Title
Tools for Single-Molecule Sequence Detection and Manipulation

Author
Zohar, Hagar

Publication Date
2011

Peer reviewed|Thesis/dissertation

Tools for Single-Molecule Sequence Detection and Manipulation

By

Hagar Zohar

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Chemical Engineering

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Susan J. Muller, Chair
Professor Harvey W. Blanch
Professor Marcin Majda

Spring 2011

Tools for Single-Molecule Sequence Detection and Manipulation

by Hagar Zohar

Abstract

Tools for Single-Molecule Sequence Detection and Manipulation

by

Hagar Zohar

Doctor of Philosophy in Chemical Engineering

University of California, Berkeley

Professor Susan J. Muller, Chair

The ability to strongly and sequence-specifically attach modifications such as fluorophores and haptens to double-stranded DNA is critical to a variety of single-molecule experiments. We present a practical guide to labeling internal sequences of double-stranded (ds) DNA molecules for single-molecule experiments. Since single-molecule experiments are diverse in their requirements, we focus on six approaches and consider the merits and drawbacks of each. By presenting a set of criteria relevant to single-molecule experiments (e.g. labeling yield, compatibility with cofactors such as Mg^{2+}), the guide provides a simple reference for selecting a labeling approach for given experimental constraints. We study one of these approaches, Peptide Nucleic Acids (PNAs), in detail. Slide stretching is used to optimize DNA-PNA binding and microfluidic trapping and stretching in a cross-slot demonstrates fluorescence-based sequence detection using PNAs on a long, roughly 50 kb, untethered DNA molecule. Optical tweezer experiments on DNA-bound PNAs further demonstrate that PNAs are versatile and robust sequence-specific tethers. Lastly, we assess the feasibility of studying DNA-protein interactions between stained DNA and fluorescently-labeled protein at the single-molecule level in the microfluidic cross-slot using the motor protein SpoIIIE.

Table of Contents

Chapter 1

Introduction

1.1 Studying single molecules

Single-molecule experiments have provided unprecedented insight into a wide array of biological systems, most notably into the mechanics of DNA and proteins.[1-6] Individual molecules can be studied via force spectroscopy, which includes optical tweezers, magnetic tweezers, and atomic force microscopy (AFM), or fluorescence and electron microscopy.[7] The molecule of interest must be immobilized,[8] trapped,[9] confined,[10] or tracked.[4] This work focuses on DNA and utilizes fluorescence microscopy and a microfluidic cross-slot (Figure 1-1) to trap and stretch individual untethered DNA molecules. Of particular interest are methods for labeling DNA sequence-specifically for single molecule experiments, the characterization of peptide nucleic acids (PNA) as sequence-specific labels and tweezer handles, and the potential for examining DNA-protein interactions in the cross-slot.

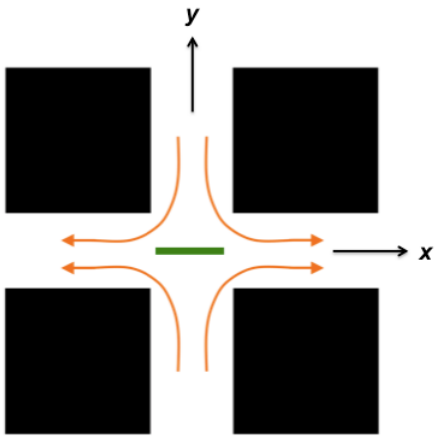


Figure 1-1. Schematic of a microfluidic cross-slot with streamlines (orange) indicating the direction of flow and a DNA molecule (green) trapped and stretched at the central stagnation point. The DNA molecule is completely untethered.

1.2 Microfluidic trapping platform

The microfluidic cross-slot provides an elegant avenue to trapping and stretching individual DNA molecules. The extent of stretching can be tuned by varying a dimensionless parameter, the Deborah number, described below. Dilute solutions of DNA molecules are introduced into the device and individual molecules are trapped and stretched by fluid flow alone so no surface tethering or surface immobilization is necessary. The two inlet streams meet at the center of the device, the region of observation, so the initial interaction of disparate reagents from separate inlet streams can be observed. The composition of the two inlet streams can be easily and independently varied throughout experiments, allowing for rapid changes in buffer conditions and reagents. Since each molecule is exposed to the same flow field, stretching at steady-state is highly reproducible from one molecule to the next.

The microfluidic cross-slot traps and stretches individual DNA molecules based on the principle that the conformation of DNA, or any polymer, in flow is governed by two timescales. The first is the relaxation time of the polymer in the fluid, often referred to as the characteristic or longest relaxation time, τ. Conceptually, τ is the time necessary for the polymer's conformation to rearrange due to Brownian motion.[11] Experimentally, τ is varied for a given polymer (of a specified molecular weight) by changing the fluid viscosity, typically by adding viscosifying agents such as sucrose or glycerol. The second timescale is a time constant, t, which for an extensional flow with strain rate $\dot{\varepsilon}$ is $t = \frac{1}{\dot{\varepsilon}}$. The strain rate, $\dot{\varepsilon}$, is a velocity gradient along the direction of flow (along the x-axis, as in Figure 1-1):

$$\dot{\varepsilon} = \frac{d\varepsilon}{dt} = \frac{dv_x}{dx} \qquad (1)$$

Experimentally, $\dot{\varepsilon}$ is varied by changing the geometry of the microfluidic device or the flow rate. The Deborah number, *De*, is a dimensionless quantity defined as the ratio of these two timescales:

$$De = \frac{\tau}{t} = \tau\dot{\varepsilon} \qquad (2)$$

At high *De*, the relaxation time of the polymer is longer than the timescale of flow. Consequently, the polymer cannot rearrange quickly enough to recover its undisturbed coiled state and is stretched. At low *De*, the relaxation time is shorter than the timescale of the flow and the polymer rearranges quickly enough to remain coiled.

The most dramatic effect of flow on polymer conformation, alluded to above, is the onset of the coil-stretch transition, studied extensively in purely extensional flow.[9,11,12] The coil-stretch transition occurs when the hydrodynamic force on the polymer overwhelms the polymer's entropic elasticity, causing extension. Intuitively, we expect the coil-stretch transition to occur at *De* of order one and this has been verified experimentally.[11] The theoretical prediction for the onset of the coil-stretch transition, using the Zimm dumbbell model, is at $De = 0.5$.[13] Therefore, *De* is of critical importance as it allows for the prediction of the conditions required to achieve stretched DNA. The stretching of DNA has implications for the DNA molecule's ability to bind probes and is also necessary for determining the positions of bound probes along the length of the DNA molecule.

The microfluidic cross-slot is a simple device that produces pure extensional flow at its center stagnation point. It has two opposing inlet and outlet streams. The extensional flow field the molecule experiences is characterized by:

$$\begin{aligned} v_x &= \dot{\varepsilon}x \\ v_y &= -\dot{\varepsilon}y \end{aligned} \qquad (3)$$

The extent of stretching of the DNA molecule can be precisely tuned in the cross-slot by varying *De*[9] and molecules can be trapped for several hours.[12] The extensional flow produced at the center of the device is characterized by the flow type parameter, ξ, defined:

$$\xi = \frac{\left|\underline{\underline{E}}\right| - \left|\underline{\underline{\Omega}}\right|}{\left|\underline{\underline{E}}\right| + \left|\underline{\underline{\Omega}}\right|} \qquad (4)$$

where $\left|\underline{\underline{E}}\right|$ is the magnitude of the deformation tensor and $\left|\underline{\underline{\Omega}}\right|$ is the magnitude of the vorticity tensor. For velocity **v,** the velocity gradient, $\nabla\mathbf{v}$, can be expressed as the sum of the deformation tensor, $\underline{\underline{E}}$, and the vorticity tensor, $\underline{\underline{\Omega}}$:

$$\nabla \mathbf{v} = \underline{\underline{E}} + \underline{\underline{\Omega}} \quad (5)$$

where $$\underline{\underline{E}} = \frac{1}{2} [\nabla \mathbf{v} + (\nabla \mathbf{v})^t] \quad (6)$$

and $$\underline{\underline{\Omega}} = \frac{1}{2} [\nabla \mathbf{v} - (\nabla \mathbf{v})^t] \quad (7)$$

Flow type is parameterized such that pure rotation corresponds to $\xi = -1$ and pure extension corresponds to $\xi = 1$. Shear, the superposition or intermediate of rotational and extensional flows, corresponds to $\xi = 0$ and other mixed flows occur at fractional values of ξ. For the cross-slot, with equal flow rates in the two opposing arms, $\xi = 1$. The microfluidic four-roll mill,[14] a more sophisticated device, can be used to create stagnation-points of varying ξ independently of De.

1.3 Overview

A discussion of several techniques that can be used to label DNA at specific internal sequences for single-molecule experiments is presented Chapter 2. These techniques include triplex-formers and more novel enzymatic-based approaches. Since single-molecule experiments are diverse in their requirements, each technique is evaluated according to a set of experimental criteria and its advantages and disadvantages are discussed. A library of such techniques is critical to single-molecule DNA studies where optical signals or attachment points to surfaces must be present at specific sequences.

One of these labeling approaches, Peptide Nucleic Acids (PNAs), is the subject of Chapter 3. The utility of PNAs in two model single-molecule experiments where individual DNA molecules are manipulated via microfluidic flow and optical tweezers, respectively, is demonstrated. The results of this work show that PNAs are versatile and robust sequence-specific tethers and their binding optimization is discussed in detail.

In Chapter 4, a new system is developed for study in the microfluidic cross-slot. The motor protein SpoIIIE, which rapidly translocates along DNA, is a fascinating protein to observe directly. However, since the conditions most conducive to microfluidic DNA trapping and visualization are not the most conducive for SpoIIIE activity (and vice versa), the feasibility of such a study had to be assessed. Through bulk and single-molecule experiments, the results of this work show that it is possible to reconcile these experimental constraints and simultaneously achieve SpoIIIE activity and the visualization of DNA in flow. Lastly, Chapter 5 summarizes major conclusions and future directions for this work.

1.4 References

1. Bustamante, C., Bryant, Z. & Smith, S. B. Ten years of tension: single-molecule DNA mechanics. *Nature* **421**, 423-427 (2003).
2. Bustamante, C., Smith, S. B., Liphardt, J. & Smith, D. E. Single-molecule studies of DNA mechanics. *Curr. Opin. Struct. Biol.* (2000).
3. Chu, S. Biology and polymer physics at the single-molecule level. *Phil. Trans. R. Soc. Lond. A* **361**, 689-698 (2003).
4. Greenleaf, W., Woodside, M. & Block, S. High-Resolution, Single-Molecule Measurements of Biomolecular Motion. *Annu. Rev. Biophys. Biomol. Struct.* **36**, 171-190 (2007).
5. Neuman, K. & Nagy, A. Single-molecule force spectroscopy: optical tweezers, magnetic tweezers and atomic force microscopy. *Nat. Methods* **5**, 491-505 (2008).
6. Roy, R., Hohng, S. & Ha, T. A practical guide to single-molecule FRET. *Nat. Methods* **5**, 507-516 (2008).
7. Walter, N., Huang, C., Manzo, A. & Sobhy, M. Do-it-yourself guide: how to use the modern single-molecule toolkit. *Nat. Methods* **5**, 475-489 (2008).
8. Visnapuu, M., Duzdevich, D. & Greene, E. C. The importance of surfaces in single-molecule bioscience. *Mol. BioSyst.* **4**, 394-403 (2008).
9. Perkins, T. T., Smith, D. E. & Chu, S. Single Polymer Dynamics in an Elongational Flow. *Science* **267**, 2016-2021 (1997).
10. Leslie, S. R., Fields, A. P. & Cohen, A. E. Convex Lens-Induced Confinement for Imaging Single Molecules. *Anal. Chem.* **82**, 6224-6229 (2010).
11. Smith, D. E. & Chu, S. Response of Flexible Polymers to a Sudden Elongational Flow. *Science* **281**, 1335-1340 (1998).
12. Schroeder, C. M., Babcock, H. P., Shaqfeh, E. S. G. & Chu, S. Observation of Polymer Conformation Hysteresis in Extensional Flow. *Science* **301**, 1515-1519 (2003).
13. Larson, R. G. & Magda, J. J. Coil-stretch transitions in mixed shear and extensional flows of dilute polymer solutions. *Macromolecules* (1989).
14. Lee, J. S., Dylla-Spears, R., Teclemariam, N. P. & Muller, S. J. Microfluidic four-roll mill for all flow types. *Appl. Phys. Lett.* (2007).

Chapter 2

Labeling Internal Sequences on Double-Stranded DNA: Methods for Single-Molecule Experiments

2.1 Introduction

The burgeoning field of single-molecule biophysics has produced new techniques that characterize individual biological molecules, allowing for observations unattainable with conventional bulk methods.[1] A common challenge for single-molecule experiments is that each experimental setup must reconcile several constraints, including requirements of (i) the molecule(s) of interest (e.g. DNA, protein) and (ii) the platform of study (e.g. AFM, optical tweezers, fluorescence microscopy).[2] Furthermore, many such experiments require the labeling of internal specific sequences on long (kb) double-stranded (ds)DNA.[3-8] Since single-molecule experiments are diverse in their requirements, there is a need for multiple labeling methods. Beyond the scope of single-molecule experiments, DNA labeling is of general interest to biotechnology as a strategy to map DNA[9-11] or achieve molecular scaffolding.[12]

This review is intended as a practical guide for experimentalists interested in specifically labeling dsDNA molecules for single-molecule experiments. We focus on six approaches demonstrated in a single-molecule context and describe relevant features of the molecular tools they utilize. The merits and drawbacks of each labeling method are considered with particular attention to the amount of specialized training and reagents required. By presenting a set of criteria relevant to single-molecule experiments (e.g. labeling yield, compatibility with cofactors such as Mg^{2+}) we provide a simple reference for selecting an approach for given experimental constraints. As DNA labeling is of great interest to many scientific disciplines, we note that several excellent reviews are devoted to this topic.[13-15] This review, however, is intended for non-specialists interested in accessible solutions to DNA labeling challenges and the approaches outlined emphasize simplicity, robustness, suitability for use by non-biologists, and utility in diverse single-molecule experiments.

2.2 Features of the molecular "toolkit"

Sequence-specific DNA labeling methods utilize the properties of several small molecules and enzymes that, together, make up a molecular "toolkit." In this context, labeling DNA refers to attaching small molecules or moieties to DNA that are not native to the DNA structure. Labels are often fluorophores[4,16-20] or haptens[3,21-24] that can create optical signals or attachment points, respectively. Fluorophores, such as the commonly used Cy dyes and Alexa Fluor dyes, allow for direct detection and localization via fluorescence microscopy. These and other fluorophores are commercially available and provide coverage over a wide range of excitation and emission spectra. Hapten labels can serve as specific attachment points for a larger protein-conjugated element, such as a microsphere, which would be challenging to attach directly to DNA.

Haptens are small molecules that form strong, non-covalent, bonds to a specific protein or antibody.[25] Commonly used haptens include biotin, digoxigenin, and fluorescein, which bind to avidin family, anti-digoxigenin, and anti-fluorescein proteins, respectively. Of the hapten-protein pairs, biotin-avidin is notable as the strongest, with nearly the strength of a covalent bond, and most widely used.[26] Since avidin proteins have four biotin binding sites, two or more biotinylated elements can be joined in the presence of a free avidin protein serving as a bridge. Hapten-conjugated nucleotides (e.g. biotin-dUTP, DIG-dUTP, fluorescein-dUTP) and protein-coated spheres (quantum dots (qdots), nano- and microspheres) are commercially available. In the following section, we describe the relevant features of tools that are critical to the sequence-specific DNA labeling methods highlighted in this review.

2.2.1 Small molecules

Oligonucleotides (Oligos) and DNA fragments: An oligo is a short, typically less than 100 base, single-stranded sequence of DNA. Oligos of any desired sequence are

commercially available and made to order via automated solid-phase synthesis.[15,27] The synthesis can accommodate 5' and 3' modifications and alternative nucleotide bases. A DNA fragment is a portion of dsDNA that is hundreds or thousands of base pairs long. Fragments are typically created via restriction enzyme digest of a given dsDNA molecule and consist of the dsDNA between two cleavage sites or a cleavage site and the end of the DNA molecule. The length and sequence of the resulting fragments are constrained by the sequence of the complete DNA molecule and the locations of the restriction sites on the DNA molecule. Online tools like NEBcutter (http://www.tools.neb.com) can be used to identify appropriate enzyme-DNA combinations for generating fragments of a particular length from a known DNA sequence and commercially available enzymes. Polymerase chain reaction (PCR) is a more flexible yet challenging method of creating fragments. A prescribed region of a known dsDNA sequence, up to several kb, can be selectively amplified and then isolated. A pair of oligo primers, each roughly 20 bases, flanking the prescribed region must be designed according to well-known guidelines.[28] Tools like Invitrogen's OligoPerfect™ Designer and Vector NTI Software (http://www.invitrogen.com) aid oligo primer pair design.

Peptide Nucleic Acids (PNAs): PNAs are synthetic molecules composed of a peptide backbone and nucleic acid bases, which confer upon them DNA sequence recognition.[29-32] They form strong and specific complexes with dsDNA, via four different binding modes, depending on their design.[33] PNAs bind optimally to target sequences roughly 8 to 10 bases in length via the strongest binding mode.[3,29,33] Due to a neutrally-charged backbone, PNA binding to dsDNA deviates substantially from that of conventional oligos with respect to binding conditions and yield. The alternative backbone also makes PNAs nuclease and protease resistant.[3] These differences are exploited in the many applications of PNAs, which are covered in a number of excellent general reviews.[34-38] PNAs are commercially available and made to order. Design considerations for PNAs include the binding mode, the length of the targeted sequence, the incorporation of modifications, and the addition of positively charged features such as lysines to enhance binding rates. Vendors typically evaluate PNA design for issues affecting synthesis feasibility and yield.

2.2.2 Enzymes

Enzymes are proteins that catalyze chemical reactions. Some enzymes require cofactors such as Mg^{2+} and adenosine triphosphate (ATP) to function. Commercially available enzymes are provided with buffer solutions containing the requisite components for optimal performance. The enzymatic reactions of interest here include those that either create or break bonds along the backbone of DNA and those that modify bases on DNA. Below we focus on a very restricted set of enzymes and their properties that are useful for preparing constructs for single-molecule experiments. All of the enzymes described below are commercially available and are typically supplied with technical protocols for optimal performance.

DNA Ligase: DNA ligase repairs nicks, which are discontinuities in one of the phosphodiester backbones of the DNA double helix, by forming a phosphodiester bond between the 5' phosphate and 3' hydroxyl termini of a DNA strand.[39] The nick-repairing action of ligase is critical to covalently joining together dsDNA fragments or an oligo to dsDNA. Fragment ends are either blunt ends lacking unpaired overhanging bases, or sticky ends consisting of single-stranded overhangs of one or more bases available to bind with other complementary overhangs. In the latter case, an annealing or hybridization step must first join together the fragments with complementary overhangs, creating a construct with nicks. Ligase then repairs these nicks and covalently joins the fragments. T4 DNA ligase is notable for also joining fragments with blunt ends. Since ligase forms the phosphodiester bond between the 5' phosphate and 3' hydroxyl termini of DNA, the 5' end must be phosphorylated. If the 5' end is a free hydroxyl group, it can be phosphorylated using polynucleotide kinase (PNK) to allow for subsequent ligation. If the 5' end is phosphorylated, the 5' phosphate can be removed with a phosphatase to prevent ligation.

Restriction Enzymes: The most commonly used restriction enzymes are type II, which typically bind to and cleave at target sites on dsDNA that are 4-8 nucleotides in length. In bacteria, they are employed to cleave and destroy foreign DNA. Binding of the enzyme to DNA, under optimized conditions, is highly specific to target sites called cognate sequences. Upon binding to target sites, most type II restriction enzymes require a cofactor such as Mg^{2+} for cleavage.[40] Restriction enzymes can cleave the two strands of DNA at staggered positions, forming short sticky-ends or cleave both DNA strands at the same location forming blunt ends. The fragments of DNA resulting from cleavage can be covalently reattached by the action of DNA ligase. Type II restriction enzymes are composed of two subunits, each of which contains an active site where cleavage of one strand of DNA occurs. In a few instances, the subunits are not identical and one subunit can be mutated and rendered inactive such that the action of the dimer forms a single-stranded break, or nick, without cleaving the dsDNA molecule.[41] In the absence of the cofactor, some enzymes will bind to the target site and not cleave. Over 600 type II restriction enzymes, with 200 distinct recognition sequences, are commercially available.[42,43]

Nicking Endonucleases (Nickases): Nickases cleave one strand of dsDNA, creating a single-stranded break or nick. Nickases, unlike restriction enzymes that cleave both DNA strands, are uncommon in nature but can be engineered.[41,44] One approach to the engineering process, mentioned above, involves starting with restriction enzymes with disparate catalytic hydrolysis sites and inactivating only one of the sites by mutation.[41] The nickase DNase I produces randomly generated nicks while other nickases form nicks at specific sequences. Over 10 nicking enzymes with specific and known recognition sites are commercially available and over 200 have been characterized.[42,43]

DNA Polymerases: DNA Polymerases use single-stranded (ss)DNA as a template to catalyze the synthesis of the complementary strand from deoxynucleotides (dNTPs).[39] The dNTPs include the four nucleoside triphosphates dATP, dCTP, dGTP, dTTP. DNA

Polymerase adds a dNTP to the free 3'hydroxyl group at the end of the DNA strand complementary to the template strand. Consequently, DNA polymerase cannot begin synthesizing a new strand unless the beginning of the growing strand is present. An oligo can be used as a primer to initiate the synthesis of the growing strand. The synthesis of the growing strand can be terminated by the presence of 2',3'-dideoxynucleoside triphosphates, collectively referred to as ddNTPs, which are nucleoside triphosphates that lack the 3' hydroxyl group. When ddNTPs are incorporated into the chain by a polymerase, chain growth is terminated because addition of the subsequent nucleotide requires a free 3' hydroxyl group.[4]

Some DNA Polymerases, such as *E. Coli* DNA Polymerase I, perform other functions including 5' to 3' exonuclease activity that translates, or moves, a nick forward toward the 3' end of DNA, usually by at least tens of nucleotides. This is called nick translation, during which a polymerase excises a series of deoxynucleotides in the DNA molecule and replaces them with free dNTPs added to the reaction.[39,45] Since some polymerases will incorporate modified dNTPs (e.g., biotin-dUTP), nick translation can be used to replace nucleotides in DNA with modified nucleotides introduced into solution.[4] The extent of nick translation can be controlled by the addition of ddNTPs. A polymerase that lacks exonuclease activity, such as Vent (exo-) Polymerase, translates nicks similarly except that it produces a displaced strand during translation instead of excising nucleotides.[16]

DNA Methyltransferases (MTases): DNA MTases are a family of enzymes that sequence-specifically transfer a methyl group to the cytosines and adenines of dsDNA.[46] Because methylated cognate sites are protected from cleavage by some restriction enzymes, bacteria protect their own DNA from cleavage through methylation and employ restriction enzymes to attack foreign DNA, which is unmethylated at cognate sites. Some MTases perform other functions, but of interest here is the subset of MTases that only perform sequence-specific methylation of DNA using S-adenosyl-L-methionine (AdoMet) as a substrate for methylation.[46] These AdoMet-dependent MTases catalyze the transfer of methyl groups from the AdoMet substrate to target sequences on DNA. When used in conjunction with modified AdoMet substrates, these MTases can be used to covalently bind chemical moieties other than methyl groups to specific sequences on dsDNA.[47-50] Over 35 DNA MTases, with over 20 distinct recognition sequences, are commercially available.[42,43]

2.3 Sequence-specific DNA labeling methods

Sequence-specific labeling of dsDNA molecules is commonly achieved by the attachment of a small modification such as a hapten or fluorophore to a target sequence. Fundamental considerations for sequence-specific labeling methods include the repertoire of target sequences and the range of available modifications. In the following section, we review the most common approaches to labeling arbitrary dsDNA for single-molecule experiments.

2.3.1 Nick translation

Site-specific labeling of dsDNA via nick translation fundamentally requires the consecutive action of two enzymes.[39] First, a nickase cuts one strand of dsDNA to produce a nick at the enzyme recognition site. Second, a polymerase then fills in that nick with nucleotide bases present in solution. If modified nucleotides are present in solution, some polymerases will incorporate them, along with modifications, starting at the nicking location of the nickase recognition sites.[4] The polymerase promotes hydrolysis and synthesis simultaneously and results in the translation of the nick along the DNA duplex in the 5' to 3' direction. To ensure that nick translation does not continue unchecked, dideoxynuclelotides (ddNTPs) can be added and their concentration controls the size of the modified region.[4] To prevent the labeling of nicks existing prior to the action of the nicking enzyme, ligase is often used as a first step to fill in any non-specific nicks. Nick translation usually also concludes with ligation to repair the translated nicks (Figure 2-1). This method has been demonstrated to specifically label individual dsDNA molecules with fluorophores such as Alexa 647 and haptens such as biotin, for characterization by optical and force microscopy.[16,20,4,51,52]

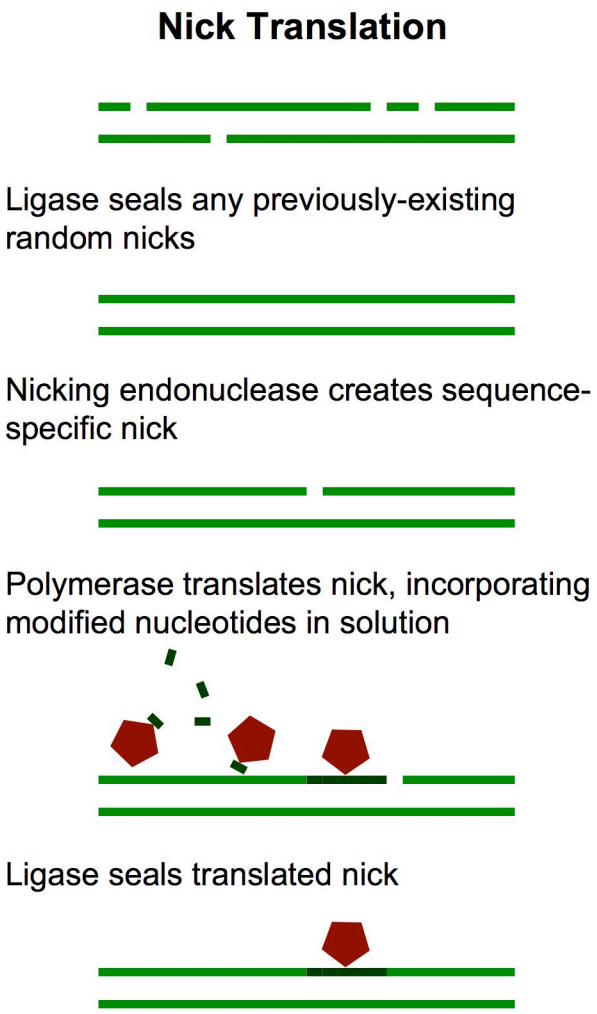


Figure 2-1. Scheme for covalently modifying DNA via nick translation. Pentagons represent any modifications available conjugated to nucleotides.

Several variations of the nick translation approach have been employed. One variation of this method uses the strand-displacement activity of an (exo-) polymerase to create single-stranded extensions beyond the nicked dsDNA site. The single-stranded extensions, created by the nucleotide chain displaced as the nick was translated, can then be hybridized with modified oligos (Figure 2-2). Das et al. used this approach in conjunction with nick translation to sequence-specifically modify DNA with the fluorophore Alexa 647 and optically map DNA.[16] The single-stranded extension flaps, which were restricted to 50 bases in length, allowed for further selective labeling of each flap sequence with unique Cy3 fluorescently-labeled oligo probes. These labeled DNA molecules were then stretched by nanochannel confinement and imaged via fluorescence microscopy, thereby creating single-molecule optical maps of genomic-length DNA.

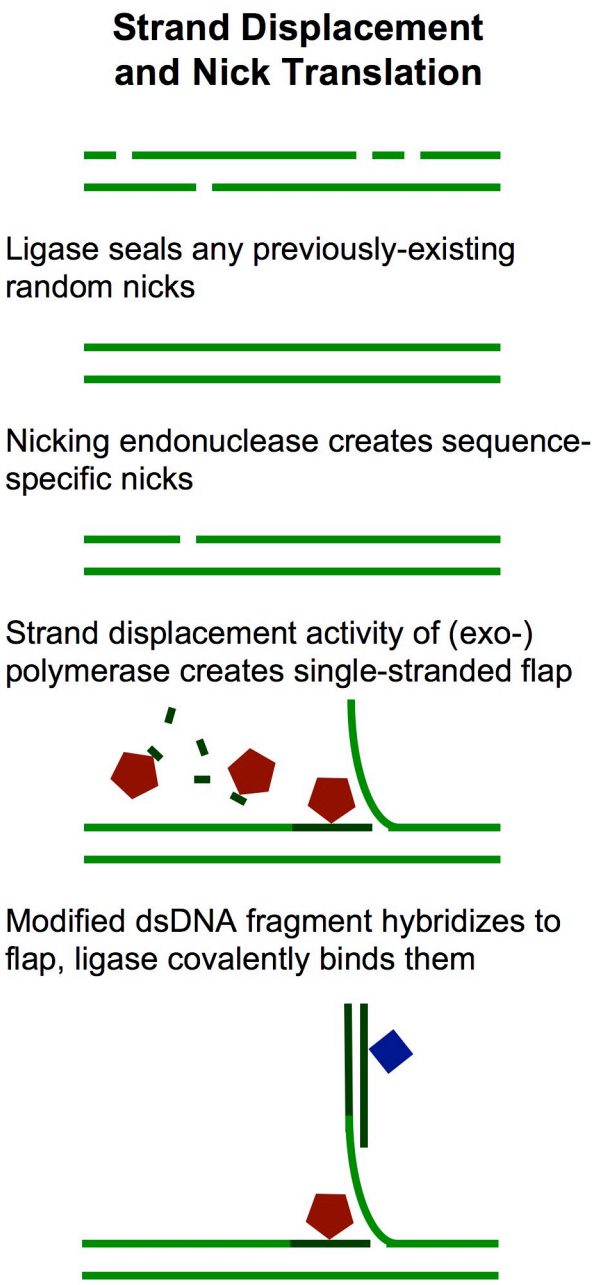


Figure 2-2. Scheme for covalently modifying DNA using a variation of nick translation. Pentagons represent any modifications available conjugated to nucleotides. The square represents any modification that can be conjugated to an oligo. This approach allows for differential dual labeling of the nicking site and flap.

Another variation uses nicking endonucleases to create vicinal nicks, nicks within 15-25 base pairs on the same strand of DNA. The single-stranded segment between the nicks can then be substituted in a strand displacement reaction with an oligo probe. If the oligo

probe is modified, the modification is covalently incorporated into the DNA upon ligation, thereby labeling DNA at a prescribed sequence (Figure 2-3).

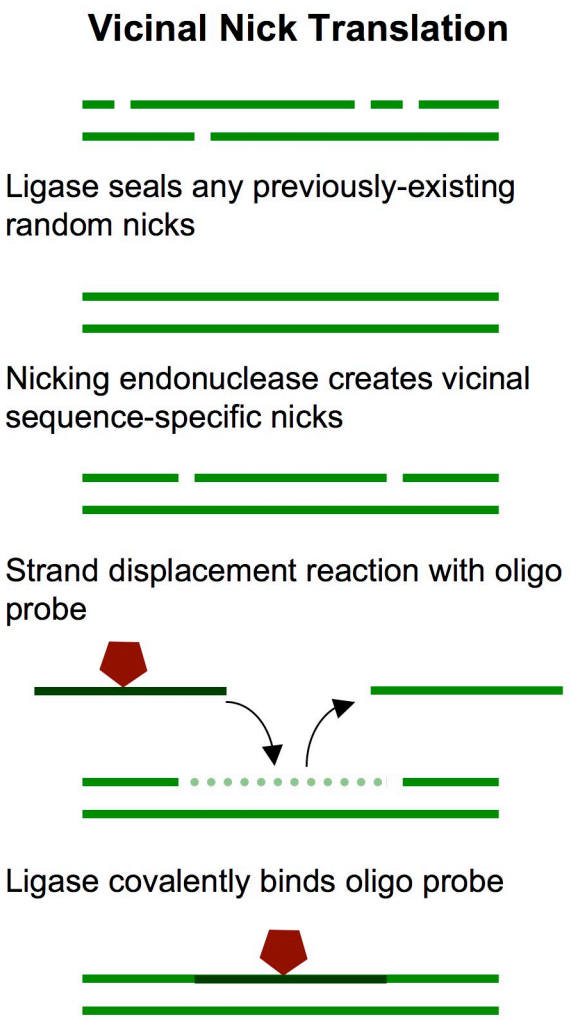


Figure 2-3. Scheme for covalently modifying DNA using a variation of nick translation. Pentagons represent any modification that can be conjugated to oligos.

In a study by Kuhn et al., the oligo probe instead contained an excess of bases at the 3' terminus creating an overhang or flap of single-stranded DNA 20 bases long.[51] The flap was used downstream as a substrate for a fluorescent reporting system, linear rolling-circle amplification (RCA), that demonstrated highly sequence-specific targeting of viral DNA. While this particular study did not image individual DNA molecules, this approach could easily be used to modify DNA for single-molecule experiments. Furthermore, different flap sequences can be designed to bind to corresponding modified oligo probes, thereby differentially labeling each flap. Luzzietti et al. developed another variation of nick translation specifically for attaching rotor-beads to internal DNA sequences for magnetic tweezer experiments.[51]

There are multiple schemes to utilize nick translation for site-specific labeling with high yield (some greater than 90%).[16,20,4,51,52] While non-specific nicks are common on DNA and could result in non-specific labeling, treating DNA with ligase first can dramatically reduce this non-specific contribution.[4] However, the repertoire of target sequences is limited to those with an associated nicking enzyme.

2.3.2 Stem-Loop Triplex Forming Oligonucleotides (TFOs)

TFOs are short single-stranded segments of DNA, typically binding to targets 10-30 bases long, that recognize and bind to the major groove of dsDNA by winding around the target sequence and forming a triple helix.[53] A special class of TFOs, stem-loop TFOs, are designed such that the center of the TFO binds to the target sequence, roughly 30 bases, and the ends of the TFO hybridize to one another and leave a single-stranded overhang, or flap.[17] In the first of two steps, a looped structure of roughly 70 bases, with a stem acting as an available flap, hybridizes around the target sequence. Lastly, a modified dsDNA fragment with an overhang complementary to the flap is hybridized to the flap and covalently attached to the flap using ligase (Figure 2-4). The attachment of fragments as long as 500 bp has been demonstrated by Escude et al. and Geron-Landre et al.[17,54] Binding of the dsDNA fragment to the TFO overhang is followed by ligation to prevent unbinding and excess TFOs can be removed later in the experimental protocol. Interestingly, Cherney et al. demonstrated in an early single-molecule electron-microscopy experiment that even TFOs 12 to 17 bases in length modified with biotin, that do not form stem-loop structures and are not ligated, can bind with initial yields exceeding 70%.[55]

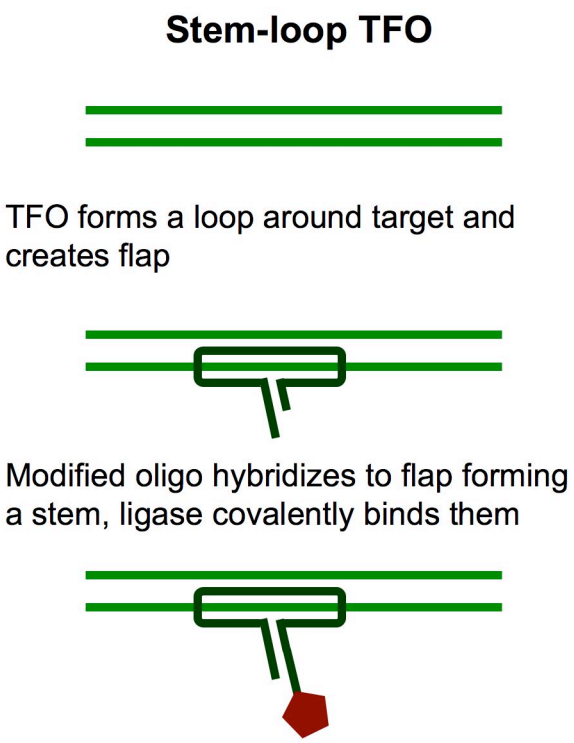


Figure 2-4. Scheme for modifying DNA using stem-loop TFOs. Pentagons represent any modification that can be conjugated to oligos.

Geron-Landre et al. demonstrated 59 base stem-loop TFOs binding to a 15 bp target with a labeling yield exceeding 60% and noted that the maximum yield was dependent on the TFO design.[56] The 500 bp fragment that was ligated to the stem flap was modified with AlexaFluor 546 and detected with single-molecule fluorescence microscopy. Escude et al. demonstrated this method with unmodified fragments of 200 and 500 bp, that were directly visualized as branched structures off of the DNA backbone using electron microscopy.[54] They reported a similar yield to Geron-Landre.

One limitation of stem-loop TFOs is that target sequences for this binding motif are restricted to homopurine sequences, which have a homopyrimidine complementary strand in dsDNA. The TFOs bind via different motifs but always to the homopurine strand of the target.[57] The pyrimidine binding motif is more stable at acidic pH when the cytosine is protonated and can facilitate Hoogsteen binding.[57] A modified cytosine base, that is protonated at neutral pH, can also be employed and there are strategies that permit the recognition of mixed homopurine and homopyrimidine sequence dsDNA at physiological pH.[17,58] An advantage to this method is that the DNA fragment ligated to the flap can contain multiple modifications (i.e. multiple fluorophores and/or haptens) and combinations of modifications. In this way, the method is similar to a variation of nick translation, where single-stranded overhangs are designed to bind to modified oligo probes. Variations on this method include two types of padlock probes, described below.

2.3.3 Padlock Probes

Padlock probes are linear oligos that are similar in design to stem-loop TFOs and also utilize the triplex-forming capability of oligos. While stem-loop TFOs are designed such that the center of the TFO binds to the target sequence and the ends of the TFO hybridize to one another, padlock probes are designed in the opposite manner such that the ends of the oligo hybridize to the target sequence and the center of the oligo connects the ends to form a ring (Figure 2-5 and Figure 2-6). The hybridizing regions are typically 20 bases each and are connected by a roughly 50 base central segment. To make a linear oligo into a closed ring, or padlock, the gaps and nicks present along the oligo backbone upon hybridization to dsDNA are sealed via polymerase and ligase, respectively.

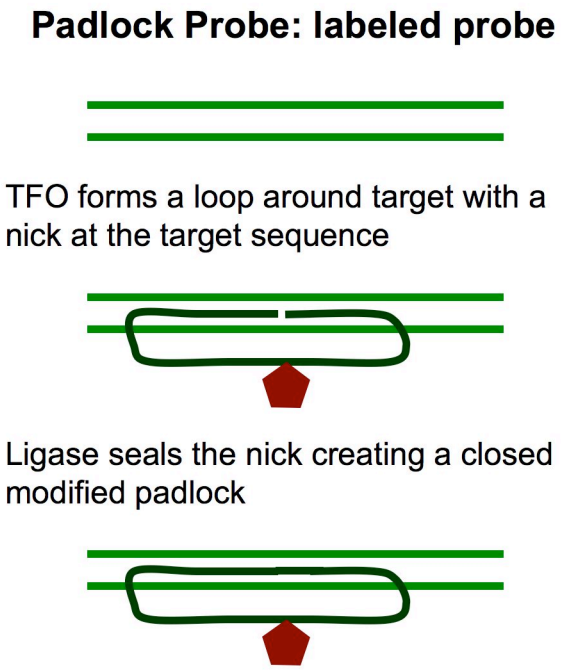


Figure 2-5. Scheme for modifying DNA using a labeled probe padlock. Pentagons represent any modification that can be conjugated to oligos.

There are two approaches to attach a modification using padlock probes, the labeled probe method and the gap-fill method.[59] In both approaches, the linear oligo goes on to form a complete modified ring. In the labeled probe method, the central segment of the oligo contains modifications and the padlock probe is designed such that there is a nick between where the two ends hybridize to the target sequence. First, the ends of the probe oligo hybridize to the target sequence creating a target-oligo triplex with a nick in the oligo strand. Second, the nick is sealed via ligase, forming a covalently-bonded modified oligo ring around the target sequence (Figure 2-5). In the gap-fill method, the padlock probe is designed such that there is a gap of at least one base between where the two ends hybridize to the target sequence. First, the ends of the probe oligo hybridize to the target

sequence leaving a gap. Second, this gap is filled in with modified dNTPs using polymerase and the remaining nick is sealed via ligation (Figure 2-6). The gap-filled ring is similar to that produced with the labeled-probe approach, with the modification placed at the target-binding segment of the oligo instead of in the central segment. Both approaches benefit from the sensitivity of ligase to correct base pairing, as mismatched oligos are not ligated efficiently.[60]

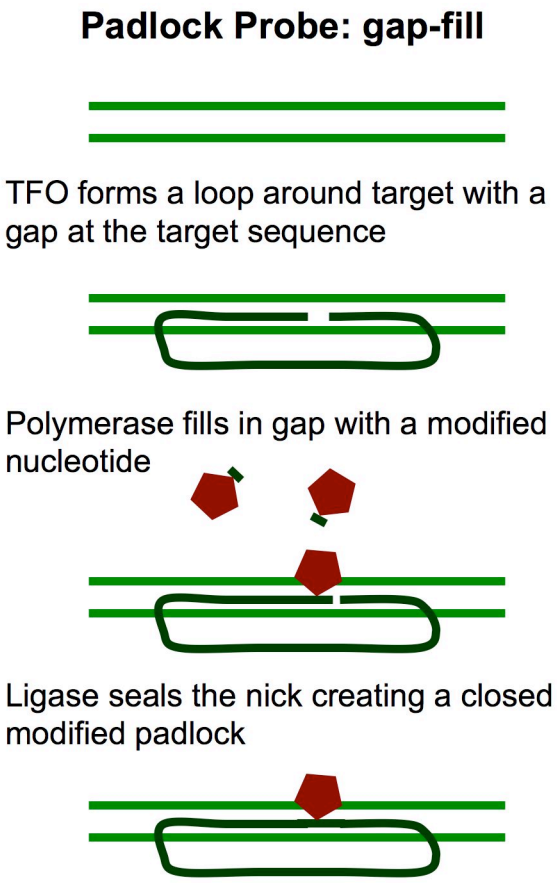


Figure 2-6. Scheme for modifying DNA using a gap-filled probe padlock. Pentagons represent any modifications available conjugated to nucleotides.

Nilsson et al. used the labeled probe approach with 90 base padlock probes and a biotin-modified linker to detect repeated sequences in human chromosome 12.[61] The locations of the biotin-modified padlock probes were visualized using fluorescence microscopy by incubating them sequentially with fluorescein-labeled avidin, biotinylated antibodies against avidin, and a second layer of fluoresceinated avidin. While these experiments did not image individual DNA molecules, this technique can be applied to single-molecule experiments. Xiao et al. used the gap-fill approach with 80 to 100 base padlock probes to achieve single-molecule haplotyping.[19] The gap-filling nucleotides were modified with Cy3 and Cy5 fluorophores such that the closed padlocks attached fluorophores to the target sites. Individual labeled DNA molecules were deposited onto a surface and the

locations of the padlocks were determined via total internal reflection fluorescence (TIRF) microscopy. The labeling efficiency for a single site was roughly 25% and the labeling efficiency is limited by competition between hybridization of the padlock probe and the native dsDNA renaturation. Variations of this approach use an additional oligo to close the padlock and achieve a multi-stranded stable structure similar to a padlock probe.[62,63]

2.3.4 Modified Proteins

Using modified proteins to sequence-specifically label DNA takes advantage of both the known sequence-specific binding properties of some proteins and the well-developed protocols and commercially available reagent kits for labeling proteins with fluorophores and haptens. In particular, type II restriction enzymes such as EcoRI and EcoRV that, in the absence of Mg^{2+}, can bind to cognate sites on dsDNA without cleaving them are well-suited for labeling DNA (Figure 2-7).[21,22,64,65] They can be labeled with reagents that react with the side chains of their cysteine or lysine residues.

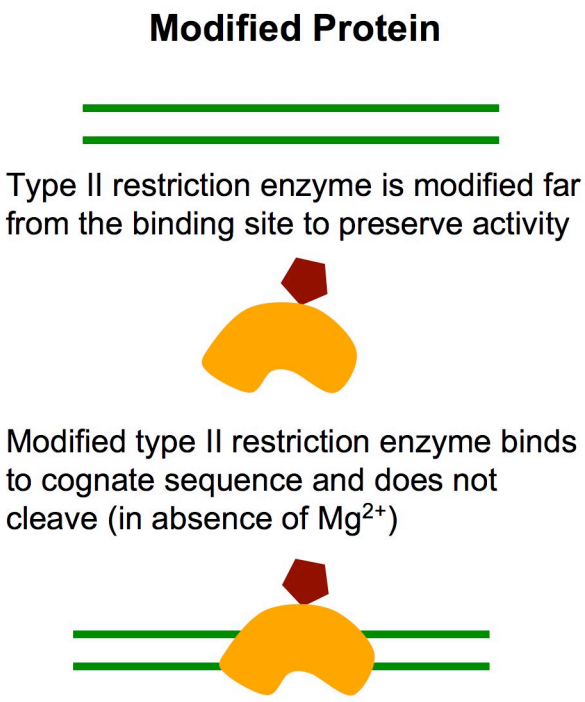


Figure 2-7. Scheme for modifying DNA using a modified protein. Pentagons represent any modification that can be conjugated to a protein.

Care must be taken so that the labeling of those residues does not hinder the binding site of the enzyme and impair binding to the DNA target sequence. Additionally, cysteine or lysine residues should be solvent-accessible to facilitate labeling. One strategy involves labeling a mutant enzyme that is engineered to have only solvent-accessible cysteines or

lysines far from the binding pocket.[21] Another strategy involves labeling a wild type enzyme while protecting those cysteines or lysines at the binding site from labeling.[22] Other proteins, such as T7 RNA Polymerase (RNAP), that would not naturally be immobilized at cognate sites, can also be labeled and covalently cross-linked to dsDNA at these sites.[66]

There have been several single-molecule demonstrations of fluorescent labeling of dsDNA using the EcoRI restriction enzyme, which has a 6 base recognition sequence, since its binding, cleavage, and structure have been extremely well-studied.[67-69] Oana et al. biotinylated EcoRI using the Sulfo-NHS-LC-Biotin reagent that reacts with the amino groups of one or more of the 22 lysine residues in EcoRI.[22] To preserve binding activity, lysines in the binding site were protected from biotinylation by binding EcoRI to heparin prior to biotinylation. The biotinylated EcoRI was incubated with rhodamine-avidin complexes, forming fluorescent EcoRI. Taylor et al. conjugated EcoRI to carboxylate functionalized fluorescent nanoparticles using the cross-linking agent 1-ethyl-3-(3-dimethyl-aminopropyl) carbodiimide (EDAC).[65] The carboxylate groups on the nanoparticle surface react with the EDAC, forming an active intermediate that is attacked by the amine side chains of the lysine residues in EcoRI. The nanoparticle-bound EcoRI demonstrates site-specific cleavage activity and, when bound to DNA in the absence of Mg^{2+}, the EcoRI nanoparticles label dsDNA at the EcoRI binding sites. Dylla-Spears et al. achieved a similar result to Taylor et al. by biotinylating mutant EcoRI, binding it to dsDNA in the absence of Mg^{2+}, and then incubating the dsDNA with Neutravidin-coated fluorescent nanospheres.[21] Here control of the biotin position with respect to the enzyme's binding pocket was achieved by biotinylating the enzyme at the mutation site, a nonessential, solvent accessible lysine that was mutated into a cysteine. Biotinylation was performed using the EZ-link Maleimide-PEO2-biotin reagent which forms a thioether bond with the cysteine. By biotinylating a single nonessential site, Dylla-Spears et al. limited the reduction in binding efficiency that can accompany modification of multiple residues.

EcoRV, another well-studied enzyme with a 6 base recognition sequence, has been labeled with qdots and the Cy3 fluorophore. While single-molecule studies of EcoRV have focused on its translocation behavior on dsDNA, it can be used to label DNA in a manner similar to EcoRI. Interestingly, EcoRV, unlike EcoRI, requires Ca^{2+} (and the absence of Mg^{2+}) to bind sequence-specifically to dsDNA without cleaving.[70] Bonnet et al. labeled an EcoRV variant that contained only a single cysteine located far from the binding site.[71] A Cy3B-Maleimide Mono-reactive pack reagent was used to directly label the cysteine with the Cy3 fluorophore. Biebricher et al. similarly biotinylated the cysteine on the same EcoRV variant using the Maleimide-PEO2-biotin reagent.[64] Streptavidin-coated qdots were then used to fluorescently label the EcoRV and the qdot-conjugated EcoRV was shown to retain its sequence specific binding activity in the presence of Ca^{2+} and the absence of Mg^{2+}.

T7 RNAP is a protein that binds to promoter sequences on the T7 bacteriophage genome and then initiates the transcription of RNA from the DNA template. Ebenstein et al.

mapped the locations of promoter sequences on individual dsDNA molecules by cross-linking T7 RNAP to DNA at cognate sites upon initiation of transcription.[66] Formaldehyde was used to form covalent bonds between proximal amino or imino groups such as a lysine residue of the T7 RNAP in contact with a cytosine base of DNA. Biotinylated T7 RNAP was created using a protein expression protocol and then fluorescently labeled with streptavidin-qdots. For dynamic single-molecule studies of T7 RNAP, Kim and Larson used a different, more accessible, approach to fluorescently label T7 RNAP.[72] T7 RNAP was labeled with the rhoadmine fluorophore by serial incubation with two commercially available reagents, T7 RNAP mouse monoclonal antibody followed by rhodamine conjugated to anti-mouse IgG.

Unless engineered enzymes, which may not be accessible to the non-specialist, are used the set of target sites accessible with modified binding proteins is limited to the cognate sites of commercially available proteins.[42,43] Since labeling methods based on type II restriction enzymes usually require modification of cytosine or lysine residues, preserving the activity of modified enzymes is non-trivial and may require a mutant enzyme, which may also not be available to the non-specialist. Furthermore, some single-molecule experiments, such as studies of motor proteins, require the presence of Mg^{2+}. Modified type II restriction enzymes would not be useful labels for such studies, since in the presence of Mg^{2+} the enzymes would cleave dsDNA instead of serving as sequence-specific labels. However, the Modrich group has developed an EcoRI mutant, EcoRI(Gln-111), that binds with high specificity and does not cleave in the presence of Mg^{2+}.[73,74] Furthermore, the large number of commercially available type II restriction enzymes and the multitude of commercially available reagent kits for biotinylating or otherwise functionalizing proteins make modified proteins a useful DNA labeling strategy for many applications.

2.3.5 Methyltransferases (MTases)

There are two DNA labeling strategies that exploit the targeted functionalization activity of MTases, both requiring synthetically prepared modified AdoMet analog cofactors.[46] In the first strategy, Sequence-specific Methyltransferase-Induced Labeling of DNA (SMILing DNA) developed by Pljevaljcic, Schmidt, and Weinhold, a modified AdoMet analogue is designed to be transferred in its entirety by the MTase to DNA.[75] A synthetically prepared modified cofactor is composed of a central moiety for cofactor binding, a reactive aziridine group, and a modification attached through a short linker. Since the final step of the AdoMet analogue synthesis is the coupling of a primary amine with various activated esters, a variety of modifications can be used and both biotin and a fluorophore have been demonstrated. The reaction creates product inhibitors, however, so the stoichiometry must be precisely controlled.[46] In the second strategy, MTase-Directed Transfer of Activated Groups (mTAG) developed by Lukinavicius et al., the methyl group of AdoMet is replaced with a different functional group of interest such as an amine group.[76] The functional group is transferred by the MTase to its recognition sites and can be used to attach fluorophores and haptens in a subsequent coupling step (Figure

2-8). The reaction rate decreases as the size of the transferable group increases but there are methods to increase transfer efficiency.[46]

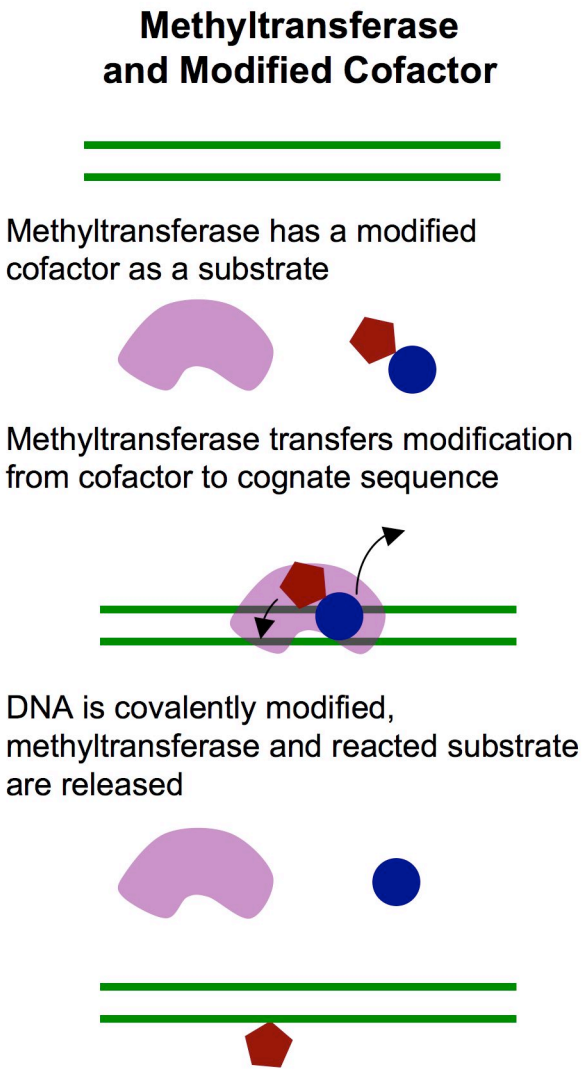


Figure 2-8. Scheme for covalently modifying DNA using Methyltransferase-directed Transfer of Activated Groups (mTAG). Pentagons represent a transferable reactive group appended to an AdoMet analog cofactor. Circles represent the remainder of the AdoMet analog cofactor.

Pljevaljcic et al. biotinylated DNA using the SMILing DNA approach with M.TaqI, which has a 4 base long cognate site.[77] A biotinylated synthetic AdoMet analogue was transferred in its entirety through the opening of its aziridine ring to the M.TaqI cognate sites. Upon treating with streptavidin, biotinylated sites on individual DNA molecules were visualized using electron microscopy. An analogous cofactor, with a Cy5 fluorophore modification instead of a biotin, was also demonstrated to work effectively with M.TaqI. Similarly, Braun et al. biotinylated DNA using M.TaqI and M.BseCI, which has a 6 base long cognate site, and a modified cofactor.[12] A biotinylated synthetic

AdoMet analogue was transferred in its entirety to the cognate sites and biotinylated gold nanoparticles were attached to biotinylated DNA sites using free streptavidin as a bridge. Labeled DNA was deposited on mica and individual molecules were directly visualized via AFM.

Neely et al. used the mTAG approach with an engineered version of the MTase M.HhaI, which recognizes a 4 base long sequence, to densely label DNA with fluorophores and construct optical DNA maps via fluorescence microscopy from individual DNA molecules deposited on a surface. First, DNA was incubated with engineered M.HhaI and a synthetically prepared modified AdoMet analog cofactor, Ado-11-amino, that contained an appended transferable amine group. With the modified cofactor as a substrate, M.HhaI covalently transferred the amine group to its recognition sequence on DNA. Second, the amine-modified DNA was incubated with an amine-reactive fluorophore Atto-647N. This process produced DNA molecules with 50-60% of the M.HhaI target sites fluorescently labeled. The efficiency of the initial amine modification by the MTase is thought to be near complete and this reduced yield is attributed to the lower efficiency of the amine-reactive fluorophore coupling in the second step.

Since the specificity of a DNA MTase is, to a first approximation, equivalent to that of its companion restriction enzyme, the functionalization activity of MTases is considered highly sequence specific.[78] The lengths of the recognition sequences in the above examples are typical as over 70% of the commercially available MTases recognize sequences with lengths between 4 to 6 bases.[42,43] Furthermore, the short length of some MTase recognition sequences allows for highly dense labeling. At least 25% of the commercially available MTases have recognition sequences with lengths of 4 bases or less and shorter recognition sequences are likely to occur more often than longer sequences on a random DNA molecule. Ultimately, however, the utility of MTase-based labeling depends on the availability or ease of synthesis of modified AdoMet analogs, which to our knowledge are not commercially available, and the library of recognition sequences of available MTases.

2.3.6 Peptide Nucleic Acids (PNAs)

Upon simple incubation, PNAs bind to dsDNA via one of four binding modes (Figure 2-9), determined by the PNA design.[33] The binding mode dictates the base content of the DNA sequences that can be targeted.

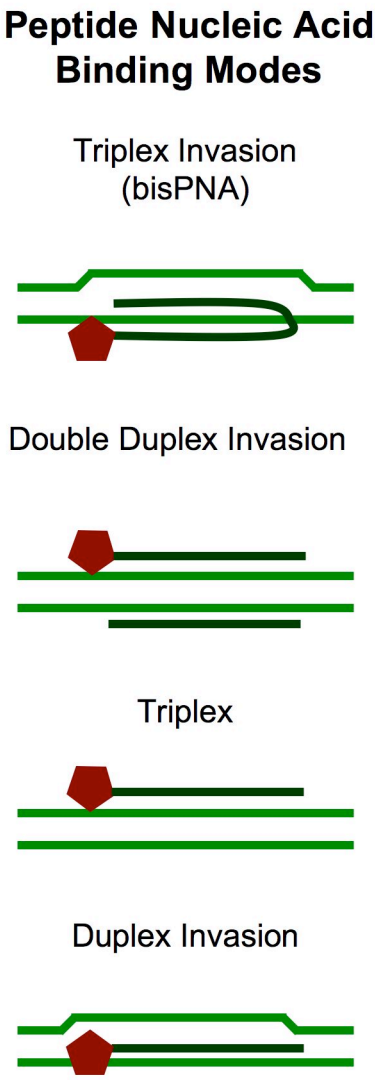


Figure 2-9. Schematics of PNA-DNA binding modes. Pentagons represent any modification that can be conjugated to a PNA. The PNA-DNA complexes are created by simple incubation and the binding mode is determined by the PNA design.

The first binding mode, triplex invasion, is the most stable and requires two homopyrimidine PNA molecules to simultaneously bind to the homopurine strand of the DNA recognition sequence. A bisPNA, which is a construct of the two triplex-forming PNA molecules covalently attached by a central linker, is often used to enhance binding rates for this binding mode. The second mode, double duplex invasion, requires two sequence complementary PNA molecules to bind to the two complementary strands of the target sequence on dsDNA. The PNAs contain sterically hindered analogs of the adenine and thymine bases so that the two sequence complementary PNA molecules will prefer to bind to a complementary DNA strand rather than each other. For the double duplex invasion binding mode, the base composition of the PNA sequence needs to have at least 50% combined cytosine and guanine base content. The third mode, triplex, which creates a triple helix by binding to the major groove of dsDNA, seems to require high cytosine content. The fourth and final mode is duplex invasion, where a single PNA molecule binds to a DNA strand by displacing the complementary DNA strand, and requires a guanine-rich homopurine PNA sequence. Even given the sequence restrictions imposed by each binding mode, PNAs can be employed to bind to an extensive repertoire of target binding sequences. Furthermore, they are commercially available and made to order with a choice of conjugated modifications including haptens and fluorophores.

Demidov et al. demonstrated physical mapping of DNA molecules using biotinylated bisPNAs with a 10 base recognition sequence.[79] Individual biotinylated PNA-DNA complexes were conjugated to streptavidin protein, which could be directly visualized via electron microscopy. The streptavidin protein served as the electron microscopy markers and appeared as 'beads' located at the target sites. Kim et al. and Qu et al. demonstrated optical mapping of DNA using PNAs with 15 and 7 base recognition sequences, respectively, modified with Alexa fluorophores.[18,80] Kim et al. determined the position of their Alexa 532 labeled PNA along the DNA molecule by concurrent scanning near-field optical and atomic force microscopy (SNOM/AFM), while Qu et al. measured the location of their Alexa 488 fluorophore-conjugated PNA on the DNA by nanometer-localized multiple single molecule (NALMS) fluorescence microscopy.

Chan et al. and Phillips et al. demonstrated mapping of DNA using bisPNAs conjugated to Alexa 546 or tetramethylrhodamine fluorophores with 7 or 8 base recognition sequences.[3,23] The PNA-DNA complexes, created with roughly 90% yield, were introduced into a continuous flow device for direct linear analysis (DLA), a scheme that uses flow to linearize the DNA and provides the spatial locations of the PNAs along the extended DNA molecule. The fluorescently labeled DNA was analyzed as it flowed past multicolor confocal fluorescence detectors with single-fluorophore sensitivity. Zohar et al. demonstrated mapping of DNA using biotinylated bisPNAs with 8 base recognition sequences that were incubated with Neutravidin-coated fluorescent nanospheres.[24] Labeled DNA was introduced into a microfluidic cross-slot where molecules were trapped and stretched in a stagnation point flow. The locations of the bisPNAs along the backbone of the DNA were determined via fluorescence microscopy. Zohar et al. further demonstrated that these same biotinylated PNAs can be used as robust and benign tethers

in conjunction with avidin-coated microspheres for optical tweezer experiments and that they bind to DNA with an average strength of roughly 60 pN.

While creating PNA-DNA complexes is relatively simple, the incubation conditions must be carefully optimized for yield and specificity for a given PNA molecule.[3,23,24,81] Incubation conditions affecting yield or specificity include incubation time and temperature,[3,82] pH,[83,84] salt concentration,[35,82,85,86] and the ratio of PNA to DNA.[3,35] Compared to other labeling approaches, however, PNAs are advantageous because they are commercially available with modifications, compatible with cofactors such as Mg^{2+}, and non-destructive to dsDNA. Furthermore, they do not require ligation and, under optimized conditions, bind with high specificity and yield.[3,23] Though PNAs are expensive relative to other commercially available reagents, a typical minimum order of roughly 40 nmol may exceed the amount needed for bulk optimization and a full course of single-molecule experiments.[24] PNAs also offer much more flexibility in the accessible target sequence than the enzyme-based labeling approaches using currently commercially available nickases, modified proteins, and MTases.

2.4 Conclusions and outlook

Single-molecule DNA experiments allow for unprecedented insights yet are challenging and diverse in their requirements. Frequently, they require the internal sequence-specific labeling of long and arbitrary dsDNA molecules. As experimentalists of different backgrounds are increasingly engaging in single-molecule DNA experiments and each new experimental system brings new constraints, there is a need for multiple labeling options that are accessible to non-specialists. Ideally, a sequence-specific dsDNA labeling approach should be straightforward, require only commercially available reagents, and offer flexibility in target sequences and modifications for labeling (Table 1). Furthermore, depending on the experimental demands, the labeling method may need to attach modifications covalently, not require ligase, be robust to cofactors such as Mg^{2+}, or exhibit high labeling yield (Table 2). It is important to note that the labeling yield refers to the initial modification made to DNA and that subsequent labeling steps, such as the attachment of a nanosphere to a modification, will result in lower overall yield.

Table 2-1. Accessibility and versatility of each labeling approach.

Approach	Commercially available modified reagents	Accessibility to non-specialist	Target Sequence Repertoire	Modification Repertoire
Nick Translation	yes	Accessible	Limited to nicking endonuclease cognate sites	Available modified nucleotides
Stem-loop TFOs	yes	Accessible but requires binding optimization	Limited to homopurine or homopyrimidine	Available custom oligo modifications
Padlock Probes	yes	Accessible but requires binding optimization	Limited to homopurine or homopyrimidine	Available modified nucleotides (gap-fill) or custom oligo (linker probe) modifications
Modified Proteins	No, but proteins and modifying kits are available	Requires mutant enzymes and/or specialized enzyme knowledge	Limited to protein cognate sites	Available modifying kits that react with lysine or cysteine residues
MTases and modified cofactors	No, modified cofactors are not available	Requires modified cofactors, otherwise accessible	Limited to methyltransferase cognate sites	Available modified cofactors
Peptide Nucleic Acids	yes	Accessible but requires binding optimization	Depends on binding mode: strongest limited to homopurine	Available custom PNA modifications

Table 2-2. Characteristics of binding and labeled DNA for each labeling approach.
[a] Modifications are on topologically constrained ligated ring-like structures

Approach	Covalently-bound modification	Requires ligase	Compatible with Mg^{2+} cofactor	Typical reported modification yield
Nick Translation	✓	✓	✓	> 90%[51]
Stem-loop TFOs	✓[a]	✓	✓	> 60 %[54,56]
Padlock Probes	✓[a]	✓	✓	25-50%[19,59]
Modified Proteins				> 50%[21]
MTases and modified cofactors	✓		✓	Quantitative, can be near complete [5,12,47,76]
Peptide Nucleic Acids			✓	80-90%[3]

When considering the six approaches described, each has advantages and disadvantages and, ultimately, the approach of choice will depend on individual experimental constraints. The approaches that use an enzyme to introduce a modification - nick translation, modified proteins, and MTases - share a limitation that the repertoire of target sequences is limited to the cognate sites of the commercially available enzymes. While nick translation and MTases produce covalently-bound modifications, modified proteins are not covalently-bound. Furthermore, modified restriction enzymes are not compatible with Mg^{2+} as it will result in enzymatic cleavage, though specially-engineered mutant enzymes may overcome this restriction. The enzymatic approaches share the advantage that the binding conditions for high yield sequence-specific activity are known for commercially available enzymes and require little to no optimization. Nick translation is particularly straightforward, MTases are easy to use with the main challenge being to secure a modified cofactor, and modifying proteins is possible with commercially available reagents but requires specialized techniques. MTases, which can have short cognate sequences, are particularly useful when high-density labeling is required. Furthermore, nick translation and MTase-based approaches produce modified DNA constructs without leaving physical obstacles on the DNA molecule.

The approaches that use a triplex-former to introduce a modification – stem-loop TFOs, padlock probes, and PNAs – require that each new probe be custom-designed and share a limitation that the binding conditions must be optimized for each designed probe. Triplex-formers, however, can access a variety of target sequences and are not limited to commercially available enzymatic cognate sites. Stem-loop TFOs and padlock probes are not directly covalently bound to the target sequence but form ligated ring-like structures that are topologically constrained to the target sequence location. PNAs are also not covalently bound to the target sequence yet have been shown to be robust and suitable as optical tweezer handles. Labeling approaches based on triplex-formers locally alter DNA structure and leave a physical obstacle on the DNA molecule. This property can be undesirable or it can be exploited, for example when obstacles for translocases are desired.

In an effort to make single-molecule DNA experiments more accessible to non-specialists, we have reviewed six demonstrated approaches to sequence-specific labeling of DNA. There are other approaches, including those based upon zinc fingers and locked nucleic acids (LNAs) and other triplex-formers that could potentially be used to label dsDNA for single-molecule experiments.[15,57,57,87-92] Our choice of examples here has been illustrative and selective, and related experiments, omitted in the interest of conciseness, may be found in the literature. Moreover, new techniques, specialized reagents, and demonstrations of variations of existing approaches are appearing with increasing speed in this burgeoning field. This review serves as an introduction to those new to the field and provides a framework for evaluating new options as they develop.

2.5 References

1. Walter, N. G., Huang, C., Manzo, A. J. & Sobhy, M. A. Do-it-yourself guide: how to use the modern single-molecule toolkit. *Nat. Methods* **5**, 475-489 (2008).
2. Ritort, F. Single-molecule experiments in biological physics: methods and applications. *J. Phys.: Condens. Matter* **18**, R531-R583 (2006).
3. Chan, E. Y. et al. DNA Mapping Using Microfluidic Stretching and Single-Molecule Detection of Fluorescent Site-Specific Tags. *Genome Res.* **14**, 1137-1146 (2004).
4. Jo, K. et al. A single-molecule barcoding system using nanoslits for DNA analysis. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 2673-2678 (2007).
5. Neely, R. K. et al. DNA fluorocode: A single molecule, optical map of DNA with nanometre resolution. *Chemical Science* **1**, 453-460 (2010).
6. Bryant, Z. et al. Structural transitions and elasticity from torque measurements on DNA. *Nature* **424**, 338-341 (2003).
7. Gore, J. et al. Mechanochemical analysis of DNA gyrase using rotor bead tracking. *Nature* **439**, 100-104 (2006).
8. Hugel, T. et al. Experimental Test of Connector Rotation during DNA Packaging into Bacteriophage **φ**29 Capsids. *PLoS Biol.* **5**, e59 (2007).
9. Olson, M. V. A time to sequence. *Science* **270**, 394-394 (1995).
10. Olson, M., Hood, L., Cantor, C. & Botstein, D. A common language for physical mapping of the human genome. *Science* **245**, 1434-1434 (1989).
11. Roberts, L. New game plan for genome mapping. *Science* **245**, 1438 (1989).
12. Braun, G. et al. Enzyme-Directed Positioning of Nanoparticles on Large DNA Templates. *Bioconjugate Chem.* **19**, 476-479 (2008).
13. Ghosh, I., Stains, C. I., Ooi, A. T. & Segal, D. J. Direct detection of double-stranded DNA: molecular methods and applications for DNA diagnostics. *Mol. BioSyst.* **2**, 551-560 (2006).
14. Kwok, P. Y. & Xiao, M. Single-Molecule Analysis for Molecular Haplotyping. *Human Mutation* **23**, 442-446 (2004).
15. Weisbrod, S. H. & Marx, A. Novel strategies for the site-specific covalent labelling of nucleic acids. *Chem. Commun.* 5675-5685 (2008).
16. Das, S. K. et al. Single molecule linear analysis of DNA in nano-channel labeled with sequence specific fluorescent probes. *Nucleic Acids Res.* **38**, e177 (2010).
17. Geron-Landre, B., Roulon, T. & Escude, C. Stem-loop oligonucleotides as tools for labelling double-stranded DNA. *FEBS J.* **272**, 5343-5352 (2005).
18. Qu, X. H., Wu, D., Mets, L. & Scherer, N. F. Nanometer-localized multiple single-molecule fluorescence microscopy. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 11298-11303 (2004).
19. Xiao, M. et al. Determination of Haplotypes from Single DNA molecules: A Method for Single-Molecule Barcoding. *Human Mutation* **28**, 913-921 (2007).
20. Xiao, M. et al. Rapid DNA mapping by fluorescent single molecule detection. *Nucleic Acids Res.* **35**, e16 (2007).

21. Dylla-Spears, R., Townsend, J. E., Sohn, L. L., Jen-Jacobson, L. & Muller, S. J. Fluorescent Marker for Direct Detection of Specific dsDNA Sequences. *Anal. Chem.* **81**, 10049-10054 (2009).
22. Oana, H., Ueda, M. & Washizu, M. Visualization of a Specific Sequence on a Single Large DNA Molecule Using Fluorescence Microscopy Based on a New DNA-Stretching Method. *Biochem. Biophys. Res. Commun.* **265**, 140-143 (1999).
23. Phillips, K. M. et al. Application of single molecule technology to rapidly map long DNA and study the conformation of stretched DNA. *Nucleic Acids Res.* **33**, 5829-5837 (2005).
24. Zohar, H., Hetherington, C. L., Bustamante, C. & Muller, S. J. Peptide Nucleic Acids as Tools for Single-Molecule Sequence Detection and Manipulation. *Nano Lett.* **10**, 4697-4701 (2010).
25. K. Landsteiner, *The Specificity of Serological Reactions*, Dover Publications, Mineola, 1962.
26. Diamandis, E. P. & Christopoulos, T. K. The biotin-(strept) avidin system: principles and applications in biotechnology. *Clin. Chem.* **37**, 625 (1991).
27. Kool, E. T. Preorganization of DNA: Design principles for improving nucleic acid recognition by synthetic oligonucleotides. *Chem. Rev.* **97**, 1473-1487 (1997).
28. C. W. Dieffenbach, G. S. Dveksler, *PCR Primer: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, 2003.
29. Nielsen, P. E. Applications of peptide nucleic acids. *Curr. Opin. Biotechnol.* **10**, 71-75 (1999).
30. Nielsen, P. E. Peptide nucleic acid: a versatile tool in genetic diagnostics and molecular biology. *Curr. Opin. Biotechnol.* **12**, 16-20 (2001).
31. P. E. Nielsen, in *PNA Technology*, ed. P. E. Nielsen, Humana Press, Totowa, 2002, vol. 208, pp. 3-26.
32. Nielsen, P. E. & Haaima, G. Peptide nucleic acid (PNA). A DNA mimic with a pseudopeptide backbone. *Chem. Soc. Rev.* **26**, 73-78 (1997).
33. Nielsen, P. E. Peptide Nucleic Acid Targeting of Double-Stranded DNA. *Methods in Enzymology* **340**, 329-340 (2001).
34. Larsen, H. J., Bentin, T. & Nielsen, P. E. Antisense properties of peptide nucleic acid. *Biochim. Biophys. Acta* **1489**, 159-166 (1999).
35. K. E. Lundin, L. Good, R. Stromberg, A. Graslund, C. I. E. Smith, Biological Activity and Biotechnological Aspects of Peptide Nucleic Acid in *Advances in Genetics*, ed. J. C. Hall, J. C. Dunlap, T. Friedman, V. van Heyningen, Elsevier, 2006, vol. 56, pp. 1-51.
36. Paulasova, P. & Pellestor, F. The peptide nucleic acids (PNAs): a new generation of probes for genetic and cytogenetic analyses. *Annales de Genetique* **47**, 349-358 (2004).
37. Ray, A. & Norden, B. Peptide nucleic acid (PNA): its medical and biotechnical applications and promise for the future. *FASEB J* **14**, 1041-1060 (2000).
38. Shakeel, S., Karim, S. & Ali, A. Peptide nucleic acid (PNA) - a review. *J. Chem. Technol. Biotechnol.* **81**, 892-899 (2006).
39. J. Sambrook, D. W. Russell, *Molecular Cloning*, Cold Spring Harbor Laboratory Press, New York, 2001.

40. Pingoud, A. & Jeltsch, A. Structure and function of type II restriction endonucleases. *Nucleic Acids Res.* **29**, 3705-3727 (2001).
41. Heiter, D. F., Lunnen, K. D. & Wilson, G. G. Site-specific DNA-nicking mutants of the heterodimeric restriction endonuclease R.BbvCI. *J. Mol. Biol.* **348**, 631-640 (2005).
42. Roberts, R. J., Vincze, T., Posfai, J. & Macelis, D. REBASE—restriction enzymes and DNA methyltransferases. *Nucleic Acids Res.* **33**, D230-D232 (2005).
43. Roberts, R. J., Vincze, T., Posfai, J. & Macelis, D. REBASE-a database for DNA restriction and modification: enzymes, genes and genomes. *Nucleic Acids Res.* **38**, D234-D236 (2010).
44. Zhang, P. et al. Engineering BspQI nicking enzymes and application of N.BspQI in DNA labeling and production of single-strand DNA. *Protein Expression Purif.* **69**, 226-234 (2010).
45. Kelly, R., Cozzarelli, N., Deutscher, M., Lehman, I. & Kornberg, A. Enzymatic Synthesis of Deoxyribonucleic Acid. *J. Biol. Chem.* **245**, 39-45 (1970).
46. Klimasauskas, S. & Weinhold, E. A new tool for biotechnology: AdoMet-dependent methyltransferases. *Trends in Biotechnol.* **25**, 99-104 (2007).
47. Liutkeviciute, Z., Lukinavicius, G., Masevicius, V., Daujotyte, D. & Klimasauskas, S. Cytosine-5-methyltransferases add aldehydes to DNA. *Nat. Chem. Biol.* **5**, 400-402 (2009).
48. Pignot, M., Pljevaljcic, G. & Weinhold, E. Efficient Synthesis of S-Adenosyl-L-Homocysteine Natural Product Analogues and Their Use to Elucidate the Structural Determinant for Cofactor Binding of the DNA Methyltransferase M.HhaI. *Eur. J. Org. Chem.* 549-555 (2000).
49. Pljevaljcic, G., Pignot, M. & Weinhold, E. Design of a New Fluorescent Cofactor for DNA Methyltransferases and Sequence-Specific Labeling of DNA. *J. Am. Chem. Soc.* **125**, 3486-3492 (2003).
50. G. Pljevaljcic, F. Schmidt, A. Peschlow, E. Weinhold, in *Sequence-specific DNA labeling using methyltransferases*, ed. C. M. Niemeyer, Humana Press Inc, Totowa, 2004, vol. 283, pp. 145-161.
51. Kuhn, H. & Frank-Kamenetskii, M. D. Labeling of unique sequences in double-stranded DNA at sites of vicinal nicks generated by nicking endonucleases. *Nucleic Acids Res.* **36**, e40 (2008).
52. Pfannschmidt, C. & Langowski, J. Superhelix Organization by DNA Curvature as Measured Through Site-specific Labeling. *J. Mol. Biol.* **275**, 601-611 (1998).
53. Knauert, M. P. & Glazer, P. M. Triplex forming oligonucleotides: sequence-specific tools for gene targeting. *Hum. Mol. Genet.* **10**, 2243 (2001).
54. Escude, C., Roulon, T., Lyonnais, S. & Le Cam, E. Multiple topological labeling for imaging single plasmids. *Anal. Biochem.* **362**, 55-62 (2007).
55. Cherney, D. I., Malkov, V. A., Volodin, A. A. & Frank-Kamenetskii, M. D. Electron Microscopy Visualization of Oligonucleotide Binding to Duplex DNA via Triplex Formation. *J. Mol. Biol.* **230**, 379-383 (1993).
56. Géron-Landre, B., Roulon, T., Desbiolles, P. & Escudé, C. Sequence-specific fluorescent labeling of double-stranded DNA observed at the single molecule level. *Nucleic Acids Res.* **31**, e125 (2003).

57. Duca, M., Vekhoff, P., Oussedik, K., Halby, L. & Arimondo, P. B. The triple helix: 50 years later, the outcome. *Nucleic Acids Res.* **36**, 5123-5138 (2008).
58. Rusling, D. A. et al. Four base recognition by triplex-forming oligonucleotides at physiological pH. *Nucleic Acids Res.* **33**, 3025-3032 (2005).
59. Gordon, M. P. Biophysical Applications of Single Molecule Fluorescence Localization. 1-86 (2006). Ph.D. Thesis, University of Illinois at Urbana Champaign, 2006.
60. Landegren, U., Kaiser, R., Sanders, J. & Hood, L. A Ligase-Mediated Gene Detection Technique. *Science* **241**, 1077-1080 (1988).
61. Nilsson, M., Helena, M., Samiotaki, M., Kwiatkowski, M. & Chowdhary, B. P. Padlock Probes: Circularizing Oligonucleotides for Localized DNA Detection. *Science* **265**, 2085-2088 (1994).
62. Shigemori, Y., Haruta, H., Okada, T. & Oishi, M. Marking of specific sequences in double-stranded DNA molecules - SNP detection and direct observation. *Genome Res.* **14**, 2478-2485 (2004).
63. Roulon, T. et al. Padlock oligonucleotides as a tool for labeling superhelical DNA. *Nucleic Acids Res.* **30**, E12 (2002).
64. Biebricher, A., Wende, W., Escudé, C., Pingoud, A. & Desbiolles, P. Tracking of Single Quantum Dot Labeled EcoRV Sliding along DNA Manipulated by Double Optical Tweezers. *Biophys. J.* **96**, L50-L52 (2009).
65. Taylor, J. R., Fang, M. M. & Nie, S. Probing Specific Sequences on Single DNA Molecules with Bioconjugated Fluorescent Nanoparticles. *Anal. Chem.* **72**, 1979-1986 (2000).
66. Ebenstein, Y. et al. Lighting Up Individual DNA Binding Proteins with Quantum Dots. *Nano Lett.* **9**, 1598-1603 (2009).
67. Halford, S., Johnson, N. & Grinsted, J. The EcoRI restriction endonuclease with bacteriophage lambda DNA. Kinetic studies. *Biochem. J.* **191**, 581 (1980).
68. Kim, Y., Grable, J. C., Love, R., Greene, P. J. & Rosenberg, J. M. Refinement of Eco RI Endonuclease Crystal Structure : A Revised Protein Chain Tracing. *Science* **249**, 1307-1309 (1990).
69. Lesser, D. R., Kurpiewski, M. R. & Jen-Jacobson, L. The Energetic Basis of Specificity in the Eco RI Endonuclease-DNA Interaction. *Science* **250**, 776-786 (1990).
70. Vipond, I. B. & Halford, S. E. Specific DNA recognition by EcoRV restriction endonuclease induced by calcium ions. *Biochemistry* **34**, 1113-1119 (1995).
71. Bonnet, I. et al. Sliding and jumping of single EcoRV restriction enzymes on non-cognate DNA. *Nucleic Acids Res.* **36**, 4118-4127 (2008).
72. Kim, J. H. & Larson, R. G. Single-molecule analysis of 1 D diffusion and transcription elongation of T 7 RNA polymerase along individual stretched DNA molecules. *Nucleic Acids Res.* **35**, 3848-3858 (2007).
73. Allison, D. P. et al. Direct atomic force microscope imaging of EcoRI endonuclease site specifically bound to plasmid DNA molecules. *Proc. Natl. Acad. Sci. U. S. A.* **93**, 8826-8829 (1996).

74. Wright, D. J., King, K. & Modrich, P. The Negative Charge of Glu-111 Is Required to Activate the Cleavage Center of EcoRI Endonuclease. *J. Biol. Chem.* **264**, 11816-11821 (1989).
75. Pljevaljcic, G., Schmidt, F. & Weinhold, E. Sequence-specific Methyltransferase-Induced Labeling of DNA (SMILing DNA). *ChemBioChem* **5**, 265-269 (2004).
76. Lukinavicius, G. et al. Targeted Labeling of DNA by Methyltransferase-Directed Transfer of Activated Groups (mTAG). *J. Am. Chem. Soc.* **129**, 2758-2759 (2007).
77. Pljevaljcic, G., Schmidt, F., Scheidig, A. J., Lurz, R. & Weinhold, E. Quantitative Labeling of Long Plasmid DNA with Nanometer Precision. *ChemBioChem* **8**, 1516-1519 (2007).
78. Lopez, O. J., Quintanar, A., Padhye, N. V. & Nelson, M. Genotyping of DNA Using Sequence-Specific Methyltransferases Followed by Immunochemical Detection. *J. Immunoassay Immunochem.* **24**, 11-28 (2003).
79. Demidov, V. V. et al. Electron-microscopy mapping of oligopurine tracts in duplex DNA by peptide nucleic acid targeting. *Nucleic Acids Res.* **22**, 5218-5222 (1994).
80. Kim, J., Hirose, T., Sugiyama, S., Ohtani, T. & Muramatsu, H. Visualizing a Hybridized PNA Probe on a DNA Molecule with Near-Field Optical Microscopy. *Nano Lett.* **4**, 2091-2097 (2004).
81. Singer, A. et al. Nanopore Based Sequence Specific Detection of Duplex DNA for Genomic Profiling. *Nano Lett.* **10**, 738-742 (2010).
82. Kosaganov, Y. N. et al. Effect of Temperature and Ionic Strength on the Dissociation Kinetics and Lifetime of PNA−DNA Triplexes. *Biochemistry* **39**, 11742-11747 (2000).
83. Hansen, M. E., Bentin, T. & Nielsen, P. E. High-affinity triplex targeting of double stranded DNA using chemically modified peptide nucleic acid oligomers. *Nucleic Acids Res.* **37**, 4498-4507 (2009).
84. Kuhn, H., Demidov, V. V., Nielsen, P. E. & Frank-Kamenetskii, M. D. An Experimental Study of Mechanism and Specificity of Peptide Nucleic Acid (PNA) Binding to Duplex DNA. *J. Mol. Biol.* **286**, 1337-1345 (1999).
85. Abibi, A., Protozanova, E., Demidov, V. V. & Frank-Kamenetskii, M. D. Specific versus Nonspecific Binding of Cationic PNAs to Duplex DNA. *Biophys. J.* **86**, 3070-3078 (2004).
86. Kuhn, H., Demidov, V. V., Frank-Kamenetskii, M. D. & Nielsen, P. E. Kinetic sequence discrimination of cationic bis-PNAs upon targeting of double-stranded DNA. *Nucleic Acids Res.* **26**, 582-587 (1998).
87. Demidov, V. V. PNA and LNA throw light on DNA. *Trends in Biotechnol.* **21**, 4-7 (2003).
88. Kauppinen, S., Vester, B. & Wengel, J. Locked nucleic acid (LNA): High affinity targeting of RNA for diagnostics and therapeutics. *Drug Discovery Today: Technologies* (2005).
89. Petersen, M. & Wengel, J. LNA: a versatile tool for therapeutics and genomics. *Trends in Biotechnol.* (2003).
90. Wahlestedt, C. et al. Potent and nontoxic antisense oligonucleotides containing locked nucleic acids. *Proc. Natl. Acad. Sci. U. S. A.* **97**, 5633-5638 (2000).

91. Xu, G. & Bestor, T. H. Cytosine methylation targetted to pre-determined sequences. *Nat. Genet.* **17**, 376-378 (1997).
92. Kiss, A. & Weinhold, E. Functional Reassembly of Split Enzymes On-Site: A Novel Approach for Highly Sequence-Specific Targeted DNA Methylation. *ChemBioChem* **9**, 351-353 (2008).

Chapter 3

Single-Molecule Sequence Detection and Manipulation:

Peptide Nucleic Acids

3.1 Introduction

Peptide nucleic acids (PNAs) are synthetic molecules composed of a peptide backbone and nucleic acid bases, which confer upon them DNA sequence recognition. They form strong and specific bonds to sequences in dsDNA, via four different binding modes, depending on their design.[1] Due to a neutrally-charged backbone, their binding to dsDNA deviates substantially from that of conventional oligonucleotides. These differences are exploited in the many applications of PNAs, which are covered in a number of excellent works.[1-9]

While a myriad of PNA applications have been developed, we seek to focus on the promising application of PNAs as tools for single-molecule experiments and to highlight

the advantages of single-molecule experiments for optimizing PNA-DNA binding conditions. Many single-molecule dsDNA experiments require a method to attach a desired modification to a specific, non-terminal location on dsDNA. The modification is usually a small molecule such as a hapten[10-12] or fluorophore.[13-15] The modification can facilitate sequence-specific detection or attachment of other components, such as nanospheres. Ultimately, such methods are being explored for a range of single-molecule biophysics experiments, for single-molecule genotyping, and for sequence-based separation methods. Since PNAs are currently commercially available with a choice of modifications including haptens, such as biotin, and fluorophores, such as TAMRA, they are excellent tools for attaching sequence-specific modifications to dsDNA.

There are several notable examples of single-molecule DNA experiments successfully employing PNAs, primarily to locate sequences on genomic-length DNA using various approaches.[10,13,14,16,17] In an early example of physical mapping of single DNA molecules, individual biotinylated PNA-DNA complexes conjugated to streptavidin were directly visualized via electron microscopy.[10] The streptavidins served as the electron microscopy markers and appeared as 'beads' located at the target sites. In a later example, high-resolution mapping of YOYO-1-stained DNA using fluorophore-conjugated PNAs was achieved by concurrent scanning near-field optical and atomic force microscopy (SNOM/AFM).[16] In an application-oriented example, high-throughput mapping of TOTO-3-stained DNA using fluorophore-conjugated bisPNAs was achieved using a continuous flow device and direct linear analysis (DLA), a scheme that provides the spatial locations of the PNAs along the extended DNA molecule.[13,14] The fluorescently labeled DNA was analyzed as it flowed past multicolor confocal fluorescence detectors with single-fluorophore sensitivity. An alternative approach to high-throughput mapping of DNA using bisPNAs was also achieved via purely electrical detection of individual DNA molecules moving through synthetic nanopores 4-5 nm in diameter.[17] The region of the DNA with a PNA bound was detected as it threaded through the pore. Here the presence of the bound PNA generated a signal due to the displacement of electrolytes from the nanopore by the PNA-DNA complex and no additional modifications to the PNA were required.

While there are other commonly used techniques to introducing site-specific modification for single-molecule DNA experiments,[15,17-22] PNAs are advantageous because they are compatible with cofactors such as Mg^{2+} and non-destructive to dsDNA. Furthermore, they do not require ligation and bind with high specificity and yield.[13] While PNAs have many favorable attributes, their binding conditions require careful optimization.

Here we describe two model experiments that demonstrate the virtues of PNAs as tools for single-molecule DNA experiments. In the first, individual PNA-DNA complexes are studied via fluorescence microscopy. The DNA backbone is fluorescently stained, the bound PNAs are labeled with fluorescent nanospheres, and the complexes are stretched either on a slide or in a stagnation-point extensional flow, which produces controlled, precise extension of the DNA (Figure 3-1). In the second, individual PNA-DNA complexes are characterized via optical tweezers. A sphere attached to one end of the

PNA-DNA complex is manipulated with optical tweezers while a second sphere, attached to the bound PNA, is held stationary by suction on a micropipette (Figure 3-2). Collectively, these single-molecule experiments with 8 bp bisPNAs demonstrate that (i) PNA-DNA binding optimization can be performed simultaneously across all target sites for an entire genomic-length DNA molecule (ii) PNAs can be used to locate specific dsDNA sequences on individual DNA molecules (iii) PNAs can serve as sequence-specific tethers for optical tweezer setups, and (iv) the PNA-DNA bond can sustain forces of roughly 60 pN on average.

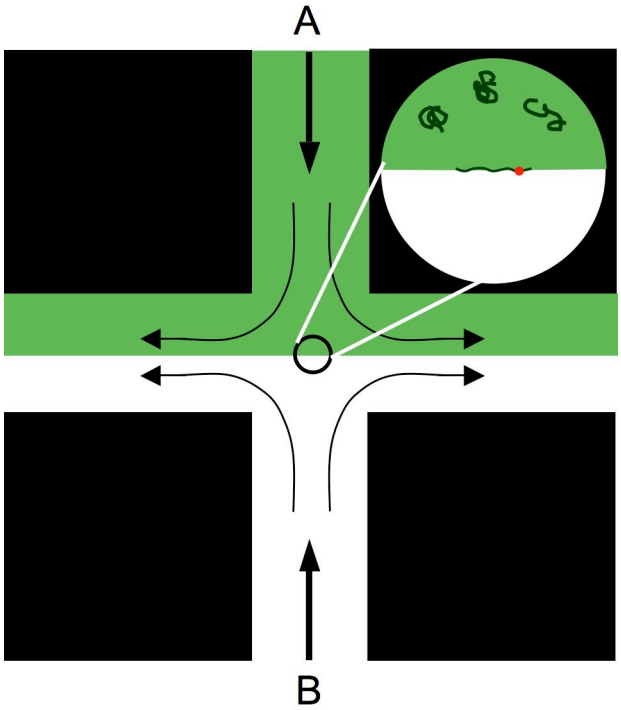


Figure 3-1. Schematic of the cross-slot flow device for single-molecule flow experiments. The green area is occupied by fluid from the top inlet arm (A) and the white area is occupied by fluid from the bottom inlet arm (B). An extensional stagnation point flow is created at the center of the device. Inset: an enlarged view of the stagnation point where a labeled, stained DNA molecule is trapped and stretched.

Figure 3-2. Optical tweezer manipulation of the twPNA-DNA-dig complex, which is attached by one end to an anti-digoxigenin-coated sphere, manipulated by the optical trap, and by the other end to a streptavidin-coated sphere, sucked onto a micropipette.

3.2 PNA specificity

The two bisPNAs used here, referred to as flPNA and twPNA for fluorescence microscopy and optical tweezer experiments, respectively, are functionalized with TAMRA at the N-terminus and biotin at the C-terminus (see Methods). The target binding sites for flPNA and twPNA are GAGAAGGA and AAGAGAAA, respectively. PNA-DNA binding conditions were first optimized for specificity because, in addition to binding to target sequences, PNAs are known to bind to sequences that have single-end mismatches (SEMM) or double-end mismatches (DEMM).[10,13,14]

The yield and specificity of PNA binding is dependent upon several factors that can be grouped into two categories: the composition of the PNA molecule and the binding conditions. Properties of the PNA that affect binding include PNA oligomer length,[23] net charge,[23-27] substitution of pseudoisocytosine for cytosine,[23,28] modifications such as haptens, fluorophores, and lysines,[24-27] and the order of the modifications.[26] The PNA binding mode is also an important factor; here we use bisPNAs that bind via triplex invasion because it is the most stable binding mode.[3] For a given PNA molecule, the incubation conditions must be optimized for yield and specificity. Incubation conditions affecting yield or specificity include incubation time and temperature,[13,25] pH,[23,28] salt concentration,[24-26,29] and the ratio of PNA to DNA.[13,29]

Some of these properties are interdependent. For example, modifications can affect the net charge of the PNA molecule. There is also an interplay between the PNA properties and incubation conditions as changes to the PNA will affect incubation conditions. The addition of a modification to a PNA, for example, is known to alter the binding yield at a given set of conditions. For example, Demidov et al. observed that modifying a PNA molecule with biotin resulted in a slightly decreased binding affinity, which they attribute

to the charge difference between the modified and unmodified PNA.[10] A *difference* in modifications on otherwise identical PNAs has also been shown to alter optimized binding conditions. For example, Chan et al. optimized binding conditions for two PNA molecules, one with a terminal TAMRA modification on one end and one with terminal TAMRA modifications on both ends. While the molecules were otherwise identical, the dual modified PNA required a 50% longer incubation time to achieve optimal binding.[13,14]

The traditional approach to determining the extent of PNA binding and specificity is a bulk gel-shift assay. Subsequent to hybridizing DNA with PNA at a given set of conditions, the mobility of fragments of DNA containing either a target site or mismatch site is compared. A representative example of this optimization process for flPNA is presented in the Methods section. While the widely-used bulk gel-shift assay is informative, it is limited in that the non-specific binding sites must be anticipated and, thus, such sites will only be discovered if the appropriate fragments are selected in advance.

3.3 Visualizing PNA binding via slide-stretching and flow-stretching

To simplify the binding optimization procedure and decrease the opportunities for non-specific binding, PNA sequences should be selected such that the occurrences of SEMM and DEMM sites, particularly when clustered, are minimized. The locations of the target, SEMM, and DEMM sites are easily determined by searching a known DNA sequence. When all of these sites are binned in the same way as the experimental data, the histograms provide the locations of potential erroneous peaks. In Figure 3-3a, the predictive binding map for the binding of flPNA on λ-DNA, there is a cluster of DEMM sites between 30 and 40 kbp. There are also clusters consisting of both SEMM and DEMM sites between 40 and 50 kbp. This type of non-specific binding, where a cluster of low-likelihood binding events appears as a substantial peak in the binding location histogram, is made obvious by the predictive binding map (Figure 3-3a) and would not have been detected with the traditional bulk gel-shift assay approach. When the binding reaction is not followed by additional heating to optimize binding, the slide-stretching assay indeed reveals extensive non-specific binding (Figure 3-3b).

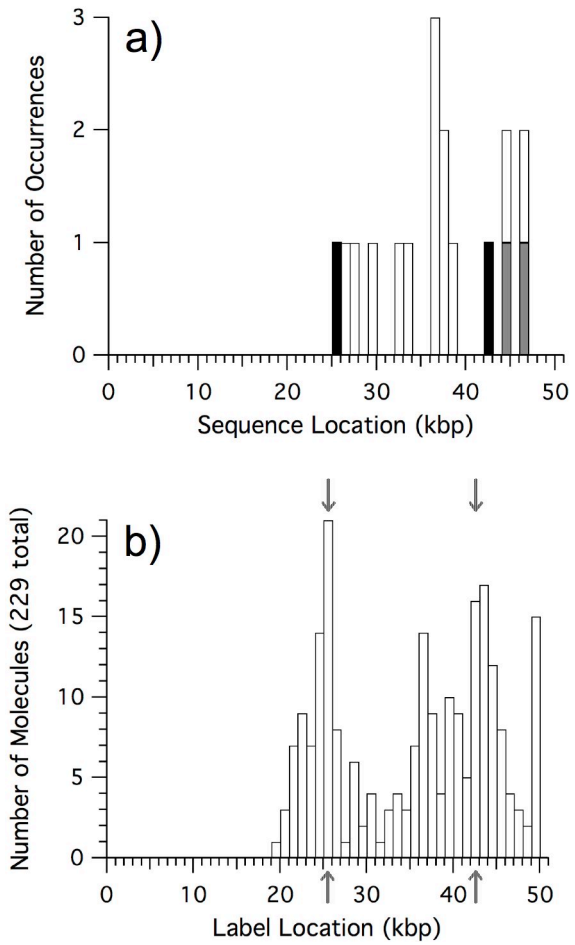


Figure 3-3. The predictive binding map (a) indicates the locations of target, SEMM, and DEMM sequences for flPNA on λ-DNA with black, gray, and white bars, respectively. A histogram of label locations (b) determined via slide-stretching (229 molecules) reveals that flPNA-DNA complexes formed at non-optimal binding conditions exhibit non-specific binding. Arrows point to locations of target sequences.

Using a simple single-molecule slide-stretching method, entire molecules, up to several Mbp in size,[30] can be assayed simultaneously for non-specific binding. The specificity of binding conditions can be evaluated by stretching the PNA-DNA complexes created at those conditions on a slide and determining the positions of the labels. The presence of a fluorescent label, as opposed to a single fluorophore, simplifies this approach by not requiring a TIRF setup or too much attention to reducing background fluorescence. The flPNA and λ-DNA complexes (flPNA-DNA) were formed in a binding reaction that was incubated at 37 °C for 25 hours. Drop dialysis was performed to remove excess flPNA. To optimize binding specificity, the recovered volume was heated to 63 °C for 15 minutes following dialysis. Immediately after heating, the NaCl concentration was brought to 100 mM to inhibit any further binding. The flPNA-DNA complexes were labeled with 40 nm NeutrAvidin-coated fluorescent polystyrene spheres and stained with

YOYO-1 prior to slide-stretching. A 5 μL drop of stained, labeled flPNA-DNA in slide-stretching buffer was deposited onto an untreated glass slide (see Methods). After 1 minute, a poly-L-lysine-coated cover slip was deposited onto the drop. Individual DNA molecules were stretched onto the positively charged surface by the receding meniscus.

For both slide-stretching and flow-stretching (discussed below), the locations of the labels on individual DNA molecules are readily determined from intensity profiles along each DNA backbone, imaged via fluorescence microscopy (Figure 3-4). The most compelling advantage of this slide-stretching approach is that mismatch sites do not need to be anticipated and all types of mismatches are detected. This is particularly important when clusters of low likelihood mismatches are present and can result in erroneous peaks.

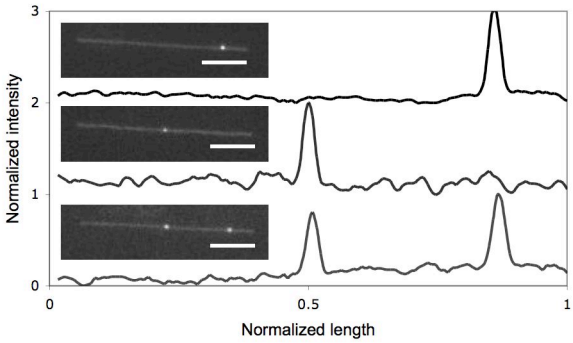


Figure 3-4. Intensity profiles of images (inset) of stained, labeled PNA-DNA complexes stretched in the cross-slot flow device labeled at each of the flPNA target binding locations. The DNA molecules are extended to 94% of their contour length. Plots are offset for clarity. All scale bars are 5 μm.

When the binding conditions are optimized, slide-stretching reveals only binding to the target sites (Figure 3-5a). The target site positions can be visualized by fitting Gaussians to the two apparent subpopulations. Note that both the predicted and experimental maps of binding locations are modified such that all sites are on one half of the molecule. This approach is consistent with procedures used in other single-molecule studies[20,30,31] and is necessary because, when there is only one label on a DNA molecule, the orientation of the DNA molecule is not known. Furthermore, both target binding sites for flPNA are known to be on the same half of the λ-DNA molecule. The label locations determined via slide-stretching (Figure 3-5a), which indicate the locations of flPNA binding sites on λ-DNA, are 23.6 ± 3.3 and 40.5 ± 3.0 kbp for the middle and end peaks, respectively. Using slide-stretching, absolute error for the middle peak is 0.8 kbp (target binding site at 24.4 kbp) and for the end peak is 1.1 kbp (target binding site at 41.6 kbp).

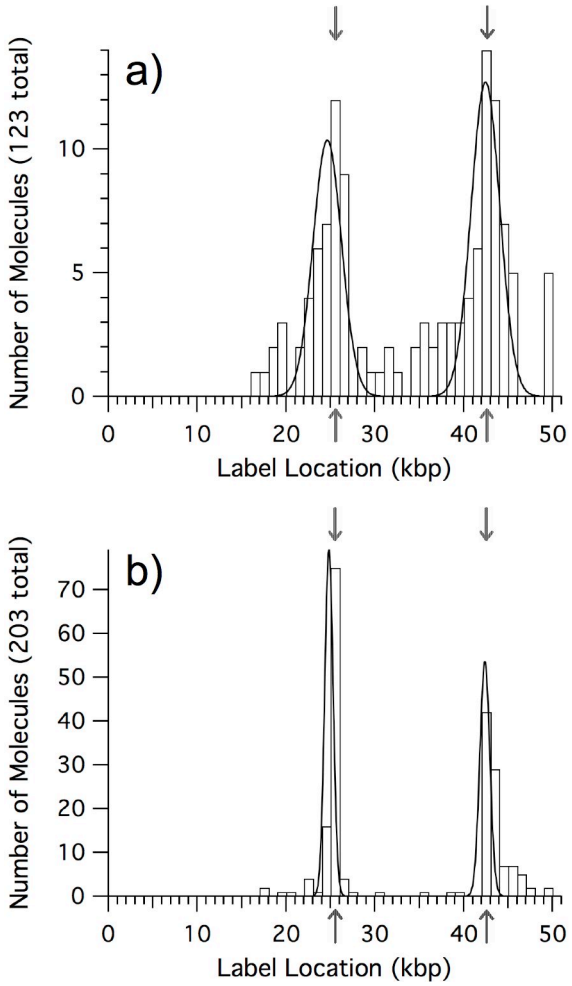


Figure 3-5. Histograms of label locations on flPNA-DNA complexes formed at optimal binding conditions determined via (a) slide-stretching (123 molecules) and (b) flow-stretching (203 molecules) where the average DNA extension is 96% of the DNA contour length. Arrows point to locations of target sequences.

After the binding conditions of the flPNA-DNA complexes were optimized for specificity using the slide-stretching assay, the flPNA-DNA binding was more rigorously determined using a flow-stretching assay, which allows for better fidelity in linearizing the DNA and better localizes the binding locations. The stained, labeled flPNA-DNA complexes were hydrodynamically trapped and stretched in a stagnation-point extensional flow generated in a microfludic cross-slot geometry (Figure 3-1), and the positions of the labels were then determined in the same manner as for slide-stretched molecules. Such microfluidic devices have been used extensively to study single DNA molecules.[32-36] The channel cross-section of the PDMS flow cell was 800 μm by 120 μm and the fabrication is described elsewhere.[31] As illustrated in Figure 3-5, when the same flPNA-DNA complexes are analyzed via slide-stretching and flow-stretching, the latter produces a more accurate and precise map. Flow-stretching increases the accuracy and precision of the label position measurements relative to slide-stretching because it results

in a narrower distribution of DNA extension and the flow eliminates "false positives" that occur in slide-stretching when free fluorescent nanospheres are coincidentally adsorbed to the slide near a stretched DNA molecule. In addition, in the present experiments, the flow-stretching conditions allow for somewhat greater mean extension of the DNA molecules. Comparing the distribution of DNA extensions, represented as percentage of the DNA contour length, slide-stretching results in 89% ± 6% while flow-stretching results in 96% ± 3%. The same label locations determined via flow-stretching (Figure 3-5b) are 24.0 ± 1.5 and 42.4 ± 1.6 kbp for the middle and end peaks, respectively. Using flow stretching, absolute error for the middle peak is 0.4 kbp and for the end peak is 0.8 kbp. Thus, utilizing labeled PNAs, the single-molecule flow-stretching assay correctly localizes target sequences to within 1 kbp, which is comparable to other single-molecule sequence detection methods.[20,31,37] This flow assay demonstrates the ability of PNAs to serve as indicators of specific sequences on dsDNA.

3.4 PNAs as tweezer handles

To demonstrate that PNAs can also be applied as tethers for optical tweezer experiments, the strength of the PNA tether was evaluated. This work was done by Craig L. Hetherington, a collaborator from the Bustamante lab at Univeristy of California, Berkeley and co-author of the version of this work published in *Nano Letters* **10**, 4697-4701 (2010). Ideally, tethers for nucleic acids in optical tweezer experiments should be able to sustain forces at least on the order of tens of pN. Forces in this range are characteristic of the folding of RNA and some proteins[38-40] and many molecular motors that act on DNA and RNA.[41-44] As these optical tweezer experiments required distinct attachment sites to two beads, a digoxigenin-modified DNA, half-λ-DNA-dig, was first formed as a precursor. TwPNA-DNA-dig complexes, which have a bound biotinylated PNA, were formed from twPNA and half-λ-DNA-dig in a binding reaction incubated at 45 °C for 5 hours, then cooled to 4 °C over 30 minutes. The NaCl concentration was brought to 50 mM to inhibit any further binding and the reaction volume was heated to 50 °C for 10 minutes to enhance binding specificity. Note that since twPNA and flPNA bind to different target sequences, their binding conditions were optimized separately, resulting in different optimal binding conditions.

Individual twPNA-DNA-dig complexes were manipulated via two spheres attached to the biotinylated PNA and the dig modification, respectively (Figure 3-2). In addition, similar control molecules that were not tethered via twPNA, called biotin-DNA-dig (see Methods), were subjected to the same routine. Control molecules did not exhibit any rupture, indicating that rupture of the twPNA-DNA-dig complex is due to the twPNA-DNA bond. The twPNA-DNA-dig complexes were repeatedly stretched and relaxed between 0 and 90 pN at a rate of 175 nm per second. Several tw-PNA-DNA-dig complexes were robust to the repeated application of 90 pN forces. For the twPNA-DNA-dig complexes that ruptured at forces below 90 pN, there was a unimodal distribution of rupture forces (Figure 3-6) with a mean rupture force of 57.3 pN and a standard deviation of 14.5 pN (sample size of 12 molecules).

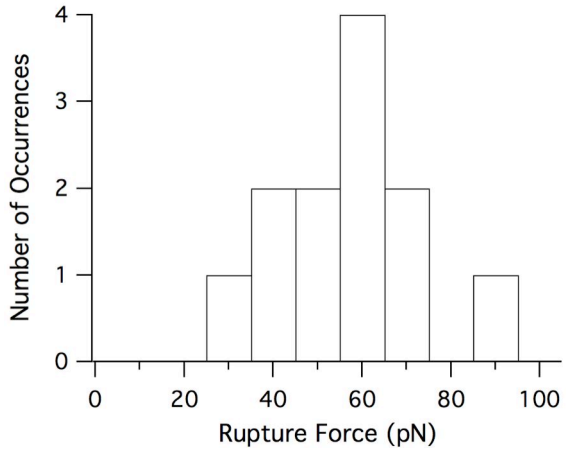


Figure 3-6. Optical tweezer manipulation of the twPNA-DNA-dig complex produces a distribution of rupture forces for those complexes that did not sustain the maximum applied force of 90 pN.

While this rupture force distribution directly characterizes only the specifically-bound twPNA used here, we expect the values to be of the same order for other 8 bp bisPNAs with similar base content. Note that we only consider here those complexes where the twPNA was specifically bound. DNA molecules with specifically-bound twPNAs are easily distinguishable in the optical tweezer experiment. When the DNA is stretched to its contour length, the distance between the two beads is known and corresponds to the distance between the end of the DNA and the twPNA binding location. The binding properties of PNAs make them attractive candidates for single-molecule dsDNA recognition and manipulation. Typical nucleic acid studies use biotin-streptavidin bonds and the weaker dig-anti-dig bonds. The mean rupture force of the PNA-DNA bond compares favorably with the rupture force of a single dig-anti-dig bond, roughly 20 pN at these loading rates,[45] making PNA-mediated tethers another excellent alternative to biotin-mediated tethers for nucleic acid manipulation. The strength of interaction for a typical 8 bp bisPNA with DNA is well-suited for optical tweezer experiments.

3.5 Conclusions

In conclusion, while the versatility of PNAs and their many uses have been discussed widely in the literature, little direct attention has been paid to their potential specifically as tools for single-molecule experiments. They are commercially available and can be synthesized with fluorophores or haptens, their binding to dsDNA is strong and specific, and they are compatible with cofactors required in many single-molecule biological assays. We present both design considerations for PNAs and an approach to optimizing their binding to dsDNA. We discuss their desirable attributes for single-molecule experiments and demonstrate this capacity with both fluorescence microscopy and optical tweezer single-molecule DNA studies. PNAs are amenable to labeling and can be used to locate specific dsDNA sequences, to act as obstacles at specific DNA sequences, or as

strong sequence-specific tethers for optical or magnetic tweezer setups. With careful attention to PNA design and optimization of binding conditions, PNAs have great potential as versatile tools for single-molecule research.

3.6 Methods

3.6.1 DNA and PNA preparation

10 μL aliquots of λ-DNA (48.5 kbp, New England Biolabs) were heated to 63 °C for 1 minute and plunged into wet ice to eliminate concatemers. The PNA sequences (Biosynthesis Inc.) from N-terminus to C- terminus are

flPNA:
TAMRA-OO-lys-lys-TCC TTC TC-OOO-JTJ TTJ JT-lys-OO-lys-biotin

twPNA:
TAMRA-OO-lys-lys-TTT CTC TT-OOO-TTJ TJT TT-lys-OO-lys-biotin

where TAMRA is carboxytetramethylrhodamine, O is 8-amino-3,6-dioxaoctanoic acid, a flexible hydrophilic tether, lys is Lysine, and J is pseudoisocytosine. These PNA sequences were modified from those used in earlier work by Chan et al.[13] Both sequences are bisPNAs and are modified with both biotin, to allow for conjugation to Avidin-coated beads, and with TAMRA, to give the PNAs UV-VIS spectra clearly distinguishable from that of DNA. Aliquots of 50 μM PNA in water were heated to 50 °C for 10 minutes and gently mixed prior to use.

3.6.2 Bulk characterization of PNA-DNA complexes

To determine the extent and specificity of the PNA-DNA binding protocols, the PNA-DNA complexes underwent a restriction enzyme digest to isolate 500 to 700 base pair fragments that contained either a target site or a site with a single end mismatch (SEMM). Binding was stopped by the addition of NaCl up to 100 mM. The target site-containing or SEMM site-containing fragments were typically run on a precast 4-15% gradient polyacrylamide gel (BioRad) in TBE buffer (89 mM Tris-borate, 89 mM boric acid, 2 mM EDTA) at 75V for 5 minutes followed by 150V for 55 minutes. DNA was stained with a 1:10,000 solution of SYBR green (Invitrogen) as per the manufacturer's instructions and imaged with a Typhoon 9200 (GE Life Sciences). The progress of fragments with a bound PNA is retarded in the gel, which produces a band shifted towards higher molecular weights relative to the fragments with no PNA bound. The relative intensities of the shifted band and unshifted band for the target site-containing fragment is an indication of the extent of specific binding. Similarly, the relative intensities of the shifted band and unshifted band for the SEMM site-containing fragment

is an indication of the extent of non-specific binding. Note that the presence of PNAs does not interfere with the activity of restriction endonucleases, provided the PNAs are not bound at the cleavage sites.

A representative binding condition optimization gel for flPNA is presented in Figure 3-7. The top gel is an assay of the target site-containing fragment. After a 25 hr incubation at 37 °C, the band is mostly shifted, indicating a high proportion of target site-containing fragments with PNAs bound, i.e., complexes. The unshifted bands represent DNA fragments without PNA bound. The complexes were heated to 45 °C for 1 hr after incubation but prior to enzyme digestion resulting in a slightly larger unshifted band. The trend continues with similar heating to 55 °C for 1 hr. The bottom gel is an assay of the SEMM site-containing fragment. When it undergoes the same heat treatment, there is a larger relative increase in the unshifted population, which is the desired outcome, since we seek to reduce PNA binding to SEMM sites via heat treatment. There is usually a trade-off between specificity and extent of binding.

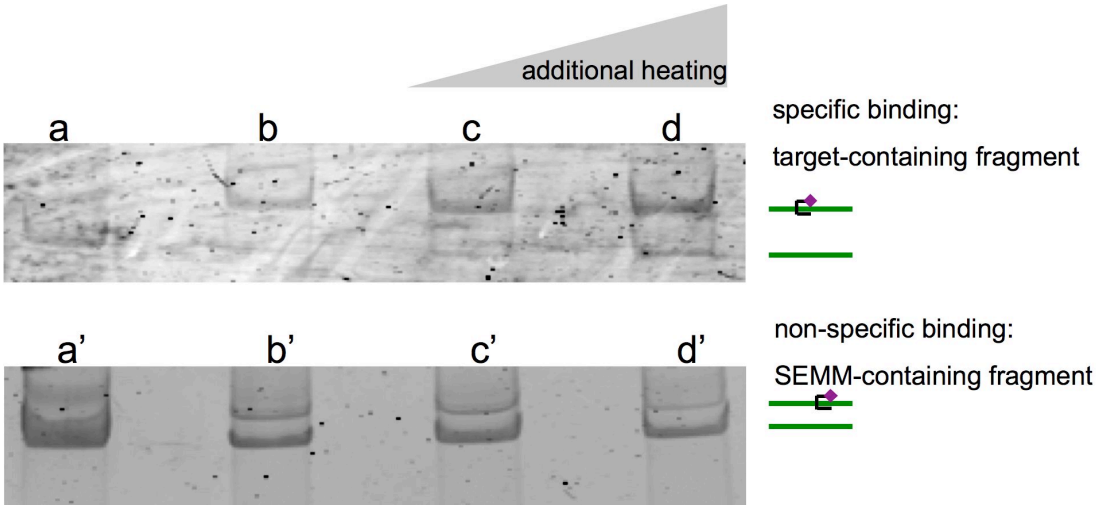


Figure 3-7. Bulk gel shift assays of flPNA binding to a target site-containing fragment (a-d) and a SEMM site-containing fragment (a'-d') of λ-DNA. Fragments that were not incubated with PNA were used as a control (a, a') and fragments with bound PNA will shift bands upwards. Complexes were first formed in a 25 hr incubation at 37 °C (b, b'). A majority of the target site-containing fragment band (b) is upwardly shifted, indicating target site binding. A minority of the SEMM site-containing fragment band (b) is upwardly shifted, indicating a small amount of SEMM site binding. After the incubation but prior to enzyme digestions, complexes were heated to 45 °C (c, c') or 55 °C (d, d') for 1 hr. Additional heating nearly eliminates SEMM site binding, indicated by the dimming of the shifted band, (c', d') and slightly decreases the extent of target site binding, indicated by a darkening of the unshifted band (c, d).

3.6.3 Single-molecule characterization via fluorescence microscopy

PNA-DNA complex formation: The flPNA and λ-DNA complexes (flPNA-DNA) were formed in a 100 μL binding reaction (pH 8) consisting of 10 mM Tris, 1 mM EDTA, 2 mM NaCl, 1.3 nM λ-DNA, and 2.5 μM PNA that was incubated at 37 °C for 25 hours. Drop dialysis was performed to remove excess flPNA. The 100 μL reaction volume was deposited onto a VSWP nitrocellulose drop dialysis membrane with a 25 nm pore size (Millipore) over 40 mL of dialysis buffer for 16 hours at 4 °C. The dialysis buffer consists of 10 mM Tris, 1 mM EDTA, and 30 mM NaCl. Following dialysis, the recovered volume was heated to 63 °C for 15 min. to enhance binding specificity. Immediately after heating, the NaCl concentration was brought to 100 mM to inhibit any further binding.

Labeling flPNA-DNA complexes with fluorescent nanospheres: For fluorescence microscopy experiments, single-molecule characterization requires that the flPNA-DNA complexes be labeled with fluorescent nanospheres (Figure 3-8). Red fluorescent spheres (excitation 580 nm/emission 605 nm) were selected to be distinguishable from the green YOYO-1 (excitation 490 nm/emission 510 nm) DNA backbone stain. Fluorescence from the nanospheres and DNA backbone could be imaged independently or together, depending on the filters used (see below).

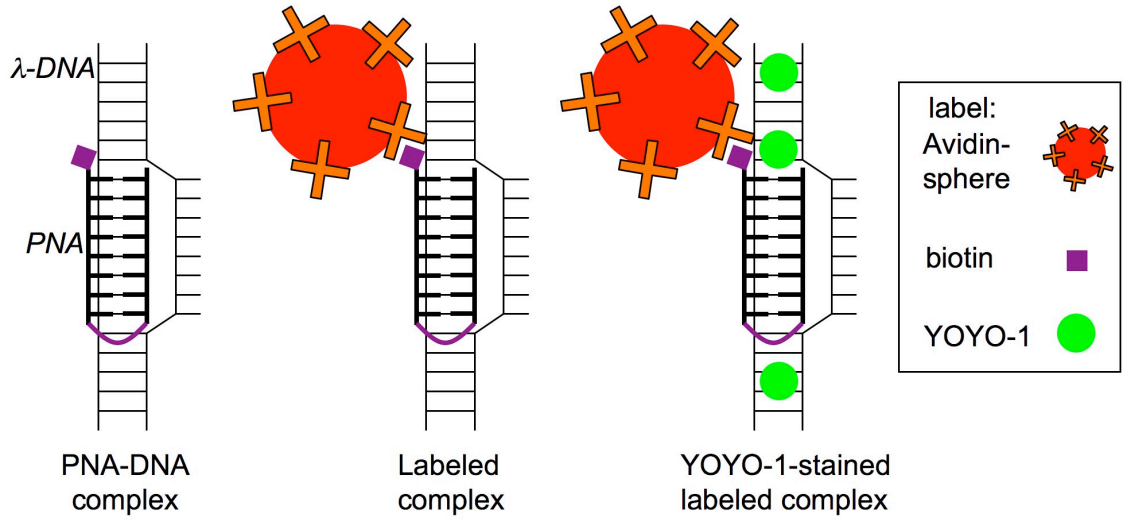


Figure 3-8. A biotinylated 8 bp bisPNA binds to λ-DNA, forming a PNA-DNA complex. The bound PNA is labeled via an Avidin-coated sphere, creating a labeled PNA-DNA complex. For fluorescence microscopy studies, the backbone of the DNA is stained with YOYO-1, resulting in a stained, labeled PNA-DNA complex.

A 200 μL labeling reaction (pH 7.5) consisting of 50 mM Tris, 1 mM EDTA, 100 mM NaCl, 5 vol% glycerol (Fisher), 0.01% Triton X-100 (Dow), 100 μg/mL BSA (Sigma-

Aldrich), 100 fM flPNA-DNA, 0.001% 40 nm NeutrAvidin-coated fluorescent polystyrene spheres (Invitrogen) was incubated at room temperature for 1 hour on a rotary mixer at 0.05 Hz.

Fluorescent stain: Labeled flPNA-DNA complexes were stained with YOYO-1 (Invitrogen) at a ratio of 1 dye molecule per 4 bp. The YOYO-1 was diluted to 10 μM in 10 mM Tris, 1 mM EDTA, 10 mM NaCl buffer, pH 8. A 100 μL staining solution consisting of 16.64 pM labeled flPNA-DNA complexes, 0.2 μM YOYO, 10 mM Tris, 1 mM EDTA, and 10 mM NaCl was incubated at room temperature for 1 hour, protected from light.

Slide-stretching: Glass cover slips were coated with a 0.1 mg/mL solution of poly-L-lysine hydrobromide (Fluka) in water. A 5 μL drop of the poly-L-lysine solution was sandwiched between two cover slips. After 1 minute, the slides were then slid apart and allowed to air dry for 20 minutes. The stained, labeled flPNA-DNA were diluted to 1.25 pM in slide-stretching buffer. Slide-stretching buffer (pH 8) consists of 10 mM Tris, 1 mM EDTA, 10 mM NaCl, and 0.0005% Triton X-100 (Dow). A 5 μL drop of stained, labeled flPNA-DNA in slide-stretching buffer was deposited onto an untreated glass slide. After 1 minute, a poly-L-lysine-coated cover slip was deposited onto the drop. Individual DNA molecules were stretched onto the positively charged surface by the receding meniscus.

The apparent presence of labels at the very end of the molecule (100% of the contour length) is likely an artifact of slide-stretching. During stretching, molecules can become tethered to the surface at the label site and the short length of DNA on one side of the label may not stretch fully. These endmost labels are neglected in further analysis.

This slide-stretching approach to binding optimization can also assay and reveal the effects of binding and buffer conditions on certain types of mismatches. Unlike gel-shift assays, which require fairly inflexible conditions to run, slide-stretching allows the data acquisition conditions to be easily adapted to the experimental conditions of interest (e.g. the presence of salts, a photoscavenging system, extreme pH, etc.). An additional benefit is that it obviates the extra step of acquiring fragments containing targets or mismatches that are traditionally purchased, made via PCR, or isolated via a restriction enzyme digest.

Flow-stretching: Individual stained, labeled flPNA-DNA complexes were trapped and stretched in a microfluidic flow cell with a cross-slot geometry. Such flow cells have been used extensively to study single DNA molecules.[32-36] The channel cross-section of the PDMS flow cell was 800 μm by 120 μm and the fabrication is described elsewhere.[31] The cross-slot geometry creates a line of pure extension where the two opposing inlet streams meet and a stagnation point, a point of zero velocity, at the center of the device. When a DNA molecule in an inlet stream arrives at the center of the device, its center of mass is trapped at the stagnation point and its opposing ends are stretched apart along the line of pure extension. The inlets and outlets of the flow cell are connected to syringes

and waste reservoirs, respectively, by Tygon tubing (0.02" ID, Cole-Parmer). The position of the stagnation point is controlled by manually adjusting the height of one of the exit reservoirs. As discussed by Schroeder et al., this allows DNA molecules to be trapped and imaged for minutes to hours.[34] The degree of extension or linearization of the trapped DNA molecule is controlled by the dimensionless product of the DNA molecule's relaxation time and the flow's extensional strain rate. The relaxation time is governed by the DNA molecule and buffer viscosity. The extensional strain rate is governed by the cross-slot geometry and the flow rate into the inlet arms. Keeping all else constant, the degree of extension of the DNA molecule can be tuned by adjusting the flow rate.

Prior to use, the tubing and flow cell were rinsed with Milli-Q filtered (Millipore) water and conditioned to prevent air bubbles and to prevent non-specific adsorption of the protein-coated fluorescent labels. To reduce surface tension and prevent the formation of air bubbles, the tubing and flow cell were incubated with a solution of 0.1 vol% Triton X-100 surfactant (Dow) in a buffer of 10 mM Tris, 1 mM EDTA, 10 mM NaCl, pH 8 for 30 min. To passivate surfaces and prevent the adsorption of fluorescent labels, the tubing and flow cell were incubated with a 10 mg/mL solution of BSA (Sigma-Aldrich) in a buffer of 10 mM Tris, 1 mM EDTA, 10 mM NaCl, pH 8 for 30 min.

Flow-stretching buffer (described below) containing stained, labeled flPNA-DNA complexes was fed through one inlet and flow-stretching buffer without DNA was fed through the opposing arm, both using the same PHD 2000 syringe pump (Harvard Apparatus). The pump delivered fluid at 2 ml/hr until the entire volume of the flow cell and tubing was filled with flow-stretching buffer, and then at 100 μL/hr for all experiments. This flow rate results in an average steady-state extension of l-DNA molecules at the stagnation point of 96% of the DNA molecule's fully extended contour length. While full extension can be achieved at higher flow rates, trapping molecules becomes somewhat more difficult. The imaging area through the camera was 80.5 μm by 80.5 μm, more than adequate to visualize a stretched λ–DNA molecule with a roughly 20 μm stained contour length.

The stained, labeled flPNA-DNA complexes were diluted to 270 fM in flow-stretching buffer with a viscosity of 50 cP. Flow-stretching buffer (pH 8) consists of 10 mM Tris, 1 mM EDTA, 100 μg/mL BSA (Sigma-Aldrich), 59 wt% sucrose (MB grade, Sigma-Aldrich), 1 wt% glucose (Sigma-Aldrich), 2 vol% β-mercaptoethanol (Sigma-Aldrich), 20 μg/mL catalase, and 100 μg/mL glucose oxidase (Roche). The sucrose increases the solution viscosity, which increases the relaxation time of the DNA and decreases the flow rate required to achieve a fixed fractional extension of the trapped DNA. The last three components act to inhibit photobleaching and photocleavage. In addition to providing more uniformly extended DNA molecules with a narrower distribution of extensions (relative to slide-stretching), the flow assay also eliminates false signals present in slide-stretching that result from labels coincidentally non-specifically adsorbed onto the surface along the DNA backbone.

Data acquisition: The stained flPNA-DNA complexes (green) labeled with fluorescent spheres (red) were visualized using a 100x, 1.4-NA oil-immersion objective on a fluorescence microscope (Leica DMIRE II). Red and green images were captured using a N2.1 Cy3/TRITC (Vashaw Scientific) and 41001 FITC (Chroma) filter set, respectively. A dual-band 51004v2 excitation/emission filter (Chroma) was used to image the red and green channels simultaneously. A cooled CCD camera (Photometrics Cascade 512b) captured images in conjunction with SimplePCI software. Images were captured at roughly 10 fps.

Data analysis: Images of slide-stretched DNA and flow-stretched DNA were analyzed with NIH Image J to extract a DNA length and label position along the length. For slide-stretched DNA, only one image is obtained per DNA molecule and only DNA molecules with a minimum extension of 16.5 μm (78% of the total stained contour length) were included in the analysis. For flow-stretched DNA, a series of images is obtained per DNA molecule as it is trapped in the stagnation point. For each DNA molecule, the DNA length and label position were determined for three images acquired after the DNA molecule reached its steady-state extension and averaged. Only DNA molecules with a minimum extension of 19 μm (90% of the total stained contour length) were included in the analysis. Note that since most complexes had only a single label attached and there was no indication of the orientation of the molecule, the label positions were always taken as farthest from the 5' end. The two flPNA target sites are present on the same half of the λ-DNA molecule, farthest from the 5' end. This is consistent with the procedures used in other single molecule studies.[20,31,34] From histograms of optimized binding positions, mean and standard deviation values for binding locations were determined. Since the flPNA used here has a target site very close to the center of the λ-DNA molecule (50.2%), data points closest to this location were randomly assigned about the center to produce a symmetric Gaussian about the target site.

Labeling efficiency determined by fluorescence microscopy: As evidenced by the gel assay in Figure 3-7, a large majority of DNA molecules form flPNA-DNA complexes following the flPNA-DNA incubation reaction. The fraction of those complexes that are successfully labeled with the NeutrAvidin-coated fluorescent nanospheres is smaller. We can estimate a labeling efficiency, the probability that a flPNA-DNA complex is labeled with a fluorescent sphere, from the observed incidence of doubly to singly labeled molecules. For the two possible labeling sites, we assume the labeling efficiency is the same for each position. The probability of labeling any position is p, so that the probability of observing x labels, where x is one or two in our case, is ${}_2C_x p^x (1-p)^{2-x}$ where ${}_2C_x$ is the binomial coefficient, "2 choose x."

We take p, the probability of any position being labeled, as the fractional yield of labeling a flPNA-DNA complex with a fluorescent sphere. This indirect approach is necessary as, for both slide-stretching and flow-stretching, it is difficult to directly observe the number of unlabeled molecules needed to calculate a direct labeling efficiency. Moreover, the labeling efficiency is important since only stained, *labeled* PNA-DNA complexes

contribute to the data acquisition. The location of PNA binding cannot be determined on stained, *unlabeled* PNA-DNA complexes.

At the optimal binding conditions, the labeling efficiency is 11% and 6% for slide-stretching and flow-stretching, respectively. We believe this disparity results from an artifact of slide-stretching, where labels can non-specifically adsorb onto the slide seemingly along a stretched DNA molecule and are indistinguishable from an attached label. This does not occur in the flow-stretching assay, where the flow sweeps such unbound labels away.

3.6.4 Single-molecule characterization via optical tweezers

Functionalizing DNA ends: For optical tweezer experiments, which required distinct attachment sites to two beads, a digoxigenin-modified DNA construct, half-λ-DNA-dig, was first formed. A 510-bp fragment from λ-DNA, comprised of base pairs 24241 to 24741, was generated via PCR using Taq polymerase (New England Biolabs). The PCR buffer contained 200 μM dATP, dCTP, and dGTP (Fermentas); 133 μM dTTP; and 66 μM dig-dUTP (Roche). After cleanup, the PCR product was digested with XbaI (New England Biolabs) and shrimp Antarctic phosphatase (New England Biolabs). Simultaneously, λ-DNA (New England Biolabs) was digested with XbaI (New England Biolabs) in NEBuffer 2 at 37 °C. The two DNA fragments were ligated overnight using T4 ligase (Fermentas) with a 2-fold excess of PCR product over digested λ-DNA. The ligation product, half-λ-DNA-dig, was then dialyzed into 10 mM Tris, 0.1 mM EDTA, pH 8.0 through a 25 nm VSWP nitrocellulose filter (Millipore) for one hour at room temperature.

Another construct, the control molecule biotin-DNA-dig, was made to serve as a basis of comparison for optical tweezer experiments performed with the twPNA-DNA-dig construct described below. Biotin-DNA-dig simply has a biotin modification on one end and a dig modification on the other end. It was produced by treating λ-DNA with Klenow fragment, exo minus (New England Biolabs) in the presence of 70 μM dGTP and dTTP; and 30 μM biotin-dCTP and biotin-dATP (Invitrogen), thus filling the cohesive ends. This was then ligated to half-λ-DNA-dig, yielding biotin-DNA-dig. It was then subjected to the same thermal and salt conditions as the twPNA-DNA-dig complex and labeled with spheres as described below for the twPNA-DNA-dig complex. Note that both the biotin-DNA-dig molecule and the twPNA-DNA-dig molecule contain multiple biotin and dig moieties, which explains the extraordinary strength of the ends.

PNA-DNA complex formation: Complexes of twPNA and half-λ-DNA-dig (twPNA-DNA-dig) were formed in a 50 μL binding reaction (pH 8) consisting of 10 mM Tris, 0.1 mM EDTA, 5 nM half-λ-DNA-dig, and 1 μM PNA that was incubated at 45 °C for 5 hours, then cooled to 4 °C over 30 minutes. The NaCl concentration was brought to 50

mM to inhibit any further binding and the reaction volume was heated to 50 °C for 10 minutes to enhance binding specificity.

Labeling PNA-DNA complexes with functionalized spheres: The twPNA-DNA-dig complexes were labeled with two distinct functionalized spheres, streptavidin-coated spheres and anti-digoxigenin spheres. The anti-digoxigenin spheres were made by hybridizing 2.88 μm Protein-G-coated polystyrene beads (Spherotech, Inc.) with anti-digoxigenin (Roche) for 3 hours in phosphate-buffered saline. The twPNA-DNA-dig was first incubated with the anti-digoxigenin spheres in a 20 μL labeling reaction (pH 7.8) consisting of 25 fM twPNA-DNA-dig, and 0.025% anti-digoxigenin spheres in 10mM Tris, 50 mM NaCl, 5 mM $MgCl_2$, and 1 mg/ml BSA (New England Biolabs) at room temperature for 30 minutes. These singly-labeled twPNA-DNA-dig complexes were introduced into an optical tweezers flow chamber, where they were trapped and manually labeled with a 2.1 μm streptravidin-coated polystyrene bead sucked onto a micropipette.

Optical tweezer manipulation: The twPNA-DNA-dig complexes labeled with anti-digoxigenin spheres were introduced into an optical tweezers flow chamber.[46] The twPNA-DNA-dig complex was stretched between an optically trapped anti-digoxigenin bead and a 2.1 μm streptravidin-coated polystyrene bead sucked onto a micropipette. The distance between the two beads was changed at a rate of 175 nm per second while the tension and the extension of the molecule were monitored. The twPNA-DNA-dig complex was stretched and relaxed in a cyclic manner between 0 and 90 pN. For those instances when the twPNA tether ruptured when subjected to forces less than 90 pN, the distribution of rupture forces was recorded.

3.7 References

1. Nielsen, P. E. Peptide nucleic acid: a versatile tool in genetic diagnostics and molecular biology. *Curr. Opin. in Biotechnol.* **12**, 16-20 (2001).
2. Nielsen, P. E. Applications of peptide nucleic acids. *Curr. Opin. in Biotechnol.* **10**, 71-75 (1999).
3. Nielsen, P. E. Peptide Nucleic Acid Targeting of Double-Stranded DNA. *Methods in Enzymology* **340**, 329-340 (2001).
4. Nielsen, P. E. & Haaima, G. Peptide nucleic acid (PNA). A DNA mimic with a pseudopeptide backbone. *Chem. Soc. Rev.* **26**, 73-78 (1997).
5. Lundin, K. E., Good, L., Stromberg, R., Graslund, A. & Smith, C. I. E. Biological Activity and Biotechnological Aspects of Peptide Nucleic Acid. *Adv. Genet.* **56**, 1-51 (2006).
6. Demidov, V. V. PNA and LNA throw light on DNA. *Trends Biotechnol.* **21**, 4-7 (2003).
7. Paulasova, P. & Pellestor, F. The peptide nucleic acids (PNAs): a new generation of probes for genetic and cytogenetic analyses. *Ann. Genet.* **47**, 349-358 (2004).
8. Shakeel, S., Karim, S. & Ali, A. Peptide nucleic acid (PNA) - a review. *J. Chem. Technol. Biotechnol.* **81**, 892-899 (2006).
9. Ray, A. & Norden, B. Peptide nucleic acid (PNA): its medical and biotechnical applications and promise for the future. *FASEB J.* **14**, 1041-1060 (2000).
10. Demidov, V. V., Cherny, D., Kurakin, A., Yavnilovich, M., Malkov, V., Frank-Kamenetskii, M., Sonnichsen, S. & Nielsen, P. E. Electron-microscopy mapping of oligopurine tracts in duplex DNA by peptide nucleic acid targeting. *Nucleic Acids Res.* **22**, 5218-5222 (1994).
11. Bryant, Z., Stone, M. D., Gore, J., Smith, S. B., Cozzarelli, N. R. & Bustamante, C. Structural transitions and elasticity from torque measurements on DNA. *Nature* **424**, 338-341 (2003).
12. Gore, J., Bryant, Z., Stone, M. D., Nollmann, M. N., Cozzarelli, N. R. & Bustamante, C. Mechanochemical analysis of DNA gyrase using rotor bead tracking. *Nature* **439**, 100-104 (2006).
13. Chan, E. Y., Goncalves, N. M., Haeusler, R. A., Hatch, A. J., Larson, J. W., Maletta, A., Yantz, G., Carstea, E., Fuchs, M., Wong, G., Gullans, S. & Gilmanshin, R. DNA Mapping Using Microfluidic Stretching and Single-Molecule Detection of Fluorescent Site-Specific Tags. *Genome Res.* **14**, 1137-1146 (2004).
14. Phillips, K. M., Larson, J. W., Yantz, G. R., D'antoni, C. M., Gallo, M. V., Gillis, K. A., Goncalves, N. M., Neely, L. A., Gullans, S. R. & Gilmanshin, R. Application of single molecule technology to rapidly map long DNA and study the conformation of stretched DNA. *Nucleic Acids Res.* **33**, 5829-5837 (2005).
15. Xiao, M., Phong, A., Ha, C., Chan, T., Cai, D., Leung, L., Wan, E., Kistler, A. L., DeRisi, J. L., Selvin, P. R. & Kwok, P. Rapid DNA mapping by fluorescent single molecule detection. *Nucleic Acids Res.* **35**, e16 (2007).

16. Kim, J., Hirose, T., Sugiyama, S., Ohtani, T. & Muramatsu, H. Visualizing a Hybridized PNA Probe on a DNA Molecule with Near-Field Optical Microscopy. *Nano Lett.* **4**, 2091-2097 (2004).
17. Singer, A., Wanunu, M., Morrison, W., Kuhn, H., Frank-Kamenetskii, M. & Meller, A. Nanopore Based Sequence Specific Detection of Duplex DNA for Genomic Profiling. *Nano Lett.* **10**, 738-742 (2010).
18. Kuhn, H. & Frank-Kamenetskii, M. D. Labeling of unique sequences in double-stranded DNA at sites of vicinal nicks generated by nicking endonucleases. *Nucleic Acids Res.* **36**, e40 (2008).
19. Jo, K., Dhingra, D. M., Odijk, T., de Pablo, J. J., Graham, M. D., Runnheim, R., Forrest, D. & Schwartz, D. C. A single-molecule barcoding system using nanoslits for DNA analysis. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 2673-2678 (2007).
20. Ebenstein, Y., Gassman, N., Kim, S., Antelman, J., Kim, Y., Ho, S., Samuel, R., Michalet, X. & Weiss, S. Lighting Up Individual DNA Binding Proteins with Quantum Dots. *Nano Lett.* **9**, 1598-1603 (2009).
21. Oana, H., Ueda, M. & Washizu, M. Visualization of a Specific Sequence on a Single Large DNA Molecule Using Fluorescence Microscopy Based on a New DNA-Stretching Method. *Biochem. Biophys. Res. Commun.* **265**, 140-143 (1999).
22. Taylor, J. R., Fang, M. M. & Nie, S. Probing Specific Sequences on Single DNA Molecules with Bioconjugated Fluorescent Nanoparticles. *Anal. Chem.* **72**, 1979-1986 (2000).
23. Hansen, M. E., Bentin, T. & Nielsen, P. E. High-affinity triplex targeting of double stranded DNA using chemically modified peptide nucleic acid oligomers. *Nucleic Acids Res.* **37**, 4498-4507 (2009).
24. Abibi, A., Protozanova, E., Demidov, V. V. & Frank-Kamenetskii, M. D. Specific versus Nonspecific Binding of Cationic PNAs to Duplex DNA. *Biophys. J.* **86**, 3070-3078 (2004).
25. Kosaganov, Y. N., Stetsenko, D. A., Lubyako, E. N., Kvitko, N., Lazurkin, Y. & Nielsen, P. Effect of Temperature and Ionic Strength on the Dissociation Kinetics and Lifetime of PNA−DNA Triplexes. *Biochemistry* **39**, 11742-11747 (2000).
26. Kuhn, H., Demidov, V. V., Frank-Kamenetskii, M. D. & Nielsen, P. Kinetic sequence discrimination of cationic bis-PNAs upon targeting of double-stranded DNA. *Nucleic Acids Res.* **26**, 582-587 (1998).
27. Nielsen, P. E., Egholm, M., Berg, R. H. & Buchardt, O. Sequence-Selective Recognition of DNA Strand Displacement with a Thymine-Substituted Polyamide. *Science* **254**, 1497-1500 (1991).
28. Kuhn, H., Demidov, V. V., Nielsen, P. E. & Frank-Kamenetskii, M. D. An Experimental Study of Mechanism and Specificity of Peptide Nucleic Acid (PNA) Binding to Duplex DNA. *J. Mol. Biol.* **286**, 1337-1345 (1999).
29. Lundin, K. E., Hasan, M., Moreno, P. M., Törnquist, E., Oprea, I., Svahn, M. G., Simonson, E. O. & Smith, C. E. Increased stability and specificity through combined hybridization of peptide nucleic acid (PNA) and locked nucleic acid (LNA) to supercoiled plasmids for PNA-anchored "Bioplex" formation. *Biomolecular Engineering* **22**, 185-192 (2005).

30. Aston, C., Mishra, B. & Schwartz, D. C. Optical mapping and its potential for large-scale sequencing projects. *Trends Biotechnol.* **17**, 297-302 (1999).
31. Dylla-Spears, R., Townsend, J. E., Jen-Jacobson, L., Sohn, L. L. & Muller, S. J. Single-molecule sequence detection via microfluidic planar extensional flow at a stagnation point. *Lab Chip* **10**, 1543-1549 (2010).
32. Smith, D. E. & Chu, S. Response of Flexible Polymers to a Sudden Elongational Flow. *Science* **281**, 1335-1340 (1998).
33. Perkins, T. T., Smith, D. E. & Chu, S. Single Polymer Dynamics in an Elongational Flow. *Science* **267**, 2016-2021 (1997).
34. Schroeder, C. M., Babcock, H. P., Shaqfeh, E. S. G. & Chu, S. Observation of Polymer Conformation Hysteresis in Extensional Flow. *Science* **301**, 1515-1519 (2003).
35. Schroeder, C. M., Shaqfeh, E. S. G. & Chu, S. Effect of Hydrodynamic Interactions on DNA Dynamics in Extensional Flow: Simulation and Single Molecule Experiment. *Macromolecules* **37**, 9242-9256 (2004).
36. Tanyeri, M., Johnson-Chavarria, E. & Schroeder, C. Hydrodynamic trap for single particles and cells. *Appl. Phys. Lett.* **96**, 224101-224103 (2010).
37. Yu, H. & Schwartz, D. C. Imaging and analysis of transcription on large, surface-mounted single template DNA molecules. *Anal. Biochem.* **380**, 111-121 (2008).
38. Liphardt, J., Onoa, B., Smith, S. B., Tinoco, I. & Bustamante, C. Reversible Unfolding of Single RNA Molecules by Mechanical Force. *Science* **292**, 733-737 (2001).
39. Kellermayer, M. S., Smith, S. B., Bustamante, C. & Granzier, H. L. Mechanical Fatigue in Repetitively Stretched Single Molecules of Titin. *Biophys. J.* **80**, 852-863 (2008).
40. Cecconi, C., Shank, E. A., Bustamante, C. & Marqusee, S. Direct Observation of the Three-State Folding of a Single Protein Molecule. *Science* **309**, 2057-2060 (2005).
41. Smith, D. E., Tans, S. J., Smith, S. B., Grimes, S., Anderson, D. L. & Bustamante, C. The bacteriophage phi 29 portal motor can package DNA against a large internal force. *Nature* **413**, 748-752 (2001).
42. Wuite, G. J. L., Smith, S. B., Young, M., Keller, D. & Bustamante, C. Single-molecule studies of the effect of template tension on T7 DNA polymerase activity. *Nature* **404**, 103-106 (2000).
43. Yin, H., Wang, M. D., Svoboda, K., Landick, R., Block, S. M. & Gelles, J. Transcription Against an Applied Force. *Science* **270**, 1653-1657 (1995).
44. Stone, M. D., Bryant, Z., Crisona, N. J., Smith, S. B., Vologodskii, A., Bustamante, C. & Cozzarelli, N. R. Chirality sensing by Escherichia coli topoisomerase IV and the mechanism of type II topoisomerases. *Proc. Natl. Acad. Sci. U.S.A.* **100**, 8654-8659 (2003).
45. Neuert, G., Albrecht, C., Pamir, E. & Gaub, H. E. Dynamic force spectroscopy of the digoxigenin-antibody complex. *FEBS Lett.* **580**, 505-509 (2006).
46. Smith, S. B., Cui, Y. J. & Bustamante, C. Optical-trap force transducer that operates by direct measurement of light momentum. *Methods in Enzymology* **361**, 134-162 (2003).

Chapter 4

Single-Molecule Translocation Studies:

SpoIIIE Motor Protein

This work was done in collaboration with Ninning Liu from the Bustamante Lab at University of California, Berkeley.

4.1 Introduction

SpoIIIE is a hexameric motor protein responsible for chromosome segregation during bacterial sporulation. It is a DNA translocase that uses energy released by ATP hydrolysis to move DNA.[1] When SpoIIIE is held stationary at an interface, it acts as a DNA pump and moves DNA across the interface. This action is critical during bacterial sporulation, when DNA must be transported through the closing septum into the budding cell, and elucidating the functions of proteins such as SpoIIIE has been the subject of much research.[2-5] Questions of what governs the directionality of the motor and whether proteins and other features are "stripped" from DNA during translocation have recently been addressed by single-molecule experiments.[6,7] A single-molecule approach can greatly contribute to the understanding of these remarkable molecular motors.[8]

The SpoIIIE monomer is composed of a membrane-anchoring domain, a flexible linker domain, and a motor domain. The motor domain is made up of the α, β, and γ subdomains and, interestingly, the γ subdomain has been shown to be responsible for sequence recognition (Figure 4-1).[7] SpoIIIE and its close homologue, FtsK, have been studied using optical and magnetic tweezers and fluorescence microscopy.[1,6,7,9-11] Properties including linear velocity (bp/s), processivity (bp) and stall force (pN) have been determined.[12,13] In tweezer experiments, where DNA is tethered between two surfaces and held stationary, these motors have been observed translocating along DNA and changing the length of DNA by forming loops or plectonemes.[10] The linear velocity of SpoIIIE, observed directly with magnetic tweezers, is roughly 4 kb/s.[7] This result is comparable to the linear velocity of FtsK, determined independently with optical and magnetic tweezers to be over 4.5 kb/s.[10,14,15]

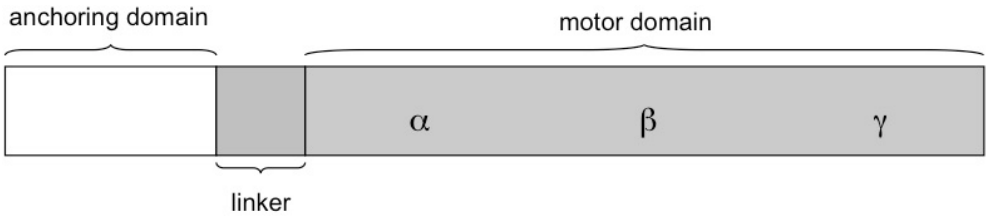


Figure 4-1. A schematic representation of a SpoIIIE monomer from N-terminus to C-terminus. A soluble derivative of the SpoIIIE monomer, which includes the portions in gray, is used in this study. The N-terminus of the linker domain is biotinylated.

While optical and magnetic tweezer-based studies of motor proteins allow for precise distance and force measurements, such studies have historically been limited to observing length changes in DNA due to translocation and looping or directly observing the movement of protein aggregates.[10] Single-molecule fluorescence-based studies of SpoIIIE could allow for direct observation of individual quantum dot (qdot) labeled SpoIIIE hexamers translocating along fluorescently stained DNA. Previous work has used a microfluidic cross-slot to trap and directly observe individual, untethered, and extended DNA molecules at a stagnation point.[16] The extent of stretching of the DNA molecule can be tuned in the cross-slot[17] and molecules can be trapped for several hours.[18,19] In principle, SpoIIIE-DNA complexes can be trapped at the stagnation point similarly to ordinary DNA, allowing for direct observation of translocation by individual motors.

There are several preliminary concerns that must be addressed for this experimental setup largely because the conditions most conducive to microfluidic DNA trapping and visualization are not the most conducive for SpoIIIE activity and vice versa. SpoIIIE activity requires the presence of Mg^{2+}, a divalent cation that interferes with the stretching

of the highly negatively charged DNA molecule. Furthermore, visualization of DNA in flow requires the presence of an intercalating dye, which may affect the translocation of SpoIIIE. Through bulk and single-molecule experiments, we demonstrate that it is possible to reconcile these experimental constraints and simultaneously achieve SpoIIIE activity and the visualization of DNA in flow. Direct observation of translocase activity using a fluorescent microfluidic platform is feasible.

4.2 Flow-stretching of DNA in SpoIIIE buffer

To determine the effects of Mg^{2+} on DNA stretching, the stretching of DNA in a microfluidic cross-slot was evaluated in a series of DNA stretching buffers containing 3 mM ATP and varying amounts of 0 to 10 mM Mg^{2+}. SpoIIIE activity is typically observed in a buffer with cofactor concentrations of 3 mM ATP and 10 mM Mg^{2+}.[6,7] The effects of these additional components on the fluorescence and photobleaching of the YOYO-1 stained DNA molecule were unknown. Furthermore, the photoscavenging system typically employed, consisting of the glucose oxidase and catalase system and β-mercaptoethanol,[16] a reducing agent, was altered. The reducing agent β-mercaptoethanol could adversely affect SpoIIIE through the reduction of disulfide bonds and was removed. Results of this series of single-molecule experiments, gathered from an ensemble of roughly 20 molecules at each set of conditions, indicate that DNA can consistently be stretched and be visualized in a buffer containing 3 mM ATP, 3 mM Mg^{2+}, and no β-mercaptoethanol. Higher Mg^{2+} concentrations result in poor stretching.

The linear velocity of SpoIIIE has been measured at over 4 kb/s in single-molecule experiments at 3 mM ATP and 10 mM Mg^{2+}.[7] Bulk colorimetric assays[20,21] were used to determine the effect of lower Mg^{2+} concentrations on SpoIIIE activity (see Methods). SpoIIIE buffer consisted of 50 mM Tris-HCl, 3 to 10 mM $MgCl_2$, and 0.1 mg/mL BSA at pH 7.4. The activity of SpoIIIE (mol ATP/sec/mol SpoIIIE monomer) on unstained DNA at 3 mM Mg^{2+} was roughly 50% of that on unstained DNA at 10 mM Mg^{2+}. Even at the lower Mg^{2+} concentration required for flow stretching, substantial SpoIIIE translocation could in principle be observed in several seconds.

4.3 SpoIIIE activity on stained DNA

After converging on a Mg^{2+} concentration that allowed for both DNA stretching and SpoIIIE activity, the effects of DNA staining on SpoIIIE activity were considered. The translocation activity of SpoIIIE on YOYO-1 stained DNA was evaluated using a series of DNA stained to various extents, ranging from no stain to complete staining, the typical ratio of 1 molecule of YOYO-1 to 4 bp of DNA.[16-18,22] A bulk colorimetric assay was used that measures the amount of phosphate released as SpoIIIE converts ATP to ADP (see Methods). During a time course, phosphate evolution is monitored by absorbance measurements (Figure 4-2).

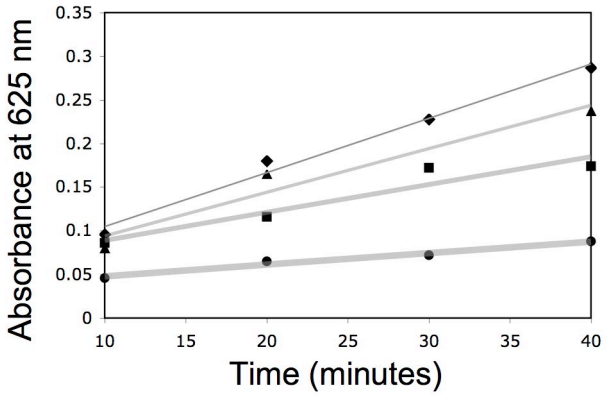


Figure 4-2. The time course of the bulk colorimetric assay shows the evolution of phosphate over time, measured as absorbance at 625 nm due to the addition of a color reagent, for DNA stained with YOYO-1 to different extents. Evolution of phosphate is later converted into ATPase activity (Figure 4-3). Diamonds represent no DNA stain and triangles, squares, and circles represent a normalized stain of 0.02, 0.1, and 1, respectively.

SpoIIIE activity is calculated from the rate of phosphate evolution. Interestingly, SpoIIIE activity decreases dramatically with increased normalized staining, where a normalized staining value of 1 corresponds to complete staining (Figure 4-3). SpoIIIE activity on completely stained DNA drops to 20% of its activity on unstained DNA. This highlights a major challenge of this experimental scheme, where molecules must be visualized at the center of the microfluidic device, roughly 60 um above the surface, thereby requiring complete or near-complete staining of DNA.

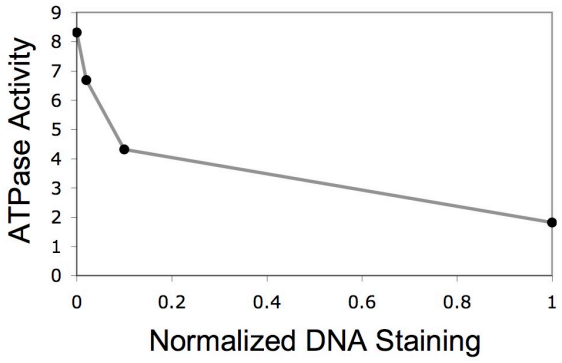


Figure 4-3. The ATPase Activity (mol ATP/sec/mol SpoIIIE monomer) of SpoIIIE was measured as a function of DNA staining. Normalized DNA staining of 1 corresponds to complete staining of the DNA backbone with 1 molecule of YOYO-1 per 4 bp of DNA. The ATPase activity of SpoIIIE declines dramatically with staining.

Other experimental schemes, where DNA is visualized close to the surface, can use total internal reflection fluorescence (TIRF) to visualize DNA with much less staining.[23-25] However, even estimating the reduction in SpoIIIE activity due to the combined effects of staining and low Mg^{2+}, translocation of 1 kb could be observed in a few seconds. Furthermore, bulk concentrations do not completely capture how species in solution interact at the single-molecule level in a prescribed local flow field.

While this bulk assay offers valuable insights, it is important to note that it is not a direct measure of translocation. The bulk assay does not reveal whether a decrease in SpoIIIE activity is due to a lower average processivity, to equally processive but slower movement, to a slower loading rate, or to some combination of these effects. Single-molecule experiments can directly observe translocation and deconvolve these effects.

4.4 Visualizing flow-stretched DNA with lower normalized staining

Since decreasing the normalized staining results in large activity gains (Figure 4-3), the visualization of individual DNA molecules stained to various extents in SpoIIIE buffer with 3 mM Mg^{2+} in the microfluidic cross-slot were evaluated. At typical imaging conditions, DNA molecules at a normalized staining of 3/4 could be visualized. DNA molecules at a normalized staining of 1/2 could be visualized by increasing the illumination intensity. Increases in illumination intensity are not desirable as they can decrease the time to photocleavage. These imaging limitations do not allow for substantial SpoIIIE activity gains by decreasing the normalized DNA staining.

4.5 Directly observing SpoIIIE activity on flow-stretched DNA molecules

Preliminary studies of DNA-SpoIIIE complexes trapped and stretched in the microfluidic cross-slot in SpoIIIE buffer with 3 mM Mg^{2+} reveal interesting behavior not observed with stained DNA alone. DNA-SpoIIIE complexes were created by separately staining λ-DNA and creating SpoIIIE-qdot conjugates (see Methods). The stained λ-DNA was incubated with SpoIIIE-qdot conjugates at room temperature for 5 minutes just prior to visualization to create DNA-SpoIIIE complexes (Figure 4-4).

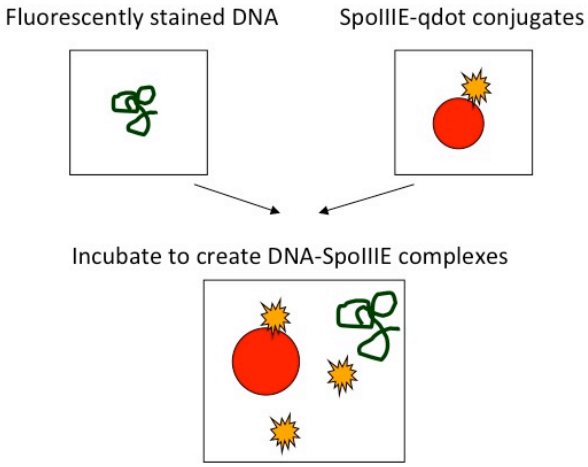


Figure 4-4. Experimental scheme for creating DNA-SpoIIIE complexes for visualization in the microfluidic cross-slot. Biotinylated SpoIIIE and Streptavidin-qdots were first conjugated and then incubated with fluorescently stained DNA to create DNA-SpoIIIE complexes.

The DNA-SpoIIIE complexes are diluted into flow-stretching buffer and introduced into one arm of the cross-slot (4-5). The opposing arm contains only flow-stretching buffer. Flow-stretching buffer consists of 10 mM Tris, 3 mM ATP, 3 mM Mg^{2+}, 100 μg/mL BSA, 59 wt% sucrose, 1 wt% glucose, 20 μg/mL catalase, and 100 μg/mL glucose oxidase at pH 7.4.

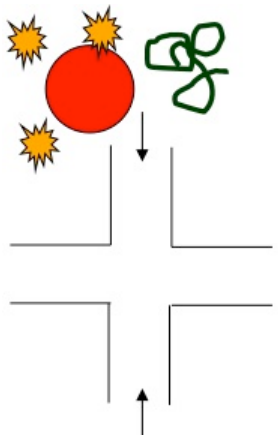


Figure 4-5. Microfluidic cross-slot used for direct observation of DNA-SpoIIIE complexes. The complexes are introduced into one arm and the buffer solutions in either inlet arm can contain varying amounts of the translocase cofactors ATP and Mg^{2+}.

Complexes are trapped and observed at the center of the device and preliminary results show an interesting behavior (Figure 4-6). Time-lapse images of a DNA-SpoIIIE complex in the microfluidic cross-slot in flow-stretching buffer show a bright aggregate moving along the DNA backbone. The bright aggregate, which is non-spherical in shape and larger than a single SpoIIIE-qdot, could be composed of stained DNA as well as several SpoIIIE-qdots. The aggregate is near the center of the molecule at time zero and moves along the DNA molecule in the direction of flow at roughly 3 kb/s until it reaches the end of the molecule and separates. Fluorescent components that are colocalized with but unaffected by DNA, as observed in control experiments, are swept away by the flow much faster, indicating that the aggregate pictured in Figure 4-6 is somehow interacting with the DNA molecule. These preliminary results may not be capturing translocation by an individual SpoIIIE hexamer but are encouraging and future work will modify conditions to reduce aggregation and increase the incidence of observable translocation events.

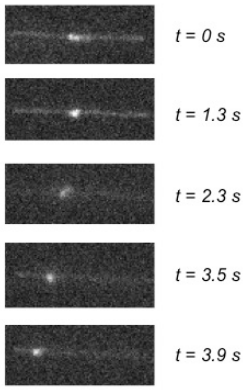


Figure 4-6. Time-lapse images of a DNA-SpoIIIE complex trapped in the microfluidic cross-slot. Image widths are 17 μm.

The cross-slot is a versatile platform and there are other possible experimental setups in this device that could be used to study SpoIIIE translocation besides the approach described above. These include (i) introducing stained DNA in one arm and SpoIIIE-qdot conjugates in the other arm, such that they meet at the center of the device, with 3 mM ATP and 3 mM Mg^{2+} in both arms (ii) introducing DNA-SpoIIIE complexes in one arm in the absence of ATP or Mg^{2+} and introducing cofactors in the opposite arm (iii) introducing stained DNA in one arm and SpoIIIE monomers (without qdots) in the other arm to observe shortening or looping of the DNA. Initial work with these other experimental setups has not been successful but we expect that improved experimental protocols developed with pre-incubated DNA-SpoIIIE complexes that demonstrate some behavior (Figure 4-6) will inform future attempts of these experimental variations.

It should be noted that in this flow-based trapping scheme, SpoIIIE is conjugated to a roughly 20 nm diameter qdot that is subject to drag forces. At the flow velocities and buffer viscosities used here, the drag force on a single qdot is less than 5 pN and drag forces on aggregates would be greater. Opposite halves of the DNA molecule are subject to flow forces in opposite directions in the cross-slot, always away from the center of mass of the molecule. When SpoIIIE moves away from the center of mass of the molecule, the qdot can act as a sail and the drag force is applied in the direction of the motor's progress. When SpoIIIE moves toward the center of mass of the molecule, the drag force on the qdot is applied against the progress of the motor. However, the stall force of SpoIIIE has been measured in single-molecule experiments at over 40 pN[5,26] and greatly exceeds the drag force on a single qdot.

4.6 Conclusions

In conclusion, we have demonstrated that it is possible to reconcile the experimental constraints of a motor protein and a single-molecule trapping platform to simultaneously achieve SpoIIIE activity and the visualization of DNA in flow. Microfluidic stretching of DNA in the cross-slot is effective at Mg^{2+} concentrations of up to 3 mM. In bulk assays, SpoIIIE activity is reduced by 50% at 3 mM Mg^{2+} relative to typically used concentrations of 10 mM Mg^{2+}. The DNA staining required for fluorescence visualization also dramatically reduces SpoIIIE activity, by up to 80% for completely stained DNA. Even though SpoIIIE activity is reduced under the conditions required to reconcile experimental constraints, significant translocation can in principle be observed given the exceptionally high velocity of the motor (> 4 kb/s) at 3 mM ATP and 10 mM Mg^{2+} on unstained DNA.

Preliminary studies of DNA-SpoIIIE complexes in the microfluidic cross-slot are promising and aggregates have been observed moving along DNA at roughly 3 kb/s. Through bulk and single-molecule experiments, we show that direct observation of translocase activity using a fluorescent microfluidic platform is feasible. Future work will focus on varying DNA-SpoIIIE complex preparation conditions to reduce aggregation and increase the incidence of observable translocation events. Other experimental schemes, such as introducing stained DNA and SpoIIIE-qdot conjugates into separate arms of the cross-slot and having them meet at the center of the device, will also be the subject of future study. Furthermore, there is great interest in what happens when SpoIIIE collides with obstacles during translocation.[6] Our previous work on attaching fluorescently-labeled Peptide Nucleic Acids (PNAs) to DNA indicate that PNAs, with a rupture force distribution centered around 60 pN,[16] may serve as model obstacles for SpoIIIE.

4.7 Methods

4.7.1 Fluorescently staining DNA

10 μL aliquots of λ-DNA (48.5 kbp, New England Biolabs) were heated to 63 °C for 1 minute and plunged into wet ice to eliminate concatemers. λ-DNA was stained with YOYO-1 (Invitrogen) at a ratio of 1 dye molecule per 4 bp for a normalized staining of 1. The YOYO-1 was diluted to 10 μM in 10 mM Tris, 10 mM NaCl buffer, pH 8. A 100 μL staining solution, at a normalized staining of 1, consists of 16.64 pM labeled λ-DNA, 0.2 μM YOYO, 10 mM Tris, and 10 mM NaCl. Staining solution was incubated at room temperature for 1 hour, protected from light. Solutions with lower normalized staining were prepared similarly but with lower concentrations of YOYO-1.

4.7.2 Biotinylating SpoIIIE monomers

Biotinylated SpoIIIE monomers were provided by Ninning Liu from the Bustamante Lab at University of California, Berkeley. The SpoIIIE monomer used in these studies only includes the linker and motor domains (portions in gray in Figure 4-1) and the N-terminus of the linker domain is biotinylated. PinPoint^{TM} Xa-1 Vector (Promega) was used to create a 13 kD fusion protein biotinylated *in vivo* by E. Coli. The biotinylated protein is attached to the N-terminus of SpoIIIE monomer and the resulting biotinylated SpoIIIE monomers contain the entire linker and motor domains.

4.7.3 Bulk SpoIIIE colorimetric activity assay

Prior to single-molecule experiments with SpoIIIE, bulk assays were performed to measure the ATPase activity of SpoIIIE under different conditions via a sensitive (nM) colorimetric assay. These experiments were performed with Ninning Liu from the Bustamante Lab at University of California, Berkeley, who provided reagents and protocols. This time-course assay measures the activity of SpoIIIE on DNA by measuring the phosphate released upon conversion of ATP to ADP.[20,21] This assay can be used to test the activity of native SpoIIIE on DNA under different conditions, of modified SpoIIIE on DNA to detect activity loss, and of SpoIIIE on modified or stained DNA. For each reaction sample, a set of time-course tubes is pre-loaded with EDTA. At each time interval, an aliquot from the reaction mixture is added to a tube pre-loaded with EDTA such that the reaction is quenched upon addition to the tube. Colorometric assay reagents are then added and UV-VIS absorption is measured for each time point in each reaction.

At low pH, malachite green forms complexes with phosphomolybdate, shifting the absorption maximum and increasing the molar absorption coefficient.[20] If ATP is converted to ADP, a noticeable color change from dark yellow to green is induced. The reaction can be quenched by adding 34% citrate solution, which inhibits SpoIIIE activity due to its acidity, or by adding excess color reagent (MG/AM). Color reagent is a 3:1 mixture of 0.045% malachite green hydrochloride (MG) and 4.2% ammonium molybdate in 4N HCl (AM). Citrate solution stabilizes the color by desensitizing color reagent to nascent phosphate. It should be noted that an alternative, and more direct, assay of SpoIIIE translocation is a triplex displacement assay. While the colorimetric assay assesses SpoIIIE activity by measuring the conversion of ATP, a triplex displacement assay measures the triplexes displaced by the action of the motor protein.

The colorimetric assay was used to evaluate the activity of SpoIIIE monomers following biotinylation. It was also used to evaluate the activity of SpoIIIE complexes at varying

concentrations of Mg^{2+} (below saturating conditions) and on DNA substrates with varying amounts of fluorescent staining. Samples, with a 250 μL total reaction volume, contained 8 nM SpoIIIE monomer, 0.2 nM DNA, and 3 mM ATP in SpoIIIE buffer. SpoIIIE buffer consists of 50 mM Tris-HCl, 10 mM $MgCl_2$, and 0.1 mg/mL BSA at pH 7.4. When the effects of Mg^{2+} were being studied, a low-Mg^{2+} SpoIIIE buffer, containing 1 to 10 mM $MgCl_2$, was used. The addition of ATP was time zero of the reaction.

For each sample, a 40 minute time-course with 10 minute intervals was tested. Each tube was pre-loaded with 180 μL EDTA at 100mM. At each time interval, 20 μL of the reaction volume was transferred to a tube pre-loaded with EDTA, which chelates the Mg^{2+} and makes this cofactor unavailable to SpoIIIE. At the end of the time course, 700 μL of color reagent and 100 μL citrate solution was added to each tube. The "blank" sample for absorbance measurements consisted of everything present in the samples except for SpoIIIE monomer. Within 30-60 minutes of the end of the time course, absorbance values were recorded at 625 nm. While YOYO-1 is present in stained DNA samples, YOYO-1 at over 5 μM, which is far greater than the concentration used, does not contribute to absorbance at 625 nm. This is consistent with YOYO-1/DNA absorbance spectra (manufacturer's literature); both the absorption and fluorescence emission spectra are negligible at 625 nm.

4.7.4 Creating DNA-SpoIIIE complexes

Creating SpoIIIE-qdot conjugates. SpoIIIE was first incubated with 20 pM Streptavidin-qdots (605 nm emission, Invitrogen) at a ratio of 30,000 SpoIIIE monomers per qdot in SpoIIIE buffer at room temperature for 30 minutes. SpoIIIE buffer consists of 50 mM Tris-HCl, 10 mM $MgCl_2$, and 0.1 mg/mL BSA at pH 7.4. This was diluted by 1000x in SpoIIIE buffer to 20 fM qdots. The manufacturer estimates that there are roughly 5 Streptavidin molecules on the 20 nm diameter qdot surface and a large excess of SpoIIIE monomers was used here as a starting point. The optimal ratio of SpoIIIE monomers to qdots is the subject of future optimization.

Incubating labeled DNA with SpoIIIE-qdot conjugates. SpoIIIE-qdot conjugates were incubated with stained λ-DNA at a ratio of 5 SpoIIIE per DNA molecule in SpoIIIE buffer at a final concentration of 13.3 pM stained DNA, 60 pM SpoIIIE monomer, and 2 fM qdots. The protein concentration was kept intentionally low to limit protein aggregation and make multiple SpoIIIE hexamers acting on a DNA molecule unlikely.

4.7.5 Flow-stretching

Individual DNA-SpoIIIE complexes were trapped and stretched in a microfluidic flow cell with a cross-slot geometry. Such flow cells have been used extensively to study single DNA molecules.[17-19,22,27] The channel cross-section of the PDMS flow cell was 800 μm by 120 μm and the fabrication is described elsewhere.[28] The cross-slot geometry creates a line of pure extension where the two opposing inlet streams meet and a stagnation point, a point of zero velocity, at the center of the device. When a DNA molecule in an inlet stream arrives at the center of the device, its center of mass is trapped at the stagnation point and its opposing ends are stretched apart along the line of pure extension. The inlets and outlets of the flow cell are connected to syringes and waste reservoirs, respectively, by Tygon tubing (0.02" ID, Cole-Parmer). The position of the stagnation point is controlled by manually adjusting the height of one of the exit reservoirs.

Prior to use, the tubing and flow cell were rinsed with Milli-Q filtered (Millipore) water and conditioned to prevent air bubbles and to prevent non-specific adsorption of the protein-coated fluorescent labels. To reduce surface tension and prevent the formation of air bubbles, the tubing and flow cell were incubated with a solution of 0.1 vol% Triton X-100 surfactant (Dow) in a buffer of 10 mM Tris, 10 mM NaCl, pH 8 for 30 minutes. To passivate surfaces and prevent the non-specific adsorption of qdots, the tubing and flow cell were incubated with a 10 mg/mL solution of BSA (Sigma-Aldrich) in a buffer of 10 mM Tris, 10 mM NaCl for 30 minutes.

Flow-stretching buffer (described below) containing DNA-SpoIIIE complexes was fed through one inlet and flow-stretching buffer without DNA was fed through the opposing arm, both using the same PHD 2000 syringe pump (Harvard Apparatus). The pump delivered fluid at 2 ml/hr until the entire volume of the flow cell and tubing was filled with flow-stretching buffer, and then at 100 μL/hr for all experiments. This flow rate results in an average steady-state extension of λ-DNA molecules at the stagnation point of 96% of the DNA molecule's fully extended contour length. The imaging area through the camera was 80.5 μm by 80.5 μm, more than adequate to visualize a stretched λ–DNA molecule with a roughly 20 μm stained contour length.

The DNA-SpoIIIE complexes were diluted to 270 fM (1.2 pM SpoIIIE monomer and 0.04 fM qdots) in flow-stretching buffer with a viscosity of 50 cP. Flow-stretching buffer (pH 7.4) consists of 10 mM Tris, 3 mM ATP, 3 mM Mg^{2+}, 100 μg/mL BSA (Sigma-Aldrich), 59 wt% sucrose (MB grade, Sigma-Aldrich), 1 wt% glucose (Sigma-Aldrich), 20 μg/mL catalase, and 100 μg/mL glucose oxidase (Roche). The sucrose increases the solution viscosity, which increases the relaxation time of the DNA and decreases the flow rate required to achieve a fixed fractional extension of the trapped DNA. The last two components act to inhibit photobleaching and photocleavage. Notably, single-molecule reactions of unlabeled EcoRI restriction enzymes cleaving flow-stretched and stained DNA have been observed.[29]

4.7.6 Data acquisition

The DNA-SpoIIIE complexes appear as the DNA backbone (green) labeled with SpoIIIE-qdot conjuagtes (red) and were visualized using a 100x, 1.4-NA oil-immersion objective on a fluorescence microscope (Leica DMIRE II). Red and green images were captured using a N2.1 Cy3/TRITC (Vashaw Scientific) and 41001 FITC (Chroma) filter set, respectively. A dual-band 51004v2 excitation/emission filter (Chroma) was used to image the red and green channels simultaneously. A cooled CCD camera (Photometrics Cascade 512b) captured images in conjunction with SimplePCI software. Images were captured at roughly 10 fps.

4.8 References

1. Bath, J., Wu, L. J., Errington, J. & Wang, J. C. Role of Bacillus subtilis SpoIIIE in DNA transport across the mother cell-prespore division septum. *Science* **290**, 995-997 (2000).
2. Allemand, J. & Maier, B. Bacterial translocation motors investigated by single molecule techniques. *FEMS Microbiol. Rev.* **33**, 593-610 (2009).
3. Greenleaf, W., Woodside, M. & Block, S. High-Resolution, Single-Molecule Measurements of Biomolecular Motion. *Annu. Rev. Biophys. Biomol. Struct.* **36**, 171-190 (2007).
4. Mehta, A. Single-Molecule Biomechanics with Optical Methods. *Science* **283**, 1689-1695 (1999).
5. Mickler, M., Schleiff, E. & Hugel, T. From biological towards artificial molecular motors. *ChemPhysChem* **9**, 1503-1509 (2008).
6. Marquis, K., Burton, B., Nollmann, M., Ptacin, J., Bustamante, C., Ben-Yehuda, S. & Rudner, D. SpoIIIE strips proteins off the DNA during chromosome translocation. *Genes Dev.* **22**, 1786-1795 (2008).
7. Ptacin, J. L., Nollmann, M., Becker, E. C., Cozzarelli, N. R., Pogliano, K. & Bustamante, C. Sequence-directed DNA export guides chromosome translocation during sporulation in Bacillus subtilis. *Nat. Struct. Mol. Biol.* **15**, 485-493 (2008).
8. Errington, J., Bath, J. & Wu, L. J. DNA transport in bacteria. *Nat. Rev. Mol. Cell Biol.* **2**, 538-544 (2001).
9. Burton, B. M., Marquis, K. A., Sullivan, N. L., Rapoport, T. A. & Rudner, D. Z. The ATPase SpoIIIE transports DNA across fused septal membranes during sporulation in Bacillus subtilis. *Cell* **131**, 1301-1312 (2007).
10. Pease, P. J., Levy, O., Cost, G. J., Gore, J., Ptacin, J. L., Sherratt, D., Bustamante, C. & Cozzarelli, N. R. Sequence-directed DNA translocation by purified FtsK. *Science* **307**, 586-590 (2005).
11. Ptacin, J. L., Nöllmann, M., Bustamante, C. & Cozzarelli, N. R. Identification of the FtsK sequence-recognition domain. *Nat. Struct. Mol. Biol.* **13**, 1023-1025 (2006).
12. Hopfner, K. & Michaelis, J. Mechanisms of nucleic acid translocases: lessons from structural biology and single-molecule biophysics. *Curr. Opin. Struct. Biol.* **17**, 87-95 (2007).
13. Seidel, R. & Dekker, C. Single-molecule studies of nucleic acid motors. *Curr. Opin. Struct. Biol.* **17**, 80-86 (2007).
14. Bigot, S., Saleh, O. A., Cornet, F., Allemand, J. & Barre, F. Oriented loading of FtsK on KOPS. *Nat. Struct. Mol. Biol.* **13**, 1026-1028 (2006).
15. Saleh, O. A., Perals, C., Barre, F. X. & Allemand, J. F. Fast, DNA-sequence independent translocation by FtsK in a single-molecule experiment. *EMBO J.* **23**, 2430-2439 (2004).
16. Zohar, H., Hetherington, C. L., Bustamante, C. & Muller, S. J. Peptide Nucleic Acids as Tools for Single-Molecule Sequence Detection and Manipulation. *Nano Lett.* **10**, 4697-4701 (2010).

17. Perkins, T. T., Smith, D. E. & Chu, S. Single Polymer Dynamics in an Elongational Flow. *Science* **267**, 2016-2021 (1997).
18. Schroeder, C. M., Babcock, H. P., Shaqfeh, E. S. G. & Chu, S. Observation of Polymer Conformation Hysteresis in Extensional Flow. *Science* **301**, 1515-1519 (2003).
19. Tanyeri, M., Johnson-Chavarria, E. & Schroeder, C. Hydrodynamic trap for single particles and cells. *Appl. Phys. Lett.* **96**, 224101-224103 (2010).
20. Hess, H. H. & Derr, J. E. Assay of inorganic and organic phosphorus in the 0.1-5 nanomole range. *Analytical Biochemistry* **63**, 607-613 (1975).
21. Lanzetta, P. A., Alvarez, L. J., Reinach, P. S. & Candia, O. A. An improved assay for nanomole amounts of inorganic phosphate. *Anal. Biochem.* **100**, 95-97 (1979).
22. Smith, D. E. & Chu, S. Response of Flexible Polymers to a Sudden Elongational Flow. *Science* **281**, 1335-1340 (1998).
23. Visnapuu, M. L. & Greene, E. C. Single-molecule imaging of DNA curtains reveals intrinsic energy landscapes for nucleosome deposition. *Nat. Struct. Mol. Biol.* **16**, 1056-1062 (2009).
24. Gorman, J., Fazio, T., Wang, F., Wind, S. & Greene, E. C. Nanofabricated Racks of Aligned and Anchored DNA Substrates for Single-Molecule Imaging. *Langmuir* **26**, 1372-1379 (2009).
25. Finkelstein, I. J., Visnapuu, M. & Greene, E. C. Single-molecule imaging reveals mechanisms of protein disruption by a DNA translocase. *Nature* **468**, 983-987 (2010).
26. Maier, B., Chen, I., Dubnau, D. & Sheetz, M. P. DNA transport into Bacillus subtilis requires proton motive force to generate large molecular forces. *Nat. Struct. Mol. Biol.* **11**, 643-649 (2004).
27. Schroeder, C. M., Shaqfeh, E. S. G. & Chu, S. Effect of Hydrodynamic Interactions on DNA Dynamics in Extensional Flow: Simulation and Single Molecule Experiment. *Macromolecules* **37**, 9242-9256 (2004).
28. Dylla-Spears, R., Townsend, J. E., Jen-Jacobson, L., Sohn, L. L. & Muller, S. J. Single-molecule sequence detection via microfluidic planar extensional flow at a stagnation point. *Lab Chip* **10**, 1543-1549 (2010).
29. Xu, W. & Muller, S. J. Exploring both sequence detection and restriction endonuclease cleavage kinetics by recognition site via single-molecule microfluidic trapping. *Lab Chip* **11**, 435 (2011).

Chapter 5

Conclusions

In this work, several aspects of single-molecule DNA experiments are explored: the labeling of specific DNA sequences, the trapping and stretching of untethered DNA molecules in microfluidic flow, and the observation of DNA-protein interactions. A practical guide to labeling internal dsDNA molecules for single-molecule experiments was developed. Focusing on six diverse approaches, this guide considers the merits and drawbacks of each. By presenting a set of criteria relevant to single-molecule experiments (e.g. labeling yield, compatibility with cofactors such as Mg^{2+}) the guide provides a simple reference for selecting an approach for given experimental constraints.

The PNA labeling approach is studied in detail in two model single-molecule experiments where individual DNA molecules are manipulated via a microfluidic cross-slot and optical tweezers, respectively. By combining microfluidic DNA stretching and a versatile DNA labeling approach, the mapping of genetic features on long (~ 50 kb) DNA molecules is achieved. Collectively, these PNA experiments demonstrate that (i) PNA-DNA binding optimization can be performed simultaneously across all target sites for an entire genomic-length DNA molecule (ii) PNAs can be used to locate specific dsDNA sequences on individual DNA molecules (iii) PNAs can serve as sequence-specific tethers for optical tweezer setups, and (iv) the PNA-DNA bond can sustain forces of roughly 60 pN on average.

Since the microfluidic cross-slot could permit direct observation of the interaction of multiple components, a more dynamic system than labeled DNA was considered. The SpoIIIE motor protein, which translocates rapidly along DNA and can be fluorescently labeled, was proposed as a model system. Several preliminary details had to be addressed since the conditions most conducive to microfluidic DNA trapping and visualization are not the most conducive for SpoIIIE activity and vice versa. SpoIIIE activity requires the presence of Mg^{2+}, a divalent cation that interferes with the stretching of the highly negatively charged DNA molecule. Furthermore, visualization of DNA in flow requires the presence of an intercalating dye, which locally changes the DNA structure and charge. Through bulk and single-molecule experiments with SpoIIIE, we demonstrate that it is possible to reconcile these experimental constraints and simultaneously achieve SpoIIIE activity and the visualization of DNA in flow. Direct observation of ATPase activity using a fluorescent microfluidic platform is feasible. Future work will focus on optimizing experimental conditions to increase the incidence of observable translocation events and exploring other experimental schemes in the cross-slot. The interaction of molecular motors with labeled DNA-bound obstacles could then also be studied. Direct observation of such systems can yield new insights into their mechanisms and function.

Appendix

Lab Protocols

CaBER Rheometer Operating Instructions
Solution Recipes
YOYO-1 iodide Best Practices
Glucose Oxidase and Catalase Best Practices
Stretching and Imaging DNA using Poly-L-Lysine coated slides
S-PNA Bulk Gel Assay
H-PNA Bulk Gel Assay
Selectively labeling the 5' sticky end of λ–DNA with fluorescein
Terminally biotinylating λ-DNA using oligos
Preparation and Visualization of λ-DNA and H-PNA Hybrids

CaBER Rheometer Operating Instructions

Instrument:
Haake CaBER 1, Thermo Electron Corporation
Fahy lab at UCSF

Directions for operating the CaBER rheometer:
(reproduced from those posted next to the rheometer in the Fahy lab)

Operating instructions:

1. Computer username and password
    a. Username: * ask Fahy lab
    b. Password: * ask Fahy lab
2. Open CaBER software (icon on desktop)
    a. Software should open without error
    b. Voltage "warning" message will show up if voltage is <5 volts (4.98-4.99 is fine, if 4.75, please note and consider calling CaBER rep)
3. To make measurements
    a. Go to "configuration" and open the pull-down menu
        i. Click on "Define Geometry" and then click on "Determine System Geometry"
            1. This step enables the instrument to figure out the proximity of the plates
            2. The top plate will lower until it touches the bottom plate and then will be raised to its "rest" position (this is NOT the initial gap position)
            3. The last step of geometry determination is setting the "initial gap." Set the gap according to the initial desired aspect ratio.
            4. If you want to change the final gap (instrument max. of final gap is about 18.63mm), the ONLY way to do so is to lower the bottom plate and re-determine the system geometry.
                a. Lower the bottom plate by adjusting the micrometer, which is accessed by the lower window.
                b. "Determine the System Geometry" again and set the initial gap.
                c. Check what the final gap is by looking in the "Define Geometry" options.
                d. Repeat steps a-c if the final gap is not the desired aspect ratio
    b. Go to "Measurement" and open the pull-down menu
        i. Click on "Define single measurement options" and make sure "high-speed measurement" is selected.

ii. Consider selecting "reduce data size" although this is not essential (difference is 2kB vs. 20-50kB)
iii. Sample rate = 1000 Hz (usually does not change)
iv. Sample duration = 10s (can change if time to break-up is unusually long)

c. Go to "Measurement" and open the pull-down menu
   i. Click on "Run Single Measurement"
   ii. The top plate will lower until it reaches the initial gap.
   iii. You will receive a prompt to name the sample run (everything you put on this sheet will end up on the final excel data file)
   iv. You will receive a prompt to load the sample and then to click "ok" to begin the run.
      1. To load the sample, move the grey side panels away from the loading dock, to maximize access to the plates.
      2. Use small diameter pipette or needle with syringe, and with bevel up, slowly discharge the sample so that it loads from the "top down" (for low-viscosity solutions, this enables the sample to wick onto the top plate without falling off the bottom plate).
      3. Make sure there is a nice meniscus of sample prior to starting the run.
d. Move the grey panels back prior to starting the run (laser beam is contained within the panels)
e. Click "ok" which will start the run
f. When finished, you will receive a prompt to save the data

If using needles with syringes, use the following protocol:

1. NEVER leave an open needle lying on a benchtop
2. NEVER recap with two hands (with one hand, slide needle back in the cap and then lock into place after the needle is covered)
3. Discard needle/syringe into sharps container (NEVER try to jam the needle into full sharps container – ask for new container if full)

Some additional notes (Courtesy of Dr. Cari S. Dutcher):

1. Turn on computer and power supply (located to the top left of the CaBER)
2. There are 4mm and 6mm plate fixtures, the plates not in use are in the drawer to the bottom left of the instrument (there is also a manual in this drawer). These screw in, to unscrew the top one, you need to hold the attached rod in place to keep it from turning with the plate.
3. You want results that are independent of aspect ratio

4. For 6mm diameter plates, starting point for the initial gap height could be 2mm and a starting point for the final gap height could be between 7-10mm.
5. The temperature can be controlled via a water bath located beneath the CaBER
6. The software has a temperature sensor, but tends to run high
7. The grey panels should only be back when loading/cleaning the plates
8. The strike time can vary from 20-50ms, it may not actually achieve 20ms, and will actually be ~23ms. Use 50ms normally.
9. You can load the sample before compression (using just a syringe) or after compression into the initial gap (using a syringe and needle). Instructions for the latter are below. If using just a syringe – first load a drop, large enough to create a hemisphere and let the plate lower. Then slowly touch small drops to the edge of the fluid until there is a concave meniscus.
10. Wait for several polymer relaxation times before taking measurements
11. The initial aspect ratio (AR_i) is defined as the initial gap height/fixture radius. You should be at $0.5<AR_i<1$. The final aspect ratio is defined as the final gap height/ fixture radius. You should vary this until consistent results are obtained.

To perform data analysis after completing the above steps:

1. Open Analysis software (icon on desktop above CaBER icon)
    a. You can view your data, overlay runs, and change axes from linear to log
    b. Go to File and open the pull-down menu, select "Export the data to excel"
    c. The software creates a multi-tab excel spreadsheet with all of the data on the last tab.
    d. Analyze data after time = 0.
    e. Make a new column of normalized diameters, D(t)/D(t=0)
    f. Fit the data to a constitutive model to extract relaxation times

Solution Recipes

65 wt% MB-grade sucrose 1xTE10 pH8, 1wt% glucose, 0.2 μm filtered
150 mL, store at RT
(Ref: HZ NB 2 pg. 16-17)

Recipe:
126.75g Molecular Biology (MB) grade sucrose
59.18g nanopure water
1.95g MB grade glucose
7.5mL 20xTE buffer pH 8 (Boston BioProducts #BM-304)
300μL of 5M NaCl

Can make this in a used 250mL Nalgene suction filter receiving bottle

Add 126.75g MB-grade sucrose
add water up to 165.5 g
add glucose up to 167.45g
add 7.5 mL 20xTE buffer pH8
add 300 μL of 5M NaCl
add water until total weight (~20.43 g more) : 195.25

The sucrose needs some coaxing to get into solution:
Cover the top with parafilm to prevent evaporation/contamination
Sonicate for 30 min.
Then add a clean stir bar and suspend over a hot plate with low heat in a shallow water bath until sucrose is in solution.

Use Nalgene suction filter (Nalgene cat #: 568-0200) and new labeled receiving bottle.

---

62 wt% MB-grade sucrose 1xTE10 pH8, 1wt% glucose, 0.2 μm filtered
45mL, store at RT
(Ref: HZ NB 2 pg. 23, 163)

Recipe:
43mL of *65 wt% MB-grade sucrose 1xTE10 pH8, 1wt% glucose, 0.2 μm filtered*
2mL of 1xTE10 *pH8, 1wt% glucose, 0.2 μm filtered (NB 2 pg.58, 04/04/08)*

1xTE10 pH8 buffer
48mL, store at RT
(Ref: HZ NB 2 pg. 52 03/19/08)

Recipe:
2.4mL of 20xTE pH8
96μL of 5M NaCl
balance to 48mL ultrapure water
filter with 10mL syringe and 0.2μm PVDF syringe filter

"TE10g" 1xTE10 pH8, 1wt% glucose (0.2 μm filtered)
45mL, store at RT
(Ref: HZ NB 2 Pg. 3, 58)

Recipe:
0.45g MB grade glucose
90μL of 5M NaCl
2.25mL 20xTE pH8
balance to 45mL ultrapure water
filter with 10mL syringe and 0.2μm PVDF syringe filter

0.1% Triton surfactant, 0.2 μm filtered in TE10
45mL, store at RT
(Ref: HZ NB 2 pg. 24, 164)

Recipe:
2.25mL 20xTE pH8
90μL of 5M NaCl
45μL Triton *
balance to 45mL ultrapure water
filter with 10mL syringe and 0.2μm PVDF syringe filter

*Note: difficult to expel from pipette tip, helps to use 200G tip and eject Triton onto (not into) water

---

1% Triton, 0.2 μm filtered
10mL, store at RT

Recipe:
Add <10mL ultra pure water
100μL Triton
fill to 10mL with ultra pure water

Use Nalgene suction filter (Nalgene cat #: 568-0200) and new labeled receiving bottle.

---

10 mg/mL BSA solution "100xBSA" (*0.2 μm filtered initially AND prior to every use)*
45mL, store in fridge for 1 month, freezer longer term
(Ref: HZ NB 1 pg. 202, NB 2 pg. 23, 50)

Recipe (and order):
Add <45mL ultrapure water
450μL of 5M NaCl
450μL of 1M Tris-HCl pH 7.4
9μL of 0.5M EDTA
mix (perhaps check pH)
2.25mL MB grade glycerol
balance to 45mL with ultrapure water
vortex mix (perhaps check pH again)
450mg BSA powder (Sigma A3059)
filter with 10mL syringe and 0.2μm PVDF syringe filter or Nalgene suction filter (Nalgene cat #: 568-0200)

Make 4 aliquots in 20mL scintillation vials. Keep one in fridge and the rest in the freezer. Record preparation date and leave a space to record thaw date. ~1 month in fridge is ok.

5μM YOYO aliquots (see YOYO best practices for more detail)
200μL, store covered with aluminum foil in freezer
(Ref: HZ NB 1 pg. 200, NB 2 pg. 34, 50, 112)

Recipe:
199μl *1XTE10 buffer pH 8*
1μl 1mM stock YOYO

make 4, 50μl aliquots to avoid many freeze-thaw cycles per aliquot

---

"Q-dot buffer" (generally good buffer for all small beads):
10mL, store in fridge
(Ref: HZ NB 2 pg. 128)

Recipe:
1.25mL of 400mM Tris/HCl pH=7.5
200μL of 5M NaCl
0.5mL glycerol
20μL of 0.5M EDTA
100μL of 1% Triton X-100
100μL of 10 mg/mL BSA
fill to 10mL with ultra pure water.

Concentrations:
50mM Tris pH 7.5
100mM NaCl
5% glycerol
1mM EDTA
0.01% Triton x-100
100 μg/mL BSA

0.1% N-Av beads 250 μL
250 μL, store covered in fridge

Recipe:
25μL of 1% 40nM N-Av bead stock (Invitrogen cat # F8770)
(invert stock before use to resuspend)
7.5μL of 5M NaCl
1.25μL of 1% Triton x-100
2.5μL of 10 mg/mL BSA
12.5μL glycerol
201μL TE50
Total: 250μL

Concentrations:
8 mM Tris
190 mM NaCl
5% glycerol
0.8 mM EDTA
0.005% Triton x-100
100 μg/mL BSA

YOYO-1 iodide Best Practices

1. General product information
2. Preparing aliquots
3. DNA staining

1. General product information:

YOYO®-1 iodide (491/509) - 1 mM solution in DMSO
Molecular Probes (Invitrogen) Catalog Number - Y3601, 200 μL
Dye should be stored at -20 °C and protected from light.

2. Preparing aliquots:

Mixing in microcentrifuge tubes (e.g.: 0.6 mL, 2 mL) should always be followed by a quick spin down.

This procedure uses ~2.5% of the stock. Keep dye covered when not in use. Upon completion, promptly return dye to -20 °C.

Make 5, 1 μL aliquots of stock YOYO in 0.6mL microcentrifuge tubes.

Each week, or whenever necessary, dilute one aliquot from 1mM to 5 μM YOYO:
Add 199 μL of a buffered solution such as TE 10 (TE buffer with 10mM salt).

Alternatively, you can dilute to 25 μM YOYO for more concentrated DNA solutions.
You can make larger aliquots if you are using more on a weekly basis.
Dye should be stored at -20 °C and protected from light.

This method will keep YOYO in its preservative solution longer. Each diluted aliquot will also endure fewer freeze-thaw cycles and less exposure to light.

3. DNA staining:

Note that when pipetting DNA, use a wide-mouth (or genomic) pipette tip with a diameter greater than 1mm and pipette slowly to minimize shearing of DNA at the tip. In the staining solution, the ratio of dye molecules to nucleic acid base pairs should be near 1:4. Allow to sit covered for at least 1 hr at room temperature.

Glucose Oxidase and Catalase Best Practices

1. General Product Information
2. Preparing Solutions
3. Viewing
4. Notes from the literature

1. Product information

Available from Sigma-Aldrich

Glucose Oxidase Type X-S, # G7141-50KU
Store at -20°C dessicated, avoid prolonged times at higher temperatures
Catalase from Bovine Liver, # C40-100mg
Store at -20°C dessicated, avoid prolonged times at higher temperatures
Glucose Sigma Ultra, # G7528-250g (substrate for enzymes)
Sucrose, molecular biology grade, # S0389-1kg (as viscosifying agent)

Glucose oxidase and catalase are enzymes, supplied as powders, which create a reducing environment in solution, thereby slowing down the photobleaching of stained DNA. Together with glucose, a necessary substrate for glucose oxidase, they form an oxygen scavenging system.

2. Preparing solutions

Warm the bottles of enzymes to room temperature before opening.

Mixing in microcentrifuge tubes should always be followed by a quick spin down.

Solutions of glucose oxidase and catalase are prepared at 1000 times the concentration used in visualization (1000X) so that 1 μL is added to ~ 1mL of the flow visualization buffer.

1000X viewing concentration corresponds to 50 mg/mL glucose oxidase and 10 mg/mL catalase. Both enzymes are dissolved in a 50%/50% (v/v) solution of TE10 (1XTE buffer with 10mM NaCl) and glycerol. Do not sonicate or vortex mix solutions containing enzymes.

<u>50 mg/mL Glucose Oxidase (1000X):</u>
Dissolve 50 mg solid catalase into 500 μL TE10 buffer
Mix by inversion of tube, should dissolve easily
Add 500 μL glycerol
Mix by inversion, then by pipette, drawing ~500 μL from the bottom of the tube and depositing at the top of the solution until no refractive index mismatch is observed.
Store at -20°C.

10 mg/mL Catalase (1000X):
Dissolve 10 mg solid catalase into 500 μL TE10 buffer
Mix by inversion of tube, should dissolve easily
Add 500 μL glycerol
Mix by inversion, then by pipette, drawing ~500 μL from the bottom of the tube and depositing at the top of the solution until no refractive index mismatch is observed.
Store at -20°C.

At these concentrations, the glucose oxidase solution will be bright yellow and the catalase solution will be dark brown.

Make a 20 μL aliquot and replace with a fresh aliquot as necessary. Store at -20°C. Avoid prolonged times at higher temperatures. Work with aliquots on ice.

3. Viewing

Viewing concentrations of 50 μg/mL glucose oxidase and 10 μg/mL catalase work well. This corresponds to 1X of the recipes above. Premix 1 μL each of 1000X glucose oxidase and catalase solutions per 1 mL of your visualization buffer prior to adding the stained DNA solution. The visualization buffer must contain some glucose, 1 wt% is sufficient. Mix by inversion and pipette and wait a few minutes to create a reducing environment before adding stained DNA. Do not sonicate or vortex mix solutions containing enzymes.

4. Notes from the literature

Schroeder, C. M. Investigating Polymer Physics with Single Molecule Experiment and Brownian Dynamics Simulation. Stanford University Dissertation (2004).

Photobleaching of the dye was substantially slowed by a factor of ~ 50, depending on the choice of dye, by adding enzymatic oxygen scavenging reagents consisting of 0.3%(v/v) glucose, glucose oxidase (0.05 μg/μL), and catalase (0.01μg/μL). Addition of 1%(v/v) β -mercaptoethanol also helped to prevent photobleaching of the dye by removing oxygen radicals from solution.

Schroeder, C. M., Teixeira, R. E., Shaqfeh, E. S. G. & Chu, S. Characteristic Periodic Motion of Polymers in Shear Flow. Physical Review Letters 95, 018301 (2005).

DNA was stained with YOYO-1 fluorescent dye (Molecular Probes) at a ratio of 1 dye per 4 base pairs. Solution buffers contained 10 mM tris/tris-HCl (pH 8.0), 10 mM NaCl, and 1 mM EDTA. Photobleaching of the dye was substantially slowed by addition of 0.3% (v/v) glucose, glucose oxidase (0.05 μg/μL), catalase (0.01 μg/μL), and 1%(v/v) β – mercaptoethanol. Experiments were conducted at 17 °C.

Stretching and Imaging DNA using Poly-L-Lysine coated slides

Materials:

Glass slide (Fisher 12-549-3)
Glass coverslip (Fisher 12-541-B)
0.1 mg/mL poly-L-lysine solution in ultrapure water

Protocol:

Pipette 5 μL of 0.1 mg/mL poly-L-lysine solution onto coverslip
Let set for ~1 min
Sandwich with second coverslip for ~30s
Separate coverslips and flip over to allow both surfaces to dry completely (~20 minutes)
Pipette 5 μL of stained DNA sample onto coated surface of coverslip
Let sit for ~30s (covered, to protect stained DNA from light)
Drop onto glass slide, coated surface down

Troubleshooting if DNA sample appears dim:

Sample:
Make a new sample using a fresh aliquot of YOYO-1 and DNA

Microscope – check:
UV shutter (pull out bar)
Beamsplitter switch eyepiece/photo/TV exit (pull out bar)
Brightness adjustment on objective lens
Field diaphragm adjustment and aperture diaphragm adjustment
Centering buttons for incident light field diaphragm
Software: 2 gains and contrast adjustment
Bulb: consult the logbook, bulb has 2000 hr lifetime.

S-PNA Bulk Gel Assay

Objective: The extent of PNA-DNA binding is determined in a bulk gel-shift assay for fragments containing target sites and single-end mismatch (SEMM) sites. The gels are run at conditions such that (1) the smallest fragment is greatly separated from the larger fragments generated during restriction and (2) there is a noticeable separation between this smallest fragment with and without a PNA bound.

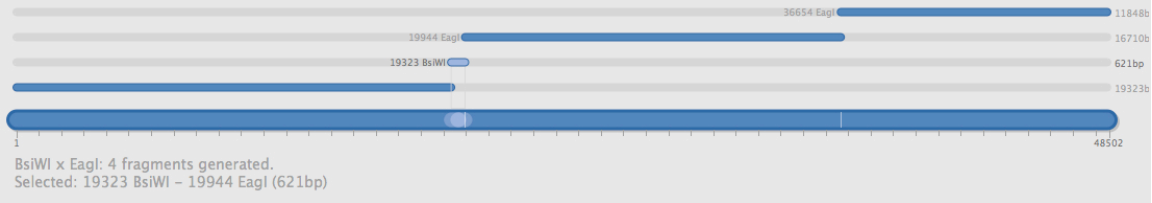
Reference: For complete protocol details, see
Zohar, H., Hetherington, C. L., Bustamante, C. & Muller, S. J. Peptide Nucleic Acids as Tools for Single-Molecule Sequence Detection and Manipulation. *Nano Letters* **10**, 4697-4701 (2010).

Maps of λ-DNA digestions (created with EnzymeX software):

S-PNA target-containing fragment (499 bp)

Target site sequence: AAGAGAAA

S-PNA SEMM-containing fragment (621 bp)



SEMM site sequence: *C*AGAGAAA

Restriction enzymes from NEB:

SalI (20,000 U/mL)
BsiWI (10,000 U/mL)
EagI (10,000 U/mL)

Control Digestion (S-PNA target site)

35.5 µL up water
5 µL 10x NEBuffer 3

0.5 μL 100x BSA
7 μL stock λ-DNA
1 μL SalI
~50 μL , 37°C 1 hr.

Control Digestion (S-PNA SEMM site)

37 μL up water
5 μL 10x NEBuffer3
7 μL stock λ-DNA
1 μL EagI
~50 μL , 37°C 1 hr.
then add
1 μL BsiWI
~50 μL , 55°C 1 hr.

S-PNA hybridization (100 μL)

69.7 μL up water
5 μL 20x TE
4 μL 50 mM NaCl
16.35 μL stock λ-DNA*
5 μL S-PNA at 50 μM**
~100 μL , 37°C 4 hr.

*before adding, heat to 65°C for 1 min, then put on ice
**before adding, heat to 50°C for 5 min

To digest S-PNA target site, add to a 50 μL hybridization:

6 μL NEBuffer 3
0.6 μL 100x BSA
1 μL SalI
~58 μL , 37°C 1 hr.

To digest S-PNA SEMM site, add to a 50 μL hybridization:

6 μL NEBuffer 3
1 μL EagI
~57 μL , 37°C 1 hr.
then add
1 μL BsiWI
~57 μL , 55°C 1 hr.

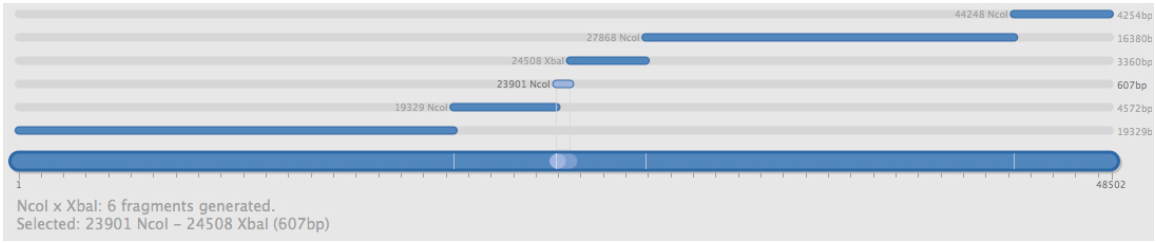
H-PNA Bulk Gel Assay

Objective: The extent of PNA-DNA binding is determined in a bulk gel-shift assay for fragments containing target sites and single-end mismatch (SEMM) sites. The gels are run at conditions such that (1) the smallest fragment is greatly separated from the larger fragments generated during restriction and (2) there is a noticeable separation between this smallest fragment with and without a PNA bound.

Reference: For complete protocol details, see
Zohar, H., Hetherington, C. L., Bustamante, C. & Muller, S. J. Peptide Nucleic Acids as Tools for Single-Molecule Sequence Detection and Manipulation. *Nano Letters* **10**, 4697-4701 (2010).

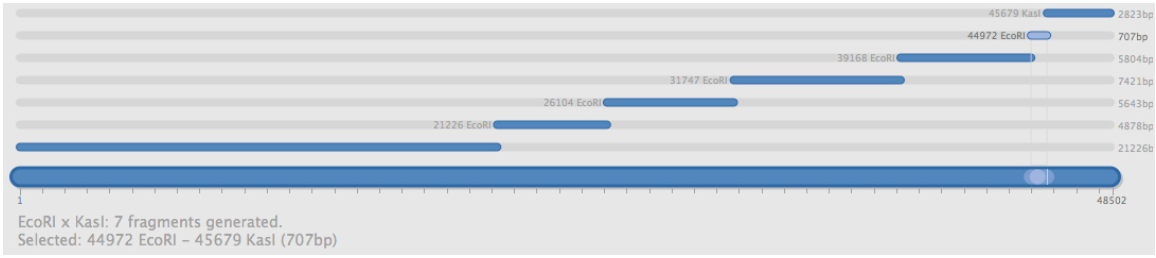
Maps of λ-DNA digestions (created with EnzymeX software):

H-PNA target-containing fragment (607 bp)



Target site sequence: GAGAAGGA

H-PNA SEMM-containing fragment (707 bp)



SEMM site sequence: GAGAAGG*T*

Restriction enzymes from NEB:

XbaI (20,000 U/mL)
NcoI (10,000 U/mL)
EcoRI (100,000 U/mL) – dilute to 10,000 U/mL in buffer
KasI (4,000 U/mL)

Control Digestion (H-PNA target site)

35.5 μL ultrapure water
5 μL 10x NEBuffer 4
0.5 μL 100x BSA
7 μL stock λ-DNA
1 μL XbaI
1 μL NcoI
~50 μL , 37°C 1 hr.

Control Digestion (H-PNA SEMM site)

35.5 μL up water
5 μL 10x NEBuffer 4
0.5 μL 100x BSA
7 μL stock λ-DNA
1 μL EcoRI
1 μL KasI
~50 μL , 37°C 1 hr.

H-PNA hybridization (100 μL)

69.7 μL up water
5 μL 20x TE
4 μL 50 mM NaCl
16.35 μL stock λ-DNA*
5 μL H-PNA at 50 μM**
~100 μL , 37°C 16 hr.

*before adding, heat to 65°C for 1 min, then put on ice
**before adding, heat to 50°C for 5 min

*To digest H-PNA target site, add to a 50 μL hybridization***:**

6 μl 10x NEBuffer 4
0.6 μl 100x BSA
1 μl XbaI @ 20,000 units/ml (> 0.1 μl)
1 μl NcoI @ 10,000 units/ml (> 0.2 μl)
~59 μl , 37°C 1 hr.

To digest H-PNA SEMM site, add to a 50 µL hybridization:

6 µl 10x NEBuffer 4
0.6 µl 100x BSA
1 µl EcoRI @ 10,000 units/ml (> 0.2 µl)
1 µl KasI @ 4,000 units/ml (> 0.5 µl)
~59 µl , 37°C 1 hr

Selectively labeling the 5' sticky end of λ–DNA with fluorescein
(can be used with other modified nucleotides)

Overview: This procedure is used to selectively label the 5' sticky end of λ-DNA with a fluorescein so that other labels on the molecule can be assigned an unambiguous position along the molecule. Once made, this labeled DNA can be frozen and used as the starting DNA for downstream labeling procedures. When incubated with anti-fluorescein quantum dots, the 5' end of the DNA molecule will fluorescently labeled.

Reagents:

dATP (Roche cat. # 11934511001)
dGTP (Roche cat. # 11934538001)
fluorescein-dUTP (Roche 11373242910)
Klenow Exo- (Fermentas EP0421)
655 qdots coated with anti-fluorescein (Invitrogen Q15421MP)

Outline of procedure:

1. Incubate λ–DNA with dATP, dGTP, fl-dUTP and Klenow Exo- (protocol 1)
2. Separate out excess nucleotides via dialysis (protocol 2)
3. Measure new DNA concentration
4. Use fl-labeled DNA as starting DNA in subsequent protocols
5. Incubate fl-lableled-DNA with anti-fluorescein qdots

Protocol 1: Incubating DNA with nucleotides

8 μL stock λ-DNA (4 μg)
2 μL 10x reaction buffer for Klenow fragment
5 μL 1mM fl-dUTP
1 μL 5mM dATP
1 μL 5mM dGTP
0.8 μL Klenow exo- at 5 units/μL (4 u, 1 u / μg DNA)
2.2 μL nuclease-free water
20 μL, Incubate at 37 °C for 15 min.

Protocol 2: Drop Dialysis

1. Dilute to sample from **Protocol 1** to 100 μL by adding 80 μL of TE10

2. Dialyze in a Petri dish over 30mL of TE10 for 3 hrs.

3. Use 100 nm nitrocellulose drop dialysis membrane (Millipore)

4. Gently stir (setting 2-3) with a small stir bar and plate, covered at room temperature

5. You should notice the color of the drop change from yellow to clear

6. Recover the sample and store in the freezer (should keep for many months)

Yield:

After incubating labeled DNA from the above procedure with qdots, this protocol produced between 10-20% yield of molecules with a qdot attached to a single end, now known to be the 5' end. Qdots were incubated in a 200 μL reaction with a 0.27nM DNA concentration and at a 2:1 q-dot:DNA ratio for 1 hr at room temperature with gentle mixing via rotation. (see HZ NB 4, pg. 107, 126, 127)

Terminally biotinylating λ-DNA using oligos

Overview

We hybridize and ligate a biotinylated oligo to each sticky end of λ-DNA sequentially, covalently biotinylating both ends of the molecule. This protocol can be adapted to biotinylate only a single end.

Reference

"Hydrodynamic Flow-stretching Assay for Single-Molecule Studies of Nucleic Acid–Protein Interactions" by Schroeder et al. in Single-Molecule Techniques – A Laboratory Manual pg. 482.

Reagents

DNA oligonucleotides (from Operon.com, those available in lab are listed)
- Oligo 1 – "terminal oligo" MW = 10521.21 g/mol:
  5'-GGGCGGCGACCTAAAAAAAAAAAAAAAAAAA-biotin-3'
- Oligo 2 – "sticky end oligo" MW = 4176.95 g/mol:
  5'-AGGTCGCCGCCC-biotin-3'
- We have Oligo 1 at 111 μM in water in -20 °C.
- We also have left over phosphorylated Oligo 1 at 11 μM from step (1)

λ DNA (~0.5 μg/ml; 15.9 nM) (New England Biolabs)
- Stored in 100 μL aliquots in -20 °C.

TE10 buffer (1XTE buffer [1mM EDTA, 10mM Tris] with 10mM NaCl at pH 8)

T4 Polynucleotide Kinase (PNK) (New England Biolabs #M0201S)
- Stored at -20 °C, should not be above -20 °C for long.

T4 ligase (supplied with ligase buffer) (New England Biolabs # M0202S)
- Stored at -20 °C.

Ultrapure (up) water

Ice (from ice machine down the hall)

GE S-400 Spin Columns

Note that Oligo 1 has a 20 A-residue spacer while Oligo 2 does not. These oligos are complimentary to the opposing sticky ends of λ DNA and, hence, to one another. They should not be incubated together because they will hybridize with one another more rapidly than with the less accessible DNA ends.

These oligos are not supplied with 5' phosphate groups so a phosphorylization step precedes the ligation step. Oligo 1 is HPLC purified and Oligo 2 is only desalted. They arrived lyophilized and were solvated, divided into aliquots, and relyophilized. These portions are resuspended to 10 μM oligo in an appropriate buffer of choice (such as 1XTE) and can be subsequently diluted. The success of the resuspension and actual

concentration can be measured with the nanodrop. Oligos can be briefly vortex mixed in solution. They are stored at -20 °C.

Equipment

PCR Thermocycler
Block heater (VWR)
Vortex mixer
Minicentrifuge for quick spin downs
0.6 mL Microcentrifuge tubes
Black Styrofoam cooler
Small beaker to fill with ice
Microcentrifuge
Ice bucket

Methods

Buffers used should be purified with 0.2 μm filter before use and mixed just prior to use. Carry out all steps on ice or at 4°C unless otherwise specified. To avoid shearing λ-DNA, use wide-tip pipette tip (1 mm or greater inner diameter) and pipette slowly when handling λ-DNA. Prepare tip using an ethanol-sterilized blade or purchase genomic tips (e.g. 200G from Molecular Bioproducts). Note that λ-DNA will shear more easily in viscous solutions. Unaltered pipette tips should be used to pipette other solutions for accuracy. Never vortex or sonicate solutions of λ-DNA. Subsequent to any type of mixing in microcentrifuge tubes or preceding any reaction, perform a quick spin down in the microcentrifuge.

Prepare ice bucket.

1. Phosphorylation

This step phosphorylates the oligos as the 5' phosphate is necessary for the subsequent ligation step.

15.5 μL of ultrapure water
2 μL of 10x ligase buffer
2 μL of 111 μM Oligo 1
0.5 μL of T4 PNK

To 15.5 μL of ultrapure water in a 0.6mL microcentrifuge tube, add 2 μL of 10x ligase buffer and 2 μL of 111 μM Oligo 1. Mix by flicking tube and by pipette. Add 0.5 μL of T4 PNK. Mix by flicking tube and by pipette and spin down. Place at 37 °C for 1-4 hours. You can use the dedicated oven, the block heater, or the PCR Thermacycler program PHOSPH. During this cycle, mix the tube by flicking, followed by a quick spin down. This is a 20 μL reaction, which diluted the oligos by a factor of 10. Store at -20°C.

Don't put away the ligase buffer, you will need it for the next step.

2. Annealing

Important: Set block heater to 70 °C 1 hr. prior to starting this step. In a hurry, you can set the block heater to 80 °C + for a faster heating rate but carefully monitor the block to avoid overshoot as cooling is slow.

In this step, λ-DNA is heated and slowly cooled in the presence of excess oligo, allowing the oligos to anneal to the terminal sticky ends. The oligos are attached to the λ DNA via hydrogen bonding, not a covalent linkage along the backbone.

50 μL of stock λ-DNA (~0.5 μg/ml; 15.9 nM) (*temperature treated as described below)
5.5 μL 10x T4 DNA ligase buffer
0.5-1 μL phosphorylated Oligo 1 (11 μM oligo)

* Heat stock DNA aliquot to 65 °C for 5-10 minutes and plunge into wet ice. The block heater will take a while to reach this temperature. The PCR Thermacycler has a program, 65HOLD, which will reach this temperature in seconds. Do this to DNA any time it is thawed from -20 °C. This breaks up concatamers of λ-DNA. If the PCR Thermacycler is not available, this is also a good time to turn on the recirculating water bath by the rheometer and set it to 16 °C for ligation.

Thaw and thoroughly vortex mix 10XT4 DNA ligase buffer. Using a wide pipette tip, add (step 1) 50 μL of temperature treated stock λ-DNA (~0.5 mg/ml; 15.9 nM) to a 0.6mL microcentrifuge tube. (Step 2) Add 5.5 μL T4 DNA ligase buffer and (step 3) 0.5 μL phosphorylated Oligo 1 (11 μM oligo). Flick to mix. This is a 56 μL reaction with [Oligo1]=100nM and [λ DNA]=14nM, so there are 7 oligos: 1 λ DNA. A ratio of 15 is good, too.

Heat to 70 °C for 2-3 minutes, cover tube with wet paper towel to facilitate heat transfer to the top of the tube. Turn off heating block and leave tube to slowly cool in block. When T = 40 °C, remove block from heater and set on counter to facilitate cooling. Start ligation when T < 30 °C. This takes about 1.5 hours. Don't put away ligase buffer, will need it for the next step.

3. Ligation

In this step, the ligase facilitates the formation of a covalent bond between the oligo and the λ-DNA backbone.

15 μL 10X T4 DNA ligase buffer
127 μL ultrapure water
2 μL T4 DNA ligase

Add to the above annealing reaction 15 μL 10X T4 DNA ligase buffer, 127 μL ultrapure water. Mix by flicking tube, spin down, then add 2 μL T4 DNA ligase. Mix slowly using wide pipette tip and spin down. Store at 16 °C in recirculating water bath or in PCR thermocycler using the PHOSPH program for 2-4 hours. Alternatively, store at 4 °C overnight. This is a 200 μL reaction with [Oligo1]=27.5nM before hybridization and [λ-DNA construct]=4nM, assuming 100% efficiency. Store constructs as 10, 20 μL aliquots at 4 °C.

4. Spin out excess oligo

In this step, we use a microcentrifuge column (GE S-400, kept at 4 °C) to remove excess (unbound) oligos. This allows the beads to bind more efficiently to constructs.

-Prepare column by spinning at 800g for 60 seconds
-Prepare receiving tube with
    1. 10 μL of 10x ligase buffer if you plan to ligate a second oligo
    OR
    2. 0.5 μL 5M NaCl if this is your last spin
(you need to somehow replenish salts or other small molecules important to the solution that were separated out by the column)
-Add 100 μL of step 3- "ligation" product to top of column, aim for center of column
-Return column so it is not disturbed (i.e. in the same orientation as before)
-Spin into prepared receiving tube at 800g for 2 min
-Right after: gently mix to achieve desired salt concentration

(You will lose ~25% of your DNA in this spin, too)

5. Anneal oligo 2

Set heating block to 70 °C 1 hr. in advance

80 μL of spun lambda-oligo 1 construct ([DNA]=~3nM)
3-6 μL of phosph. oligo 2 (10μM, phosphorylated in the same way as oligo 1 in step 1)
gently mix and spin down

Heat to 70 °C for 3 min. and allow to cool in block heater covered with wet paper towel

6. Ligate oligo 2

Add to the above
15 μL of 10x ligase buffer
138 μL of ultra pure water
2 μL ligase

genty mix and spin down
Total : 237 μL ([DNA]=~1nM)

Mix slowly using wide pipette tip and spin down. Store at 16 °C in recirculating water bath or in PCR thermocycler using the PHOSPH program for 2-4 hours.

6.5 Inactivate Ligase

If this is your last ligation step and your construct is complete: heat the sample to 65°C for 10 minutes to **inactivate the ligase** enzyme

7. Spin out excess oligo 2

Repeat step 4, using 50 μL of product from step 6- "Ligate oligo 2" and option 2 for receiving tube
([DNA]=~0.75nM)

8. Incubate lambda-oligo1,2 construct with N-Av beads
(Ref: HZ NB 2 pg. 137)

86 μL of "Q-dot buffer"
14 μL of product from step 7- "Spin out excess oligo 2"
- mix
Total: 100 μL
Add: 4 μL 0.1 % 40nm N-Av beads
([DNA]=~0.1nM, beads: 0.004%)

9. Stain incubated construct

220 μL TE10
25 μL of product from step 8 –"Incubate…" ([DNA]=~0.1nM, beads: 0.004%)
10 μL of 5μM YOYO
Total: 255 μL
-mix
-incubate 1 hr. in the dark at room temperature
([DNA] = ~10 pM, beads: 4e-4%)

10. Viewing solution for slide

970 μL of buffer: 10mM tris/HCl (pH 7.5), 10 mM NaCl, 1 mM EDTA, 0.5% glucose
10 μL of 10 mg/mL BSA
10 μL of betamercaptoethanol
1 μL of 1000x glucose oxidase
1 μL of 1000x catalase

- mix to provide reducing environment
15 μL of product from step 9 –"Stain incubated…" ([DNA] = ~10 pM, beads: 4e-4%)
([DNA]=~150 fM, beads: 5.9e-6%)

11. Viewing solution for microfluidic device

(Ref: HZ NB 2 pg. 141)
850 μL of 65 wt% sucrose, 1wt% glucose in 1xTE with 10 mM NaCl (TE10)
10 μL of 10 mg/mL BSA
10 μL of betamercaptoethanol
2 μL of 1000x glucose oxidase
2 μL of 1000x catalase
- mix to provide reducing environment
125 μL of product from step 9 –"Stain incubated…" ([DNA] = ~10 pM, beads: 4e-4%))
([DNA]=~1.25 pM, beads: 4.9e-5%)

Viscosity-matched buffer
850 μL of 65 wt% sucrose. 1wt% glucose, TE10
150 μL TE10

Notes

When adding PCR programs, have the program end with a hold at 4 °C, this is handy if you aren't around to remove the tubes and is equivalent to transferring the tubes to the fridge.

Handy worksheets

1. Spinning out excess oligo worksheet

Spin ID: _____
Date: _____
Spin column type: _____

Purpose:_____

Column prep _____ x g / rpm
_____ sec
(replace tube such that
column is not disturbed)
UP water rinse x _____
_____ μL
_____ x g / rpm
_____ sec

Salt replenishment: _____

Sample: _____
_____ μL
_____ x g / rpm
_____ sec
New sample label: _____

2. Hybridization worksheet

Hybridization ID: _____ Date: _____

Purpose:_____

Sample label: _____ BSA-treated tube?_____

Recipe:
_____μL beads _____ wt % 40 nm red fl. Fluospheres (02/01/08)
_____μL construct _____nM DNA/oligo____ (__/__/__) Spun?_____
_____μL dilution buffer _____
_____μL Total Mixing?_____

DNA _____nM Beads ____ wt% Temp: _____ Time: _____

Add stain:
_____μL _____μM YOYO (__/__/__)
_____μL dilution buffer _____
_____μL Total Mixing?_____ Covered?_____

DNA _____nM Beads ____ wt% Temp: _____ Time: _____

Post-Staining Spin ID: _____ to remove excess beads

Preparation and Visualization of λ-DNA and H-PNA Hybrids

Day 1 (afternoon)

1) H-PNA hybridization (100 μL)

78 μL ultrapure water
5 μL 20xTE
4 μL 50 mM NaCl
8.2 μL stock λ-DNA*
5 μL H-PNA at ~50 μM**
~100 μL , 37°C 25 hr.
[DNA]= 1.3 nM
[PNA]=2.5 μM
PNA:DNA=1900

* before adding, heat to 65°C for 1 min, quickly plunge into wet ice
** to resolubilize a new aliquot: add 20.4 μL ultrapure water to vial, pipette up and down, vortex, heat on block heater 40-50°C for 10 min., vortex, recover 20.4 μL.

Day 2 (overnight)

1) Dialysis with 25nm-pore Millipore nitrocellulose membrane, 16 hrs over 40mL of 1XTE with 30 mM NaCl (TE30) in refrigerator

Day 3 (early morning)

1) Remove sample from dialysis

- flick to mix and spin down
- make as many full 10 μL batches with wide-tip pipette as recovered sample allows
- with remaining sample. take UV-VIS to determine new DNA concentration (1-3 nM)
    For λ-DNA: [DNA] nM = Absorbance (260 nm)* 15.3
    e.g. A = 0.15, [DNA] = 2.3 nM
    This is based on the Beer-Lambert law and the known molar absorptivity of dsDNA at 260nm: 13.2 mL/(μmol*cm*bp)

2) Heat to enhance binding specificity

- e.g. 63 °C for 15 min. The time and temperature for this step are PNA construct specific, so optimization will be required for new PNA sequences.
- add 0.2 μL of 5M NaCl to each aliquot to achieve 100mM NaCl

3) Bead Incubation

The following recipe can be scaled by adjusting the volume of DNA constructs to achieve the same final DNA concentration, using as many 10.2 μL aliquots as needed.

The volume of bead solution can then remain constant and only the buffer need be adjusted to balance to 200 μL

177.3 μL q-dot buffer
20.4 μL DNA hybrids at 1.35nM
2.3 μL 0.1% 40nm N-Av beads
~200 μL, 1 hr. rotating at 0.05 Hz at room temperature, covered
[DNA]=0.14 nM
beads: 0.001%
4 beads: 1 DNA

This solution should be refrigerated and is good for ~3 days.

4) Stain bead incubation (begin (5) below)

The following recipe should be adjusted to maintain the final DNA and YOYO concentrations.

86 μL TE10
2 μL YOYO-1 at 10 μM
12 μL bead inc. DNA at 0.14 nM
100 μL, 1 hr. at room temperature, keep covered
[DNA]=16.64 pM
[YOYO-1]=0.2 μM
1 mol yoyo: 4 mol bp DNA

This solution should be discarded at the end of the day.

5) Pre-treat device

In order,
- flow in 2mL IPA through a filter at 2 mL/hr
- flow in 2mL water through a filter at 2 mL/hr
- flow in 0.5 mL 0.1% triton X-100 in 1xTE with 10 mM NaCl (TE10), let sit for 20-30 min, then flow through remaining 0.5 mL of this solution
- flow in 0.5 mL 10 mg/mL BSA "100xBSA" solution (see recipe), let sit for 20-30min, flow through remaining 0.5 mL of this solution

The first step of the experiment will be to slowly flow through the visualization sample

until it occupies the entirety of the devices and tubing and then establish flow prior to data acquisition, thereby effectively flushing the device of the pre-treatment.

6) Optional : Viewing solution for slide

184 μL TE 10 pH 8
15 μL stained DNA at 16.64 pM
1 μL 0.1% triton X-100 in 1xTE with 10 mM NaCl (TE10)
~200 μL
[DNA]=1.25 pM

7) Viewing solution for flow

Prior to each experiment, the combined sucrose and glucose wt% of the viscosified buffer stock should be determined via the refractometer as the concentration may change with time. The index of refraction is related to wt% and wt% to viscosity via available graphs, the volume of viscosified buffer in the visualization solution can be adjusted to achieve the desired final viscosity and the volume of TE10 should be adjusted to maintain a total 1mL volume. In the viscosity-matched buffer solution, the volume of stained product is replaced with TE10.

890 μL "65 wt%" sucrose buffer
10 μL 10mg/mL BSA (100X)
60 μL TE10
2 μL glucose oxidase 1000X
2 μL catalase 1000X
20 μL β-mercaptoethanol
16 μL stained DNA sample at 16.64 pM
1 mL

Note:
"65 wt%" sucrose buffer contains ~64 wt% sucrose. ~1 wt% glucose as a substrate for the glucose oxidase and 1xTE10, pH8.

This solution should be used promptly as the glucose oxidase and catalase activity will drop the pH over time. Exchange visualization solutions roughly every 2 hours.