ABSTRACT

Title of Dissertation: DNA-PROTEIN NANOTECHNOLOGY: DEVELOPING UNIQUE BIOLOGICAL NANOSTRUCTURES AND BIOLOGICAL TOOLS

Michael Andrew Morgan, Doctor of Philosophy, 2005

Dissertation directed by: Associate Professor Jason D. Kahn
Department of Chemistry and Biochemistry

Lac repressor (LacI) when bound to two operator DNA sites loops the intervening DNA sequence, enhancing the efficiency of transcriptional repression. Previous results from footprinting, electrophoretic mobility shift and ring closure experiments on *lac* operator/A-tract DNA constructs suggest that the LacI C-terminal 4-helix bundle probably acts as a hinge allowing the protein to adopt multiple conformations. One DNA molecule under study, designated "9C14," was proposed to be capable of forming two different geometries, with either a closed (V-shaped) LacI or a more open, extended form of the repressor. Previous bulk Fluorescence Resonance Energy Transfer (FRET) results confirmed the existence of a closed loop form (giving a net 70 % efficiency of transfer) but did not definitively confirm the presence of an open form complex or whether there existed populations with

intermediate LacI-DNA geometries. We applied the technique of single molecule FRET to Cy3-Cy5 labeled 9C14-LacI DNA loops freely diffusing in solution in order to further assess looping geometries. Through careful consideration of photophysical bleaching effects, which can affect both single molecule and ensemble FRET measurements, we have shown that our results clearly demonstrate that LacI-9C14 exists strictly as a closed loop exhibiting nearly 100 % energy transfer (ET) efficiency.

In a second project, we set out to design and characterize DNA-LacI self-assembled 2D and 3D nanostructures, applying four critical design rules. First, DNA provides us with a building material that is stiff, stable and easily manipulated. Second, we use circular DNA to build our molecules because it adopts a more restricted range of conformations then its linear counterpart. Third, the positioning of curved A-tracts (short runs of $(dA)_{4-6} \bullet (dT)_{4-6}$) at the corners of designed "squares" and "triangles" allows the sides of the "squares" and "triangles" to remain straight, with the majority of the bending localized at the corners of the molecule, making it compact and rigid. Fourth, the formation of three proposed large nanostructures is dependent on the orientation of the *lac* operator sequence relative to the plane of the DNA. By strategically altering the DNA sequence we can control the rotational phasing of the pseudo-dyad axis of the *lac* operator, allowing for the cooperative assembly of the three nanostructures. Characterization of DNA "squares" and "triangles" has successfully been carried out using Atomic Force Microscopy (AFM) tapping mode in air and liquid. Current efforts are focused on AFM imaging of the proposed 2D and 3D DNA-protein nanostructures.

DNA-PROTEIN NANOTECHNOLOGY: DEVELOPING UNIQUE BIOLOGICAL NANOSTRUCTURES AND BIOLOGICAL TOOLS

by

Michael Andrew Morgan

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 the
Doctor of Philosophy
2005

Advisory Committee

Associate Professor Jason D. Kahn, chair
Assistant Professor Douglas English
Associate Professor Sheryl Erhman
Professor George Lorimer
Professor Steven Rokita

Dedication

In loving memory of my grandmother Mary Ross

Acknowledgments

I would like to first acknowledge GOD for pointing me in the right direction throughout my life. GOD has given me the strength to overcome all of life's obstacles and without that guidance I would not be the person that I am today. I would like to thank all those who have played a very special role in my life. To my grandmother Gertrude, mother Helen, father Kevin, brother Dan, sister Amy and all of my aunts, uncles and cousins and friends. You worked hard to help me achieve my goals and to you I will forever be thankful. I would like to say a special thank you to Triscel N. Webb, thank you for giving me energy and happiness. I'm truly honored by all of your tireless efforts to help me succeed throughout life. Without your constant guidance and love I would be half the man I am today, you have all been a constant source of support and encouragement for me and for that I'm truly blessed. I would like to thank my high school teachers Jodi Millette, Philip Nash and Robert Devantry for giving me the opportunity to learn and foster an appreciation for science and writing. I would like to thank my former undergraduate mentors Dr. Philip N. Borer and Dr. James C. Dabrowiak for all of their support and for allowing me to develop skills and confidence as a scientist. I would like to thank the members of my graduate school committee, Dr. Steven Rokita, Dr. Sheryl Erhman and Dr. George Lorimer. I would like to thank Dr. Dorothy Beckett and Dr. David Jollie for their constant support and willingness to teach. I would like to thank Dr. Doug English, for without his support, encouragement and tireless efforts in training me, I would not be where I am today. Lastly, I would like to thank Dr. Jason Kahn for his efforts in

training me to be a competent scientist and individual. Thank you for allowing me to grow as a scientist for that I will always be truly grateful.

Table of Contents

List of Figures:

List of Tables and Equations

Tables

Equations

List of Abbreviations

AFM	atomic force microscopy
APTES	aminopropyltriethoxysilane
APD	avalanche photodiode
bp	base pair
BSA	bovine serum albumin
CAP	catabolite activator protein
DCLP	dichroic long pass filter
DNA	deoxyribonucleic acid
DTT	dithiolthreitol
E. coli	*Escherichia coli*
EDTA	ethylenediaminetetraacetate
ERP	extended reverse primer
ET	energy transfer
EtBr	ethidium bromide
EtOH	ethanol
FRET	fluorescence resonance energy transfer
HEPES	N-[2-Hydroxyethyl]piperazine-N′-[2-ethanesulfonic acid)
IPTG	isopropyl-β-D-thiogalactoside
LacI	Lac repressor protein
NaOAc	sodium acetate
NEB	New England Biolabs
NP40	nonident P 40 detergent
nt	nucleotide
ONPF	orthonitrophenylfucoside
PDB	Protein data bank
PCR	polymerase chain reaction
PVA	poly vinyl alcohol
RNAP	RNA polymerase
SMFS	single molecule fluorescence spectroscopy

SMS	single molecule spectroscopy
STM	scanning tunneling microscopy
TBE	Tris Borate EDTA
TE	Tris EDTA
TEMED	N, N, N′, N′-teramethylethylenediamine
Tris	tris(hydroxymethyl)aminomethane

INTRODUCTION I.

Nanotechnology

> As manufacturing methods are perfected and scaled up, nanotechnology is expected to soon pervade, and often revolutionize, virtually every sector of industrial activity, from electronics to warfare, from medicine to agriculture, from the energy we use to drive our homes to the water we drink and the food we eat. Nanotechnology is today's version of the space race, and countries around the globe are enthusiastically pouring billions of dollars into support of research, development, and commercialization. (Hood, 2004)

Nanotechnology, also known as molecular manufacturing, involves the application, creation and manipulation of individual or groups of molecules. It is a technology that is defined on the nanometer length scale, typically ranging from sizes of 1 to 100 nanometers, and involves the inexpensive manipulation and control of the chemical, physical and electrical properties of atoms. Nanotechnology draws ideas and concepts from disciplines including engineering, physics, chemistry, biology, mathematics and computer science. The central dogma of the "bottom-up" version of nanotechnology is the notion of self-assembly, which is the spontaneous assembly of materials into predetermined ordered structures or complexes. Although this may sound like a novel concept, we are constantly surrounded by the notion of self-assembly. For example, the self-assembly of nucleic acid-protein complexes in HIV-1 virion particle assembly is a highly ordered, sequential and structured process which is essential to HIV infection. More importantly, one can extrapolate this notion to countless other cellular processes in mammals and bacteria, lending credibility to the importance and abundance of self-assembly.

Because nucleic acid and protein structures are on the nanometer scale, the relatively new field of bio-nanotechnology has become interesting to nanotechnologists. Nucleic acids' intrinsic structural characteristics provide nanotechnologists with a building material that offers several levels of design hierarchy. DNA's basic structural components start with the four nucleic acid building blocks, which have the ability to form three different helical formations (A, B and Z-form). Hydrogen-bond mediated base-pairing such as Watson and Crick, Hoogsteen pairing of a third base in the major groove, and water mediated hydrogen bonding offer several more degrees of variability in DNA structure. Because of DNA's high precision base-pairing and recognition capabilities, DNA offers researchers with a potential starting material that could be used in the design, engineering and development of countless nano-materials, nano-structures and nano-machines.

Presented here are two representative examples from a large field of ongoing research in bio-nanotechnology that have utilized proteins and nucleic acids to build novel structures and devices. Nadrian Seeman of New York University has successfully designed and synthesized nanoscale self-assembly systems and self-assembled molecular computing devices (Sa-Ardyen *et al*., 2003). One such device designed by Seeman (Mao *et al*., 1999) works by utilizing the B-Z transition of double crossover DNA molecules (DX molecules) as shown in Figure 1B. DX molecules are comprised of two Holiday junctions that are connected via two double helical arms. There are several types of DX molecules, which differ in the relative

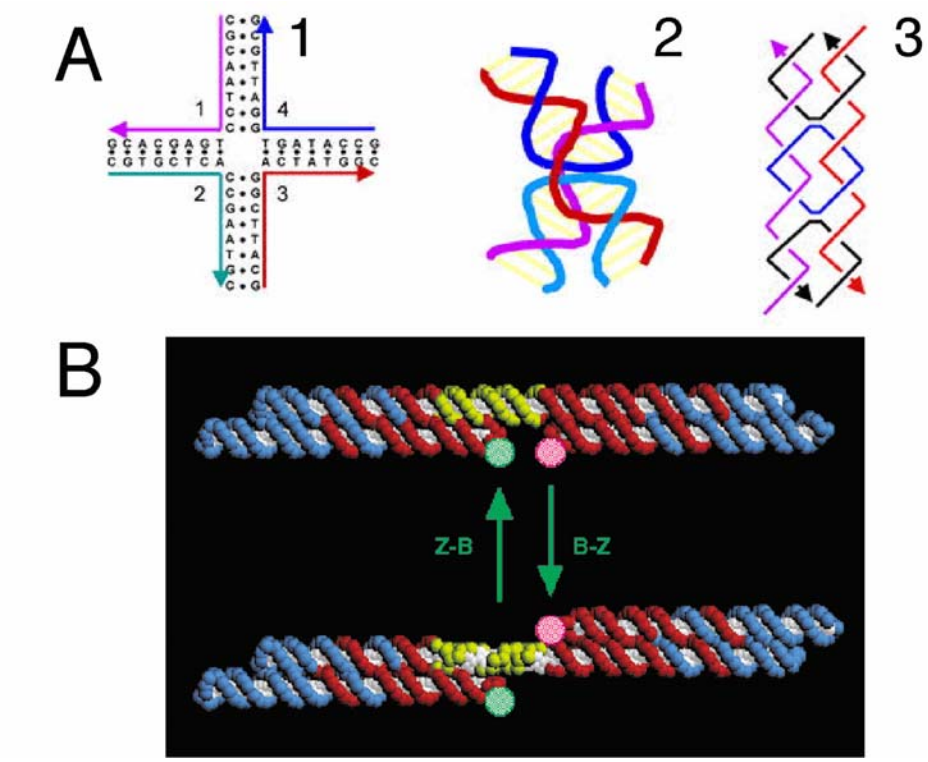


Figure 1A. Schematic of Double Crossover (DX) DNA. 1B. DNA nanomechanical device. A1. Structure of Holliday junction. A2. Structure of the 4-arm junction utilized in construction of DX DNA. A3. DX DNA contains two DNA duplexes that lie side by side to one another. The two DNA duplexes are joined by two crossovers, which prevent any duplex from twisting against its neighbor duplex, making the molecule very stable. Figure 1B. The device contains two double-crossover DNA molecules represented as blue and red strands. Each nucleotide is shown as two spheres, a colored one for the backbone and a white one for the base. The DNA molecules contain three unique strands, one central strand indicated in red containing a yellow segment (a 20 bp sequence of proto-Z DNA in a B-form conformation) and two flanking blue strands that are triply catenated to the red strand. Fluorophores attached to free hairpins at the center of the molecule and are represented as green (fluorescein) and magenta (Cy3) spheres, and are used to monitor the transition from B-Z form by FRET. Upon addition of Hexaamminecobalt (III) a B-Z transformation occurs and the yellow segment becomes left-handed Z DNA. This transition is reversible (Figure and legend adapted from Mao *et al*., 1999, 2004).

orientations of their helical axes (parallel or antiparallel) and by the number of double helical half-turns between the two crossover junctions (Sa-Ardyen *et al*., 2003). The two double crossover molecules in the B-Z sensor contain a "bridge" sequence that undergoes a conformational change upon addition of Hexaamminecobalt (III). This conversion of a right-handed B-DNA to left-handed Z-DNA is monitored by FRET. In the B-form orientation the two helices of the double crossover segments along with the fluorophores are aligned on the same side of the helix containing the bridge, resulting in high FRET efficiency. The switch to the Z-form places the two helices and the fluorophores on opposing sides of the DNA helix, causing the FRET efficiency to decrease. One could imagine this type of DNA device being used in the design of molecular switches or motors.

Nature has already utilized the intrinsic abilities of peptides and proteins to act as scaffolds or building blocks to produce macromolecular structures such as collagen, pearl, coral and calcite (Zhang, 2003). Todd Yeates and colleagues at University of California, Los Angeles have already shown the ability to design and engineer proteins that self assemble into large symmetrical nanomaterials including molecular cages, filaments, layers, and porous materials. To design these novel structures Yeates utilized the intrinsic self-assembling and high precision recognition capabilities of proteins. By fusing two different proteins together, both of which form naturally self-assembling oligomers, they build protein self-assemblies made from the identical copies of oligomers into designed nanohedral particles and materials. The design strategy is demonstrated through the successful engineering and characterization of two fusion proteins, one of which is a 49-kDa protein designed to

assemble into a cage-like structure 15 nm in diameter, and the other a 44-kDa protein designed to assemble into long filaments approximately 4 nm wide (Padilla *et al*., 2001).

Our lab has set out to design innovative and unique DNA-Lac repressor tools and macromolecular complexes that utilize the self-assembling properties of DNA and protein to build novel nanostructures and biological tools.

Lac repressor

In 1961, Jacob and Monod introduced a general model for gene regulation based on experiments they conducted on *Escherichia coli's* (*E. coli*) use of lactose (Jacob & Monod, 1961). Today, the lactose operon is one of the most widely studied and best understood paradigms for gene regulation. The LacI protein, the product of the *lacI* gene, negatively regulates the transcription of genes responsible for lactose metabolism by binding to *lac* operator O1 which blocks the promoter (Müller-Hill *et al*., 1998). In order to obtain full repression of the wild type promoter there must be at least one additional operator sequence (O2 or O3) positioned in close proximity and on the same side of the DNA helix as the O1 operator (Müller-Hill *et al*., 1998). The role of the secondary operator is to increase the concentration of LacI in the local area of the primary operator (Mossing & Record, 1986). The binding of LacI to the *lac* operator is made stronger by the formation of DNA loops (Brenowitz *et al.*, 1991). The formation of DNA loops is believed to play an important role in mechanisms involving the efficient repression of gene expression in both prokaryotic and eukaryotic systems.

Crystal structures of the intact LacI, a complex of the LacI with two operator DNA fragments, and LacI bound to the synthetic inducer isopropyl-β-D-thiogalactoside (IPTG) were elucidated in 1996 (Lewis *et al*., 1996). A model for the loop formed between the O1 and O3 *lac* operators is shown in Figure 2; note that many of the details of the model are incorrect (Perros, Steitz, Fried and Hudson, technical comments in Science, 1996).

In 2001 the crystal structure of the Lac repressor bound to a 22 bp DNA sequence containing the natural operator sequence O1 and anti-inducer orthonitrophenylfucoside (ONPF) was published (Bell & Lewis, 2001). A year later the NMR structure of the LacI DNA-binding domain dimer bound to natural operator O1 was published (Kalodimos *et al*., 2002). The information gathered from the NMR and crystal structures has provided scientists with a clear depiction of the overall LacI structure.

The Lac repressor is composed of four 360 amino acid subunits that associate into a homotetramer of approximately 154 kD. The Lac repressor monomer is usually described as five distinct domains. The DNA binding domain (residues 1-45) folds into three α-helices which form a helix-turn-helix motif. The hinge region (residues 46-62) connects the amino terminal region to the core of the protein. The inducer-binding domain, also known as the core region (residues 63-357) can be subdivided into two sections, the N-terminal and the C-terminal sub-domain of the core. The tetramerization helix domain (residues 340-357) is crucial to the formation of the tetrameric molecule (Lewis *et al*., 1996). It is proposed that the tetramerization domain is attached to the body of the protein by flexible linkers, which may give the

protein the ability to adopt different conformations (Freidman *et al*., 1995; Lewis *et al*., 1996).

Because the Lac repressor has strong dimer and weak tetramer interactions, the spatial arrangement of the overall structure is a tethered dimer (Lewis *et al*., 1996). In fact the Lac repressor should not be viewed as a homotetramer but rather as a dimer of dimers. The two-fold axes that relate monomers to each other within the dimers are not parallel to the two-fold axis that relates the dimers within the tetramer, creating a molecule that is more or less V-shaped (Lewis *et al*., 1996).

We have utilized structural and functional aspects of LacI to serve as our protein component for the development of unique DNA-protein tools and nanostructures.

Figure 2. Shown is crystal structure of the V-shaped *lac* repressor tetramer (violet) bound to double stranded duplex DNA (red) containing Osym operator sites (O1 and O3). The 4-helix bundle is believed to play an important role in allowing the protein to adopt multiple conformations. A superimposed model of a 93 bp DNA Wrapping Away (WA) repression loop that corresponds to *lac* operon sequence -82 to +11 is extended off of the O1 and O3 operator sites. The catabolite activator protein (CAP) and 30 bp blue DNA segment were modeled into the loop structure from PDB coordinates (PDB entry 1LBG, Lewis *et al*., 1996). However, at the time this model was published there was a lot of controversy surrounding the placement of the modeled catabolite activator protein (CAP) inside a positive supercoiled loop that wrapped away from the repressor (Perros & Steitz, 1996). The overall curvature of the modeled DNA loop is consistent with the curvature of the DNA observed in the CAP structure and that observed for the *lac* repressor (Lewis *et al*., 1996).

Design Rationale: A-Tract and *lac* Operator DNA Sequences

Utilizing our previous knowledge of protein induced DNA binding, bending and twisting our lab has designed two research projects, both of which rely heavily on the ability to design DNA molecules that are predisposed to form unique geometric shapes and structures upon LacI binding. One project exploits our ability to form different protein-induced DNA loop geometries to study mechanisms of DNA looping and their importance to gene regulation. The second project draws on our ability to design small, compact, and rigid DNA minicircles that self-assemble into predetermined 2D and 3D nanostructures upon LacI binding. Both projects use highly bent regions of DNA to form novel DNA geometries and shapes for our experiments.

Bent DNA was first discovered in the DNA of the kinoplast body of *Leishmania tarantolae* in the early 1980s (Marini *et al*., 1983). It was later determined by Hagerman (1985), Diekmann (1986), and Koo *et al*. (1986) that the intrinsic bending ability of the DNA was localized at regions containing multiple repeats of homopolymeric dA • dT. These regions of DNA were named A-tracts, which are defined as short runs of $(dA)_{4-6} \bullet (dT)_{4-6}$. To date there is disagreement about the true mechanism involved in DNA bending (junction *versus* wedge model), however most agree on the degree and direction of the bend for a single A-tract region (Crothers *et al*., 1990; Goodsell *et al*., 1994). Hydrodynamic experiments (Levene *et al*., 1986c) and DNA cyclization studies (Koo *et al*., 1990) on A-tract DNA gave a value of 18 º of bend per A-tract. By studying the helical phasing of an A-tract bend against the bending of DNA wrapped around the CAP protein, Zinkel &

Crothers (1987) concluded that the bend direction of an A-tract is directed towards the minor groove of the DNA at the center of the A-tract. In order to generate a large global bend in a DNA molecule, A-tracts must be positioned in phase with each other every helical repeat. The small bend incurred at each individual A-tract, an 18 º bend, can then be added constructively to give the DNA a large global curvature. For example, three A-tract repeats positioned in phase with each other every helical turn will result in a 54 º bend in the DNA as shown in the sequence below.

5 ' –TTTTTTGCCCGTTTTTTGCCGTTTTTT– 3'
AAAAAACGGGCAAAAAACGGCAAAAAA

Another vital component to our DNA constructs is the *lac* operator sequence, which is responsible for efficient binding of LacI to our DNA constructs. The *lac* operator sequence called Osym is a 20 bp palindromic sequence that differs from the wild type sequence in that it lacks a central base pair. This minor change in sequence results in LacI binding ten times more tightly then the wild type operator (Sadler *et al.*, 1983; Simons *et al*., 1984).

The design of the grid, cube and sandwich nanostructures (Figure 3) relies heavily on the notion of self-assembly, which is mediated through non-covalent interactions between the DNA and protein components. The formation of proposed LacI-DNA nanostructures should be a cooperative process, by analogy to the stabilization of a looped complex between two operator sites by other proteins that induce bending (Brenowitz *et al*., 1991; Lobell & Schleif 1991; Santero *et al*., 1992). Just as binding makes it energetically favorable for the DNA and protein to self-assemble, we can use localization via "sandwich complex" formation, in which LacI bridges two DNA molecules. For example, in order to form the grid structure shown

in Figure 3, four DNA molecules must bridge four Lac repressor molecules. When four DNA molecules are linked by three Lac repressor proteins, the binding of the fourth repressor molecule can bind to two sites simultaneously without an additional loss in entropy. There is no additional loss in entropy because the remaining two unoccupied *lac* operator sites (located on different DNA molecules) are in close proximity to each other. The cooperative self-assembly of the nanostructures offers the potential for the self-assembly of the grid, sandwich and cube DNA-LacI nanostructures.

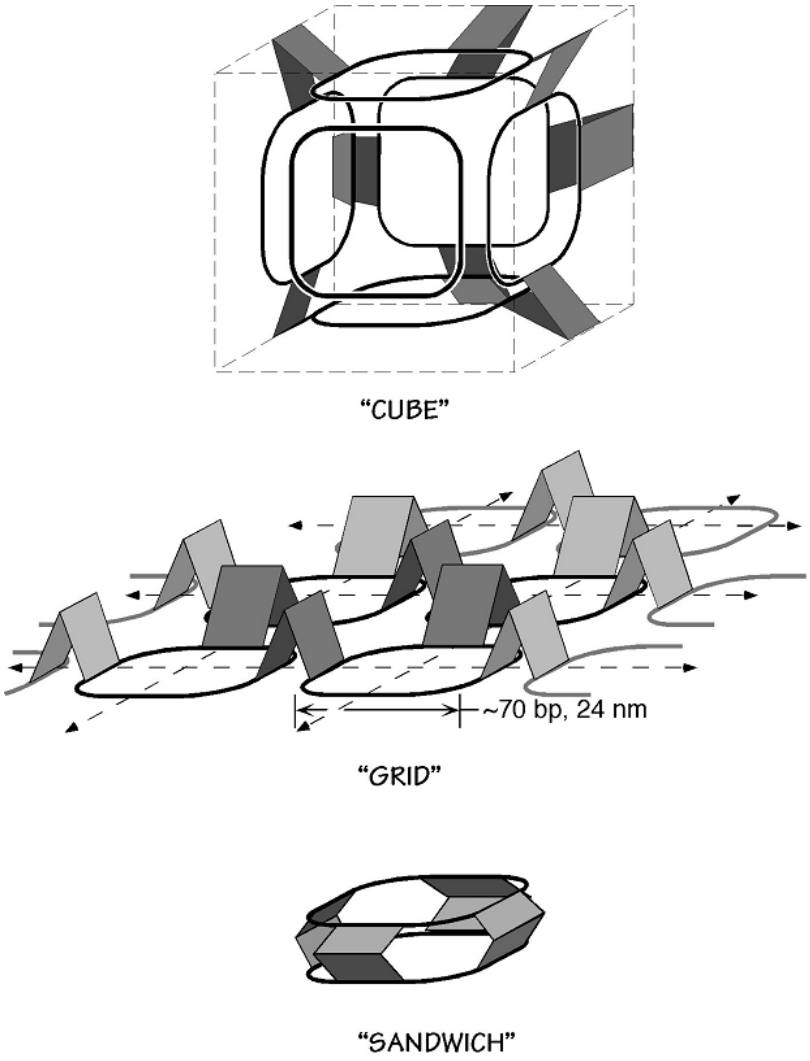


Figure 3. Illustration of proposed cube, grid and sandwich DNA-LacI nanostructures. The Lac repressor is indicated by a V-shape structure while the DNA is indicated by solid lines. The structural representations are roughly to scale, with the protein occupying ~ 20 base pairs per site and each DNA construct being 288 bp. For simplicity the four *lac* repressors at the front of the cube are removed.

The use of circular DNA allows the number of configurations the DNA minicircles can adopt to be limited. A-tracts at the bent corners of designed 4x-OP-ATR and 3x-OP-ATR constructs, which we propose will form "square" and "triangle" geometries upon T4-DNA ligase mediated ring closure, allows for facile cyclization of the constructs (Koo *et al*., 1990). The A-tracts also allow the sides of the proposed "square" and "triangle" constructs to remain straight, with the majority of the bending at the corners: this should stabilize the binding of proteins that do not induce large bending at their binding sites. This offers us the ability to design stable and compact minicircles.

The formation of the three different proposed nanostructures is dependent on the orientation of the *lac* operators. By placing A-tracts at the corners of the "squares" we are able to lock the rotational phase of the operator pseudo-dyad axis relative to the fixed plane of the circle.

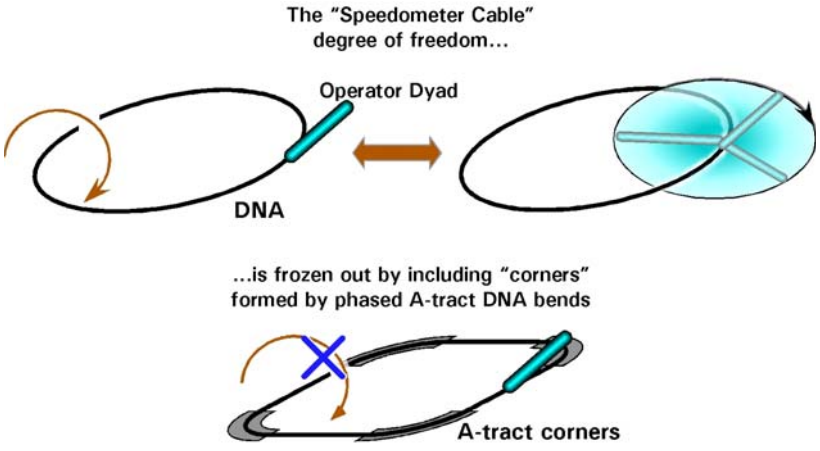


Figure 4. Schematic representing the importance of positioning A-tracts sequences at the corners of the minicircle "square" constructs. A DNA minicircle that lacks A-tract sequences allows free rotation of the pseudo-dyad axis of the operator (blue) with respect to the DNA helix axis, in the same way a speedometer cable has 360 º of freedom. When A-tracts are positioned at the corners of the molecules the rotational phase of the operator sequence is locked, because rotation of the DNA about its helix axis would require inverting the A-tract bends.

This allows the angle at which the LacI binds to the "square" to be controlled. Without the A-tracts, the operator dyad axis would be allowed to rotate freely about the DNA helix axis of the DNA minicircle as shown in Figure 4.

A former graduate student in the Kahn lab, Huazhen Chen, utilized computer simulation to assist in the design process for the nanostructure constructs. Kahn adapted programs written by Levene & Crothers (1983, 1986) that were originally designed for Monte Carlo simulations of DNA to predict and visualize the shapes of the circular molecules. The programs also carry out Monte Carlo simulations to provide a set of possible conformations for the corresponding DNA minicircles. The design process runs as follows: A "Display" script (Kahn & Crothers, 1998) is used to deliver input to other programs that simulate geometries of the three different bent linear DNA constructs. The program works by building DNA chains by adding vectors that represent the helix axis of that particular base pair, one base pair at a time. Rotation matrices that relate the direction of the helix axis from one base pair to the next are generated from Gaussian distributions for roll, tilt and twist angles (Kahn & Crothers, 1998). As a result, a large number of molecules with slightly different shapes are generated for each DNA construct. This assortment of computer-simulated molecules that have different shapes is then filtered for agreement with previously determined cyclization conditions. Several randomly selected cyclized molecules for each proposed shape along with their parameters are then input into the previously described "Display" program to rebuild the DNA constructs and generate PDB files (Chen, 1998). DNA Strider (Marck, 1992) was used for manipulation of DNA sequences and restriction enzyme sites.

An optimal length of 288 bp was chosen for all three DNA constructs. The "corners" of the squares are comprised of four phased A-tracts that give the DNA an overall bend of approximately 72 ° at each corner (Chen, 1998). This should allow the molecules to form stable and flat structures upon T4 DNA ligase mediated ring closure. The "sides" of the squares are composed of the *lac* operator site Osym. By adding or removing one or two base pairs from either side of the *lac* operator we can correctly orient the pseudo-dyad axis. For example, the DNA sequence for the grid construct lacks the two base pairs on the 5 prime side of the *lac* operator and contains two base pairs located on the 3 prime side. In the cube construct there is 1 bp located on the 5' side of the *lac* operator and 1 bp located on the 3' end, while in the sandwich construct there are 2 bp located on the 5' side and none located on the 3' side. Shifting the operator sequence towards one end of the construct, base-pair-by-base-pair, while keeping the length of the side of the square constant alters the phasing of the *lac* operator by approximately 36 ° per base pair (Chen, 1998). By having control over the phasing of the DNA helix, we can orient the binding of the Lac repressor differently for each construct, which is essential to giving each construct its own unique shape (Figure 5).

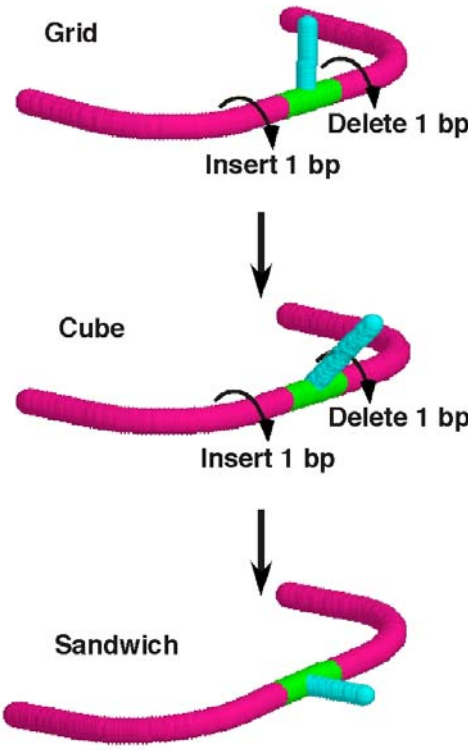


Figure 5. Schematic displaying the control of phasing between the corners and the *lac* operator sequence, which is important for the proper self-assembly of three unique nanostructures. Shown is one side of a DNA minicircle containing A-tract DNA (pink) and the *lac* operator binding site (green). The pseudo-dyad axis is represented in turquoise. By removing and adding base pairs at the beginning and end of the *lac* operator sequence while keeping the spacing between the A-tracts constant we can change the orientation of the pseudo-dyad axis. The operator is moved towards one end of the construct base-pair-by-base-pair, going from grid structure to the sandwich structure.

To date, as discussed in Chapter I, I have successfully established and optimized purification protocols for the grid, cube and sandwich DNA constructs and imaged all three constructs p02, p11 and p20 using both Tapping mode AFM in liquid and air. Currently efforts are underway to image LacI bound to 4x-OP-ATR and 3x-OP-ATR minicircles in order to determine whether DNA-LacI complexes form our proposed nanostructures.

INTRODUCTION II. Fluorescence Resonance Energy Transfer (FRET)

Many processes that occur within the cell require an extensive series of cascade and feedback mechanisms that are involved in biochemical processes such as diffusion, catalysis and binding-induced conformational changes (Jares-Erijman & Jovin, 2003). One way to develop a global map of cellular structure and function is to measure intra- and intermolecular distances and diffusion rates of various cellular components. FRET has the inherent ability to measure *in vivo* and *in vitro* intra- and intermolecular distances of 1-10 nm which has allowed scientists to study structural and chemical dynamics of cellular processes that could not be studied by classical biochemical and biophysical techniques. FRET has become a widely used tool for elucidating structural information on a variety of materials such as cell surfaces, nucleosomes, chromatin, nucleic acids and proteins. It has also been used to study dynamic events in real-time such as human immunodeficiency virus (HIV-1) protease activity and actin assembly (Selvin and references therein, 1995).

FRET, initially described by Theodor Förster in the 1940s, is described by theory based on principles of electromagnetism and quantum mechanics. In general the phenomenon of FRET is best described as a photophysical event that involves the transfer of energy in a nonradiative process (via long-range dipole-dipole coupling) from an excited donor molecule to an acceptor molecule (Cantor & Schimmel, 1980). In order for FRET to occur there must be sufficient overlap between the absorption spectra of the acceptor fluorophore and the emission spectra of the donor fluorophore (Figure 6A-B). The donor and acceptor transition dipole orientations must also be

approximately parallel as shown in Figure 6D (see page 23), and the distance between the donor/acceptor pair must be short enough (Figure 6C).

In steady-state FRET experiments efficiency of energy transfer (E) is defined as the fraction of donor molecules de-excited via energy transfer to the acceptor. There are several different experimental and theoretical ways used to determine E for experimental systems. The first and most commonly used method for determining E is to measure the decrease in the fluorescence intensity of the donor (due to some of the energy being transferred to the acceptor molecule):

$$E = \left(1 - \frac{I_{DA}}{I_D}\right)\left(\frac{1}{f_A}\right) \quad \text{Equation 1}$$

where I_{DA} is the donor fluorescence intensity in the presence of the acceptor, I_D is the donor fluorescence intensity in the absence of the acceptor, and f_A is the fractional labeling with acceptor (Lakowicz, 1999).

A second way to experimentally determine E is by monitoring the enhancement of fluorescence emission for the acceptor as shown in equation 2:

$$E = \frac{I_{AD} \times (q_D / q_A)}{I_{DA} + I_{AD} \times (q_D / q_A)} \quad \text{Equation 2}$$

where I_{AD} is the integrated area of the sensitized emission of the acceptor, I_{DA} is the integrated area under the donor emission curve in the presence of the acceptor, and q_D and q_A are the quantum yields of the donor and acceptor. The labeling efficiencies of the DNA molecules are hidden in the q_D and q_A terms for this calculation as shown in Equation 3.

$$\frac{q_D}{q_A} = \frac{I_D}{I_A} \times \frac{\varepsilon^A(\text{X nm}) \times f_A}{\varepsilon^D(\text{X nm}) \times f_D} \quad \text{Equation 3}$$

where I_D and I_A are the integrated fluorescence intensities of donor and acceptor in the absence of quenching or energy transfer. ε^D and ε^A are their extinction coefficients, and f_A, f_D are the labeling efficiencies of the fluorophores. X indicates the extinction wavelength used in the experiments, typically chosen to be the donor excitation maximum. Since we cannot determine the absolute quantum yield of the donor and acceptor we determine the ratio of (q_A/q_D), and from these values can substitute them back into equation 2 to determine accurate energy transfer efficiency values.

The third way is to measure a decrease in the fluorescence lifetime of the donor in the presence of the acceptor, given by the following equation:

$$E = \left(1 - \frac{\tau_{DA}}{\tau_D}\right)\left(\frac{1}{f_A}\right) \quad \text{Equation 4}$$

where τ_{DA} is the donor lifetime in the presence of the acceptor, τ_D is the donor lifetime in the absence of the acceptor, and f_A is the fractional labeling with acceptor (Lakowicz, 1999). Once the efficiency of energy transfer has been calculated, the distance between the two fluorophores can be determined using equation 5:

$$E = \frac{R_0^{\,6}}{R_0^{\,6} + r^6} \quad \text{Equation 5}$$

where r is the distance between the donor and acceptor molecules, and R_0 is the Förster distance, the distance at which 50 % efficiency of energy transfer occurs. R_0 has a constant value for each donor and acceptor pair, and is calculated from several different parameters as shown in Equation 6:

$$R_0 = 2.11 \times 10^{-2} \bullet [\kappa^2 \bullet J(\lambda) \bullet \eta^{-4} \bullet Q_D]^{1/6} \quad \text{Equation 6}$$

where $J(\lambda)$ is the overlap integral in the region of the donor emission and acceptor absorbance spectra (with the wavelength expressed in nanometers), Q_D is the quantum yield of the donor, n is the index of refraction (typically 1.3-1.4), and κ^2 is a geometric factor related to the relative angle of the two transition dipoles (Selvin, 1995). For most FRET systems it is assumed that the fluorophores are oriented randomly, which gives a κ^2 value that averages to 2/3. For this reason FRET has the ability to measure only relative and not absolute distances. In our experimental system we assume that the fluorophores are randomly oriented and are allowed to rotate freely due the length of the C6 linker between the nucleotide base of attachment and the fluorophore. Energy transfer is very sensitive to distance changes. When r is approximately 50 % of R_0, the FRET efficiency is near the maximum and shorter distances cannot be reliably determined. When the donor-acceptor distance exceeds the R_0 value by 50 %, the slope of the efficiency vs. distance curve is so shallow that longer separation distances are not resolved, as exhibited in Figure 6C. Because of the strong sixth-power dependence of transfer efficiency on distance, measurements of the donor-acceptor separation distance are only reliable when the donor-acceptor distance is within a factor of 2 of the Förster distance.

Previous experiments conducted in the Kahn lab (Edelman *et al*., 2001) utilized bulk FRET to assess different DNA looping geometries of DNA-LacI complexes. Results confirmed the existence of a "closed" loop form (giving a net 70 % efficiency of transfer) but did not definitively confirm the presence of an "open" loop form or whether there existed populations with intermediate LacI-

DNA geometries. We applied single-molecule FRET (SM-FRET) to the DNA-LacI loops in order to further assess looping geometries.

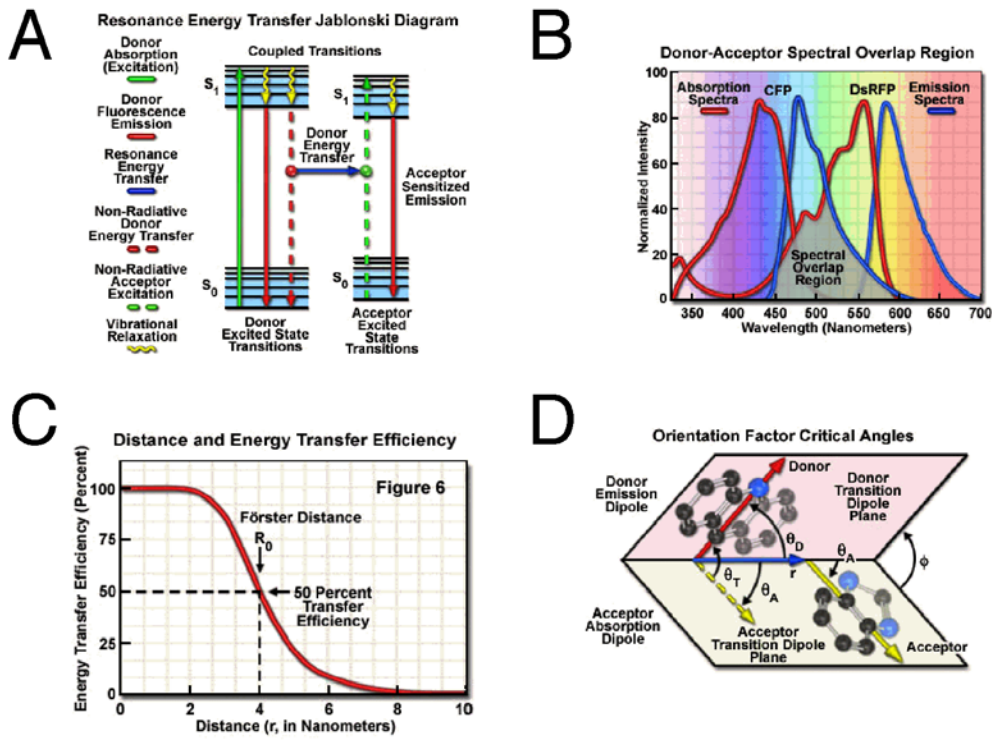


Figure 6. A Jablonski diagram illustrating interaction between donor emission and acceptor absorbance in FRET. Absorption (green arrow) and emission (red arrow) spectra are represented by straight vertical lines, while vibrational relaxation is indicated by yellow arrows. In the presence of an acceptor, the donor fluorophore can transfer excited state energy directly to the acceptor without emitting a photon (blue arrow). B. The absorption (red) and emission (blue) spectra of cyan fluorescent protein (CFP) and red fluorescent protein (RFP) are shown above. The region of overlap between the emission spectra of CFP and absorption spectra of RFP is represented in gray. C. Illustrates the relationship between energy transfer efficiency and the distance separating the donor and acceptor fluorophores. The efficiency rapidly increases to 100 percent as the separation distance decreases below R_0, and conversely, decreases to zero when r is greater than R_0. When r is approximately 50 % of R_0, the resonance energy transfer efficiency is near the maximum and shorter distances cannot be reliably determined. D. The dependence of the orientation factor $\kappa^2 = (\sin\theta_D \sin\theta_A \cos\phi - 2\cos\theta_D \cos\theta_A)^2$, on the relative orientations of the donor emission dipole and the acceptor absorption dipole. θ_D and θ_A are the angles between the donor emission dipole and the acceptor absorption dipole and the vector joining the donor and acceptor, and ϕ is the angle between the planes containing the two transition dipoles. (Figure and legend adapted from www.olympusfluoview.com)

Single Molecule Spectroscopy

Our understanding of molecular interactions in chemistry is usually described at the molecular level. These understandings have historically been derived from ensemble-averaged measurements on billions of molecules, due in part to experimental limitations in instrumentation (Xie & Trautman, 1998). Not until recently (within the last decade or so) has the high-impact introduction of single-molecule spectroscopy (SMS) allowed researchers to study individual molecules within a heterogeneous mixture of trillions of other molecules. Single Molecule Fluorescence Spectroscopy (SMFS) has given scientists the capability to measure chemical and physical properties, as well as intra- and intermolecular distances of individual molecules with a detection limit of one molecule $\approx 1.66 \times 10^{-24}$ moles (a guacomole) of molecules of interest. In SMFS experiments a single molecule can act as a reporter of its nano-environment, providing valuable information on the orientation of functional groups, atoms, ions and electrostatic charges (Moerner & Fromm, 2003). Both SMS and SMFS offer several advantages over classical ensemble techniques. It is widely known that bulk FRET measurements give a great deal of information regarding average values for a given parameter. However, SMS removes the need for ensemble-averaging of data allowing one to construct a histogram of the actual distribution of values for a given parameter. By studying distributions researchers are able to discover the existence of separate populations of molecules within a heterogeneous sample (Moerner & Fromm, 2003). SMS experiments allow for the measuring of single molecule trajectories, which provide valuable information regarding statistical and dynamic information on both

equilibrium and non-equilibrium systems. SMS has an added benefit of being able to study time-dependent processes in non-equilibrium environments without the need for synchronization of a population of molecules (Garcia-Parajo *et al*., 2001). The benefits of SMS outlined above have made this technique a valuable tool for studying biological systems that could not be previously studied by classical ensemble-averaged techniques.

There are several different types of SMS scanning methods, including confocal microscopy, near-field scanning optical microscopy, wide field scanning microscopy and confocal fluorescence microscopy. We have utilized confocal fluorescence microscopy to study our specific system because it offers several benefits over other techniques. Confocal fluorescence microscopy has a high degree of sensitivity for measuring fluorescence. With very low background noise this technique offers noninvasive detection of molecules immobilized on a surface or interface, immobilized in an agarose gel matrix or freely diffusing in solution tens to hundreds of microns above the surface plane. Having the ability to filter out out-of-focus light as well as undesired excitation light with a variety of commercially available filters allows for the detection of single photons. These features, together with high sensitivity and experimental simplicity, have made confocal fluorescence a powerful microscopy tool for studying complex biological systems.

In a standard confocal fluorescence microscope setup a laser beam is reflected off a dichroic beamsplitter through the microscope objective and is focused into a liquid sample with the focal spot located ~10 μm from the glass liquid interface. Emitted fluorescence (in the form of fluorescence bursts) is observed when molecules

traverse the beam. Photons emitted in a downward direction are collected by the same microscope objective and directed through a holographic notch filter to remove scattered excitation light, only allowing the fluorescence emission light to pass through. This light is then directed to a dichroic mirror. The long wavelength emission from the acceptor and the shorter wavelength emission from the donor are transmitted to two separate avalanche photodiodes (APDs), allowing for the number of photons from both the donor and acceptor to be counted. The two data streams generated are stored and analyzed on a PC.

Our goal for using SMS was to assess the proposed looping geometries for the 9C14-LacI molecule freely diffusing through, or immobilized onto or above a surface. By allowing the 9C14-LacI complexes to freely diffuse through solution, a large population of molecules can be studied in a short amount of time. This allowed us to study an unbiased equilibrium distribution of molecules. Conversely, by immobilizing the 9C14-LacI complexes, one individual molecule can be studied for an extended period of time, allowing us to study the dynamic processes of individual molecules.

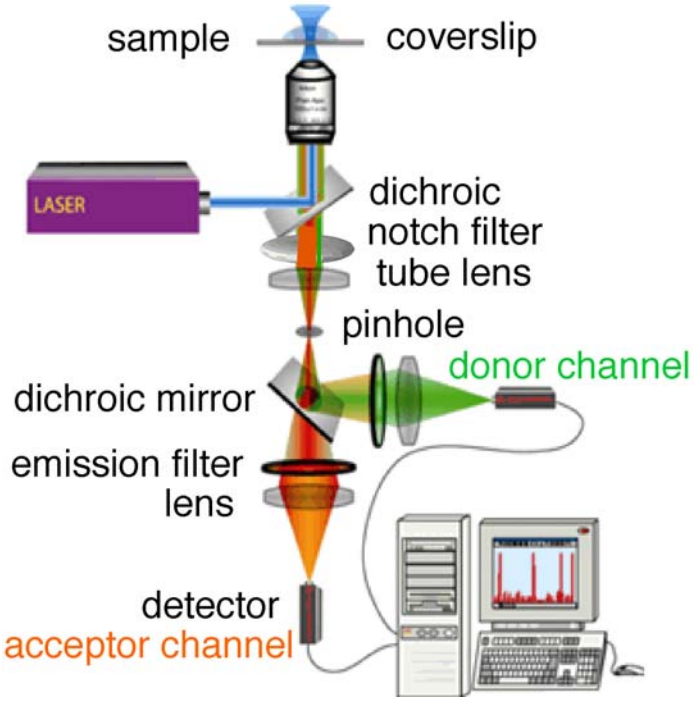


Figure 7. A basic single molecule microscope setup is as follows. A laser is reflected off a dichroic beamsplitter into the back of a high numerical aperture (n.a.=1.2) microscope objective lens. The laser passes through the objective lens where it is focused ~10 μm above the liquid glass interface into the sample. Molecules diffusing into the focal volume emit photons that are collected by the same objective. The scattered excitation light is passed through a notch filter, allowing only fluorescent emission light to pass through. The fluorescence photons are separated according to wavelength by a second dichroic beamsplitter and directed to one of two single-photon-counting avalanche photodiodes. One APD detects photons from the donor fluorophore and one detects photons from the acceptor fluorophore. These two detectors create two data streams that are stored and analyzed on a PC (Figure adapted with permission from http://www.public.asu.edu/~sulin/setup.html.)

AFM

Scanning tunneling microscopy (STM) was developed in the early 1980s by G. Binnig and H. Rohrer (Binnig & Rohrer, 1987). The introduction of STM revolutionized the way that scientists could visualize macromolecules, and subsequently earned the Nobel Prize for its discovery. STM utilizes properties of electrically conducting samples to image at atomic resolution (Hansma *et al.*, 1988). A small bias voltage is applied between an atomically sharp tip and the sample. When the tip is brought very close (≤10 Å) to the sample a tunneling current flows across the gap between the tip and the sample. The tunneling current is monitored by a feedback loop that keeps the current, and hence the distance, between tip and sample constant. When the tunneling current exceeds its preset value, the distance between tip and sample is increased. When the current falls below its preset value, the feedback decreases the distance. As the tip move across the surface the change in current is monitored to provide an image of the surface topography.

Several years after the introduction of the STM, the Atomic Force Microscope (AFM) was introduced. Using AFM, non-conducting inorganic and biological samples can be imaged in real time at molecular and atomic resolution (Ratneshwar & Scott, 1994). Carlos Bustamante, a leading researcher in the field of biological AFM, has pioneered new and exciting AFM techniques that have been used to provide valuable information on a number of different molecular and cellular mechanisms. For example, Bustamante and colleagues visualized the elongation complexes of *E. coli* RNAP and *Saccharomyces cerevisiae* RNAP III in order to address conflicting theories on the extent of wrapping and bending of DNA. These

studies showed that the DNA is wrapped against the polymerase during transcription elongation, and that the growing RNA transcript is extruded from the protein away from the DNA (Rivetti *et al*., 2003). However, it is important to note that there are still conflicting opinions on the extent of DNA wrapping.

A schematic representation of how the AFM works is shown in Figure 8. The tip, which is attached to a cantilevered spring, moves across the surface of the sample and is deflected by the interaction forces between tip atoms and sample atoms. Due to the specific force constants associated with cantilevers, the cantilever can respond to small forces exerted by individual atoms of the sample. The forces sensed by the cantilever are transduced to generate molecular images of the sample topography (Ratneshwar & Scott, 1994).

The AFM is comprised of several major components. The AFM head component houses the controls for the photodiode adjust. The four quadrant photodiode enables both the normal and lateral motions of the tip to be differentiated. It's responsible for measuring tip displacement by comparing the relative intensity of the reflected laser light in each quadrant. The laser *x* and *y*-axis adjust controls are used to align the laser location onto the tip. The head *x* and *y*-axis adjusts allow for the proper positioning of the sample in reference to the laser position. The piezo scanner is responsible for moving the sample in three orthogonal directions *x*, *y* and *z*, which allows the sample to be in constant contact with the AFM tip (Morris, Kirby & Gunning, 1999).

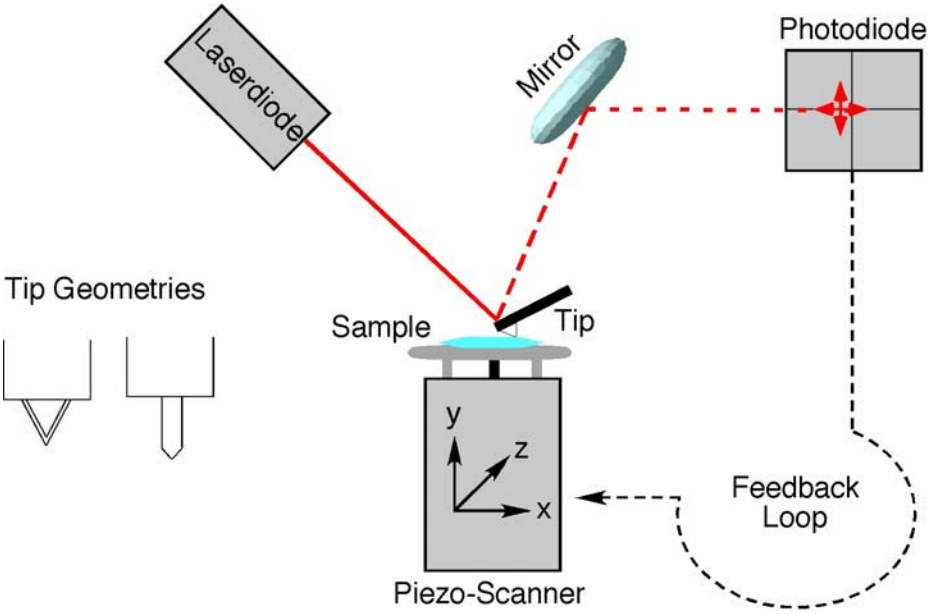


Figure 8. Atomic Force Microscope setup. In tapping mode AFM, a laser beam is focused onto an oscillating cantilever (tip) with a characteristic frequency (generally on the order of hundreds of kHz). The two different common tip geometries are displayed, the triangular geometry (left) and the beam geometry (right). When the oscillating tip comes in contact with the sample surface it bends and the laser beam is reflected off into a photodiode detector. The photodiode is split into four sections enabling both normal and lateral motions of the tip to be differentiated. A measurement of tip displacement can be achieved by comparing the relative intensity of the reflected laser light in each quadrant. The feedback loop is essential for maintaining the appropriate amplitude between the tip and the sample. The *x, y* and *z* displacements of the piezoelectric scanner are recorded and displayed to produce an image of the sample surface (adapted from Morris, Kirby & Gunning, 1999).

The most common types of tip materials used for AFM imagining are silicon, silicon nitride or diamond. There are generally two basic cantilever geometries for imaging samples. The first is a triangular geometry, which is used to minimize torsional motions when imaging. The second tip is commonly referred to as a beam geometry tip (Figure 8), and is used for imaging frictional properties of a sample (Morris, Kirby & Gunning, 1999).

There are several different modes of detection used in AFM experiments. Contact mode AFM requires that a tip be scanned across the sample making sure to keep in contact with the surface all the time. The cantilever is brought into contact with the sample surface where a piezo scanner induces a repulsive force with a mean value of 10^{-9} N on the cantilever (tip). By driving the cantilever against the sample in the *z* direction the deflection of the cantilever, which is induced from the tip coming in contact with a sample surface, is sensed and compared in a feedback amplifier to a desired deflection value set by the experimenter. If the measured deflection value varies from the preset value, the feedback amplifier applies a voltage to the piezo scanner to raise or lower the sample relative to the cantilever in order to restore the preset value. The voltage that the feedback amplifier applies to the piezo scanner is a measure of the height of features on the sample surface. The image is generated by plotting the *z* correction signal from the feeback loop against *x* and *y* (Morris Kirby & Gunning, 1999). Contact mode is generally not used for imaging biological or soft materials due to its inherent ability to damage or alter the topography of the sample. For this reason, we have employed tapping mode AFM,

which allows for high resolution imaging of biological samples (Hansma & Hoh, 1994).

During tapping mode, the cantilever oscillation amplitude is maintained constant by a feedback loop. When the tip passes over a bump in the surface, the cantilever has less room to oscillate and the amplitude of the oscillation decreases. Conversely, when the tip passes over a depression, the cantilever has more room to oscillate and the amplitude increases. The oscillation amplitude of the tip is measured by a photodiode detector. The photodiode is a semiconductor device that turns the laser light into an electrical signal, as the incident light becomes brighter the electrical signal increases. The photodiode, which is split into four quadrants, enables both the normal and lateral motions of the tip to be differentiated. By comparing the relative intensity of the reflected laser light in each quadrant, a value for the tip displacement can be obtained. These values are then used to generate an image of the sample surface (Morris, Kirby & Gunning, 1999).

There are several different types of image maps produced for each scan. The phase map, which is measurement of the change in the phase of oscillation, is used to discriminate between samples that contain different elastic properties. The amplitude map is used to display the amplitude of the cantilever at each point in the scan, which can yield topographic information about the sample. A topography map displays the topography of the sample (PicoSPM II AC Mode User's Manual).

We have utilized the ability of the AFM to characterize the DNA-LacI nanostructures.

CHAPTER I: Synthesis and Characterization of DNA Constructs

Design Rationale for Construction of DNA Components of Nanostructures

The experimental design for the self-assembly of two-dimensional and three-dimensional LacI-DNA nanostructures is based on several important concepts. First, using dsDNA minicircles allows us to limit the number of possible geometries the molecule can adopt compared to its linear counterpart. This provides us with a stiff, stable, and well-characterized starting material. Second, the formation of the LacI-DNA nanostructures should be a cooperative process, by analogy to the stabilization of a looped complex between two operator sites when one site is localized by other proteins that induce bending (Brenowitz *et al*., 1991; Lobell & Schleif, 1991; Santero, *et al*., 1992). The last important design concept is that the formation of the nanostructures is dependent on the ability to control the orientation of the *lac* operators with respect to the plane of the DNA in order to form the three different nanostructures. Utilizing these key design aspects offers the potential for the self-assembly of the grid, cube and sandwich.

Previous students of the Kahn lab, Huazhen Chen and Katrina Pei, were both instrumental in the design, assembly and cloning of the three linear precursor DNA molecules denoted p02, p11 and p20, which correspond to the grid, cube, and sandwich constructs respectively. pBluescript II KS(+) is a widely used 2961 bp phagemid derived from pUC19 (Stratagene). Briefly the construction of the linear precursor DNA molecules was as follows: Sequences designated LU (left unique sequence), OP (operator containing sequence), ATR (A-tract region) and UATR

(unique restriction site and A-tract sequence) were designed as individual oligonucleotides to be ligated together by T4 DNA ligase. This resulted in a linear product which contained the desired 288 bp nanostructure sequence with 5' and 3' flanking regions containing Bam HI and Hind III sites for cloning (Figure 9). A sequence designated NOP (not operator) was designed to fit into the same position as the OP sequence, but does not have the ability to bind the LacI. In principle, ligating the NOP sequence into the linear precursor molecule would permit the formation of squares that lacked binding sites, allowing for further control over the overall global size of the 2D and 3D nanostructures.

Previous transformation and maxiprep attempts performed by Pei, unexpectedly resulted in the generation of a heterogeneous mixture of four different plasmids for each construct. Similarly, PCR amplification attempts also generated a heterogeneous mixture of products for each construct. My initial introduction to this project was focused on trying to establish successful protocols to amplify the desired DNA constructs.

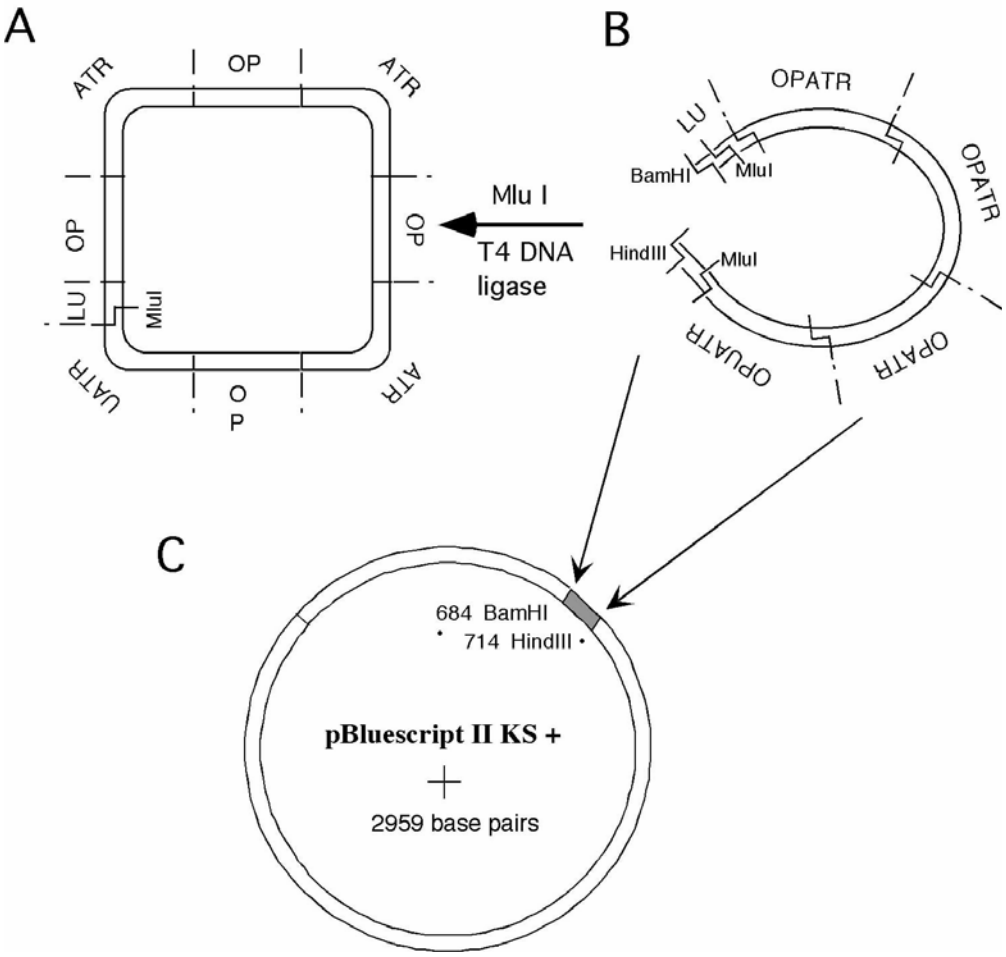


Figure 9. Design of minicircle DNA constructs. (A) Each square contains four alternating A-tract containing regions and operator containing regions, with a unique Mlu I site for cyclization. (B) The linear precursor of the circular molecule shown in A is cloned into the Bam HI and Hind III sites of the vector and transformed into *E. coli* for amplification. (C) pBluescript II KS (+) with insert representing the three plasmids p02 (grid), pll (cube) and p20 (sandwich) (Chen, 1998).

Materials

Oligonucleotides were purchased from Integrated DNA Technologies, Inc. (Coralville, IA). A clone expressing Lac repressor was generously supplied by Dr. Michael Brenowitz (Albert Einstein College of Medicine), and the protein was expressed and purified by Larry Edelman as described by Brenowitz *et al*. (1991). Unless otherwise specified all reagents were purchased from Fisher Scientific (Pittsburgh, PA). All restriction enzymes and T4 DNA ligase were purchased from New England Biolabs Inc. (Beverly, MA). The FailSafe™ PCR kit was purchased from Epicentre Technologies (Madison, WI).

Methods

Transformations

Transformations of plasmid clones containing desired nanostructure linear precursor molecules were performed into several different *E. coli* cell lines including ElectroMAX™ Stbl4™ (Invitrogen), XL-1 Blue (Strategene) and SURE® (Strategene) competent cells lines. Transformations were performed via standard electroporation protocols. Approximately 50 ng of DNA was mixed with 50 μL of electrocompetent cells and pipetted into a prechilled cuvette that had been incubated on ice for 5 minutes. The cuvette was inserted into the electroporator and pulsed at 17 kV/cm. After the cuvette was pulsed 1 mL of SOC medium is added (SOC is made by supplementing 10 mL of SOB buffer (2 % w/v bacto-tryptone, 0.5 % w/v bacto-yeast, 10 mM NaCl, 2.5 mM KCl, 10 mM $MgCl_2$, and 20 mM $MgSO_4$) with 100 μL of 1M $MgCl_2$, 100 μL of 1M $MgSO_4$, and 100 μL of 2 M glucose. The mixture is gently pipetted to mix, placed into a glass culture tube and shaken for 37 °C for 45 minutes.

The cells are then plated onto LB agar containing appropriate antibiotics and grown overnight at 37 °C for 14-16 hours.

Miniprep of DNA

Standard minipreps for all cloning procedures were performed using the QIAprep® Spin miniprep kit. The special applications purification protocol was used for the purification of low-copy number plasmids. The purification procedure consists of three basic steps. The first step involves lysing the bacterial cells with a modified alkaline lysis buffer and clearing the lysates by centrifugation. In the second step DNA is bound to a silica-gel membrane in a high-salt buffer and eluted in a low-salt buffer. Optimized buffers in the lysis step combined with efficient binding of DNA to the silica-gel membrane ensure that only DNA is absorbed to the membrane while RNA and cellular proteins are washed out in the flow-through volume. The final step involves washing the membrane with a buffer to remove endonucleases and other proteins, followed by elution of the purified plasmid DNA in a low-salt buffer. According to the manufacturer, bacterial cells from 10 mL of overnight culture are spun down in a centrifuge to pellet the cells. The supernatant is decanted and the pellet is resuspended in 250 µL of Buffer P1 containing RNase A (buffer conditions are not given in the manufacturers documentation). 250 µL of Buffer P2 are added to the solution and gently inverted 4-6 times to mix (vortexing of solution results in shearing of chromosomal DNA) followed by addition of 350 µL of buffer N3 and gentle mixing. The solution is centrifuged at 13,000 rpm for 10 minutes, and the supernatant is added to the spin column and centrifuged for an additional 60 seconds. The spin column is washed with 0.5 mL of buffer PB (to

remove trace amounts of nuclease) and centrifuged for 60 seconds. The column is washed by adding 0.75 mL of buffer PE centrifuged for 60 seconds and the flow-through is removed and discarded. The column is then centrifuged for an additional minute to remove residual wash buffer. The miniprep column is placed in a new eppendorf tube and 50 µL of buffer EB (10 mM Tris-Cl, pH 8.5) or water heated to 70 °C is added to the column, let sit for 1 minute and then centrifuged for 1 minute to give purified plasmid.

Phenol/Chloroform Extraction

An equal volume of phenol: chloroform (1:1), pH 7.9, was added to the DNA sample, vortexed and then centrifuged at 12,000 rpm for 3 minutes. The top aqueous layer was removed and placed into a new eppendorf tube. An equal volume of chloroform: isoamylalcohol (24:1) was added to the aqueous layer and the process was repeated.

Ethanol Precipitation

Sodium acetate (NaOAc), pH 7.0 was added to a DNA sample at a final concentration of 150 mM. 2.5 volumes of 100 % ethanol (EtOH) was then added to the DNA sample and placed at -80 °C for 1 hour. The sample was then centrifuged at 12,000 rpm for 20 minutes. The EtOH was decanted and 200 µL of 70 % EtOH was added to the tube, and centrifuged again for an additional 15 minutes. The EtOH was decanted and the tube was placed in a Speedvac concentrator for 15-20 minutes under vacuum. The DNA pellet was then rehydrated in an appropriate volume of 1X TE (10 mM Tris-Cl, 1 mM EDTA, pH 7.4) or double distilled water (ddH_2O).

Polyacrylamide Gel Electrophoresis

Polyacrylamide gel electrophoresis was performed as follows: A 6 % gel stock containing 40:1 or 75:1 (acrylamide: bis-acrylamide) in 1X TBE (50 mM Tris, 50 mM boric acid and 1 mM EDTA, pH 8.4), 0.05 % APS and 0.005 % TEMED was prepared. 10 % denaturing gels were prepared with 8 M urea. Long gels were prepared (20 cm x 40 cm, 0.75 mm in thickness) and were run at approximately 21 Watts (50 W for 10 % denaturing gels) for an appropriate amount of time. Gels were stained in either ethidium bromide (0.5 μg/mL) or 1X SYBR® Gold nucleic acid gel stain, Molecular Probes (Eugene, OR). Appropriate bands were excised and eluted.

Gel Elution

400 μL of crush and soak buffer (50 mM NaOAc, 1mM EDTA, pH 7.0) was added to crushed gel slices and the tube was placed on a shaker at 180 rpm overnight, making sure to cover the sample with aluminum foil if the gel slice was stained in ethidium bromide or contained fluorescent primers. The eluant was then transferred into a new tube and concentrated in the speedvac under vacuum. An additional 200 μL of crush and soak buffer was added to the original tube and shaken for an additional hour, and the second eluant was then added into the tube in the speedvac. This gives a final volume of ~ 200 μL at 150 mM NaOAc. The DNA is precipitated as above.

Polymerase Chain Reaction (PCR)

PCR reactions were run on an Ericomp Twin Block System Easy Cycler Series instrument. The FailSafe™ PCR System (Epicentre) was used to perform PCR amplification of our desired constructs. The FailSafe™ PCR system includes a blend of thermostable enzymes, a set of 12 Failsafe PCR buffer premixes (each of which contains dNTPs, buffer, and varying amounts of $MgCl_2$), and the Failsafe PCR enhancer, which contains betaine to significantly enhance PCR performance. Templates for the PCR reactions were obtained from a restriction digestion of the plasmids followed by gel purification. The bands containing the templates were excised, eluted, phenol/chloroform extracted and ethanol precipitated. Templates were rehydrated in 1X TE and diluted to a concentration of 10^8 molecules/µL. Each PCR reaction contained 1 µL of template, 1 µL each of 50 µM primer, 1 µL of Failsafe PCR Enzyme Mix, 25 µL of Premix G and 21 µL of ddH_2O. The PCR cycling conditions were as follows: 94 °C for 30 sec, 55 °C for 30 sec, 60 °C for 1 sec, 66 °C for 1 sec, 72 °C for 1 min. The cycle was repeated 25 times, followed by a final extension step of 72 °C for 5 minutes.

Restriction Digestions

Mlu I Restriction

Mlu I restriction digests of approximately 1-2 µg of each PCR products (P02, P11 and P20) were performed as follows: 1 µL of Mlu I (10,000 units/mL), 10 µL of 10X NEB Buffer # 3, (final concentrations 100 mM NaCl, 50 mM Tris-HCl, 10 mM $MgCl_2$, 1 mM dithiothreitol, pH 7.9 at 25 °C), 1 µL 100X NEB BSA, 38 µL of 1X TE buffer were incubated at 37 ºC for 2 hours. The restriction digests were

phenol/chloroform extracted, ethanol precipitated and then analyzed by PAGE. The bands containing the desired products were excised, eluted, phenol/chloroform extracted and ethanol precipitated.

Pvu II Restriction

Pvu II restriction digest of each plasmid p02, p11 and p20 was performed as follows: 1 µL of Pvu II (10,000 units/mL), 10 µL of 1X NEB Buffer # 2 (final concentrations 50 mM NaCl, 10 mM Tris-HCl, 10 mM $MgCl_2$, 1 mM dithiothreitol, pH 7.9 at 25 °C), 1 µL 100X NEB BSA, 37 µL of 1X TE buffer and 50 µL of 1 µL of 2 µg/µL of each plasmid. The samples were incubated at 37 ºC for 2 hours and were then phenol/chloroform extracted, ethanol precipitated and analyzed by PAGE. The bands containing the desired products were excised, eluted, phenol/chloroform extracted and ethanol precipitated.

Ligation

Ligation reactions to form our desired minicircles were performed as follows: 1 µL of T4 DNA ligase (400,000 units/mL), 2 µL 10X T4 DNA Ligase Reaction Buffer (1X Buffer contains 50 mM Tris-HCl, pH 7.5, 10 mM $MgCl_2$, 1 mM dithiothreitol, 1 mM ATP, 25 µg/mL BSA, and 17 µL of purified PCR product (~ 10 ng/µL). Ligations were carried out at room temperature for 20 minutes followed by phenol/chloroform extraction and ethanol precipitation.

Atomic Force Microscopy (AFM)

Visualization of DNA "squares" and "triangles" as well as 9C14 DNA was performed on either a Digital Instruments Multimode™ or a Molecular Imaging PicoPlus™ Atomic Force Microscope. Imaging was performed using tapping mode in both air and in liquid. Samples were prepared several different ways. Typically for tapping mode in air experiments, 30 μL of a sample containing 1-10 μg/mL of DNA in TE buffer containing 1-15 mM $MgCl_2$ was deposited on a freshly cleaved mica disc and allowed to adsorb at room temperature for 10 minutes. Several washes with ddH_2O were performed to remove unbound DNA and proteins, and samples were then dried with a stream of N_2. For experiments using tapping mode in liquid, samples were prepared the same way except that after the drying step the samples were submerged in ddH_2O or 1X Lac buffer (25 mM bis-Tris, 5 mM $MgCl_2$, 100 mM KCl, 2 mM DTT) for protein imaging.

Several other protocols for adhering DNA to mica involve functionalizing the mica surface with aminopropyltriethoxysilane (APTES). In general, 500 μL of APTES is placed into a cap or vial inside a desiccator with freshly cleaved mica discs for approximately 1 hr. The APTES evaporates into the closed environment, functionalizing the mica surface with a silane bearing an amino group. Further functionalization of an APTES coated surface by glutaraldehyde (Sigma Aldrich) exposure was accomplished by adding approximately 200 μL of 1 mM glutaraldehyde solution in water onto the APTES modified surface for 10 minutes and then washed with ddH_2O. Tapping mode in air was performed with a variety of different Veeco NanoProbe™ tips including NPS (sharpened silicon nitride probes,

spring constant of 0.06-0.58 N/m), and NPST (oxide-sharpened silicon nitride probes with twin tips, 0.06-0.58 N/m). Tapping mode in liquid was performed with OTR4 (sharpened silicon nitride probes, 0.02 - 0.08 N/m).

Results

Substantial effort was devoted to cloning as the primary amplification approach for our nanostructures because it offers several benefits over PCR amplification. Cloning reduces the risk of introducing mutations or mismatches into the desired DNA because of the cell's high precision proofreading enzymes. DNA inserted into a vector can be manipulated by standard mutagenesis protocols. Cloning offers the potential for large-scale amplification and purification of the desired DNA constructs.

Previous transformation and maxiprep attempts by Chen and Pei for each individual plasmid (p02, p11 and p20) unexpectedly resulted in a heterogeneous mixture of four different plasmids for each construct. To establish whether it was the vector or the strain or a combination of both that was the responsible for the cloning problems, several unsuccessful attempts were made to reclone the insert into pACYC184, a 4245 bp low copy plasmid (MBI Fermentas). Several unsuccessful attempts were also made to transform p11 into ElectroMAX™ Stbl4™ cells (Invitrogen) and SURE® cells (Stratagene), both of which are strains used for high efficiency cloning of unstable inserts, including repetitive DNA and retroviral sequences. One possibility is that the source of cloning errors could involve LacI bridging between two *lac* operator sites located on two different plasmids, resulting in the formation of unique DNA geometries that could potentiate recombination or interfere with replication. Several unsuccessful attempts were made to clone the plasmids in the presence of the inducer IPTG in an attempt to avoid forming the putative bridged plasmids.

It is well known that the migration of A-tract DNA is retarded in PAGE due to its large global bend. Constructs that are generally on the order of 300 bp tend to migrate at ~700 bp on a 6% 75:1 native gel. Therefore, DNA ladders can only be used as a reference point for migration of DNA containing A-tracts, not as a true indication of size. Another very interesting phenomenon is the different migration of the same sized products. For example, all the square constructs are 288 bp in length and differ only in phasing of the *lac* operator binding site, but they migrate slightly different. This suggests a change in the global bend, either due to small bends in the *lac* operator DNA or else changes in twist that alter the phasing of the "corners."

In general, one would expect that if the full length plasmids p02, p11 and p20 containing the desired insert (3255 bp) were digested with Pvu II restriction enzyme (our plasmids have two Pvu II restriction sites) the restriction product would give two unique bands, a small fragment containing the desired product (742 bp) and a larger band containing the remaining linearized plasmid (2513 bp). However, when p02, p11 and p20 are restricted with Pvu II there are five distinct bands per plasmid that appear on the gel instead of the expected two bands (Figure 10). Further analysis of the bands suggests that the top band corresponds to the linearized portion of the plasmid (2531 bp), while the second band from the top is believed to be the full-length desired construct (742 bp) that contains the desired 288 bp sequence (four-OP-ATR repeats). The third band from the top of the gel is believed to be the full-length product minus a ~ 72 bp section of DNA that contains one Osym operator site and one A-tract repeat (three-OP-ATR repeats). The discovery of a 216 bp DNA construct containing three Osym operators and three A-tract repeats gives us three

new constructs (P02, triangle, P11, triangle and P20, triangle) that have the ability to form triangular geometries upon T4-DNA ligase mediated ring closure. The fourth band from the top of the gel is believed to be the full-length product minus two Osym operators and two A-tract repeats (two-OP-ATR repeats), while the fifth band is believed to only contain 1 Osym operator sequence and 1 A-tract repeat (one-OP-ATR repeat) as seen in Figure 10. The discovery of the 216 bp (three-OP-ATR repeats) construct provided us with three more unique constructs P02 (triangle) P11 (triangle) and P20 (triangle), each with different phased Osym operator sequences, potentially allowing for the formation of different 2D and 3D macromolecular structures with triangle based geometries. It was apparent at this time that the cloning approach was not working as expected, so it was decided to focus on a PCR amplification approach to obtain the desired DNA constructs P02, P11 and P20 (Figure 11).

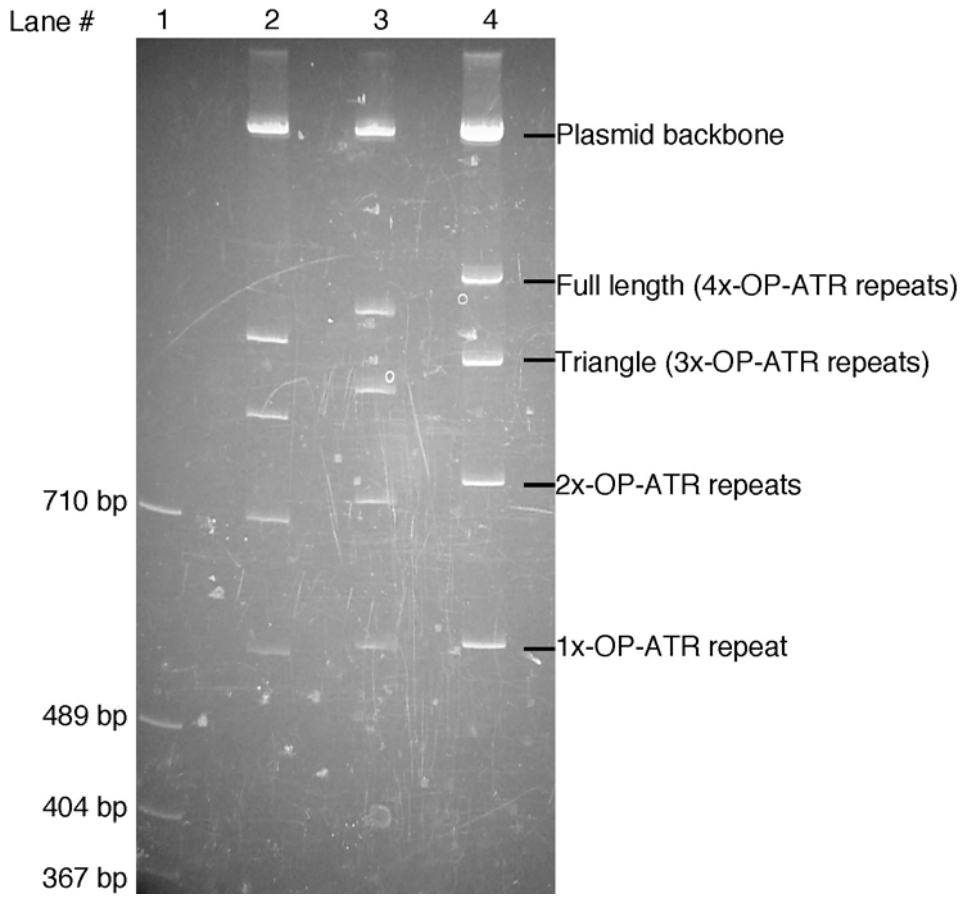


Figure 10. Plasmids p02, p11 and p20 restricted with Pvu II. Lane 1: DNA marker ladder. Lane 2: p02. Lane 3: p11. Lane 4: p20. Each of the plasmids was restricted with Pvu II, and five bands are observed instead of expected two bands. The top band in each lane corresponds to the restricted 2513 bp linear fragment. The full-length construct (288 bp) contains 4 *lac* operator and A-tract regions (4-OP-ATR repeats). The triangle construct (216 bp) contains 3-OP-ATR repeats. The fourth band from the top of the gel (144 bp) contains 2-OP-ATR repeats, while the band furthest down the gel (72 bp) contains 1-OP-ATR repeat.

P02 square DNA PCR Product

gtaatacgactcactatagggcgaattggagctccaccgcggtggcggccgctctagaactagtggatc
CCTG**ACGCGT**CGGG**AATTGTGAGCGCTCACAATT**GAAGTCTTTTTTGCCCGTTTTTTGCCGTTTTTTGC
CCGTTTTTTGATCCTCGGG**AATTGTGAGCGCTCACAATT**GAAGTCTTTTTTGCCCGTTTTTTGCCGTTT
TTTGCCCGTTTTTTGATCCTCGGG**AATTGTGAGCGCTCACAATT**GAAGTCTTTTTTGCCCGTTTTTTGC
CGTTTTTTGCCCGTTTTTTGATCCTCGGG**AATTGTGAGCGCTCACAATT**GAAGTCTTTTTTGCCCGTTT
TTTGCCGTTTTTTGCCCGTTTTTT**ACGCGT**Aagcttggccatcctatcgataccgtcgacctcgatcga
ggggggggcccggtacccagcttttgttccctttagtgagggttaattgcgcgct**tggcgtaatcatggt**
catagctgtttc

P11 square DNA PCR Product

gtaatacgactcactatagggcgaattggagctccaccgcggtggcggccgctctagaactagtggatc
CCTG**ACGCGT**CGGG**TAATTGTGAGCGCTCACAATTA**AGTCTTTTTTGCCCGTTTTTTGCCGTTTTTTGC
CCGTTTTTTGATCCTCGGG**TAATTGTGAGCGCTCACAATTA**AGTCTTTTTTGCCCGTTTTTTGCCGTTT
TTTGCCCGTTTTTTGATCCTCGGG**TAATTGTGAGCGCTCACAATTA**AGTCTTTTTTGCCCGTTTTTTGC
CGTTTTTTGCCCGTTTTTTGATCCTCGGG**TAATTGTGAGCGCTCACAATTA**AGTCTTTTTTGCCCGTTT
TTTGCCGTTTTTTGCCCGTTTTTT**ACGCGT**Aagcttggccatcctatcgataccgtcgacctcgatcga
ggggggggcccggtacccagcttttgttccctttagtgagggttaattgcgcgct**tggcgtaatcatggt**
catagctgtttc

P20 square DNA PCR Product

gtaatacgactcactatagggc*gaattggagctccaccgcggtggcggccgctctagaactagtggatc*
*CCTG**ACGCGT**CGGG**TC**AATTGTGAGCGCTCACAATT**AGTCTTTTTTGCCCGTTTTTTGCCGTTTTTTGC*
*CCGTTTTTTGATCCTCGGG**TC**AATTGTGAGCGCTCACAATT**AGTCTTTTTTGCCCGTTTTTTGCCGTTT*
*TTTGCCCGTTTTTTGATCCTCGGG**TC**AATTGTGAGCGCTCACAATT**AGTCTTTTTTGCCCGTTTTTTGC*
*CGTTTTTTGCCCGTTTTTTGATCCTCGGG**TC**AATTGTGAGCGCTCACAATT**AGTCTTTTTTGCCCGTTT*
*TTTGCCGTTTTTTGCCCGTTTTTT**ACGCGT**Aagcttggccatcctatcgataccgtcgacctcgatcga*
*ggggggggcccggtacccagcttttgttccctttagtgagggttaattgcgcgct**tggcgtaatcatggt***
catagctgtttc

Figure 11. Sequences of P02, P11 and P20 PCR products. Shown are the top strands only from 5' to 3'. PCR products were amplified from a Pvu II restriction fragment of each plasmid (p02,p11 and p20) using primers T7 (indicated in red) and ERP, which is complimentary to this sequence (indicated in blue). The first base corresponds to position 625 in the plasmid sequence. The *lac* operators are in bold. Capital letters indicate the 288 bp desired product and lower case letters represent the remaining sequence. The Mlu I restriction recognition sites are represented in green. The addition or removal of a nucleotide base from the 5' and 3' sides of the *lac* operator is highlighted in magenta. Pushing the operator sequence towards one end of the construct, base-pair-by-base-pair, while keeping the length of the side of the square constant alters the phasing of the *lac* operator by approximately 36 °.

Initially the PCR approach was also quite challenging, as PCR amplification of the desired constructs with Pfu DNA polymerase did not give reproducible results. I then utilized the Failsafe™ PCR system which is designed for amplification of any PCR template, including difficult ones that contain high GC % composition or repetitive sequences. The Failsafe kit includes a blend of thermostable DNA polymerases that contain a 3'→ 5' proofreading enzyme to ensure high fidelity amplification, as well as a set of 12 different reaction buffers that span a range of experimental conditions. We tested all twelve buffers and for our system buffer G was chosen to carry out all PCR reactions. The template DNA for each plasmid p02, p11 and p20 was obtained by restricting each of the plasmids with Pvu II followed by PAGE purification. The bands including the 4-OP-ATR repeats (742 bp PvuII fragment) and 3-OP-ATR repeats (670 bp fragment) were excised and the DNA was eluted, and used as the templates for PCR. PCR followed by Mlu I restriction digestion of all six products P02 (square), P02 (triangle), P11 (square), P11 (triangle), P02 (square) and P02 (triangle) was successful, as shown in Figure 12. The Mlu I restricted products (288 bp and 216 bp respectively for the squares and triangles) were excised, and the DNA was eluted, phenol/chloroform extracted, and ethanol precipitated. The PCR gives deleted and also longer products similar to those seen in cloning experiments. The remaining two bright bands that migrate near the bottom of the gel correspond to the smaller flanking 5' and 3' regions outside of the Mlu I restriction sites. Note that these DNA's migrate with the mobility of fragments approximately double their actual size, as expected for highly bent molecules. Their mobilities vary depending on subtle changes in sequence.

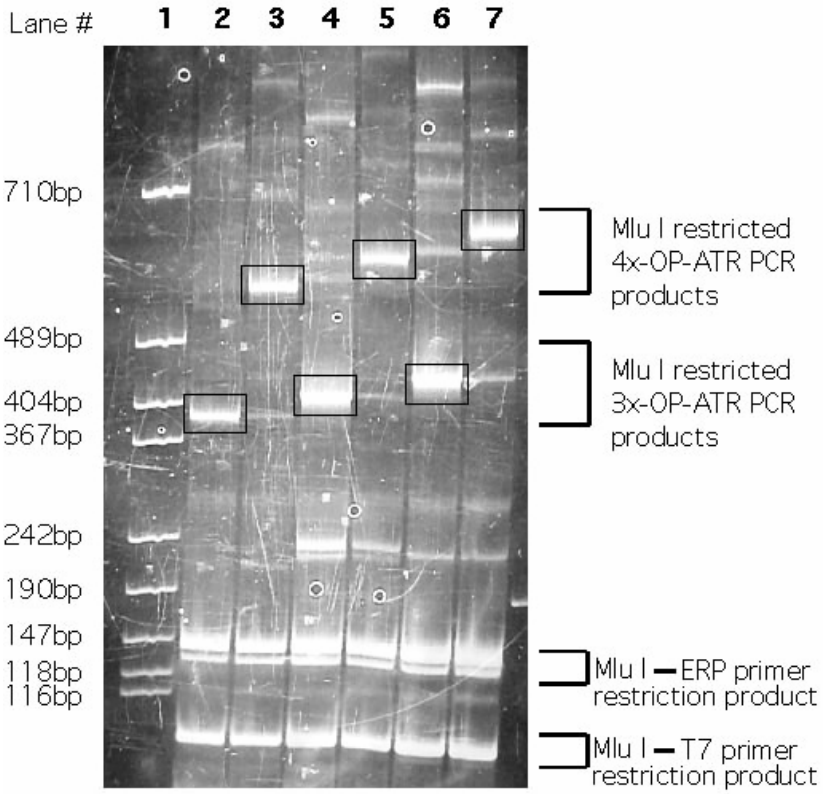


Figure 12. Mlu I restriction digest of PCR products. Lane 1: DNA marker ladder. Lane 2: P02(3-OP-ATR repeats). Lane 3: P02(4-OP-ATR repeats). Lane 4: P11(3-OP-ATR repeats). Lane 5: P11(4-OP-ATR repeats). Lane 6: P20(3-OP-ATR repeats). Lane 7: P20(4-OP-ATR repeats). The restricted PCR products were purified by native PAGE 6 % (75:1) at 21 W for 1 hr and stained in ethidium bromide. The bright bands at the bottom of the gel are the fragments generated from the Mlu I restriction site to the T7 and ERP primer locations. The faint bands located in each lane are truncated products or unrestricted PCR products generated in the PCR.

The 288 bp and 216 bp molecules (4x-OP-ATR and 3x-OP-ATR respectively) were treated with T4 DNA ligase to form the minicircle 4x-OP-ATR and 3x-ATR products. Since these linear DNA constructs are highly bent, cyclization proceeds efficiently as long as the DNA ends are undamaged. A small aliquot was taken from each sample and analyzed by PAGE to make sure that efficient ligation of the 4x-OP-ATR and 3x-ATR products had occurred (Figure 13).

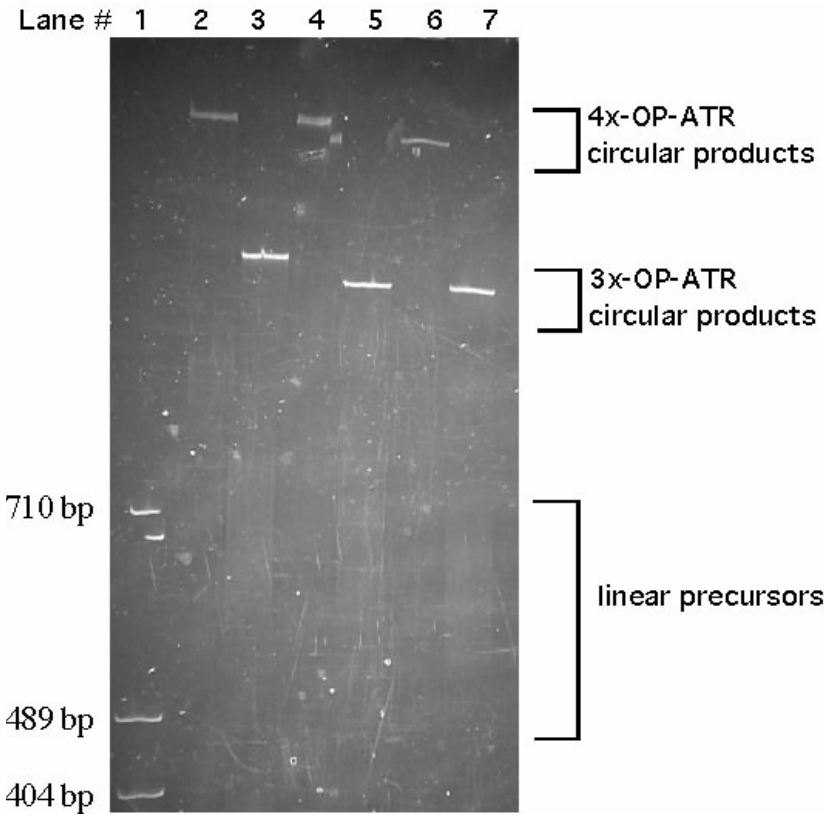


Figure 13. T4 DNA ligase mediated ring closure of 4x-OP-ATR and 3x-ATR products. Lane 1: DNA marker ladder. Lane 2: P20(4x-OP-ATR repeats). Lane 3: P20(3x-OP-ATR repeats). Lane 4: P11(4x-OP-ATR repeats). Lane 5: P11(3x-OP-ATR repeats). Lane 6: P02(4x-OP-ATR repeats). Lane 7: P02(3x-OP-ATR repeats). Ligated products were run by native PAGE 6 % (75:1) at 21 W for 105 minutes.

Development and optimization of PCR and purification protocols was time consuming, and there were additional delays in learning AFM well enough to successfully image DNA and proteins in both air and liquid using tapping mode. Several initial attempts to image our 4x-OP-ATR and 3x-OP-ATR DNA by Transmission Electron Microscopy (TEM) failed, probably because the DNA deposition protocols were optimized for larger DNA molecules.

Much of the development and optimization of protocols for AFM imaging revolved around having enough purified DNA to image. Initial attempts to image plasmid DNA demonstrated the importance of having pure samples, meaning they were free of residual polyacrylamide or aggregates of TBE or $MgCl_2$ (essential components to most deposition buffers for biological imagine using AFM). Otherwise huge aggregates were observed. A variety of different buffers containing different combinations and concentrations of $MgCl_2$, Tris-HCl, HEPES and NaCl were used in order to determine optimal conditions for DNA deposition on mica surfaces. These deposition buffers worked without great success, which coincides with anecdotal evidence that suggests this is typical for others as well. The main problem is with reproducibility: a given set of conditions will often give good images less than 50% of the time. The optimized protocol for imaging both DNA and protein involved the treatment of mica surfaces with aminopropyltriethoxysilane (APTES). APTES is used to functionalize the mica surface with a silane bearing an amino group. The amino group is sometimes methylated after silane binding to mica. In water the positively charged ammonium groups bind to the negatively charged DNA (Lyubchenko *et al*., 1992a). The main advantage of this immobilization technique is

that much more reproducible imaging of DNA-protein complexes can be done, under a broad range of environmental conditions (Shlyakhtenko *et al*., 1999).

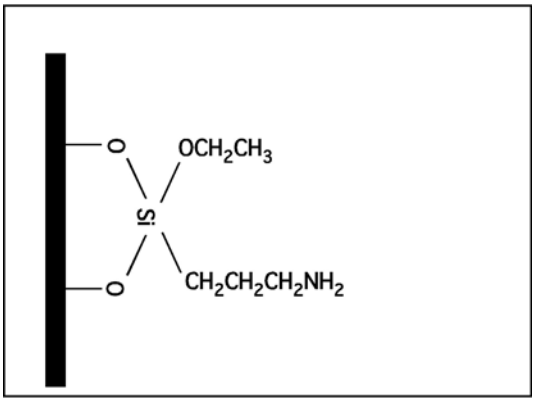


Figure 14. APTES modified mica. APTES is evaporated onto freshly cleaved mica surface. In water positively charged ammonium groups bind to the negative backbone of DNA.

For imaging of the *lac* repressor, glutaraldehyde was added to a modified APTES surface to cross-link the protein to the surface.

An image of the 4x-OP-ATR molecule is shown below in Figure 14. The sides of the molecule are approximately 30 nm in length. The height of the DNA molecules for all experiments depending on sample preparation was ~ 1.5 nm.

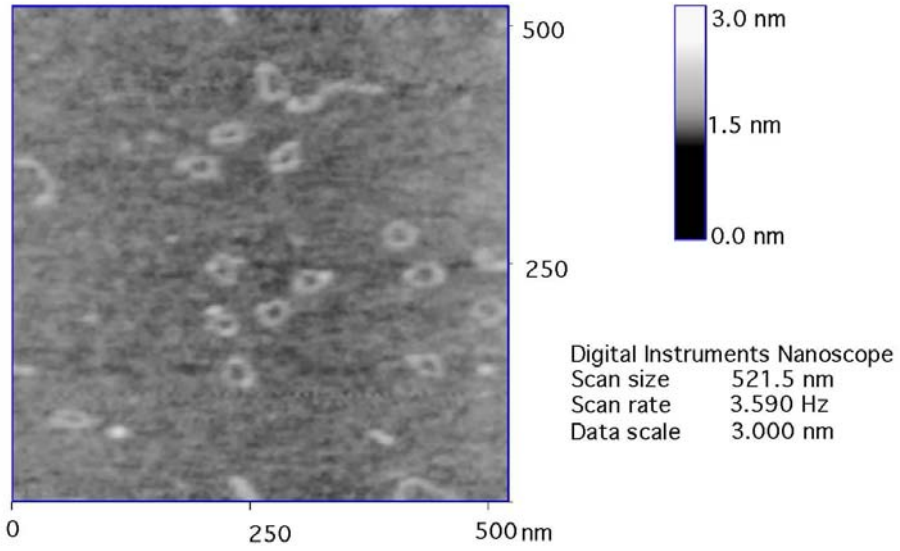


Figure 15. AFM image of 4x-OP-ATR DNA in 1 mM $MgCl_2$ buffer deposited on freshly cleaved mica surface using tapping mode AFM in liquid. Control images of freshly cleaved mica only showed no circular features (data not shown). Images were taken on Larry Sita's AFM)

AFM images of DNA from other labs often show a similar height, which is less than the 2.5 nm expected for a DNA helix. The AFM images shown in Figure 15 (4x-OP-ATR molecules) and Figure 16 (3x-OP-ATR molecules) show the DNA has adopted a variety of conformations including circular and square-like shapes. Some square images have four visible corners. The variety of shapes could be due to several causes, with the most likely being that deposition of DNA onto modified mica surfaces may disturb or alter the DNA shape. A second plausible explanation could

be that the tip itself could be altering the shape of individual molecules. In addition, the molecules may not be as rigid and as expected.

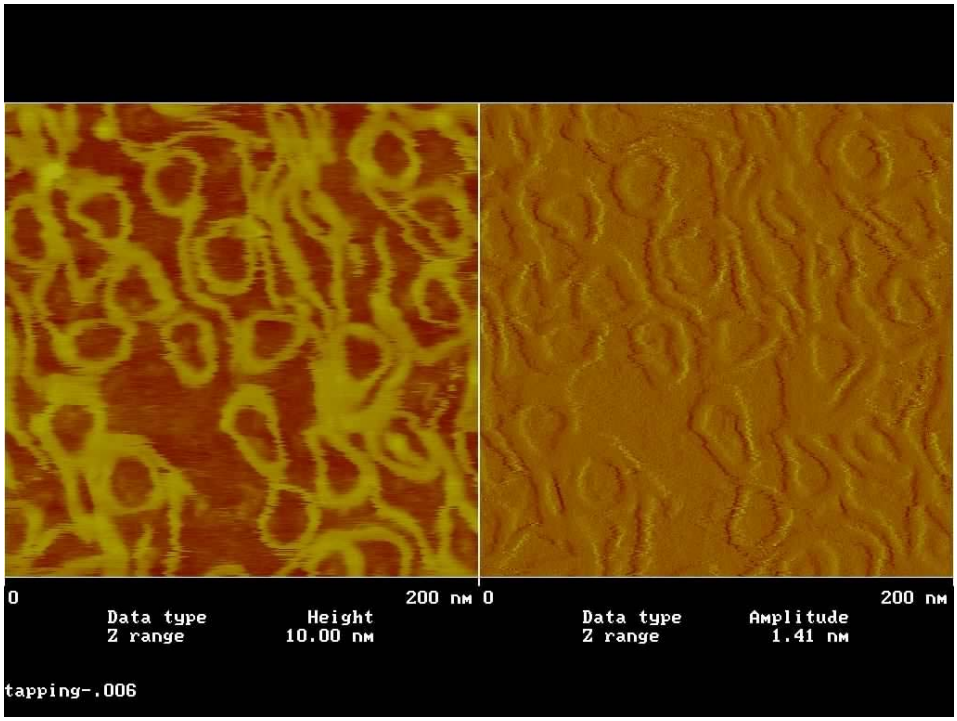


Figure 16. Left image: AFM height map of 3x-OP-ATR circular products. Right image: Amplitude map of the same scan. As the oscillating cantilever scans the surface, the amplitude of the oscillation is reduced. This reduction in oscillation is recorded as a change in amplitude and an image is generated. The 3x-OP-ATR circular DNA product in 1 mM $MgCl_2$ buffer was deposited on APTES modified mica and image using tapping mode AFM in liquid. Control scans of mica showed no features (data not shown).

Any combination of these situations could make it more difficult to obtain or observe uniform self-assemblies of the desired nanostructures. In both the pictures of the DNA minicircle it looks as if there are some unligated linear molecules as well. Gel elution of minicircle DNA is extremely inefficient because the majority of the circular DNA cannot be eluted out of the gel slice. Hence, we do not gel purify our minicircle DNA, and for this reason we see some unligated DNA molecules.

The AFM images of the Lac repressor are also not as revealing as hoped, as shown in Figure 17. One might expect to see the V-shaped Lac repressor geometry, however experiments conducted by Sankar Adhya and colleagues on 599 bp DNA-

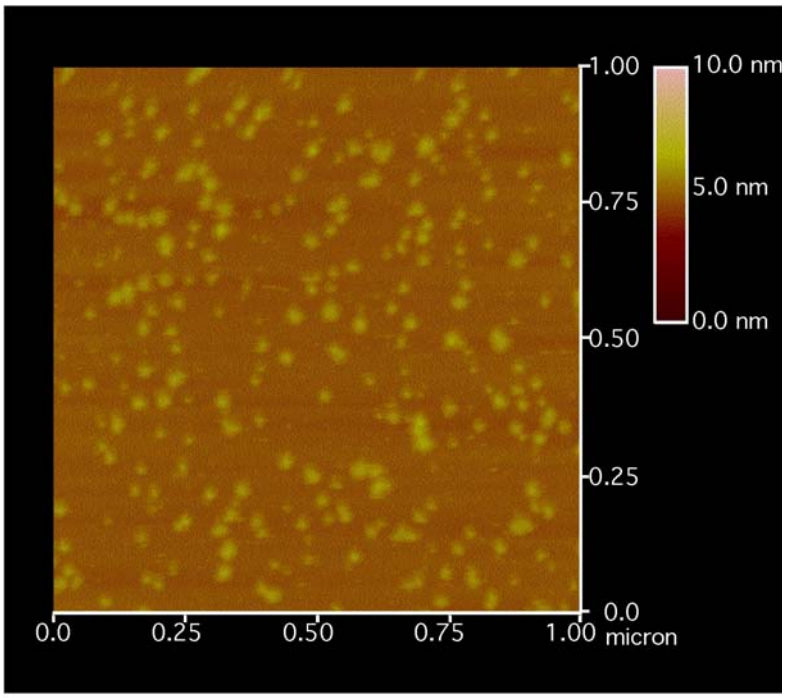


Figure 17. AFM image of LacI in 1X LacI buffer (25 mM bis-Tris, 5 mM $MgCl_2$, 100 mM KCl, 2 mM DTT, pH 7.0) deposited on a glutaraldehyde modified APTES treated mica surface, imaged using tapping mode AFM in liquid.

LacI minicircle loops showed that the overall V-shape of the LacI was not resolvable (Virnik *et al*., 2003), however the relative heights of the spots shown in figure 17 correlate well with the heights of LacI imaged by Adhya. This is most likely due to the radius of curvature of the tip silicon nitride tips used, which is to large to resolve the topography of LacI.

At the present time, we are refining AFM imaging protocols and working to obtain images of DNA-LacI complexes and nanostructures. In view of the difficulty of obtaining minicircles for imaging, we are also pursuing the use of synthetic DNA nanostructures.

Discussion

The notion of DNA-LacI molecules self-assembling into pre-defined 2D and 3D macromolecular nanostructures is unique and exciting. This project relies heavily on the ability to orient the dyad axis of the *lac* operator in order to form three possible different nanostructures. By utilizing A-tract repeats we can control the rotational phase of the helix (Lavigne *et al.* 1992) to form the three different proposed nanostructures as well as design molecules that are small, stable and stiff. The difficulties in sample preparation have thus far prevented us from testing our designs. A lot of effort was spent trying to clone plasmids containing the desired molecules, establish and optimize PCR conditions, and define effective and efficient purification protocols for the 288 bp and 216 bp DNA constructs.

As previously stated, I believe that many of the problems with amplification of our desired DNA molecules correlate to the abundance of A-tracts in the sequences. In general, it is known that cloning DNA constructs containing AT and GC rich regions is difficult (Ishida *et al*., 2002). Although the exact reasons behind this are unclear, several explanations have been suggested. One rationalization is that AT rich sequences can behave as transcriptional promoters in *E. coli*, causing transcription driven by the insert to proceed into the vector, which could interfere with replication or expression of the drug resistant genes (Godiska *et al*., 2001). Another possible explanation is that the Lac repressor could bind two unique operators on different plasmids. This bridged structure could interrupt the essential mechanisms involved in replication; however, we would have expected that cloning in the presence of IPTG would eradicate this problem Another possible explanation

for the cloning problems is the formation of double or single crossover intermediates via intra- or intermolecular recombination.

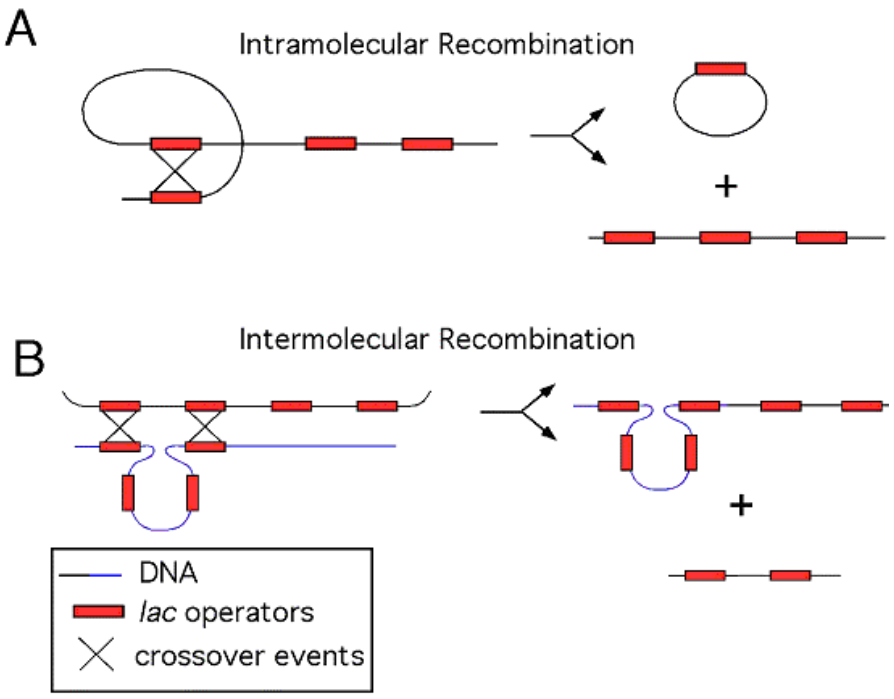


Figure 18. Potential mechanisms for cloning induced errors. A. An example of intramolecular recombination is shown here. A crossover event (X) occurs between adjacent *lac* operators (red boxes) located on the same DNA strand, which results in recombinants with one *lac* operator on one strand and three located on the other. B. An intermolecular recombination event occurs between adjacent operators located on separate DNA strands, which results in recombinants with two *lac* operators located on one strand and four on the other (Figure adapted from Weaver, 2005).

For example, an intermolecular recombination event, where a pair of Holliday junctions is formed between adjacent operators on separate plasmids could result in recombinants with two operators on one plasmid and four operator sequences on the other, or other combinations. An intramolecular recombination event where a

Holliday junction is formed between adjacent operators on the same strand could result in recombinants with one operator on one strand and three on the other, or other combinations (Figure 18). Presumably recombinants with greater than four operators are rapidly lost due to defective replication.

The problems associated with the cloning approach were more serious than those associated with the PCR approach, but there were still several issues that had to be resolved to make the PCR approach work reproducibly and efficiently. First, simply identifying the desired DNA constructs on a gel was difficult due to their retarded migration. In principle, the DNA in my experiments should have been purified after each step, i.e. after PCR, restriction digestion, and cyclization for a total of three gels. However, in the protocol design I wanted to limit the number of gel purification steps, because each entails a significant loss of DNA. Therefore, PCR products were digested with Mlu I before purification, and cyclized products were not purified after ring closure. As a result of this, the purification gel (Figure 12) showed not only the desired 288 bp and 216 bp products but also multiple undesired products from aberrant PCR and/or incomplete restriction digestion.

Obtaining a sufficient quantity of DNA that was pure enough for AFM imaging was a challenge as well. A calculated theoretical yield for the quantity of DNA obtained from a 50 μL reaction is ~10 μg, however the actual yield is much lower, about 1 μg for a 50 μL reaction based on gel band intensity. Because my PCR products were subjected to one gel purification and phenol/chloroform extraction (which resulted in the loss of ~20-30 % of product per round of purification) it is my best estimate that there was on the order of approximately 700-800 ng of circular

DNA for AFM imaging. In order to overcome this obstacle we scaled up our PCR by a factor of four so that PCR was done in 200 μL reactions instead of 50 μL reactions. Although scaling up the PCR provided us with more starting material it also led to more undesired incomplete PCR reaction and restriction digest products.

One possibility for future work on this project utilizes the advances made in the design, engineering and manufacturing of synthetic DNA. Linear DNA molecules that are over 100 bp long can be prepared at a relatively low cost and high purity. Synthesizing 288 bp DNA's from 50-100 bp DNA gives us another potential approach to build our desired DNA constructs. However, problems could arise using this method because it relies on annealing two long oligonucleotide strands together that contain homologous A-tract and operator-repeats. This could potentially result in the annealing of misaligned sequences, which would yield highly undesirable products.

To date, nucleic acids have served as essential building blocks for nanomechanical devices (Mao *et al*., 1999), self-assembled DNA graphs for design of DNA based computers (Sa-Ardyen *et al*., 2003) and ultrasensitive DNA detection devices (Elghanian *et al*., 1997). Recent advances in nanotechnology have utilized DNA-protein complexes to strategically position enzymes, antibodies and gold nanoclusters on ssDNA and dsDNA molecules (Wacker *et al*., 2004). The hope is that novel bio-molecules will one day be used to generate advanced hybrid materials for the design of bio-nanoelectronic circuits.

CHAPTER II. Protein-DNA Loop Characterization Using Single Molecule FRET (SM-FRET).

The Lac repressor when bound to *lac* operator site O_1 and secondary operator sites O_2 or O_3 negatively regulates the transcription of the genes responsible for lactose metabolism in *E. coli*. This mechanism is mediated by the formation of DNA-LacI loops. When the crystal structure of LacI bound to two DNA duplex DNA segments containing two Osym *lac* operator sequences was published in 1996, there was a lot of controversy surrounding the placement of the modeled catabolite activator protein (CAP) inside a positive supercoiled loop that wrapped away from the repressor (Perros, Steitz, Fried, and Hudson in the same issues of science, 1996). Almost a decade earlier Kramer *et al*. (1988) and Whitson *et al*. (1987) had shown that DNA loops responsible for transcriptional repression and LacI binding were stabilized by negative supercoiling and not positive supercoiling. Based on their structure of the core domain of LacI, Friedman *et al.* (1995) had previously proposed the loop geometry that was most likely the actual loop geometry in solution, was the Wrapping Towards (WT) or the Simple Loop (SL) geometry, and not the Wrapping Away (WA) geometry that Lewis and colleagues had later suggested.

In order to further investigate proposed DNA-LacI loop geometries, Ruchi Mehta, a former graduate student in the Kahn lab, designed three DNA constructs that would potentially stabilize different looping geometries. The molecules, denoted WA 9C14, SL 11C12, and WA 7C16 (a shorter WA loop), were designed to mimic the loop geometries proposed by the Lewis and Steitz groups as shown in Figure 19. The three constructs contain two *lac* operator binding sites that bracket a phased A-tract

sequence-directed region, which give the molecules their large overall global bent shape. The symmetry axes of the *lac* operators, which are directed through the center of the LacI dimers, are pointed in different orientations relative to the global bend of the molecule in order to form different stable WA and SL loop geometries. By designing DNA molecules that are predisposed to have a large degree of bending we influence but do not necessarily control the overall global shape of the final loop, depending on the free energy needed to deform the DNA *versus* the LacI protein (Mehta & Kahn, 1999). Therefore the DNA molecule whose initial structure most closely resembles the optimal structure preferred by the LacI protein will form the most stable looped complex. It is important to note that the stability of the LacI-DNA loop will inevitably be affected by influencing the overall loop geometry.

Mehta conducted a series of experiments on the 11C12 and 9C14 constructs in order to assess the ability of the molecules to form different stable looped geometries upon LacI binding. Results from Electrophoretic Mobility Shift Assay (EMSA) experiments on both the 9C14 and 11C12 molecules showed single bands with different mobilities for each construct. EMSA is a technique commonly used to study gene regulation and works by monitoring DNA-protein interactions by gel electrophoresis. Competition experiments with excess cold DNA showed that both molecules were hyperstable with a half-life on the order of days. These experiments suggest the presence of one unique loop geometry for each individual molecule.

DNA cyclization kinetics studies on 9C14 and 11C12 molecules that were extended by ~200 bp to enable cyclization, suggested the presence of a closed loop form (giving a positive supercoil) for 9C14 as well as an open loop form (giving a

relaxed topoisomer) for both the 11C12 and 9C14 molecules. These results suggested that there are actually two possible forms for the 9C14 molecule, that it can form both an open loop and closed loop geometry.

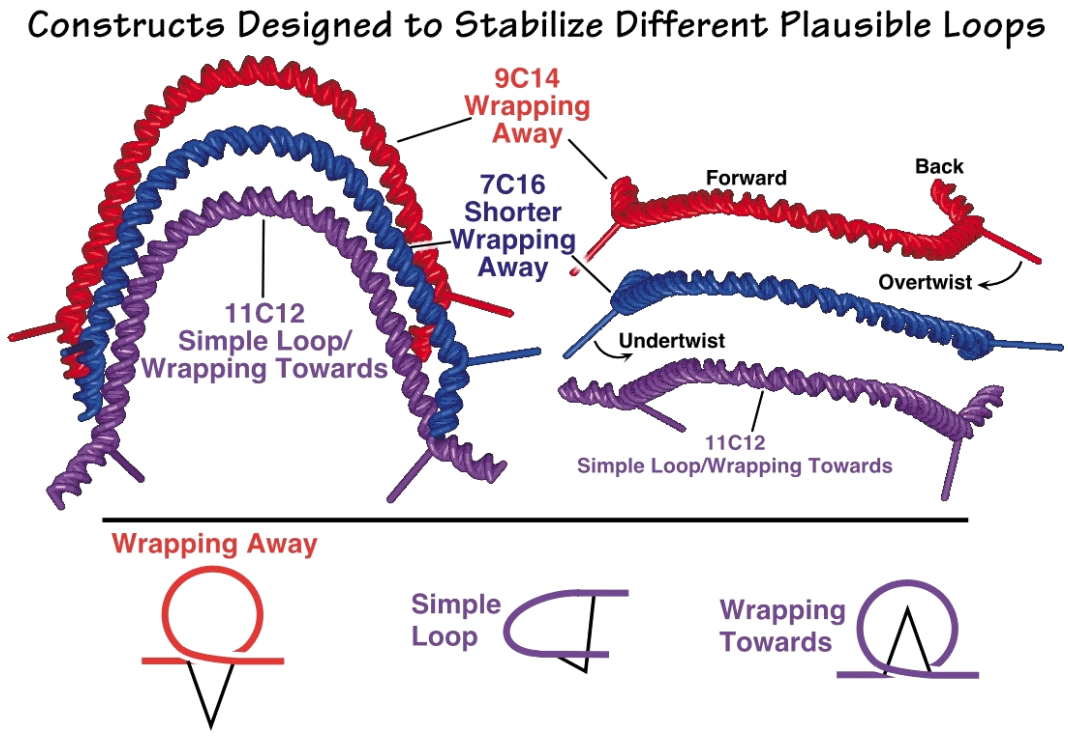


Figure 19. Design of loop stabilizing DNA constructs. Computer models for equilibrium structures of DNA bound to LacI dimer at each operator were generated as previously described (Mehta & Kahn, 1999). Note that these are not the anticipated shapes of the loops. The long cylinders represent the dyad axes of the DNA operator and protein dimer. The left and the right halves of the figure represent two views of the same set of molecules. The twist directions indicated in the figure are necessary to deform WA 9C14 and WA 7C16 into simple loops (Mehta & Kahn *et al*., 1999).

Results from both sets of experiments suggest the presence of an open form geometry for the LacI-11C12 loop. However, results from experiments on the 9C14 molecule were conflicting. EMSA experiments suggest the presence of one unique loop geometry for the 9C14-LacI complex, while DNA cyclization kinetics studies suggest the presence of both an open and closed loop form for the complex. At this point in time the lab sought a more direct method to assess the loop geometries and protein shape of both the 11C12-LacI and 9C14-LacI complexes. Bulk FRET measurements on both the 9C14- and 11C12-LacI complexes were performed by a former undergraduate student in the lab, Larry Edelman (Edelman *et al*., 2003). Bulk FRET results for the 11C12-LacI complex gave energy transfer efficiency values of 11 % and 8 % based on the decrease in quantum yield of the donor and the sensitized emission of the acceptor respectively. These values yielded a calculated interdye-distance of ~80 Å, which was interpreted to be either a uniform looped complex population or a rare closed form population exhibiting efficient transfer. Bulk FRET measurements on the 9C14-LacI complex gave an energy transfer efficiency value of 74 % based upon the quantum yield of the donor, which corresponded to a calculated interdye-distance of 46 Å. A lower efficiency of 38 % was calculated based on the sensitized emission of the acceptor, which corresponded to a calculated interdye-distance of 60 Å. The discrepancy between the two values is probably caused by an unexpected donor dependent quenching observed for the acceptor fluorophore (Edelman *et al*., 2003). The steady state efficiency of ~70 % could be due to 100 % of the molecules exhibiting 70 % ET, which would indicate the presence of a single species of molecules in a closed loop geometry. The second possibility is that 70 %

of the molecules are exhibiting 100 % ET with the remaining 30 % of the molecules exhibiting little to no ET, which would suggest the presence of a mixture of open and closed loop forms. Any intermediate of the two previous cases is also possible. It is important to note that the bulk FRET measurements did confirm the existence of a closed form (yielding a net 70 % efficiency of energy transfer), but it was not determined whether a second form existed, whether intermediate geometries existed, or the time scales of any interconversion processes. Bulk FRET measurements also confirmed the dominance of an open looped geometry for the 11C12 complex.

At this point a collaborative effort was started between the Kahn group and Doug English's research group in order to study the 9C14-LacI system using single molecule spectroscopy. We employed single-molecule FRET (SM-FRET) on Cy3, Cy5 labeled 9C14-LacI loops freely diffusing in solution to further assess looping geometries. Here I present results from SM-FRET studies aimed at providing a quantitative description of DNA-loop conformational distributions for the 9C14 sequence. Through detailed consideration of photophysical bleaching effects, which can affect both single molecule and ensemble FRET measurements, we demonstrate that LacI-9C14 exists strictly as a closed loop. This finding is presented as a critical result in reconciling previous discrepancies among results between DNA cyclization kinetics studies and EMSA experiments.

Materials

Oligonucleotides labeled with either Cy3 or Cy5 were purchased from Gene Link (Hawthorne, NY). A clone expressing Lac repressor was generously provided by Dr. Michael Brenowitz (Albert Einstein College of Medicine), and the protein was expressed and purified by Larry Edelman essentially as described by Brenowitz *et al*. (1991). Unless otherwise specified all reagents were purchased from Fisher Scientific (Pittsburgh, PA). All restriction enzymes and T4 DNA ligase were purchased from New England Biolabs Inc. (Beverly, MA). The FailSafe^{TM} PCR kit was purchased from Epicentre Technologies (Madison, WI).

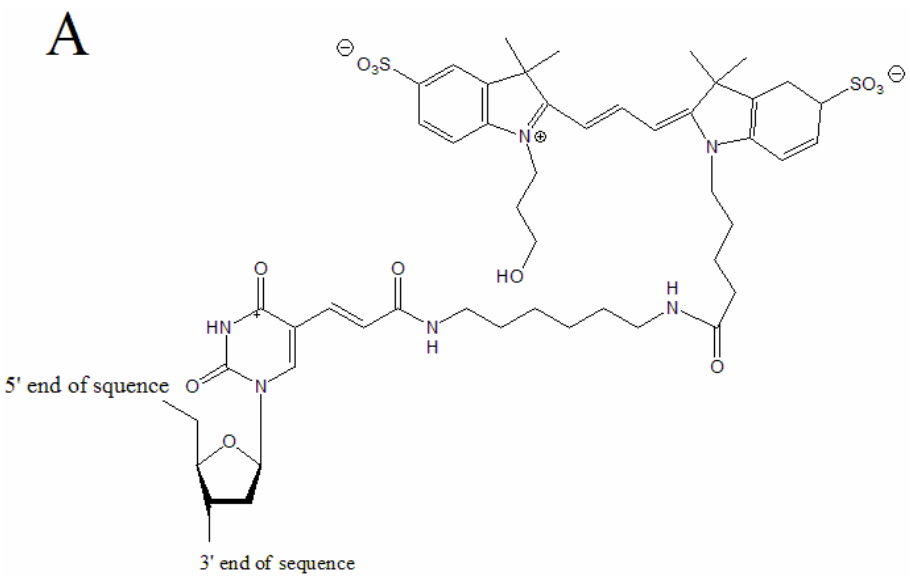
Methods

Synthesis of fluorescently labeled 9C14

Two 56 nt fluorescently-labeled oligonucleotides purchased from Gene Link (Hawthorne, NY) were used as PCR primers for synthesis of the 9C14 DNA with a Cy3 (donor) and Cy5 (acceptor) fluorophore, as shown in Figure 19. Edelman used the molecular modeling program *Insight II* (Molecular Simulations Inc., 1998) to determine that the optimal labeling position for the fluorophores was two base pairs outside the operator sequences. We utilized the same optimal position to label the 9C14 molecule with Cy3/Cy5 fluorophores. Strategic positioning of the fluorophores ensures that upon loop formation, the fluorophores are oriented away from the face of the DNA bound by the protein and that they are separated by only 35 Å in a closed loop form. Orientation and separation are both important factors for ensuring efficient energy transfer in the closed loop form.

Oligonucleotides were synthesized on a 1 μmole scale with an amino dT C-6 internal modification which was then conjugated to the fluorophore. Oligonucleotides were synthesized and purified by as follows: Oligonucleotides were purified on a 12 % polyacrylamide 8.3 M urea gel. Full-length primers were excised and the DNA was eluted and ethanol precipitated. A conjugation reaction between the purified oligonucleotides and the Cy3 or the Cy5 dye was used to attach the fluorophores to the oligonucleotides. The fluorescently labeled oligonucleotides were subjected to another round of gel purification.

Fluorescently labeled DNA-looping constructs were synthesized using the Failsafe PCR system (Epicentre Technologies, Madison, WI) with optimized premix buffer J. Template for the 9C14 DNA, previously described (Mehta & Kahn, 1999) was restricted with BstN I and diluted to 10^8 molecules/μL. PCR reactions contained 1 μL of template solution, 1 μL of both 50 μM Cy3 and 50 μM Cy5 labeled primers, 1 μL of Failsafe PCR Enzyme Mix, 25 μL of Premix J and 21 μL of ddH_20. PCR cycling conditions were as follows: 90 ºC for 30 sec, 55 ºC for 30 sec, 72 ºC for 1 min. The cycle was repeated 25 times, followed by a final extension step of 72 ºC for 5 minutes. PCR products were purified on a 6% polyacrylamide native gel [75:1 (w/w) acrylamide:bis-acrylamide] and eluted by shaking overnight in 400 μL of 50 mM NaOAc (pH 7.0), 1 mM EDTA followed by phenol-chloroform extraction and ethanol precipitation.



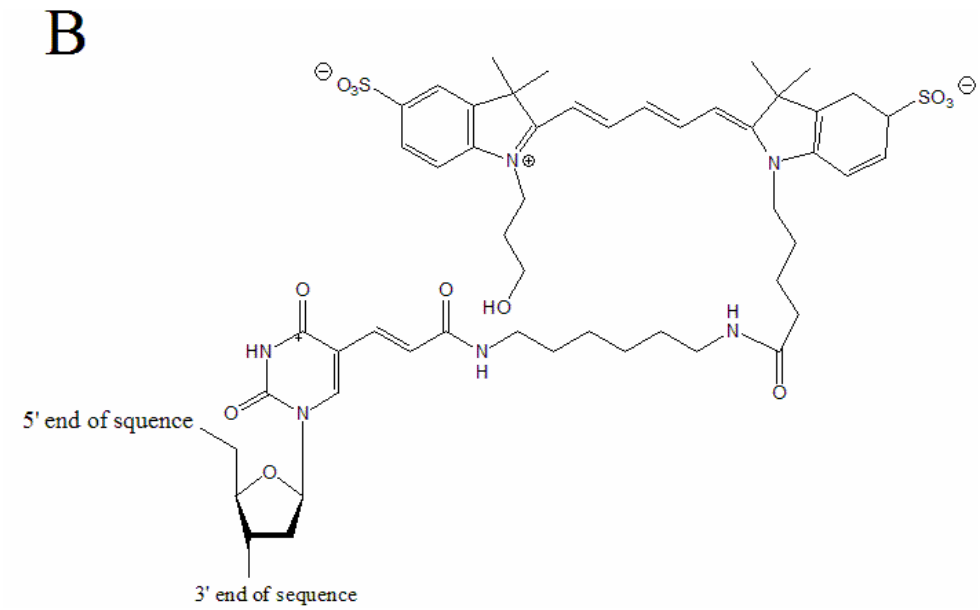


Figure 20. Structures of Cy3 and Cy5 fluorophores. A. Chemical structure of the Cy3 (Donor) fluorophore attached to the 5 position of a thymine base via a six carbon linker. B. Chemical structure of Cy5 (Acceptor), also attached to the 5 position of a thymine base via a six carbon linker.

9C14 DNA PCR Product – Labeled with Cy3 and Cy5.

Shown is the top strand only from 5' to 3'. **F** represents a Cy3 dT modification. **Z** represents an A that is complementary to a Cy5 dT modification on the bottom strand. This PCR product is amplified from a BstN I fragment from pRM9C14 (Mehta, 2000) using primers Cy3TOP9C14 and Cy5BOT9C14. The first base corresponds to position 2213 in the plasmid sequence.

CTGCAGGTCAGTCTAGGTAATTGTGAGCGCTCACAATTAFATCTCAATTCGTACGGA
TCCGGTTTTTTGCCCGTTTTTTGCCGTTTTTTGCCCGTTTTTTGCCGTTTTTTGCCC
GTTTTTTGCCGTTTTTTGCCCGTTTTTTGCGCTGAACGCGTCCTAGAATCGAAGCTA
GZTAATTGTGAGCGCTCACAATTCGTTGATGGTAAAGCTTG

Cy3TOP9C14 56 nt Oligonucleotide – Labeled with Cy3

Shown is the Cy3TOP9C14 56 nt oligonucleotide from 5' to 3'. **F** represents a Cy3 dT modification.

CTGCAGGTCAGTCTAGGTAATTGTGAGCGCTCACAATTAFATCTCAATTCGTACGG

Cy5BOT9C14 56 nt Oligonucleotide – Labeled with Cy5

Shown is the Cy5BOT9C14 56 nt oligonucleotide from 5' to 3'. **Y** represents a Cy5 dT modification.

CAAGCTTTACCATCAACGAATTGTGAGCGCTCACAATTAYCTAGCTTCGATTCTAG

Figure 21. Sequences of 9C14 PCR product and the 56 nt oligonucleotide primers used in SM-FRET experiments. The *lac* operator sequences are in magenta.

Determination of Fluorophore Labeling Efficiencies

The primers used for PCR reactions are synthesized with modified thymine bases at internal positions in the oligonucleotide sequence. In a subsequent reaction step the fluorescent dyes are then conjugated to the modified base via a coupling reaction. However, there remains a fraction of oligonucleotides that are not labeled in the reaction, so it is imperative to determine the percent of dye-labeled oligonucleotides in order to determine an accurate efficiency of energy transfer. In order to determine the fraction of molecules labeled with a fluorophore, absorption spectra for each labeled oligonucleotide are used, as shown in Figure 22.

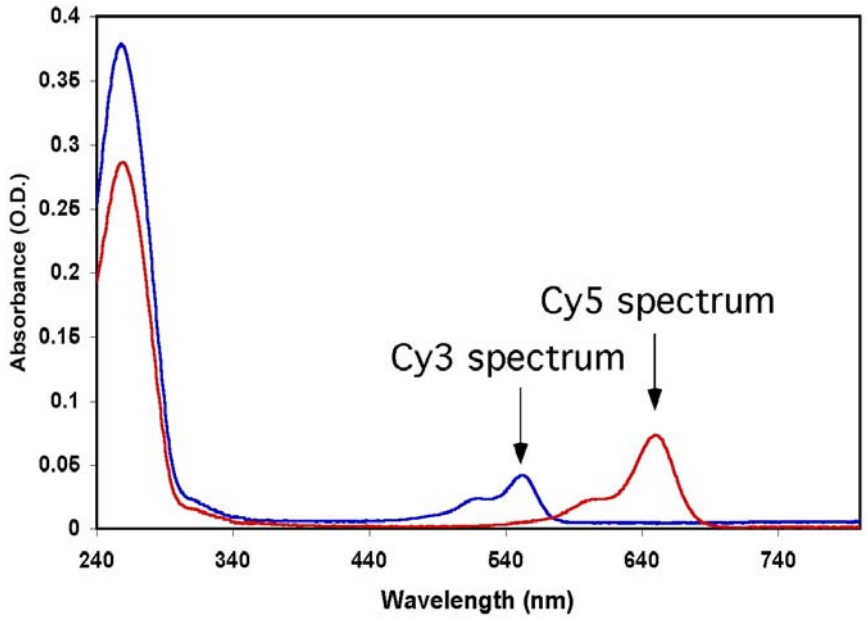


Figure 22. Absorption spectra of Cy3-labeled (blue) and Cy5-labeled (red) oligonucleotides in 1X LacI buffer (25 mM bis-Tris, 5 mM $MgCl_2$, 100 mM KCl, 2 mM DTT, 0.02 % (v/v) nonidet P40 (NP40).

Labeling efficiencies were determined by calculating the [dye]/[DNA] of the Cy3 and Cy5 56 nt labeled oligonucleotides. The concentrations of dyes were determined using Beer's Law and the absorbance at 552 nm (Equation 7 below) for Cy3 labeled DNA, or 643 nm for Cy5 labeled DNA, using extinction coefficients reported by Gene Link (Table 1). The contribution of the dyes to absorption at 260 nm was then calculated based on their concentrations and their extinction coefficients at 260 nm (Equation 8). The absorbance of the dye at 260 nm is then subtracted from the total absorbance and the concentration of DNA in each sample is determined from the remainder (Equation 9). The labeling efficiencies were calculated for Cy3TOP9C14 and Cy5BOT9C14 (Equation 10) and were determined to be 38 % for the former and 54 % for the latter.

	Absorbance Maximum(nm)	Emission Maximum(nm)	Extinction coefficient at 260nm (M^{-1} cm^{-1})	Extinction coefficient at absorbance maximum (M^{-1} cm^{-1})
Cy3 Donor	552	570	4930	150000
Cy5 Acceptor	643	667	10000	250000

Table 1. Spectroscopic properties of Cy3 and Cy5 fluorophores as reported by Gene Link (Hawthorne, NY).

Shown below are the calculations for determining the labeling efficiency for Cy3 labeled DNA oligonucleotide:

Cy3 Concentration

$$[\text{Cy3}] = \frac{A_{552}}{\varepsilon_{552} \cdot 1\,\text{cm}} \quad \text{Equation 7}$$

Cy3 Contribution at 260nm

$$\text{Cy3 Abs}_{260} = [\text{Cy3}] \cdot \varepsilon_{260} \cdot 1\,\text{cm} \quad \text{Equation 8}$$

DNA concentration

$$[\text{DNA}] = \frac{A_{260} - [\text{Cy3} - A_{260}]}{\varepsilon_{260} \cdot 1\,\text{cm}} \quad \text{Equation 9}$$

Labeling efficiency

$$\text{Labeling efficiency} = [\text{Cy3}]/[\text{DNA}] \cdot 100 \quad \text{Equation 10}$$

Electrophoretic Mobility Shift Assays (EMSA)

Electrophoretic Mobility Shift Assays were conducted in order to confirm the ability of the 9C14 complex to bind Lac repressor at stoichiometric binding conditions (data not shown). 2.5 nM 9C14 DNA was incubated with varying concentrations of protein (0.0-10.0 nM) in 1X LacI buffer, pH 7.9 (25 mM bis-Tris, 5 mM $MgCl_2$, 100 mM KCl, 2 mM DTT, 0.02 % (v/v) nonidet P40) for 15 minutes on ice and equilibrated at room temperature for 10 minutes. 8 % (v/v) glycerol was added to the samples and they were loaded onto a 6 % 75:1 native polyacrylamide gel thermostatted at 21 ºC and run for 1 hour at 21 W. The gel was imaged using a Molecular Dynamics Storm Imager 860 using the red excitation wavelength of 635 nm.

Excitation and Emission spectra of Cy3 and Cy5 labeled oligonucleotides

Excitation and emission spectra were taken of both Cy3- and Cy5-labeled oligonucleotides using a Jobin Yvon Horiba Fluorolog-3 spectrofluorometer. Excitation and emission spectra were obtained at 140 nM [DNA] in 1X LacI buffer. The excitation wavelength for emission spectra was λ_{ex} = 525 nm for Cy3 ssDNA and λ_{ex} = 600 nm for Cy5 ssDNA. The excitation spectra for Cy3 ssDNA and Cy5ss DNA were monitored at λ_{em} = 650 nm and λ_{em} = 800 nm respectively. Slit widths for both measurements were set at 5 nm.

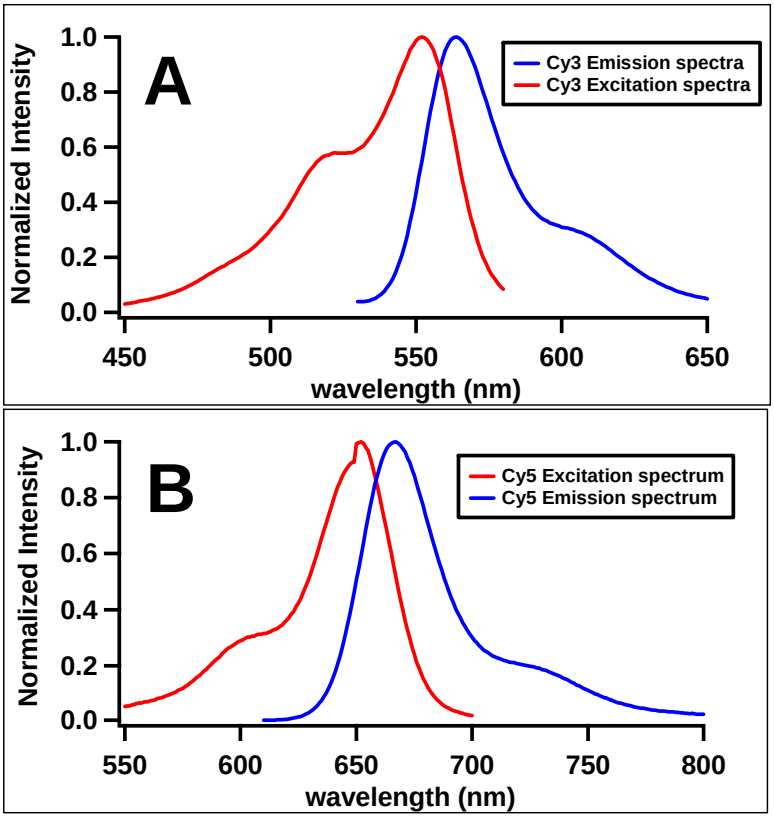


Figure 23. Normalized excitation and emission spectra for Cy3 (A) and Cy5 (B) labeled oligonucleotides. Experiments were conducted using 138.5 nM DNA in 1X LacI buffer. Emission spectra for Cy3 and Cy5 were monitored at (λ_{ex} = 525 nm) and λ_{ex} = 600 nm respectively. Excitation spectra for Cy3 and Cy5 were monitored λ_{em} = 650 nm and λ_{em} = 800 nm respectively.

Single molecule FRET measurements

All single molecule FRET measurements were conducted at 1.25 nM DNA concentration with varying LacI concentrations in 1X LacI buffer, 50 μg/mL bovine serum albumin (BSA) and 1X Oxygen Scavenger (100X 0_2 scavenger stock: 450 mg/mL glucose, neat beta-mercaptoethanol diluted 1:200 final, glucose oxidase 43.2 mg/mL buffer (24 mM pipes pH 6.8, 4 mM $MgCl_2$, 2 mM EGTA) and catalase 7.2 mg/mL buffer. Samples were allowed to equilibrate at room temperature (20 °C) for at least 10 minutes before conducting FRET experiments.

The single molecule experiments were conducted on freely diffusing molecules in solution following the methods of Weiss (1999, 2000). Briefly, a microscope objective (Fluar, 100X, numerical aperture (n.a.) = 1.3, Carl Zeiss, Oberkochen, Germany) is used to focus a He-Ne laser beam into a DNA solution with the focal spot located 10 μm from the glass liquid interface formed by a microscope coverslip and the DNA solution. Fluorescence bursts are detected when molecules traverse the beam and the emission is collected by the same microscope objective and directed through a holographic notch filter (Kaiser Optical, Ann Arbor, MI) to remove scattered excitation light. Experiments were conducted with both 514 nm and 543 nm excitation. Typical radiant flux densities were on the order of 2-5 $kWcm^{-2}$ in the laser focal volume, which is commonly found to afford single molecule excitation rates sufficient for producing detectable bursts during the brief transit time (< 1 ms) of a single molecule in the detection volume. The emitted light is then directed to a dichroic long pass beam splitter (625 DCLP, Chroma, Rockingham, VT), which transmits long wavelength emission from the acceptor fluorophore onto one

avalanche photodiode and reflects shorter wavelength emission (< 625 nm) from the donor fluorophore onto a separate avalanche photodiode. Photon counts from the two photodiodes are recorded in 1 ms time bins on separate channels of a counter/timer board (PCI 6602, National Instruments, Austin, TX) controlled by a personal computer with Labview 6.0 software (National Instruments). Burst data is analyzed by calculating the observed energy transfer efficiency for every burst with a combined donor and acceptor intensity above a set threshold. The calculated values of energy transfer efficiency are then used in constructing a histogram describing the probability distribution of FRET values. The energy transfer efficiency value of each burst falling above the threshold is determined from the following equation:

$$E = \frac{S_A}{S_A + \gamma S_D} \quad \text{Equation 11}$$

where γ is a correction factor to account for differences in the dye quantum yields and photodiode detection efficiencies. As in previous studies using Cy3 and Cy5 we consider γ to be unity (Cosa *et al.* 2004). S_A and S_D are the acceptor and donor signals corrected for background counts. S_A is also corrected for the presence of donor-emitted photon counts appearing in the acceptor channel:

$$\begin{aligned} S_D &= I_D - B_D \\ S_A &= I_A - B_A - \alpha S_D \end{aligned} \quad \text{Equation 12}$$

where I_A and I_D are the acceptor and donor intensities respectively and B_A and B_D are the background counts in the acceptor and donor channels. The correction term α is used to correct for donor photons that "leak" into the acceptor channel as a result of spectral overlap. α was experimentally determined to have a value of 0.361 by collecting bursts from a donor-only labeled DNA molecule. Correcting the acceptor

channel for donor leak-through improves the FRET efficiency accuracy but degrades the precision since the noise in the donor channel is added to the acceptor channel as a result the zero-peak has a width that is about twice the expected value at our signal levels.

Immobilization of DNA-LacI complexes

Immobilization of DNA-LacI complexes in poly vinyl alcohol (PVA).

Approximately 5-10 μL of 20 nM 9C14 + 40 nM Lac I was added to 150 μL of 0.8-8.0 % (w/v) PVA. The sample was then pipetted onto a piranha cleaned coverslip and spincoated at 2000 rpm for 30 seconds. Coverslips were prepared as follows: Glass coverslips were cleaned with piranha solution (3:1 sulfuric acid:hydrogen peroxide) for 45 minutes, rinsed with ddH_2O, and dried in a stream of N_2.

Immobilization of DNA-LacI complexes using aminopropyltriethoxysilane (APTES).

Briefly, piranha cleaned coverslips were placed into a dessicator containing 500 μL of APTES in a cap or vial for approximately 1 hr. The APTES evaporates into the closed environment to functionalize the mica surface with a silane bearing an amino group. Typically 50 μL of sample containing 2.5 nM 9C14 + 5.0 nM LacI in 1X Lac buffer was pipetted onto the glass surface for imaging.

Immobilization of DNA-LacI complexes using glutaraldehyde

Typically a solution of 2.5 % glutaraldehyde in 1 mM phosphate buffer (pH 7.0) is pipetted onto the APTES treated glass surface immediately after removal from the dessicator, and incubated for 10 minutes. A solution containing 2.5 nM 9C14 + 5.0 nM LacI, 1X Lac buffer, 0.02% NP40 is pipetted onto the APTES-glutaraldehyde glass coverslip and imaged.

Immobilization of DNA-LacI complexes in Agarose

Typically 1.5 mL of low melt agarose (1-2% w/v) is heated to the melting temperature of the gel and slowly cooled back down until it reaches approximately 37 ºC. Approximately 10-20 μL of a 20 nM DNA + 40.0 nM LacI in 1X LacI buffer was added to approximately 100 μL of agarose and pipetted onto a pirannah cleaned coverslip and sandwiched with another piranha cleaned coverslip placed very gently on top. The agarose sandwich was allowed to cool to room temperature and then imaged.

Results

Previous experiments including cyclization kinetics, electrophoretic mobility shift assays and ensemble FRET measurements (Edelman *et al.,* 2003; Mehta & Kahn, 1999) have shown the 9C14 molecule forms a hyperstable looped construct with $t_{1/2} > 24$ hr. At the time these experiments were conducted there was still uncertainty about the geometry of the loop. DNA cyclization studies of the 9C14 and 11C12 molecules with extended DNA tails were carried out by Mehta & Kahn (1999). Results from these experiments suggested that 9C14-LacI complex could adopt both and open and closed loop form, while the 11C12-LacI complex adopted only an open loop form. However, DNA cyclization results conflicted with results from electrophoretic mobility shift assay experiment on the 9C14-LacI loop, in which only one conformer was resolved. Based on its mobility it was believed to be the closed loop construct. Observations from the EMSA experiments could be reconciled with the idea of two populations stabilizing a closed loop geometry, or if the proposed conformers are in rapid exchange in the gel with the closed loop form migrating very rapidly. In an effort to better determine the conformational make-up of the 9C14-LacI complex and resolve the apparent discrepancies between previous experiments, a fluorescently labeled version of 9C14 was synthesized for use in steady-state and time-resolved FRET measurements (Edelman *et al.,* 2003). The outcome of ensemble FRET measurements was the definitive verification of the closed-loop complex by observation of significant energy transfer in accordance with the predicted model. Due to complications including possible non-Förster quenching of the donor molecule as well as donor-dependent quenching of the acceptor, the ensemble FRET data was

insufficient to unambiguously determine the number and relative abundance of 9C14 conformations. For this reason we have undertaken single molecule experiments to directly resolve the equilibrium composition of the LacI/9C14 complex.

Single molecule 9C14 DNA FRET results upon titration with LacI

The 9C14 molecule was originally designed to mimic the modeled wrapping away (WA) loop geometry proposed by Lewis and colleagues (Lewis *et al*., 1996). The molecule was designed so that the dyes were placed ~130 bp apart from each other (~ 400 Å apart along the contour of the DNA) as shown in Figure 24. The fluorophores are positioned approximately 1 helical turn away from the LacI binding site in order to minimize interference of efficient FRET upon LacI binding. We expect the 9C14 molecule should undergo a large change in observed FRET efficiency upon formation of a closed loop. The predicted interdye-distance for the closed loop form of the complex is 35 Å, which yields a FRET efficiency of 0.9 based on the predicted Förster distance of 60 Å. In the open loop form the interdye-distance of ~100 Å results in a FRET efficiency of 0.46. Hence the closed-loop geometry should be easily distinguished from unbound 9C14 or from the open-loop complex.

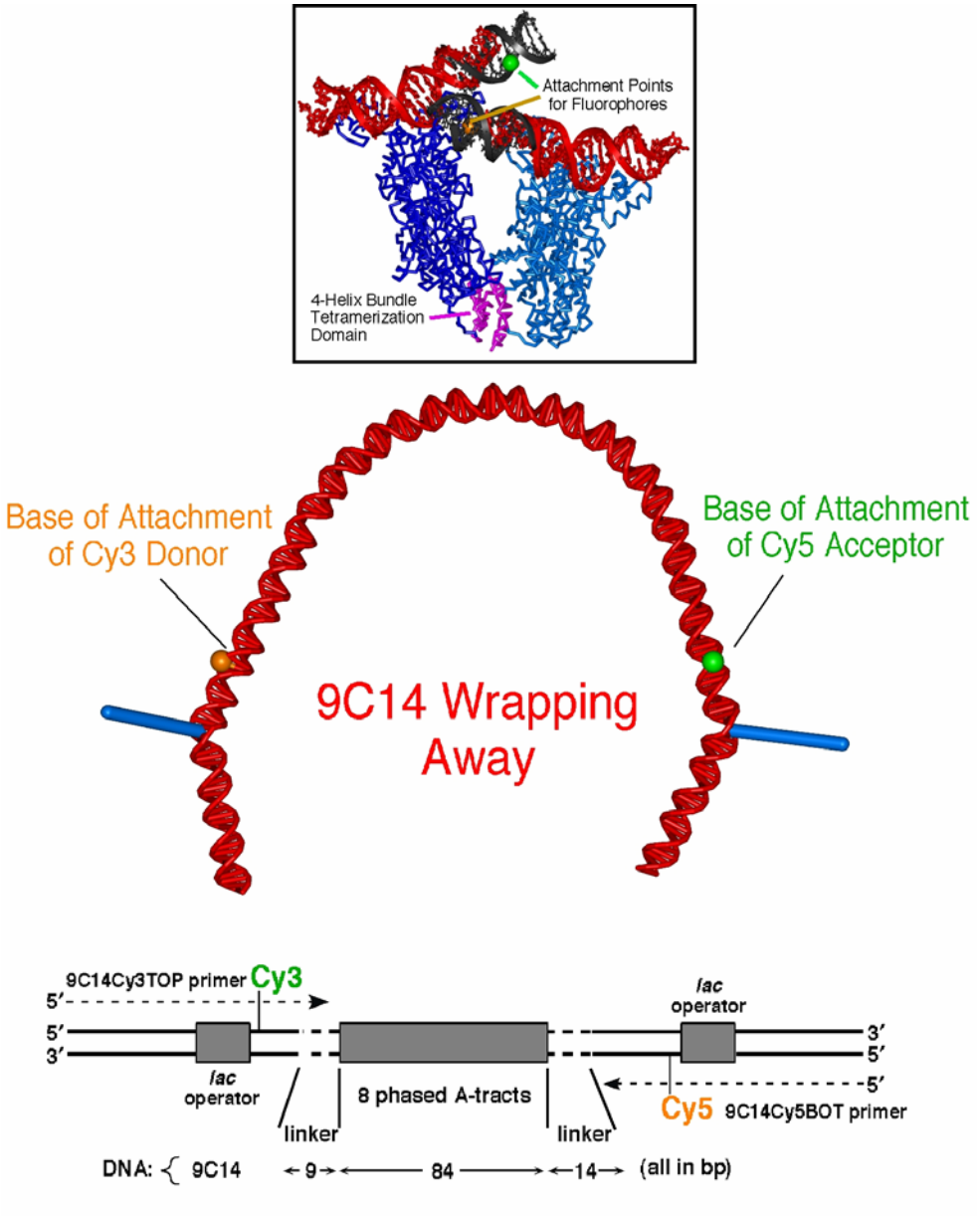


Figure 24. The cocrystal structure of the LacI tetramer DNA "sandwich complex" is shown (PDB) with the protein in blue, the tetramerization domain in magenta, and the DNA in red. The black DNA is a modeled B-DNA extension, and the fluorophore attachment points at the thymine C5 position is indicated. The fluorophores project into the major groove, away from the protein. A model for the looping construct shows the dyad axis of the DNA operators, which lie along the symmetry axis of the LacI dimers, and represented as long blue cylinders. The base of the attachment of the Cy3 is shown as a green ball and the base of attachment of the Cy5 acceptor as an orange ball on both molecules (a distance of ~130 bp separates the fluorophores). The dimensions in base pairs of the molecules and the attachment sites for the Cy3 donor and Cy5 acceptor are shown.

Figure 25A shows actual experimental data obtained from a 1 ms acquisition of the fully titrated 9C14-LacI molecules freely diffusing through solution. Fluorescence bursts generated from 9C14-LacI complexes traversing through the focal volume of the laser are indicated by black arrows in figure 25A. It was subsequently determined that a single green burst, which appears principally in the donor channel, belongs to a donor-only or a mislabeled 9C14 molecule (a detailed discussion of these findings will follow later in the chapter). A single high intensity red burst that appears principally in the acceptor channel indicates a 9C14-LacI closed loop complex.

Figures 25B and 26A show the results from titration experiments conducted with 1.25 nM 9C14 freely diffusing through solution. The addition of LacI results in a decrease in the number of molecules showing zero FRET efficiency, and a concomitant increase in molecules exhibiting high energy transfer. The population with high FRET efficiency corresponds to the formation of a closed-loop 9C14-LacI complex possessing a FRET efficiency of ~90%, in good agreement with predicted values for the formation of a closed-loop geometry. The trend is illustrated in Figure 26B, which plots the fraction of closed loop population (defined as the portion of events with FRET efficiency greater than 50 %) as a function of LacI concentration. The trend saturates at 2.5 nM LacI, where LacI concentration is approximately 2 times greater than 9C14 concentration. At saturation, the proportion of closed loop molecules was typically recorded as 60 %. Occasionally there were samples measured which showed a lower abundance of the closed loop form, and this was attributed to variability in the PCR. However there remained a substantial population

(~ 40%) that was centered at ~0 % FRET efficiency. This result initially appeared to correlate very well with the bulk FRET results, suggesting that 60 % of the molecules were exhibiting nearly 100 % ET with the remaining 40 % of the molecules exhibiting little to no ET. This suggested that presence of both an open loop and closed loop geometry for the 9C14-LacI complex. However, further analysis of the 40% of molecules centered around 0 % ET efficiency would suggested that this was not the case. The 0% ET efficiency peak ("zero-peak") will be discussed in further detail later in the chapter.

In all experiments, the population with high energy transfer efficiency was absent before titration began and grew in with increasing LacI concentration until saturation was attained, after which a slight decrease in the population was observed. This saturation and subsequent break-up of loops agrees with results from gel-shift assays (Mehta & Kahn, 1999). An obvious advantage gained with single molecule experiments is the ability to track both populations throughout the titration process rather than observing an average value. In bulk steady-state FRET titration experiments a weighted average of the two populations is observed throughout the titration process (Edelman *et al*., 2003).

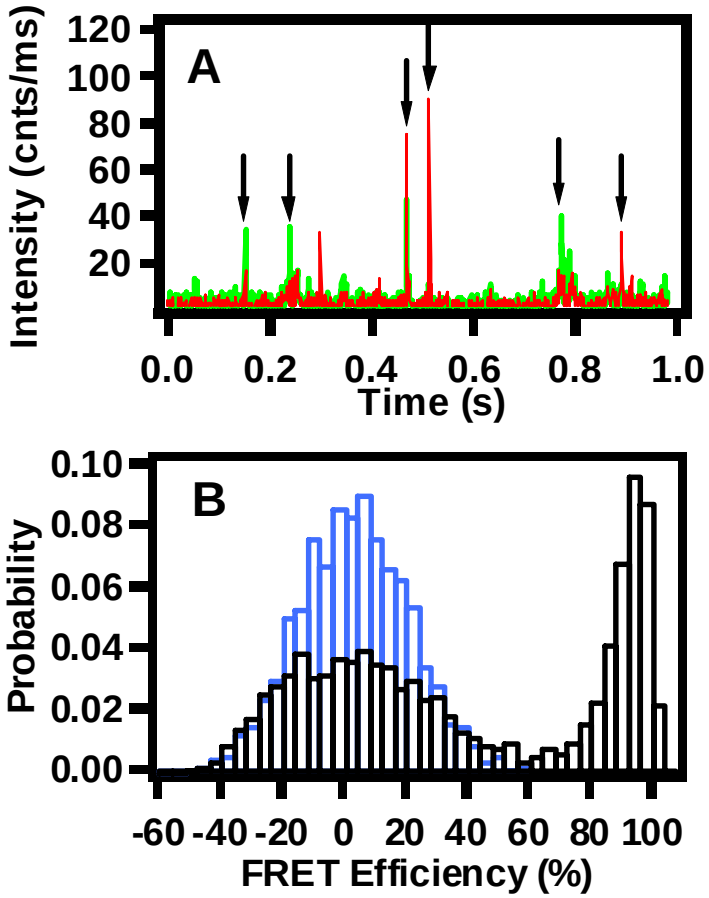


Figure 25. SM-FRET data from diffusing 9C14 molecules. *(A)* A typical burst sequence acquired with the 9C14 DNA construct (1 nM) excited at 514 nm. The acceptor counts are shown in red and donor counts are in green. The LacI concentration is 2.5 nM and a mixture of molecules exhibiting both high and low energy transfer values are observed. The binning time for the counts is 1 ms. Arrows denote bursts that exceed the intensity threshold for analysis. The bursts located near 0.5 s appear principally in the acceptor channel and originate from a LacI/9C14 complex forming a closed loop with an energy transfer efficiency of nearly 1.0. Bursts which appear principally in the donor channel, such as the one near 0.15 s belong to untitrated or mislabeled 9C14 molecules. *(B)* Typical FRET histograms acquired with 514 nm laser excitation for samples without LacI (blue) and saturated with LacI (black). Negative efficiency values in the zero-peak are due to bursts in which the acceptor channel intensity was less than the average background. The width of the zero-peak is further broadened by the subtraction of donor counts from the acceptor channel to correct for non-ideal optical filtering as described in the experimental section.

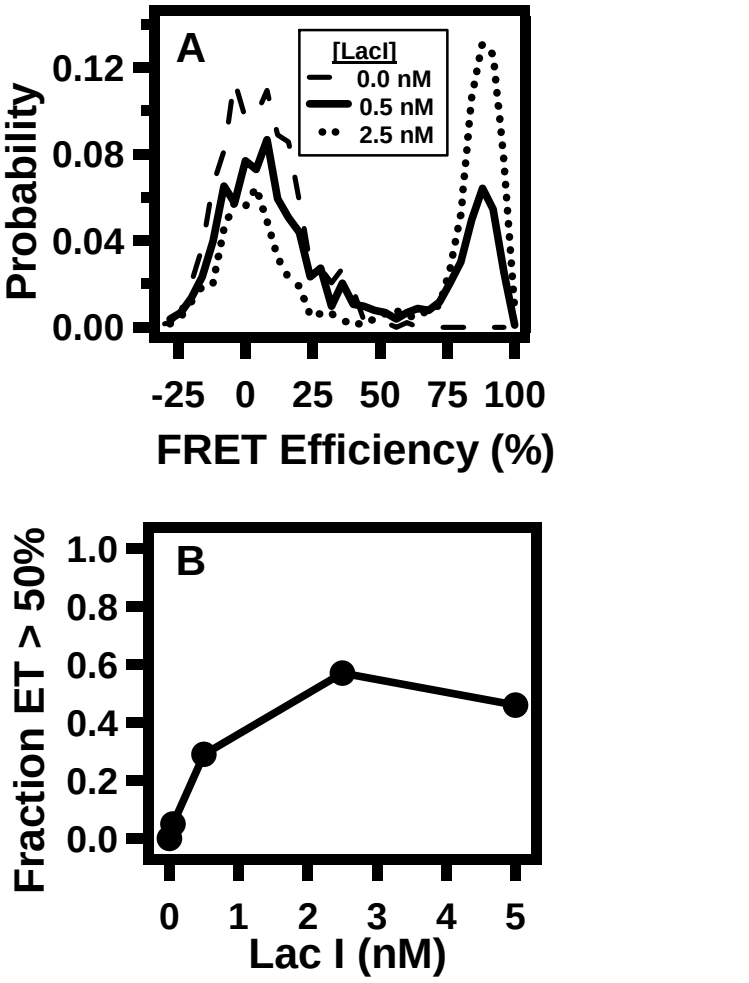


Figure 26. SM-FRET results from titration experiments performed with 543 nm excitation. *(A)* Histograms acquired with three different LacI concentrations illustrating the decrease in the zero-peak accompanied by simultaneous increase in the proportion of molecules appearing with nearly 100% energy transfer efficiency. Without LacI (dashed line), only the zero peak is observed, while integration of the fully titrated histogram (dotted line) reveals that 60% of bursts appear with high energy transfer efficiency and originate from the closed-loop 9C14 conformer. *(B)* Graph displaying the fraction of bursts appearing with high energy transfer efficiency as a function of LacI concentration. This shows the direct relationship between closed loop formation and LacI concentration. Note that at 5.0 nM LacI a slight decrease in the fraction of high energy transfer is observed which may be due to quenching of the acceptor by nonspecific binding of LacI to 9C14.

Comparison of single molecule FRET results with ensemble results

Steady-state bulk FRET experiments had previously established an apparent average energy transfer efficiency of 74 % based on donor-quenching for the fully titrated 9C14 complex (Edelman *et al.*, 2003). This placed a lower limit on the FRET efficiency for the closed loop, since the bulk experiments report a weighted average of all values present in the sample. It should be noted that the value of 74 % obtained from bulk measurements is in reality a highly unlikely value: individual molecules never really display this value, as shown by the histograms shown in Figure 25B and Figure 26A. Our single molecule results definitively establish 90 % as the closed-loop energy transfer efficiency observed by monitoring sensitized emission; without immobilizing molecules we cannot observe the energy transfer efficiency based on donor quenching for single molecules. The single molecule results also establish a lower limit of 60 % as the fraction of molecules appearing with 90 % energy transfer efficiency in the fully bound sample. It is important to emphasize that 60 % is only a lower limit due to the uncertainty surrounding the identity of the molecules which compose the zero-energy transfer peak (henceforth the "zero-peak"). Molecules displaying zero energy transfer could arise from either 9C14-LacI complexes with an open-loop form or from molecules with missing, photobleached or inactive acceptor fluorophores. To investigate the composition of the zero-peak we conducted experiments to determine the role of laser-induced photobleaching in addition to experiments conducted at two different excitation wavelengths to reveal compositional heterogeneity.

Role of photobleaching and the composition of the zero-peak

The intense excitation necessary to produce detectable bursts from individual diffusing molecules can lead to undesired excited-state processes. The most serious effect is irreversible photobleaching. Irreversible photobleaching of the donor fluorophore has negligible effect since this process renders the molecule undetectable. Photo-induced processes involving the acceptor fluorophore present a more serious difficulty since they introduce artifacts in the SM-FRET histogram (Dahan *et al.*, 1999). When a molecule's acceptor fluorophore undergoes irreversible photobleaching, the molecule is detected (from donor emission) with an energy transfer efficiency value of zero. If photobleaching of the acceptor fluorophore occurs during the acquisition of a single-molecule burst, the *apparent* energy transfer efficiency will be a weighted average of the real energy transfer efficiency value (pre-photobleaching) and zero (post-photobleaching). It is also possible for the acceptor fluorophore to undergo irreversible photobleaching prior to reaching the observation volume, in which case the apparent energy transfer efficiency will always be zero, regardless of the molecule's actual conformation. In addition to *irreversible* photobleaching an acceptor fluorophore may enter a transient dark-state through intersystem crossing or by assuming a non-emissive conformation. These transient processes will also introduce artifacts in the SM-FRET histogram. The excited state photophysics of the acceptor fluorophore are only relevant when there is efficient energy transfer, as otherwise the acceptor remains in its ground state.

We conducted excitation-power dependence studies to investigate the role of photosensitized bleaching of Cy5 or other excited-state processes on the observed

SM-FRET histogram. To help evaluate the effect of oxygen, experiments were conducted on samples of fully titrated 9C14, both under ambient conditions and also in reduced-oxygen system. Cy5 photobleaching should appear as a decrease in the fraction of molecules showing efficient energy transfer and an increase in the size of the zero-peak. The results from these studies are summarized in Figure 27, where the percentage of total molecules with energy transfer efficiency above 50% is plotted against excitation laser power.

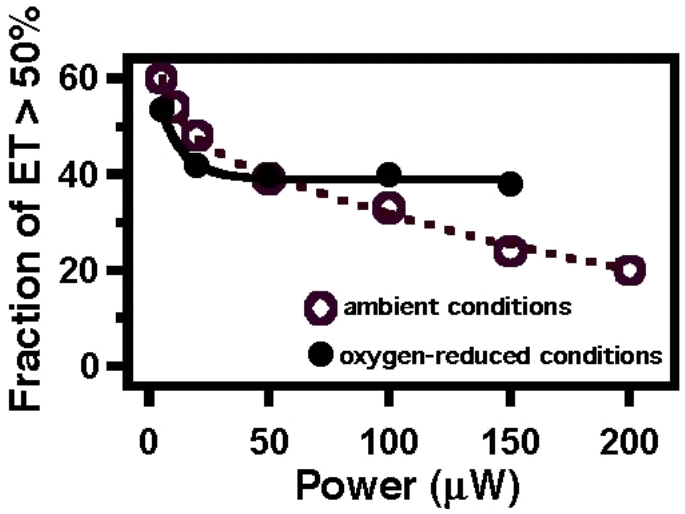


Figure 27. Plot illustrating the laser excitation power dependence of observed SM-FRET. The fraction of molecules in the fully titrated sample with energy transfer efficiency above 0.5 is plotted as a function of laser excitation power under ambient conditions (open circles) and under reduced-oxygen conditions (closed circles). Under ambient conditions, as the power is increased there is a steep, non exponential power dependence. With the addition of the glucose oxidase based oxygen scavenger the power dependence is significantly diminished but still present.

There is clearly a strong power dependence under ambient conditions (empty circles). The addition of oxygen scavenger (solid circles) markedly decreases this dependence; however, even with oxygen scavenger present a weak power dependence remains.

Under reduced-oxygen conditions it is likely that FRET-induced formation of the acceptor fluorophore triplet state is responsible for the observed power dependence. This is due to the acceptor in the triplet state being nonemissive. Under ambient conditions triplet state lifetimes are on the order of 10 μs and do not appreciably affect the FRET efficiency value measured, since the observation time of 9C14 in the laser beam is several hundred microseconds. However, at reduced oxygen concentrations it is well known that cyanine fluorophores such as Cy5 form long-lived triplet states. Previous single molecule studies with the carbocyanine dye $DiIC_{18}$ showed that triplet lifetimes exceed 100 ms under vacuum (English *et al*., 2000a; English *et al*., 2000b; Tinnefeld et al., 2003; Weston *et al*., 1999). Sauer and coworkers have shown that Cy5 can display millisecond triplet lifetimes under a nitrogen-purged environment (Tinnefeld *et al*., 2003). Hence, under the conditions used here it is reasonable to assume that the triplet lifetime will be on the millisecond time scale.

The power dependence data can all be reconciled by considering the long-lived triplet acceptor state ^{3}A. We also must assume that it acts as an energy transfer acceptor, as shown previously (Tinnefeld, 2003), but with R_0 for the D-^{3}A pair greater than that for D-A. At very low excitation powers, the time spent in the singlet manifold (i.e. before triplet formation) will be longer than the observation time due to the low quantum yield of triplet formation, and hence the triplet state will have no

effect on the FRET histogram. Even if the triplet is formed, the resulting D-^{3}A state is unlikely to absorb a second photon and therefore will not be detected. As the laser intensity increases the excitation rate increases linearly and so does the rate of triplet formation. As the population of the nonemissive triplet state increases, decreased efficiency of ET is observed because donor excitation of D-^{3}A leads to some donor emission. The plateau in the power dependence suggests one additional property of the triplet: if the triplet state persisted during the entire observation period for each single molecule, then we would expect a continued decrease in observed efficiency as power increases. However, if the triplet lifetime is also power dependent then a photostationary state is established, in which the singlet and triplet manifolds of the acceptor are in dynamic equilibrium. The establishment of a steady state requires that Cy5 undergo a reverse intersystem crossing process that is induced by energy transfer from the Cy3 singlet state.

In summary, we propose that the depletion of efficient energy transfer observed at low power in the absence of oxygen is due to formation of a single-molecule photostationary state at laser powers above 40 μW. The observed *E* is the weighted average of the efficiencies of the singlet and triplet states of the acceptor. At lower power where the triplet is not created or not re-excited during the time the molecule is in the beam, we observe the *E* characteristic of the singlet manifold of the acceptor. These results illustrate the necessity for using oxygen scavenging to reduce irreversible photobleaching but also show that the resulting increased triplet lifetimes can impact the FRET histogram.

Excitation wavelength dependence studies

As previously mentioned a substantial population of fully titrated 9C14-LacI molecules (~ 40 %) remain centered at 0 % ET efficiency. However, it was unclear as to whether this population of molecules was from 40 % of the 9C14-LacI complexes existing in an open loop geometry exhibiting little to no energy transfer, or whether there remained a substantial amount of donor-only labeled molecules that were incapable of undergoing energy transfer. In order to resolve this issue experiments were conducted at two different wavelengths. Single molecule experiments conducted with multiple excitation wavelengths have proven useful for elucidating a variety of photophysical processes (English *et al.,* 2000b; Fukaminato *et al.,* 2004; Kulzer *et al.,* 1997). More recently, multi-color excitation has been demonstrated as a powerful technique for sorting freely diffusing single molecules and complexes in solution (Kapanidis *et al.,* 2004; Li *et al.,* 2004). Here we have conducted single-molecule titration studies at two separate excitation wavelengths, 514 nm and 543 nm. Both wavelengths efficiently excite the donor fluorophore, Cy3, but only excitation with 543 can significantly excite the acceptor fluorophore, Cy5. With 543 nm light Cy5 is excited approximately 40 times more efficiently than with 514 as shown in Figure 28.

Our current approach lacks the technical sophistication of modulated two-color experimental schemes (English *et al.,* 2000b; Kapanidis *et al.,* 2004), but it is easily implemented and comparison of data acquired at the two excitation wavelengths allows us to resolve the composition of the zero-peak. To illustrate this we have constructed difference histograms obtained by subtracting the histogram of

free DNA (un-titrated) from that of the fully titrated sample. The difference histogram from data acquired at 514 nm is shown in Figure 29A and represents the amount of DNA which appears at high energy transfer values in the titrated histogram.

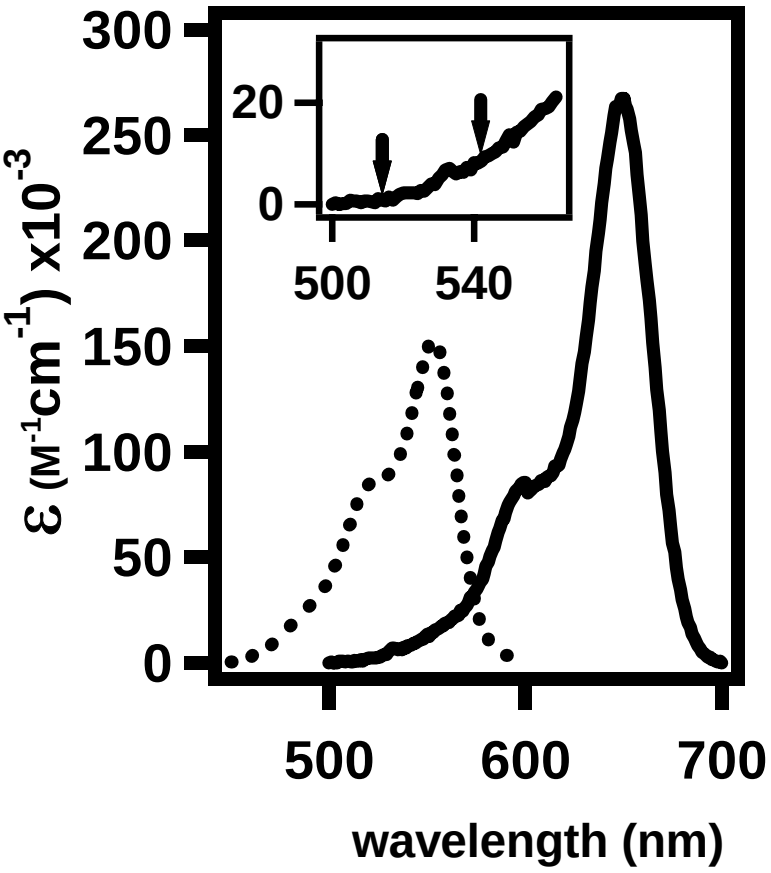


Figure 28. Corrected excitation spectra of the Cy3 donor (dotted line) and Cy5 acceptor (solid lines). The inset provides an expanded view of the acceptor excitation spectrum in the region near the two laser wavelengths to illustrate the large increase in absorbance (an order of magnitude) at 543 nm relative to 514 nm for Cy5.

The peak is symmetric about zero, as expected indicating homogeneous depletion of the zero-peak with no preferential removal of a discernable subpopulation. The difference histogram of data acquired with 543 nm excitation is shown in Figure 29B, and in contrast shows an asymmetric depletion of the zero-peak. The preferential removal of the population from the right edge of the zero-peak is detectable when using 543 nm light and is indicative of an inhomogeneous composition within the zero-peak. When 543 nm excitation is employed, weak excitation of the acceptor fluorophore results in a small shift from zero FRET efficiency for DNA molecules possessing both fluorophore, even though no actual energy transfer is occurring. This results in asymmetric broadening of the zero-peak before titration. This heterogeneity is masked under 514 nm excitation. This indicates that the subpopulation appearing on the right side of the zero-peak with 543 nm excitation is composed of molecules containing an unbleached acceptor fluorophore. Weak direct excitation of the acceptor fluorophore leads to a larger apparent energy transfer because there is emission in the acceptor channel.

Our approach is achieved by using a donor-only labeled DNA sample to determine the correction factor α used in equation 2. This ensures that molecules labeled with donor fluorophores only or with bleached acceptors will be distributed near zero in the FRET histogram.

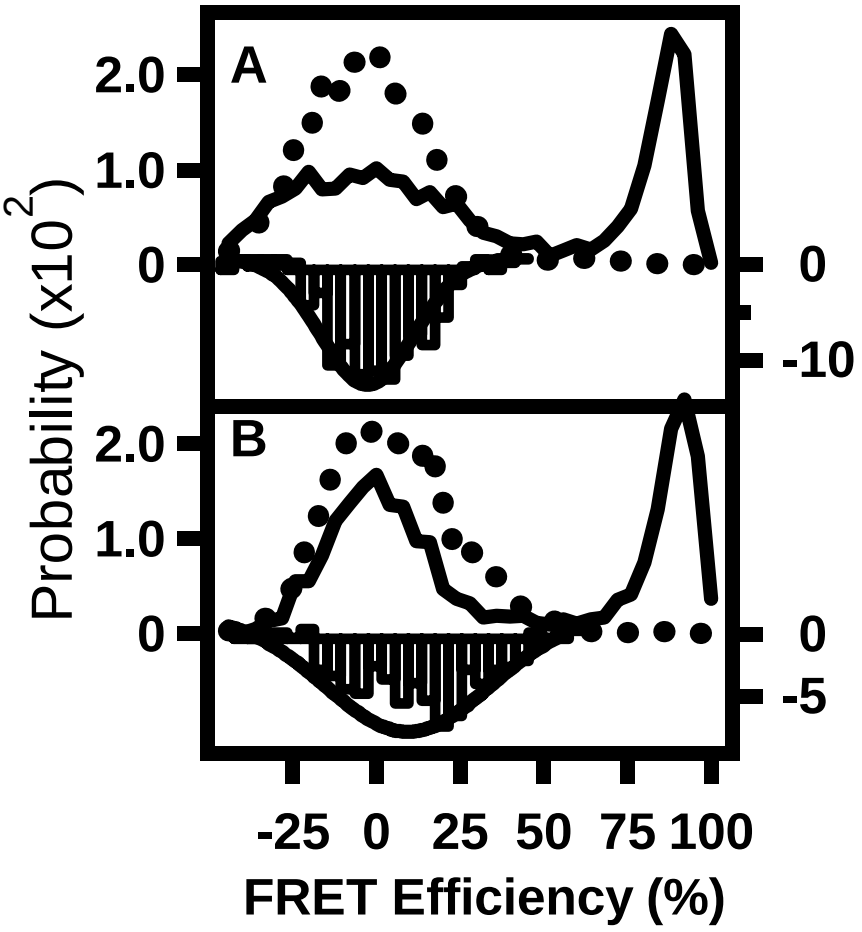


Figure 29. Histogram analysis of the FRET data to determine the zero-peak composition. SM-FRET histograms from the untitrated (dots) and fully titrated (solid line) samples acquired with (A) 514 nm and (B) 543 nm excitation. A difference histogram was constructed for the zero-peak at each wavelength by subtracting the histogram of the untitrated sample (dotted line) from that of the fully titrated sample (solid line). The difference histograms (negative bars) for each wavelength were fit with Gaussian functions (negative solid lines). With 514 nm excitation the difference histogram is symmetrically positioned around zero. In contrast, the difference histogram obtained with 543 nm excitation is shifted to positive values due to the asymmetric way in which the zero-peak is depleted. The parameters obtained from the Gaussian fit in (B) along with parameters obtained from fitting data acquired with donor-only labeled DNA were used to model the zero-peak distribution at all LacI concentrations. The results are displayed in Table 2.

To determine the relative contributions to the zero-peak from DNA molecules with functional acceptor fluorophores and from those with damaged fluorophores, we have used the position and widths of both the difference peak and the donor-only peak to provide basis for modeling the zero-peak as a distribution described by a sum of two Gaussian functions (Equation 13). A Gaussian distribution with an average energy transfer of 9 % and a standard deviation of 22 % is shown superimposed on the calculated difference histogram in Figure 30B. These fit parameters were used in conjunction with parameters taken from the donor-only histogram in describing the double Gaussian distribution used to fit the zero-peak throughout the titration (Table 2). The two Gaussian distributions represent 9C14 molecules in which the acceptor fluorophores are present (apparent FRET efficiency = 9 %) and those in which the acceptor fluorophore was absent or non-emitting (apparent FRET efficiency = 0 %). The existence of Cy5 fluorophores which are nonemissive when attached to nucleic acids is documented, but not explained in the literature (Ha, 2001).

Figure 30A shows the resulting fit (red line) to the zero-peak (histogram) of the untitrated sample. This fit was obtained by holding the widths and means of the two component Gaussians (black lines) constant and allowing a best-fit to be determined by varying the amplitudes of the two individual Gaussians. The result shows that the zero-peak is well described by the linear combination of the two distributions. Omitting either of the Gaussians from the fit results in a significant increase in the χ^2 value.

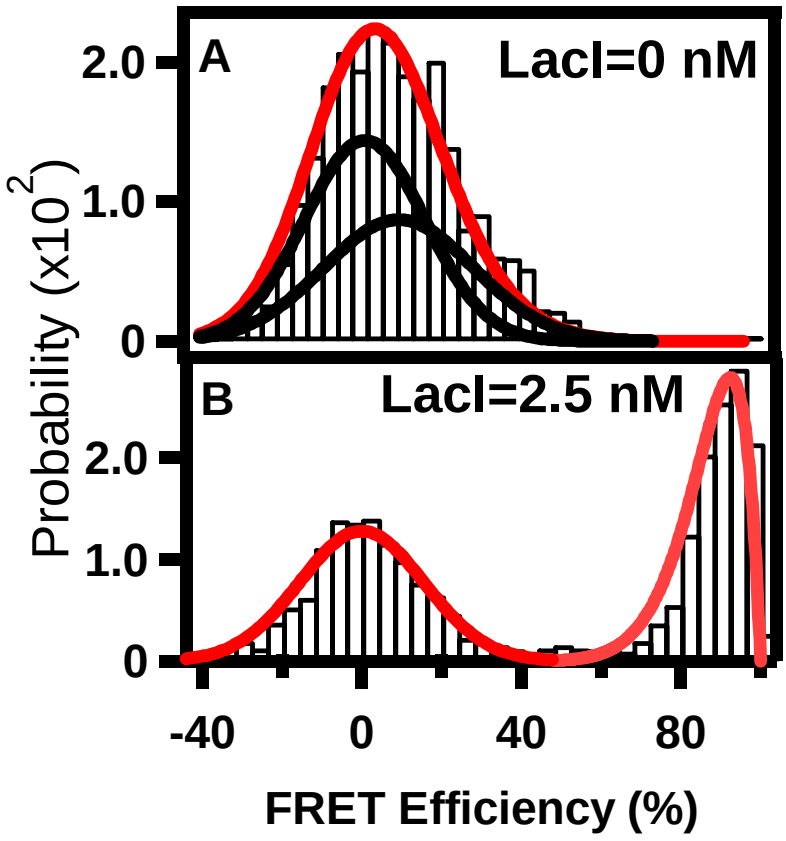


Figure 30. (A) SM-FRET histogram of free 9C14 (bars) acquired with 543 nm excitation. The zero-peak is fit to a sum of two Gaussian distributions using the width and mean parameters determined from fits to histograms of donor-only labeled 9C14 primer and from the difference histogram in Figure 29B. B. Histogram of the fully-titrated 9C14 construct acquired using 543 nm excitation. The distribution is fit with a linear combination of a single Gaussian distribution and a β-function. The Gaussian's position and width are identical to those required to obtain a good fit to a FRET histogram acquired with a donor-only labeled sample.

$$P(Eff) = A_1 e^{-(Eff - <Eff>_1)^2 / \sigma_1^2} + A_2 e^{-(Eff - <Eff>_2)^2 / \sigma_2^2} \quad \text{Equation 13}$$

[LacI] (nM)	A_1	A_2	χ^2 x10^5
0.0	0.014	0.0087	5.8
0.05	0.014	0.0086	6.0
0.5	0.012	0.0054	4.3
2.5	0.014	0.0003	2.6

Table 2. Summary of parameters used to fit the zero-peak distribution at different LacI concentrations. The zero-peak was fit by the sum of two Gaussian distributions (Equation 13). $<Eff>_1$ and $<Eff>_2$ were held constant at 0% and 9% respectively, and σ_1 and σ_2 were held constant at 22% and 28% respectively.

A summary of the parameters obtained from fitting the zero-peak for a range of LacI concentrations is given in Table 2. The amplitude of the Gaussian component centered at zero is constant throughout the range of LacI concentrations. This trend is expected since the population at zero is composed of molecules lacking an unbleached acceptor fluorophore. The contribution of the second Gaussian, describing the population with an intact, emissive acceptor fluorophore steadily decreases with addition of LacI protein. The disappearance of this population is accompanied by the appearance of molecules with high energy transfer as shown in the titration experiment in Figure 25B and Figure 26A. Figure 30B gives the results from fitting the histogram of the fully titrated sample. The β-distribution function has been used to fit the peak at 90 % efficiency and only the Gaussian distribution centered at zero is used to fit the zero-peak. The *β*-function has been used previously

to describe shot-noise limited FRET data (Dahan *et al.*, 1999). Removing the second Gaussian from the fit changes the reduced χ^2 only marginally, from 2.6 x 10^{-5} to 3.2 x 10^{-5}. In comparison, a single Gaussian fit to the zero-peak of the untitrated 9C14 sample gives a reduced χ^2 of 7 x 10^{-5}. The excellent fit obtained with a single Gaussian for the fully-titrated sample is possible because the zero-peak is symmetrically centered about zero, unlike the zero-peak of the untitrated sample which is asymmetric.

From these results we assign the zero-peak of the fully-titrated sample's histogram to a population of 9C14 molecules with either bleached or missing acceptor fluorophore. These molecules, while still capable of forming a "closed" complex, cannot display any energy transfer. Thus all of the 9C14 molecules capable of energy transfer appear in the peak at 90 % ET. From this result we conclude that the 9C14 molecule behaves as predicted for a pure WA configuration: i.e. the titrated species is a closed loop stabilized by a V-shaped LacI tetramer.

Immobilization of 9C14-LacI complexes

During the same time period when the SMS experiments were conducted we also set out to characterize the anticipated dynamic equilibrium between the open and closed loop forms of 9C14. Previous results from bulk FRET, DNA cyclization and EMSA experiments suggested the presence of an open and closed loop geometry for the 9C14 molecule. Our lab had previously proposed that in order for two different 9C14 DNA loop geometries to exist, complex formation would need to incur two distinctly different DNA mechanical strains. To form an open loop with the 9C14 molecule, LacI would have to induce twisting strain on the DNA. However, to form

the closed loop geometry, LacI binding would induce bending strain, as shown in Figure 31. An interconversion process between the two forms would require the release of at least one LacI dimer from the operator sequence.

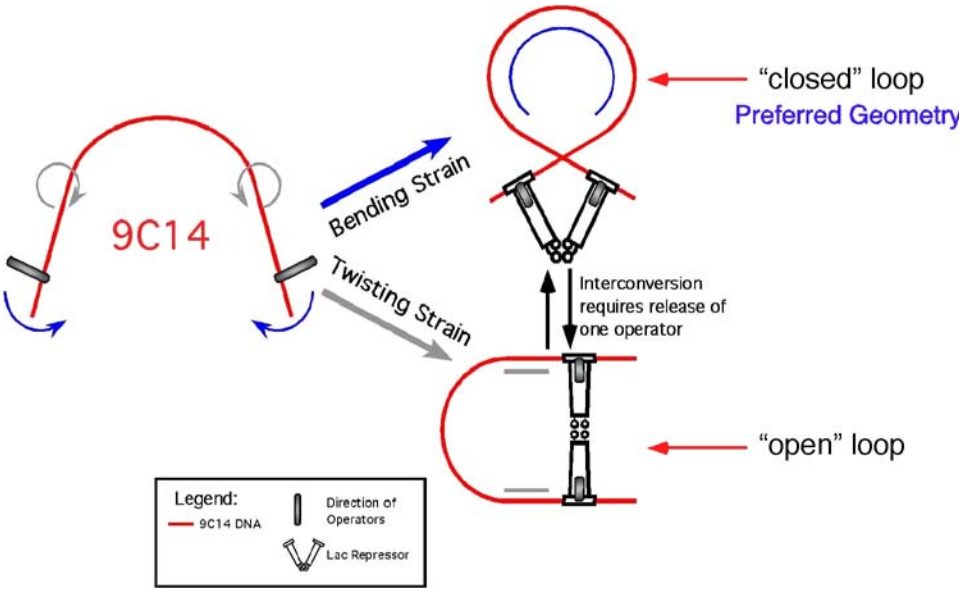


Figure 31. Bending versus twisting, and interconversion. Shown on the left is the unbound 9C14 complex with pseudo dyad axis operators represented as rods extending off of the DNA (red). When LacI binds to the 9C14 complex to form the preferred closed loop, bending strain on the DNA is incurred by the LacI. In order to form the open loop geometry twisting strain is incurred on the DNA by LacI. Interconversion between the two proposed forms would require release of LacI dimer from one operator sequence.

In order to characterize the equilibrium and interconversion kinetics between the two forms we conducted SMS experiments of 9C14-LacI complexes immobilized using agarose, polyacrylamide, poly vinyl alcohol (PVA), aminopropyltriethoxysilane (APTES), and glutaraldehyde. Immobilization in agarose and polyacrylamide gel matrices has proven to be a useful tool for studying DNA-protein complexes. (Segers-Nolten *et al*., 2002 and references therein). In general, small DNA molecules will freely diffuse through an agarose gel medium, where proteins will remain immobile. In principle when a DNA-protein complex is formed the complex will remain immobilized. The confinement of the DNA-protein complex allows one to focus the laser beam onto this molecule for an extended period of time. By using water-based agarose gels the functionality of the molecules is preserved and influence from interactions with a surface is avoided. This allows one to visualize and quantify biomolecular interactions under physiological and equilibrium conditions.

Experiments were conducted using a variety of different immobilization techniques, but results from these experiments did not determine whether an equilibrium existed between the two possible geometries of 9C14-LacI. Immobilization of the 9C14-LacI complex using APTES or PVA were somewhat successful, but we never observed a clear population of molecules exhibiting 100 % ET. We generally saw complexes exhibiting anywhere from 10 %-60 % ET as shown in Figure 32. It is unclear why this was the case; however one possible explanation is that the fluorophores themselves were interacting unfavorably with the immobilization medium. Immobilization using polyacrylamide was unsuccessful. We are currently trying to immobilize the 9C14-LacI complex in agarose. Initial

experiments conducted on the 9C14-LacI complex immobilized in agarose suggested that this approach is probably the best approach. We are currently trying to optimize protocols to immobilize 9C14-LacI complexes in agarose.

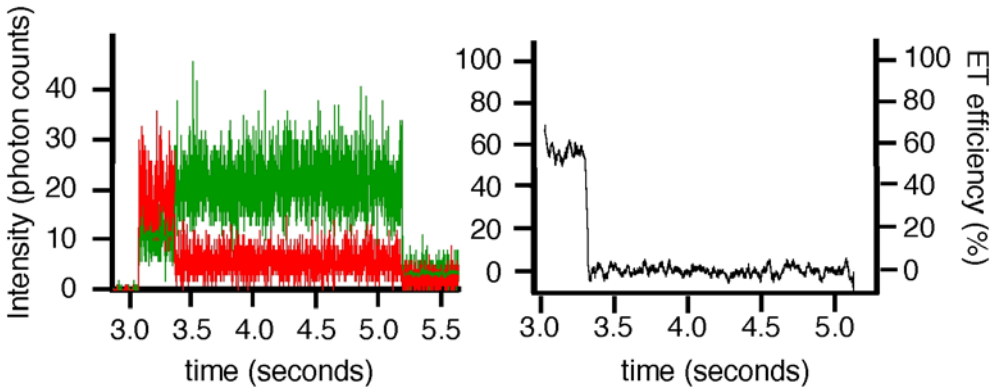


Figure 32. 9C14-LacI immobilized in poly vinyl alcohol (PVA). Left panel. Experimental data collected on 9C14-LacI complex immobilized in 1 % PVA. Note the efficient ET (~60 %), which is indicative from the dominating acceptor signal in red. After approximately 3.25 seconds the acceptor fluorophore most likely becomes photobleached. Right panel. Shows an approximate value of 60 % ET efficiency for this individual 9C14-LacI complex.

Discussion

Previous DNA cyclization kinetics, EMSA (electrophoretic mobility shift assay) and bulk FRET studies conducted on the 9C14-LacI complex gave conflicting results, which could not definitively determine whether the 9C14-LacI complex existed as a closed loop or a mixture of open and closed loop conformations. EMSA results suggested the presence of one unique conformational population that was most likely a closed loop (Mehta & Kahn, 1999). However, DNA cyclization kinetics studies on the 9C14-LacI complex (extended by ~200 bp to enable cyclization) suggested that the complex existed in two forms, a closed loop form (giving a positive supercoil) and an open loop form (giving a relaxed topoisomer). In order to resolve the disparities between the two sets of results, bulk FRET measurements on the 9C14-LacI complex were conducted. Results from these experiments gave a steady state efficiency of ~70 % ET, but this value could be interpreted in two different ways. Either 100 % of the molecules exhibit 70 % ET efficiency, which suggests the presence of a unique closed loop form, or 70% of the molecules exhibit 100 % ET efficiency and the remaining 30 % of the molecules exhibit little or no ET, which suggests the presence of both open and closed loop forms. Any intermediate of the two previous cases is also possible. Thus, bulk FRET measurements did confirm the existence of a closed form but could not determine whether a second form existed or whether intermediate geometries existed (Edelman *et al*., 2003).

SMS experiments reported here on the 9C14-LacI complex definitively confirmed the existence of a single closed loop population characterized by > 90 % ET efficiency. The single molecule results were also able to establish a lower limit of

60 % as the fraction of molecules appearing with > 90 % energy transfer efficiency in the fully bound sample. It is important to emphasize that 60 % is only a lower limit due to the uncertainty surrounding the identity of the molecules which compose the zero-energy transfer peak, and the studies we reported using different excitation wavelengths strongly suggest the actual fraction of double-labeled molecules at high ET is nearly 100%.

Where do the discrepancies lie among the different sets of experiments? First, SM-FRET allows one to determine an unbiased equilibrium distribution of conformations for the 9C14-LacI complex whereas bulk FRET measurements give a weighted average of all values present in the sample. Determining weighted average values can sometimes mask the appearance of the true distribution between different populations in the sample. Also, bulk FRET measurements rely on calculating ET efficiencies from the absolute dye labeling efficiencies, whereas the dual-wavelength experiments allow us to measure SM-FRET efficiencies without knowing labeling efficiencies beforehand. In the ensemble FRET measurements this is one of the most likely sources of error. The labeling efficiencies were determined upon arrival of the fluorescently labeled oligonucleotides, but subsequent processing of fluorescently labeled DNA via PCR and gel purification, Storm imaging, repeated freeze-thaw cycles, exposure to room light and so forth probably caused some irreversible photobleaching of the fluorophores. At the start of a titration experiment there is no true indication as to what percentage of the fluorophores are actually functional, since the absorbance of the sample is typically too small to measure. In SMS experiments, molecules that have mislabeled or bleached acceptors contribute to the zero-peak and

can be quantified using two-color excitation. We suggest that bulk FRET on 100% fully functional double-labeled samples would have shown a uniform population with very high ET, and that the discrepancy between bulk and SM-FRET is due to the overestimation of the f_A (fractional labeling of the acceptor) by Edelman *et al*. (2003). In the case that Edelman *et al.* (2003) had experimentally determined a f_A value of 0.61 he would have calculated a value of 100% ET.

DNA cyclization kinetics studies suggested that the 9C14-LacI complex existed as a mixed population, a closed loop form and an open loop form. This conclusion was based on the appearance of both relaxed and positive topoisomers in cyclization reactions. The most likely source of the discrepancy between these experiments and our results was the addition of ~ 200 bp tails to the ends of the 9C14 molecule. If the tails destabilize the closed form shape, this could result in a perturbing effect on the stability of the different conformers.

In conclusion, SMS has allowed us to definitively determine the existence of one unique population exhibiting nearly 100% ET efficiency for the 9C14-LacI complex. In the future, SMS experiments investigating DNA-protein geometry and topology will provide information on mixed DNA-protein conformational populations.

Concluding Thoughts

Does the research outlined in this dissertation have any impact on the scientific community? My belief is yes. We set out to design, construct and characterize biological tools in order to study complex biological problems. Gene regulation via DNA looping mechanisms has been studied for over a quarter of a century, and yet there is still controversy surrounding the mechanisms involved these processes. I believe we have successfully shown the ability to design DNA-protein complexes to study particular proposed DNA looping geometries. The obvious next set of experiments would be to try and perform SM-FRET on wild type sequences from the Lac operon in order to assess the "true" DNA loop geometry. I believe that SMS-FRET is the technique that will eventually resolve the ongoing dispute of the actual LacI-DNA geometry in solution.

Has the development of unique DNA-LacI self-assembled nanostructures project impacted the scientific community and its research efforts? My belief is that this ongoing project will eventually impact nanotechnology research. We are very close to determining whether our design scheme is going to work or not. This type of research could potentially open a whole new area in the field of bionanotechnology that could lead to the development of novel nanomaterials and devices.

REFERENCES

Bell, C.E. and Lewis, M. (2001). Crystallographic analysis of Lac repressor bound to natural operator O_1. J. Mol. Biol. 312: 921-926.

Binnig, G. and Rohrer, H. (1987). Scanning tunneling microscopy from birth to adolescence. Rev. Mod. Phys. 59: 615-625.

Brenowitz, M., Pickar, A. and Jamison, E. (1991). Stability of a Lac repressor mediated "looped complex". Biochemistry 30: 5986-5998.

Bukrinskaya, A.G. (2004). HIV-1 assembly and maturation. Arch. Virol. 149:1067-1082.

Cantor, C.R. and Shimmel, P.R. (1980). Biophysical Chemistry. San Francisco: W.H. Freeman and Company.

Chen, H. (1998). Design and Construction of Self-Assembled Protein-DNA Nanostructures. Masters thesis University of Maryland, College Park.

Cosa, G., Harbron, E.J., Zeng, Y.N., Liu, H.W., O'Connor, D.B., Eta-Hosokawa, C., Musier-Forsyth, K. and Barbara, P.F. (2004). Secondary structure and secondary structure dynamics of DNA hairpins complexed with HIV-1NC protein. Biophys. J. 87(4):2759-2767.

Crothers, D.M., Haran, T.E. and Nadeau, J.G. (1990). Intrinsically bent DNA. J. Biol. Chem. 256 (13):7093-7095.

Dahan, M., Deniz, A.A., Ha, T.J., Chemla, D.S., Schultz, P.G. and Weiss, S. (1999). Ratiometric measurement and identification of single diffusing molecules. Chemical Physics 247(1):85-106.

Deniz, A.A., Dahan, M., Grunwell, J.R., Ha, T.J, Faulhaber, A.E., Chemla, D.S., Weiss, S. and Schultz, P.G. (1999). Single-pair fluorescence resonance energy transfer on freely diffusing molecules: Observation of Förster distance dependence and subpopulations. Proc. Natl. Acad. Sci. 96(7):3670-3675.

Diekmann, S. (1986). Sequence specificity of curved DNA. FEBS Lett. 95:53-56

Edelman, L.M., Cheong R. and Kahn J.D. (2003). Fluorescence resonance energy transfer over ~130 basepairs in hyperstable lac repressor-DNA loops. Biophys. J. 84(2, pt. 1):1131-1145.

Elghanian, R., Storhoff, J.J., Mucic, R.C., Letsinger R.L. and Mirkin, C.A. (1997). Selective colorimetric detection of polynucleotides based on the distance-dependent optical properties of gold nanoparticles. Science 277:1078.

English, D.S., Furube, A. and Barbara P.F. (2000a). Single-molecule spectroscopy in oxygen-depleted polymer films. Chem. Phys. Lett. 324(1-3):15-19.

English, D.S., Harbron, E.J. and Barbara, P.F. (2000b). Probing photoinduced intersystem crossing by two-color, double resonance single molecule spectroscopy. J. Phys. Chem. A. 104(40):9057-9061.

Fried, M.G. and Hudson, J. M. (1996). Technical comment: DNA looping and lac repressor-CAP interaction. Science. 274:1929-1930.

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

Fukaminato, T., Sasaki, T., Kawai, T., Tamai, N. and Irie, M. (2004). Digital photoswitching of fluorescence based on the photochromism of diarylethene derivatives at a single-molecule level. J. Am. Chem. Soc. 126(45):14843-14849.

Garcia-Parajo, M. F., Veerman, J. A., Kuipers, L. and van Hulst, N.F. (2001). Looking at the photodynamics of individual fluorescent molecules and proteins. Pure Appl. Chem. 73(3):431-434.

Godiska, R., Melodee, R., Schoenfeld, T., Sheets, L., Derr, A. and Mead, D. (2001). The Clonesmart advantage: easy cloning of difficult DNA targets. Elucidations (Epicentre) 1:1-3.

Goodsell, D.S., Kaczor-Grzeskowiak, M. and Dickerson, D.E. (1994). The crystal structure of C-C-A-T-T-A-A-T-G-G. Implications for bending of B-DNA at TA steps. J. Mol. Biol. 239(1):79-96.

Ha, T.J. (2001). Single-molecule fluorescence resonance energy transfer. Methods 25(1):78-86.

Hagerman, P.J. (1985). Sequence dependence of the curvature of DNA: A test of the phasing hypothesis. Biochemistry 24:7033-7037.

Hansma, P.K., Elings, V.B. Marti, O. and Bracker, C.E. (1988). Scanning tunneling microscopy and atomic force microscopy-application to biology and technology. Science 242:209-216.

Hansma, H.G. and J.Hoh. (1994). Biomolecular imaging with the atomic force microscope. Annu. Rev. Biophys. 23: 115-139

Hood, E. (2004). Nanotechnology: Looking as we leap. Environmental Health Perspectives 112(13):A741-A749.

Ishida, M., Oshima, T. and Yutani, K. (2002) Overexpression in *Escherichia coli* of the AT-rich trpA and trpB genes from the hyperthermophilic archaeon *Pyrococcus Furiosus*. FEMS Microbiology Letters 216:179-183.

Jacob, F. and Monod, J. (1961). Genetic regulatory mechanisms in the synthesis of proteins. J. Mol. Biol. 3: 318-356.

Jares-Erijman, E. A., and Jovin, T.M. (2003). FRET imaging. Nature Biotechnology 21(11):1387-1395.

Kahn, J. D. and Crothers, D.M. (1992). Protein-induced bending and DNA cyclization. Proc. Natl. Acad. Sci. USA 89:6343-6347.

Kahn, J. D. and Crothers, D.M. (1998). Measurement of the DNA bend angle induced by the catabolite activator protein using Monte Carlo simulation of cyclization kinetics. J. Mol. Biol. 276:287-309.

Kalodimos, C.G., Bonvin, A.M.J.J., Salinas, R.K., Wechselberger, R.B., and Kaptein, R. (2002). Plasticity in protein-DNA recognition: *lac* repressor interacts with its natural operator O1 through alternative conformations of its DNA-binding domain. EMBO J. 21(12):2866-2876.

Kapanidis, A.N., Lee, N.K., Laurence, T.A., Doose, S., Margeat, E. and Weiss S. (2004). Fluorescence-aided molecule sorting: Analysis of structure and interactions by alternating-laser excitation of single molecules. Proc. Natl. Acad. Sci. USA 101(24):8936-8941.

Kasche, V. and Lindqvist, L. (1964). Reactions between the triplet state of fluorescein and oxygen. J. Phys. Chem. 68(4):817-23.

Koo, H.S., Wu, H. and Crothers, D.M. (1986). DNA bending at Adenine•Thymine tracts. Nature 320:501-506.

Koo, H.S., Drak, J., Rice, J.A. and Crothers, D.M. (1990). Determination of the extent of DNA bending by an Adenine-Thymine tract. Biochemistry 29:4227-4234.

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

Kulzer, F., Kummer, S., Matzke, R., Brauchle, C. and Basche, T. (1997). Single-molecule optical switching of terrylene in p-terphenyl. Nature 387(6634):688-691.

Lakowicz, J.R. (1999). Principles of fluorescence spectroscopy: Second Edition. New York: Kluwer Academic/Plenum Publishers.

Lavigne, M., Kolb, H.M. and Buc, H. (1992). Upstream curved sequences influence the initiation of transcription at the Escherichia coli galactose operon. J. Mol. Biol. 220:293-306.

Levene, S.D. and Crothers, D.M. (1983). A computer graphics study of sequence-directed bending in DNA. J. Mol. Biomol. Struct. Dyn. 1:429-435.

Levene, S.D. and Crothers, D.M. (1986a). Ring closure probabilities for DNA fragments by Monte Carlo simulation. J. Mol. Biol. 189:61-72.

Levene, S.D. and Crothers, D.M. (1986b). Topological distributions and the torsional rigidity of DNA: A Monte Carlo study of DNA circles. J. Mol. Biol. 189:73-83.

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

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

Li, H., Ren, X., Ying, L., Balasubramanian, S. and Klenerman, D. (2004). Measuring single-molecule nucleic acid dynamics in solution by two-color filtered ratiometric fluorescence correlation spectroscopy. Proc. Natl. Acad. Sci. USA 101(40):14425-14430.

Lobell, R.B. and Schleif, R.F. (1991). AraC-DNA looping: Orientation and distance-dependent loop breaking by cyclic AMP receptor protein. J. Mol. Biol. 218(1):45-54.

Lyubchenko, Y.L., Gall, A.A., Shlyakhtenko, L.S., Harrington, R.E., Jacobs, B.L., Oden, P.I. and Lindsay, S.M. (1992a). Atomic force microscopy imaging of double stranded DNA and RNA. J. Biomol. Struct. Dyn. 10(3):589-606.

Mao, C. (2004). The emergence of complexity: Lessons from DNA. PLoS Biology 2 (12): 2036-2038.

Mao, C., Sun, W., Shen, Z., and Seeman, N.C. (1999). A nanomechanical device based on the B-Z transition of DNA. Nature 397:144-146.

Marck, C. (1992). DNA Strider, v. 1.2. Gif sur Yvette, France.

Marini J.C., Levene S.D., Crothers, D.M., and Englund, P.T. (1983). A bent helix in kinetoplast DNA. Cold Spring Harbor Symp. Quant. Biol. 47(1) 1: 279-83.

Mehta, R.A. and Kahn, J.D. (1999). Designed hyperstable lac repressor•DNA loop topologies suggest alternative loop geometries. J. Mol. Biol. 294(1):67-77.

Moerner, W.E. and Fromm, D.P. (2003). Methods of single-molecule fluorescence spectroscopy and microscopy. Scientific Instruments 74(8):3597-3619.

Morris, V.J., Kirby, A.R. and Gunning, A.P. (1999). Atomic force microscopy for biologists. Imperial College Pres. Covent Garden, London.

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

Müller-Hill, B. (1998). The function of auxiliary oporators. Mol. Microbiol. 29:13-18.

Müller, J., Barker, A., Oehler, S. and Müller-Hill, B. (1998). Dimeric lac repressors exhibit phase-dependent co-operativity. J. Mol. Biol. 284:851-857.

Padilla, J.E., Colovos, C. and Yeates, T.O. (2001). Nanohedra: Using symmetry to design self assembling protein cages, layers, crystals, and filaments. Proc. Natl. Acad. Sci. USA 98(5):2217-2221.

Perros, M. and Steitz, T.A. (1996). Technical comment: DNA looping and lac repressor-CAP interaction. Science. 274:1929-1930.

Ratneshwar, L. and Scott, J. A. (1994). Biological applications of atomic force microscopy. Am. J. Physiol. 266:C1-C21.

Rivetti, C., Codeluppi, S., Dieci, G. and Bustamante, C. (2003). Visualizing RNA extrusion and DNA wrapping in transcription elongation complexes of bacterial and eukaryotic RNA polymerases. J. Mol. Biol. 5:1413-1426.

Sa-Ardyen, P., Jonoska, N. and Seeman, N.C. (2003). Self-assembling DNA graphs. DNA-based computers VIII, LNCS 2568, Springer-Verlag, Berlin, 1-9.

Sadler, J.R., Sasmor, H. and Betz, J.L. (1983). A perfectly symmetric lac operator binds the lac repressor very tightly. Proc. Natl. Acad. Sci. USA. 80(22): 6785-6789.

Santero, E., Hoover, T.R., North, A.K., Berger, D.K., Porter, S.C. and Kustu, S. (1992). Role of integration host factor in stimulating transcription from the sigma 54-dependent nifH promoter. J. Mol. Biol. 227:602-620

Segers-Nolten G.M., Wyman, C., Wijgers, N., Vermeulen, W., Lenferink, A.T., Hoeijmakers, J.H., Greve, J. and Otto, C. (2002). Scanning confocal fluorescence microscopy for single molecule analysis of nucleotide excision repair complexes. Nucleic Acids Res. 21:4270-4277.

Selvin, P.R. (1995). Fluorescence resonance energy transfer. Methods in Enzymology 246:301-334.

Shlyakhtenko, L.S., Gall, A.A., Weimer, J.J., Hawn, D.D. and Lyubchenko, Y.L. (1999). Atomic force microscopy imaging of DNA covalently immobilized on a functionalized mica substrate. Biophys. J. 77:568-576.

Simons, A., Tils, D., von Wilcke-Bergmann, B. and Muller-Hill, B. (1984). Possible ideal lac operator: Escherichia coli lac operator-like sequences from eukaryotic genomes lack the central G X C pair. Proc. Natl. Acad. Sci. USA 81(6):1624-1628.

Tinnefeld, P., Buschmann, V., Weston, K. and Sauer, M. (2003). Direct observation of collective blinking and energy transfer in a bichromophoric system. J. Phys. Chem. A. 107(3):323-327.

Turro, N.J. (1991). Modern Molecular Photochemistry. Mill Valley, CA: University Science Books.

Virnik, K., Lyubchenko, Y.L., Karymov, M.A., Dahlgren, P., Tolstorukov, M.Y., Semsey, S., Zhurkin, V.B. and Adhya, S. (2003). Antiprallel DNA loop in gal repressosome visualized by atomic force microscopy. J. Mol. Biol. 334:53-63.

Wacker, R., Schröder, H., and Niemeyer, C. M. (2004). Performance of antibody microarrays fabricated by either DNA-directed immobilization, direct spotting, or streptavidin-biotin attachment: a comparative study. Anal. Biochem. 330:281-287.

Weaver, R.F. (2005). Molecular Biology: Third Edition. Boston: McGraw-Hill Company.

Weiss, S. (1999). Fluorescence spectroscopy of single biomolecules. Science. 283:1676-1683.

Weiss, S. (2000). Measuring conformational dynamics of biomolecules by single molecule fluorescence spectroscopy. Nat. Struct. Biol. 7:724-729.

Weston, K.D., Carson, P.J., DeAro, J.A. and Buratto, S.K. (1999). Single-molecule detection fluorescence of surface-bound species in vacuum. Chem. Phys. Lett. 308:58-64.

Whitson, P. A., Hsieh, W.-T., Wells, R. D. and Mathews, K. S. (1987). Influence of supercoiling and sequence context on operator DNA binding with lac repressor. J. Biol. Chem. 262:14592-14599.

Xie, X.S, and Trautman, J. K. (1998). Single-molecule optical studies at room temperature. Ann. Rev. Phys. Chem. 49:441-480.

Zhang, S. (2003). Fabrication of novel biomaterials through molecular self-assembly Nat. Biotechnol. 21(10):1171-1178.

Zinkel, S.S. and Crothers, D.M. (1987). DNA bend direction by phase sensitive detection. Nature 328: 178-181.