'L0'; 'LL'; 'LC'; 'CL'; 'Q0'; 'LQ';];
% This is a list of the named constructs (con_names) padded to 5
character places.
con_names = ['05C10'; '05C12'; '05C14'; '05C16'; '05C18'; '07C10';
'07C12'; '07C14'; '07C16'; '07C18';'09C10'; '09C12'; '09C14'; '09C16';
'09C18'; '11C10'; '11C12'; '11C14'; '11C16'; '11C18'; '13C10'; '13C12';
'13C14'; '13C16'; '13C18';];
% Here is the list of fluorophore postition variants (FPV_names) padded
to 4 character spaces.
FPV_names = ['DIAI'; 'DIAE'; 'DEAI';];
% LacI was shown to enhance the donor signal upon edition for donor
only 9C14 by 1.05, the acceptor-only sample showed no effects and is
equal to 1.  However, this produces a slight bias to the loop shape of
9C14, therefore this generically applied to represent all FPVs.
time_char = ['0015'; '0030'; '0046'; '0062'; '0079'; '0095'; '0112';
'0141'; '0157'; '0173'; '0189'; '0205'; '0221'; '0239'; '0264'; '0285';
'0362'; '0426'; '0484'; '1063'; '1169'; '1256'; '1285'; '1301'; '1360';
'1376'; '1392'; '1408'; '1426'; '1442';];
time_course(1,:) = str2num(time_char);
for i = 1:ETMnumvars;
ETMvar = var_names(i,:);
for j = 1:ETMnumcons;
ETMcon = con_names(j,:);
for k = 1:ETMnumFPVs;
ETMFPV = FPV_names(k,:);
for l = 1:timenum;
time = time_char(l,:);
if -exist(['ET_' ETMcon '_' ETMFPV '_' ETMvar '_D'],'var');
else
assignin('base',['ET_' ETMcon '_' ETMFPV '_' ETMvar '_D'],NaN(size
(ET_05C10_DIAI_00)));
end
if -exist(['ET_' ETMcon '_' ETMFPV '_' ETMvar '_A'],'var');
else
assignin('base',['ET_' ETMcon '_' ETMFPV '_' ETMvar '_A'],NaN(size
(ET_05C10_DIAI_00)));
end
if -exist(['ET_' ETMcon '_' ETMFPV '_' ETMvar ''],'var');
else
assignin('base',['ET_' ETMcon '_' ETMFPV '_' ETMvar ''],NaN(size
(ET_05C10_DIAI_00)));
end
```

```
evalin('base',['ET_' ETMFPV '_' ETMvar '_landscape_' time '(j,:) = ET_'
ETMcon '_' ETMFPV '_' ETMvar '(l,2)']);
end
end
end
end
for i = 1:ETMnumvars;
ETMvar = var_names(i,:);
for j = 1:ETMnumcons;
ETMcon = con_names(j,:);
for k = 1:ETMnumFPVs;
ETMFPV = FPV_names(k,:);
for l = 1:timenum;
time = time_char(l,:);
evalin('base',['ET_' ETMFPV '_' ETMvar '_landscape_' time ' = reshape
(ET_' ETMFPV '_' ETMvar '_landscape_' time ',5,5)']);
end
end
end
end
```

3.4 Exponential Fit to LacI•DNA FRET Decay in Competition Studies.

```
% This is script is how the 1/2 lives were calculated from the
exponential decay of FRET.
% Since The LacI•DNA FRET efficiency decays overtime, the LacI•DNA +
xsDNA competitior needs to normalized to the observed decay rate for
LacI•DNA only sample. This was performed as Final - Initial for
LacI•DNA only samples and then subtracted from LacI•DNA + xsDNA as
below. We want to perform this for all FPVs therefore, we are using the
generic for loop to process each FPV in the workspace.
FPV_names = ['DIAI'; 'DIAE'; 'DEAI';];
ETMnumFPVs = size(FPV_names,1);
time_char = ['0015'; '0030'; '0046'; '0062'; '0079'; '0095'; '0112';
'0141'; '0157'; '0173'; '0189'; '0205'; '0221'; '0239'; '0264'; '0285';
'0362'; '0426'; '0484'; '1063'; '1169'; '1256'; '1285'; '1301'; '1360';
'1376'; '1392'; '1408'; '1426'; '1442';];
time_course(1,:) = str2num(time_char);
time_LC = ((time_course(:,8:21)-125)./60);
for k = 1:ETMnumFPVs;
ETMFPV = FPV_names(k,:);
evalin('base',['norm_vec_ET_' ETMFPV '_L0 = (ET_column_' ETMFPV
'_L0_landscape(:,8:21) - repmat(ET_column_' ETMFPV '_L0_landscape(:,8),
1,14))']);
evalin('base',['norm_vec_ET_' ETMFPV '_LC = (ET_column_' ETMFPV
'_LC_landscape(:,8:21) - norm_vec_ET_' ETMFPV '_L0)']);
% Additionally, we want to set the plateau values for each LacI•DNA+ xs
DNA FRET efficiencies, which is equal to the minimal values observed
for DNA + xsDNA + LacI (the end point). All the matrices are being
transposed so that each column corresponds to a different construct (Y)
```

```
and the rows correspond to a different time series (X) for entering
into "Prism" by transposing the matrix.
evalin('base',['min_ET_' ETMFPV '_CL = (min(ET_column_' ETMFPV
'_CL_landscape(:,8:21),[],2))']);
evalin('base',['span_FPV(:,k) = max(ET_column_' ETMFPV '_LC_landscape
(:,8:21),[],2) - min_ET_' ETMFPV '_CL']);
end
% 5C16-DIAI had an artifact that gave it substantial FRET throughout
the experiment. Therefore this is taken care of below.
over_ET_5C12_DIAI = min_ET_DIAI_CL(4,1);
min_ET_DIAI_CL(4,1) = min_ET_DIAI_CL(4,1) - over_ET_5C12_DIAI;
span_FPV(4,1) = span_FPV(4,1) - min_ET_DIAI_CL(4,1);
norm_vec_ET_DIAI_LC(4,:) = norm_vec_ET_DIAI_LC(4,:) - (repmat
(over_ET_5C12_DIAI,1,(size(norm_vec_ET_DIAI_LC,2))));
% the plateau for 11C12 also was miscalculated due to an plate artifact
and therefore the minimal and the span are corrected
span_FPV(17,2) =  min_ET_DIAE_CL(17,1);
min_ET_DIAE_CL(17,1) = 0;
half_lives_DIAI = zeros(25,1);
for j = 1:size(span_FPV,1)
x = time_LC;
span = abs(span_FPV(j,1));
plateau = min_ET_DIAI_CL(j,1);
exp_vec = norm_vec_ET_DIAI_LC(j,:);
if exp_vec(1,1)  .05
    exp_vec = zeros(1,size(norm_vec_ET_DIAI_LC,2));
end
% This sets the initial estimate for k that will be minimzed by
bounding k between 0 and 10.
k = 1;
lowbnd = 0;
uppbnd = 10;
resid = @(k)(xponential_resid(exp_vec,span,plateau,x,k));
options = optimset('MaxFunEvals',10000000,'MaxIter',1000000);
fans = fmincon(resid,k,[],[],[],[],lowbnd,uppbnd,[],options);
half_lives_DIAI(j,:) = 0.69./fans;
end
DIAI_half_lives = reshape(half_lives_DIAI,5,5);

half_lives_DIAE = zeros(25,1);
for j = 1:size(span_FPV,1)
x = time_LC;
span = abs(span_FPV(j,2));
plateau = min_ET_DIAE_CL(j,1);
exp_vec = norm_vec_ET_DIAE_LC(j,:);
if exp_vec(1,1)  .05
    exp_vec = zeros(1,size(norm_vec_ET_DIAI_LC,2));
end
k = 1;
lowbnd = 0;
uppbnd = 10;
resid = @(k)(xponential_resid(exp_vec,span,plateau,x,k));
```

```
options = optimset('MaxFunEvals',10000000,'MaxIter',1000000);
fans = fmincon(resid,k,[],[],[],[],lowbnd,uppbnd,[],options);
half_lives_DIAE(j,:) = 0.69./fans;
end
DIAE_half_lives = reshape(half_lives_DIAE,5,5);
half_lives_DEAI = zeros(25,1);
for j = 1:size(span_FPV,1)
x = time_LC;
span = abs(span_FPV(j,3));
plateau = min_ET_DEAI_CL(j,1);
exp_vec = norm_vec_ET_DEAI_LC(j,:);
if exp_vec(1,1)  .05
    exp_vec = zeros(1,size(norm_vec_ET_DIAI_LC,2));
end
k = 1;
lowbnd = 0;
uppbnd = 10;
resid = @(k)(xponential_resid(exp_vec,span,plateau,x,k));
options = optimset('MaxFunEvals',10000000,'MaxIter',1000000);
fans = fmincon(resid,k,[],[],[],[],lowbnd,uppbnd,[],options);
half_lives_DEAI(j,:) = 0.69./fans;
end
DEAI_half_lives = reshape(half_lives_DEAI,5,5);
 z_max = roundn(max(max([DIAI_half_lives DIAE_half_lives
DEAI_half_lives])),2);
DA = DIAI_half_lives;
DA1 = DIAE_half_lives;
D1A = DEAI_half_lives;
zDA = NaN(2*(size(D1A)));
zD1A = NaN(2*(size(D1A)));
zDA1 = NaN(2*(size(D1A)));
z_max = max(max([DA DA1 D1A]));
% spaces the data landscapes so that there are positioned in every
other column/row.
for j = 1:size(DA,1);
    for i = 1:size(DA1,1);
    zDA((j*2-1),(i*2-1)) = DA(j,i);
    zDA1((j*2-1),(i*2-1)) = DA1(j,i);
    zD1A((j*2-1),(i*2-1)) = D1A(j,i);
    end
end
scrsz = get(0,'ScreenSize');
figure('Position',[1 scrsz(4)/1 scrsz(3)/4 scrsz(4)/1]);
subplot(3,1,1);
hold on;
stem3(zDA,'Marker','.','MarkerSize',15,'MarkerFaceColor','auto',...
    'LineStyle','none','Color',[0 0 0]);
hold on
axis ([0.5 10.5 0.5 10.5 -z_max/4 z_max]);
y = interpft(interpft(DA,10,1),10,2);
surfc(y,'FaceColor','interp'),colormap('jet');
h = findobj('Type','patch');
```

```
set(h,'LineWidth',4);
view(-30,45);
grid on;
caxis([0,z_max]);
set(gca,'color','none','xtick',1:10,'ytick',1:10,'ztick',-z_max/
4:z_max/4:z_max,...
    'ZTickLabel',[],'YTickLabel',[],'XTickLabel',[]);
alpha(0.8);
subplot(3,1,2);
hold on;
stem3(zDA1,'Marker','.','MarkerSize',15,'MarkerFaceColor','auto',...
    'LineStyle','none','Color',[0 0 0]);
hold on
axis ([0.5 10.5 0.5 10.5 -z_max/4 z_max]);
y = interpft(interpft(DA1,10,1),10,2);
surfc(y,'FaceColor','interp'),colormap('jet');
h = findobj('Type','patch');
set(h,'LineWidth',4);
view(-30,45);
grid on;
caxis([0,z_max]);
set(gca,'color','none','xtick',1:10,'ytick',1:10,'ztick',-z_max/
4:z_max/4:z_max,...
    'ZTickLabel',[],'YTickLabel',[],'XTickLabel',[]);
alpha(0.8);
subplot(3,1,3);
hold on;
stem3(zD1A,'Marker','.','MarkerSize',15,'MarkerFaceColor','auto',...
    'LineStyle','none','Color',[0 0 0]);
hold on
axis ([0.5 10.5 0.5 10.5 -z_max/4 z_max]);
y = interpft(interpft(D1A,10,1),10,2);
surfc(y,'FaceColor','interp'),colormap('jet');
h = findobj('Type','patch');
set(h,'LineWidth',4);
view(-30,45);
grid on;
caxis([0,z_max]);
alpha(0.8)
set(gca,'color','none','xtick',1:10,'ytick',1:10,'ztick',-z_max/
4:z_max/4:z_max,...
    'ZTickLabel',[],'YTickLabel',...
    {'10','','12','','14','','16','','18',''},...
    'XTickLabel',{'5','','7','','9','','11','','13',''});
set(gcf,'PaperPositionMode','auto');
print('-djpeg','-r200',filename);
```

Bibliography

Bahl, C. P., Wu, R., Stawinsky, J., Narang, S. A. (1977). Minimal length of the lactose operator sequence for the specific recognition by the lactose repressor. *Proceedings of the National Academy of Sciences, U.S.A., 74*(3), 966-970.

Balci, H., Ha, T., Sweeney, H. L., Selvin, P. R. (2005). Interhead distance measurements in myosin VI via SHRImP support a simplified hand-over-hand model. *Biophysical Journal, 89*(1), 413-417.

Barkley, M. D., Riggs, A. D., Jobe, A., Burgeois, S. (1975). Interaction of effecting ligands with lac repressor and repressor-operator complex. *Biochemistry, 14*(8), 1700-1712.

Becker, N. A., Kahn, J. D., Maher, L. J., 3rd. (2005). Bacterial repression loops require enhanced DNA flexibility. *Journal of Molecular Biology, 349*(4), 716-730.

Becker, N. A., Kahn, J. D., Maher, L. J., 3rd. (2007). Effects of nucleoid proteins on DNA repression loop formation in Escherichia coli. *Nucleic Acids Research, 35* (12), 3988-4000.

Berg, O. G., Blomberg, C. (1976). Association kinetics with coupled diffusional flows. Special application to the lac repressor--operator system. *Biophys Chem, 4*(4), 367-381.

Beyreuther, K., Adler, K., Geisler, N., Klemm, A. (1973). The amino-acid sequence of lac repressor. *Proceedings of the National Academy of Sciences, U.S.A., 70*(12), 3576-3580.

Blattner, F. R., Plunkett, G., 3rd, Bloch, C. A., Perna, N. T., Burland, V., Riley, M., et al. (1997). The complete genome sequence of Escherichia coli K-12. *Science, 277* (5331), 1453-1462.

Boezi, J. A., Cowie, D. B. (1961). Kinetic studies of beta-galactosidase induction. *Biophysical Journal, 1*, 639-647.

Bond, L. M., Peters, J. P., Becker, N. A., Kahn, J. D., Maher, L. J., 3rd. (2010). Gene repression by minimal lac loops in vivo. *Nucleic Acids Research, 38*(22), 8072-8082.

Borowiec, J. A., Zhang, L., Sasse-Dwight, S., Gralla, J. D. (1987). DNA supercoiling promotes formation of a bent repression loop in lac DNA. *Journal of Molecular Biology, 196*(1), 101-111.

Bourgeois, S., Riggs, A. D. (1970). The lac repressor-operator interaction. IV. Assay and purification of operator DNA. *Biochemical and Biophysics Reseach Communications, 38*(2), 348-354.

Brenowitz, M., Mandal, N., Pickar, A., Jamison, E., Adhya, S. (1991). DNA-binding properties of a lac repressor mutant incapable of forming tetramers. *Journal of Biological Chemistry, 266*(2), 1281-1288.

Brunger, A. T., Strop, P., Vrljic, M., Chu, S., Weninger, K. R. (2011). Three-dimensional molecular modeling with single molecule FRET. *Journal of Structural Biology, 173*(3), 497-505.

Chen, J., Matthews, K. S. (1994). Subunit dissociation affects DNA binding in a dimeric lac repressor produced by C-terminal deletion. *Biochemistry, 33*(29), 8728-8735.

Choi, P. J., Cai, L., Frieda, K., Xie, X. S. (2008). A stochastic single-molecule event triggers phenotype switching of a bacterial cell. *Science, 322*(5900), 442-446.

Clark, D. J., Marr, A. G. (1964). Studies on the Repression of Beta-Galactosidase in Escherichia Coli. *Biochimica et Biophysica Acta, 92*, 85-94.

Cloutier, T. E., Widom, J. (2005). DNA twisting flexibility and the formation of sharply looped protein-DNA complexes. *Proceedings of the National Academy of Sciences, U.S.A. 102*(10), 3645-3650.

Crothers, D. M., Drak, J., Kahn, J. D., Levene, S. D. (1992). DNA bending, flexibility, and helical repeat by cyclization kinetics. *Methods in Enzymology, 212*, 3-29.

Crothers, D. M., Haran, T. E., Nadeau, J. G. (1990). Intrinsically bent DNA. *Journal of Biological Chemistry, 265*(13), 7093-7096.

Culard, F., Maurizot, J. C. (1981). Lac repressor - lac operator interaction. Circular dichroism study. *Nucleic Acids Research, 9*(19), 5175-5184.

Daber, R., Sharp, K., Lewis, M. (2009). One Is Not Enough. *Journal of Molecular Biology, 392*(5), 1133-1144,

Daber, R., Stayrook, S., Rosenberg, A., Lewis, M. (2007). Structural analysis of lac repressor bound to allosteric effectors. *Journal of Molecular Biology, 370*(4), 609-619.

Delon, A., Wang, I., Lambert, E., Mache, S., Mache, R., Derouard, J., et al. (2010). Measuring, in solution, multiple-fluorophore labeling by combining fluorescence correlation spectroscopy and photobleaching. *Journal of Physical Chemistry B, 114*(8), 2988-2996.

Donner, J., Caruthers, M. H., Gill, S. J. (1982). A calorimetric investigation of the interaction of the lac repressor with inducer. *Journal of Biological Chemistry, 257* (24), 14826-14829.

Dunaway, M., Droge, P. (1989). Transactivation of the Xenopus rRNA gene promoter by its enhancer. *Nature, 341*(6243), 657-659.

Edelman, L. M., Cheong, R., Kahn, J. D. (2003). Fluorescence resonance energy transfer over approximately 130 basepairs in hyperstable lac repressor-DNA loops. *Biophysical Journal, 84*(2 Pt 1), 1131-1145.

Elf, J., Li, G. W., Xie, X. S. (2007). Probing transcription factor dynamics at the single-molecule level in a living cell. *Science, 316*(5828), 1191-1194.

Falcon, C. M., Matthews, K. S. (2001). Engineered disulfide linking the hinge regions within lactose repressor dimer increases operator affinity, decreases sequence selectivity, and alters allostery. *Biochemistry, 40*(51), 15650-15659.

Farabaugh, P. J. (1978). Sequence of the lacI gene. *Nature, 274*(5673), 765-769.

Fickert, R., Müller-Hil, B. (1992). How Lac repressor finds lac operator in vitro. *Journal of Molecular Biology, 226*(1), 59-68.

Finkelstein, I. J., Visnapuu, M. L., Greene, E. C. (2010). Single-molecule imaging reveals mechanisms of protein disruption by a DNA translocase. *Nature, 468* (7326), 983-987.

Flashner, Y., Gralla, J. D. (1988). Dual mechanism of repression at a distance in the lac operon. *Proceedings of the National Academy of Sciences, U.S.A., 85*(23), 8968-8972.

Förster, T. (1948). Zwischenmolekulare Energiewanderung und Fluoreszenz. *Annalen der Physik, 437*, 55–75.

Friedman, A. M., Fischmann, T. O., Steitz, T. A. (1995). Crystal structure of lac repressor core tetramer and its implications for DNA looping. *Science, 268*(5218), 1721-1727.

Furini, S., Domene, C., Cavalcanti, S. (2010). Insights into the sliding movement of the lac repressor nonspecifically bound to DNA. *Journal of Physical Chemistry B, 114*(6), 2238-2245.

Galande, S., Purbey, P. K., Notani, D., Kumar, P. P. (2007). The third dimension of gene regulation: organization of dynamic chromatin loopscape by SATB1. *Current Opinions in Genetic Developement, 17*(5), 408-414.

Gilbert, W., Gralla, J., Majors, J., Maxam, A. (1975). Lactose operator sequences and the action of *lac* repressor. *In symposium on Protein-Ligand Interactions*. Sund, H. and Blauer, G. (eds): 193-206

Gilbert, W., Maxam, A. (1973). The nucleotide sequence of the lac operator. *Proceedings of the National Academy of Sciences, U.S.A., 70*(12), 3581-3584.

Gilbert, W., Müller-Hill, B. (1966). Isolation of the lac repressor. *Proceedings of the National Academy of Sciences, U.S.A., 56*(6), 1891-1898.

Gilbert, W., Müller-Hill, B. (1967). The lac operator is DNA. *Proceedings of the National Academy of Sciences, U.S.A., 58*(6), 2415-2421.

Gimenes, F., Takeda, K. I., Fiorini, A., Gouveia, F. S., Fernandez, M. A. (2008). Intrinsically bent DNA in replication origins and gene promoters. *Genetics and Molecular Research, 7*(2), 549-558.

Goyal, S., Lillian, T., Blumberg, S., Meiners, J. C., Meyhöfer, E., Perkins, N. C. (2007). Intrinsic curvature of DNA influences LacR-mediated looping. *Biophysical Journal, 93*(12), 4342-4359.

Hagerman, P. J., Ramadevi, V. A. (1990). Application of the method of phage T4 DNA ligase-catalyzed ring-closure to the study of DNA structure. I. Computational analysis. *Journal of Molecular Biology, 212*(2), 351-362.

Hahn, S., Dunn, T., Schleif, R. (1984). Upstream repression and CRP stimulation of the Escherichia coli L-arabinose operon. *Journal of Molecular Biology, 180*(1), 61-72.

Han, L., Garcia, H. G., Blumberg, S., Towles, K. B., Beausang, J. F., Nelson, P. C., et al. (2009). Concentration and length dependence of DNA looping in transcriptional regulation. *PLoS ONE, 4*(5), e5621.

Hieb, A. R., Halsey, W. A., Betterton, M. D., Perkins, T. T., Kugel, J. F., Goodrich, J. A. (2007). TFIIA changes the conformation of the DNA in TBP/TATA complexes and increases their kinetic stability. *Journal of Molecular Biology, 372*(3), 619-632.

Jacob, F., Monod, J. (1961). Genetic regulatory mechanisms in the synthesis of proteins. *Journal of Molecular Biology, 3*, 318-356.

Jacobson, H., Stockmayer, W. H. (1950). *J. Chem. Phys., 18*, 1600.

Jobe, A., Bourgeois, S. (1972). The lac repressor-operator interaction. VII. A repressor with unique binding properties: the X86 repressor. *Journal of Molecular Biology, 72*(1), 139-152.

Kalodimos, C. G., Biris, N., Bonvin, A. M., Levandoski, M. M., Guennuegues, M., Boelens, R., et al. (2004). Structure and flexibility adaptation in nonspecific and specific protein-DNA complexes. *Science, 305*(5682), 386-389.

Kania, J., Müller-Hill, B. (1977). Construction, isolation and implications of repressor-galactosidase - beta-galactosidase hybrid molecules. *European Journal of Biochemistry, 79*(2), 381-386.

Kaptein, R., Zuiderweg, E. R., Scheek, R. M., Boelens, R., van Gunsteren, W. F. (1985). A protein structure from nuclear magnetic resonance data. lac repressor headpiece. *Journal of Molecular Biology, 182*(1), 179-182.

Koo, H. S., Drak, J., Rice, J. A., Crothers, D. M. (1990). Determination of the extent of DNA bending by an adenine-thymine tract. *Biochemistry, 29*(17), 4227-4234.

Koo, H. S., Wu, H. M., Crothers, D. M. (1986). DNA bending at adenine-thymine tracts. *Nature, 320*(6062), 501-506.

Kramer, H., Amouyal, M., Nordheim, A., Müller-Hill, B. (1988). DNA supercoiling changes the spacing requirement of two lac operators for DNA loop formation with lac repressor. *EMBO Journal, 7*(2), 547-556.

Kramer, H., Niemoller, M., Amouyal, M., Revet, B., von Wilcken-Bergmann, B., Müller-HillMüller-Hill, B. (1987). lac repressor forms loops with linear DNA carrying two suitably spaced lac operators. *EMBO Journal, 6*(5), 1481-1491.

Kugel, J. F. (2008). Using FRET to measure the angle at which a protein bends DNA. *Biochemistry and Molecular Biology Education, 36*(5), 341-346.

La Penna, G., Perico, A. (2010). Wrapped-around models for the lac operon complex. *Biophysical Journal, 98*(12), 2964-2973.

Levandoski, M. M., Tsodikov, O. V., Frank, D. E., Melcher, S. E., Saecker, R. M., Record, M. T., Jr. (1996). Cooperative and anticooperative effects in binding of the first and second plasmid Osym operators to a LacI tetramer: evidence for contributions of non-operator DNA binding by wrapping and looping. *Journal of Molecular Biology, 260*(5), 697-717.

Levene, S. D., Wu, H. M., Crothers, D. M. (1986). Bending and flexibility of kinetoplast DNA. *Biochemistry, 25*(14), 3988-3995.

Lewis, M. (2005). The lac repressor. *Comptes Rendus Biologies, 328*(6), 521-548.

Lewis, M., Chang, G., Horton, N. C., Kercher, M. A., Pace, H. C., Schumacher, M. A., et al. (1996). Crystal structure of the lactose operon repressor and its complexes with DNA and inducer. *Science, 271*(5253), 1247-1254.

Lillian, T. D., Goyal, S., Kahn, J. D., Meyhöfer, E., Perkins, N. C. (2008). Computational analysis of looping of a large family of highly bent DNA by LacI. *Biophysical Journal, 95*(12), 5832-5842.

Lin, S., Riggs, A. D. (1975). The general affinity of lac repressor for E. coli DNA: implications for gene regulation in procaryotes and eucaryotes. *Cell, 4*(2), 107-111.

Lobell, R. B., Schleif, R. F. (1991). AraC-DNA looping: orientation and distance-dependent loop breaking by the cyclic AMP receptor protein. *Journal of Molecular Biology, 218*(1), 45-54.

Loverdo, C., Benichou, O., Voituriez, R., Biebricher, A., Bonnet, I., Desbiolles, P. (2009). Quantifying hopping and jumping in facilitated diffusion of DNA-binding proteins. *Physics Review Letters, 102*(18), 188101.

Maizels, N. M. (1973). The nucleotide sequence of the lactose messenger ribonucleic acid transcribed from the UV5 promoter mutant of Escherichia coli. *Proceedings of the National Academy of Sciences, U.S.A., 70*(12), 3585-3589.

Manning, R. S., Maddocks, J. H., Kahn, J. D. (1996). A continuum rod model of sequence-dependent DNA structure. *The Journal of Chemical Physics, 105*(13), 5626-5646.

Marini, J. C., Levene, S. D., Crothers, D. M., Englund, P. T. (1982). Bent helical structure in kinetoplast DNA. *Proceedings of the National Academy of Sciences, U.S.A., 79*(24), 7664-7668.

Markiewicz, P., Kleina, L. G., Cruz, C., Ehret, S., Miller, J. H. (1994). Genetic studies of the lac repressor. XIV. Analysis of 4000 altered Escherichia coli lac repressors reveals essential and non-essential residues, as well as "spacers" which do not require a specific sequence. *Journal of Molecular Biology, 240*(5), 421-433.

Mehta, R. A., Kahn, J. D. (1999). Designed hyperstable Lac repressor.DNA loop topologies suggest alternative loop geometries. *Journal of Molecular Biology, 294* (1), 67-77.

Monod, J., Wyman, J., Changeux, J. P. (1965). On the Nature of Allosteric Transitions: A Plausible Model. *Journal of Molecular Biology, 12*, 88-118.

Morgan, M. A., Okamoto, K., Kahn, J. D., English, D. S. (2005). Single-molecule spectroscopic determination of lac repressor-DNA loop conformation. *Biophysical Journal, 89*(4), 2588-2596.

Mossing, M. C., Record, M. T., Jr. (1986). Upstream operators enhance repression of the lac promoter. *Science, 233*(4766), 889-892.

Mrazek, J. (2010). Comparative analysis of sequence periodicity among prokaryotic genomes points to differences in nucleoid structure and a relationship to gene expression. *Journal of Bacteriology, 192*(14), 3763-3772.

Muller, J., Oehler, S., Müller-Hill, B. (1996). Repression of lac promoter as a function of distance, phase and quality of an auxiliary lac operator. *Journal of Molecular Biology, 257*(1), 21-29.

Müller-Hill, B. (1998). The function of auxiliary operators. *Molecular Microbiology, 29* (1), 13-18.

Müller-Hill, B., Crapo, L., Gilbert, W. (1968). Mutants that make more lac repressor. *Proceedings of the National Academy of Sciences, U.S.A., 59*(4), 1259-1264.

O'Gorman, R. B., Dunaway, M., Matthews, K. S. (1980). DNA binding characteristics of lactose repressor and the trypsin-resistant core repressor. *Journal of Biological Chemistry, 255*(21), 10100-10106.

Oehler, S., Alberti, S., Müller-Hill, B. (2006). Induction of the lac promoter in the absence of DNA loops and the stoichiometry of induction. *Nucleic Acids Research, 34*(2), 606-612.

Oehler, S., Amouyal, M., Kolkhof, P., von Wilcken-Bergmann, B., Müller-Hill, B. (1994). Quality and position of the three lac operators of E. coli define efficiency of repression. *EMBO Journal, 13*(14), 3348-3355.

Oehler, S., Eismann, E. R., Kramer, H., Müller-Hill, B. (1990). The three operators of the lac operon cooperate in repression. *EMBO Journal, 9*(4), 973-979.

Ohshima, Y., Mizokoshi, T., Horiuchi, T. (1974). Binding of an inducer to the lac repressor. *Journal of Molecular Biology, 89*(1), 127-136.

Pfahl, M., Gulde, V., Bourgeois, S. (1979). "Second" and "third operator" of the lac operon: an investigation of their role in the regulatory mechanism. *Journal of Molecular Biology, 127*(3), 339-344.

Record, M. T., Jr., Lohman, M. L., De Haseth, P. (1976). Ion effects on ligand-nucleic acid interactions. *Journal of Molecular Biology, 107*(2), 145-158.

Reznikoff, W. S., Winter, R. B., Hurley, C. K. (1974). The location of the repressor binding sites in the lac operon. *Proceedings of the National Academy of Sciences, U.S.A. 71*(6), 2314-2318.

Rippe, K., von Hippel, P. H., Langowski, J. (1995). Action at a distance: DNA-looping and initiation of transcription. *Trends in Biochemical Sciences, 20*(12), 500-506.

Romanuka, J., Folkers, G. E., Biris, N., Tishchenko, E., Wienk, H., Bonvin, A. M., et al. (2009). Specificity and affinity of Lac repressor for the auxiliary operators O2 and O3 are explained by the structures of their protein-DNA complexes. Microscopy Research and Technique*Journal of Molecular Biology, 390*(3), 478-489.

Ruben, G. C., Roos, T. B. (1997). Conformation of Lac repressor tetramer in solution, bound and unbound to operator DNA. *Microscopy Research and Technique, 36* (5), 400-416.

Rutkauskas, D., Zhan, H., Matthews, K. S., Pavone, F. S., Vanzi, F. (2009). Tetramer opening in LacI-mediated DNA looping. *Proceedings of the National Academy of Sciences, U.S.A. 106*(39), 16627-16632.

Sadler, J. R., Sasmor, H., Betz, J. L. (1983). A perfectly symmetric lac operator binds the lac repressor very tightly. *Proceedings of the National Academy of Sciences, U.S.A., 80*(22), 6785-6789.

Saecker, R. M., Record, M. T., Jr. (2002). Protein surface salt bridges and paths for DNA wrapping. *Current Opinion in Structural Biology, 12*(3), 311-319.

Sasse-Dwight, S., Gralla, J. D. (1988). Probing co-operative DNA-binding in vivo. The lac O1:O3 interaction. *Journal of Molecular Biology, 202*(1), 107-119.

Schlax, P. J., Capp, M. W., Record, M. T., Jr. (1995). Inhibition of transcription initiation by lac repressor. *Journal of Molecular Biology, 245*(4), 331-350.

Schmitz, A., Schmeissner, U., Miller, J. H. (1976). Mutations affecting the quaternary structure of the lac repressor. *Journal of Biological Chemistry, 251*(11), 3359-3366.

Semsey, S., Virnik, K., Adhya, S. (2005). A gamut of loops: meandering DNA. *Trends in Biochemical Sciences, 30*(6), 334-341.

Shore, D., Baldwin, R. L. (1983). Energetics of DNA twisting. I. Relation between twist and cyclization probability. *Journal of Molecular Biology, 170*(4), 957-981.

Simons, A., Tils, D., von Wilcken-Bergmann, B., Müller-Hill, B. (1984). Possible ideal lac operator: Escherichia coli lac operator-like sequences from eukaryotic genomes lack the central G X C pair. *Proceedings of the National Academy of Sciences, U.S.A., 81*(6), 1624-1628.

Spronk, C. A., Slijper, M., van Boom, J. H., Kaptein, R., Boelens, R. (1996). Formation of the hinge helix in the lac repressor is induced upon binding to the lac operator. *Nature Structural Biology, 3*(11), 916-919.

Stenger, J. E., Tegtmeyer, P., Mayr, G. A., Reed, M., Wang, Y., Wang, P., et al. (1994). p53 oligomerization and DNA looping are linked with transcriptional activation. *EMBO Journal, 13*(24), 6011-6020.

Swigon, D., Coleman, B. D., Olson, W. K. (2006). Modeling the Lac repressor-operator assembly: the influence of DNA looping on Lac repressor conformation. *Proceedings of the National Academy of Sciences, U.S.A. 103*(26), 9879-9884.

Villa, E., Balaeff, A., Schulten, K. (2005). Structural dynamics of the lac repressor-DNA complex revealed by a multiscale simulation. *Proceedings of the National Academy of Sciences, U.S.A. 102*(19), 6783-6788.

Wedel, A., Weiss, D. S., Popham, D., Droge, P., Kustu, S. (1990). A bacterial enhancer functions to tether a transcriptional activator near a promoter. *Science, 248*(4954), 486-490.

Wiggins, P. A., van der Heijden, T., Moreno-Herrero, F., Spakowitz, A., Phillips, R., Widom, J., et al. (2006). High flexibility of DNA on short length scales probed by atomic force microscopy. *Nature Nanotechnology, 1*(2), 137-141.

Wilson, C. J., Das, P., Clementi, C., Matthews, K. S., Wittung-Stafshede, P. (2005). The experimental folding landscape of monomeric lactose repressor, a large two-domain protein, involves two kinetic intermediates. *Proceedings of the National Academy of Sciences, U.S.A. 102*(41), 14563-14568.

Wong, O. K., Guthold, M., Erie, D. A., Gelles, J. (2008). Interconvertible lac repressor-DNA loops revealed by single-molecule experiments. *PLoS Biology, 6*(9), e232.

Yagil, G., Yagil, E. (1971). On the relation between effector concentration and the rate of induced enzyme synthesis. *Biophysical Journal, 11*(1), 11-27.

Yasmin, R., Yeung, K. T., Chung, R. H., Gaczynska, M. E., Osmulski, P. A., Noy, N. (2004). DNA-looping by RXR tetramers permits transcriptional regulation "at a distance". *Journal of Molecular Biology, 343*(2), 327-338.

Yildiz, A., Selvin, P. R. (2005). Fluorescence imaging with one nanometer accuracy: application to molecular motors. *Accounts of Chemical Research, 38*(7), 574-582.

Zhan, H., Camargo, M., Matthews, K. S. (2010). Positions 94-98 of the Lactose Repressor N-Subdomain Monomer-Monomer Interface Are Critical for Allosteric Communication. *Biochemistry*.

Zhang, Y., McEwen, A. E., Crothers, D. M., Levene, S. D. (2006). Analysis of in-vivo LacR-mediated gene repression based on the mechanics of DNA looping. *PLoS One, 1*, e136.