SINGLE MOLECULE STUDIES OF CHROMATIN DYNAMICS AND TRANSCRIPTION COUPLED REPAIR

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Ming Li

February 2016

SINGLE MOLECULE STUDIES OF CHROMATIN DYNAMICS AND TRANCRIPTION COUPLED REPAIR

Ming Li, Ph. D.

Cornell University 2016

Biological systems create designs that respond to the need to perform specific functions. In particular, protein-DNA complexes form unique structures to maintain the stability of genetic information and yet the dynamics for necessary cellular processes. Motor proteins translocate along, and rotate around, DNA molecules to separate DNA strands, carry out polymerization reactions, resolve topological issues, repair DNA damage, and modify DNA-binding proteins. By investigating one molecular complex at a time, single molecule techniques provide controlled and quantitative approaches to measure and manipulate the protein-DNA interactions as well as visualize the function of motor proteins in real time. These techniques have now made it possible to address many problems that are difficult or impossible to study with traditional assays

In this dissertation, we first introduce DNA unzipping as a powerful tool to study protein-DNA interactions at the single-molecule level. In particular, we detail protocols for preparing an unzipping template, constructing and calibrating the instrument, and acquiring, processing, and analyzing unzipping data. We also summarize major results from utilizing this technique in the studies of nucleosome structures and dynamics. After that, we use DNA unzipping to systematically investigate the interplay between nucleosome remodeling and the binding of transcription factors. The results provide direct evidence for a novel mechanism for both nucleosome positioning regulation by bound TFs and TF regulation via dynamic repositioning of nucleosomes. In the last chapter, we elaborate the single molecule unzipping tracker technique

and its application in understanding the function of the bacterial transcription coupled repair factor Mfd. The results provide important insights into the role of Mfd beyond the scope of transcription coupled repair and significantly contribute to the understanding of Mfd function in the larger context of transcription.

BIOGRAPHICAL SKETCH

Born and raised in China, Ming Li started his academic training at Beijing Institute of Technology, where he earned a Bachelor's of Engineering in Material Science and Engineering 2004. Upon receiving a master of science at National University of Singapore, 2008. Ming started his PhD training by joining the lab of Dr. Michelle D. Wang in 2008. Ming graduated with a Master of Science in 2010 and a Ph.D. in 2015 at Cornell University.

Ming works at the interface among physics, chemistry and biology, by utilizing novel physics instruments to answer fundamental questions in the biological systems. His research has made significant impact in understanding cancer development, aging mechanisms and neurological disorder at a fundamental level. Ming also has a strong passion of bridging science and technology developed in research institutes and universities to the industries to improve people's life.

To Tianyan Deng!

ACKNOWLEDGMENTS

It is very exciting to write this acknowledgement page because it causes me to spend much time reflecting back the life in the past seven years. I have really enjoyed doing my graduate studies in Cornell University, because I have spent this valuable time with my supportive colleagues and friends.

My seven years of graduate studies seems long, but I can still vividly remember a lot that happened when I first joined the lab. I still remember that seven years ago, my first meeting with Michelle happened at the 3rd floor classroom of Clark Hall on the second day of my arrival, and how she has instilled in me with her passion in single molecule biophysics and influenced me with her diligent working spirit as a dedicated scientist ever since. Over the past seven years, Michelle has been a great mentor in science, a strong support in research, a passionate motivator and an easygoing friend in the lab. Countless times, she encouraged me with sound and practical advices, sincere attitude as well as great senses of humor. In addition, she has given me a lot of opportunities to express myself and the freedom to explore and learn from textbooks and experience. For all that she has done, I am eternally grateful for her guidance and help and am truly honored to be one of her students.

The Wang lab is filled with gifted scientists with a wide spectrum of knowledge and skills, as well as friends with kind hearts. At the beginning of my PhD study, I was extremely honored to get help and instructions from Dr. Shanna M. Fellman, Dr. James T. Inman, Dr. Michael A. Hall, Dr. Jing Jin, Dr. Scott Forth, Dr. Benjamin Y Smith, and Dr. Maxim Y Shenin, all of who kindly and patiently taught me the use of sophisticated instruments and complicated experiment procedures. During my PhD study, I also overlapped with Dr. Jie Ma, Dr. Bo Sun, Dr. Yi Yang, Dr. Mohammod Soltani, Dr. Robert A. Forties, Dr. Jun Lin, Dr. Chuang Tan and had a very good time working with them. All of the people at Wang Lab have offered me the generous friendship

and have provide most of the memories I will carry with me for a long period of time.

The interdisciplinary feature of my research projects created me great opportunities to interact constantly with prestigious laboratories all over the world. Over the years, I was involved in extensive collaboration with Dr. Karolin Luger's lab at Colorado State University, Dr. Gregory Bowman's lab at Johns Hopkins University, Dr. Blaine Bartholomew's lab at University of Texas MD Anderson Cancer Center, Dr Alexandra Deaconescu at Brown University, as well as Dr. Jeffrey R. Roberts' lab at Cornell Biotechnology. The knowledge that I have accumulated through these collaboration is enormous. I greatly appreciate their assistance in multiple projects that I have accomplished.

Last but not least, my acknowledgement goes to Department Chemistry and Chemical Biology, Cornell University, especially to my thesis committee chair Dr. Peng Chen, my thesis committee member Dr. Hening Lin (who also acted as my director of graduate studies for a long time), as well as Dr. Brian Crane, for granting me the generous support in my first year at Cornell, and enough freedom to explore my passion in science and technology at Cornell.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

UNZIPPING TO STUDY NUCLEOSOME STRUCTURE AND DYNAMICS

DNA unzipping is a powerful tool to study protein-DNA interactions at the single-molecule level. In this chapter, a detailed and practical guide to performing this technique are provided with an optical trap, using nucleosome studies as an example. We detail protocols for preparing an unzipping template, constructing and calibrating the instrument, and acquiring, processing, and analyzing unzipping data. We also summarize major results from utilization of this technique for the studies of nucleosome structure, dynamics,.

INTRODUCTION

As the fundamental units of eukaryotic chromatin, nucleosomes are responsible for packaging the genome into the nucleus and regulating access to genetic information during various cellular processes. The nucleosome core particle consists of 147 bp of DNA wrapped ~1.7 times around a histone octamer, containing two copies of H2A, H2B, H3 and H4 (Luger et al., 1997). The non-uniform distribution of histone-DNA interactions within a nucleosome governs its dynamic role in regulating access to nucleosomal DNA during transcription, replication and DNA repair (Andrews and Luger, 2011; Korber and Becker, 2010). The position of nucleosomes along the genome are not only influenced by the properties of the underlying DNA sequence (Kaplan et al., 2009), but are also regulated by various histone chaperones (Das et al., 2010; Park and Luger, 2008; Ransom et al., 2010), ATP-dependent chromatin remodeling complexes (Bowman, 2010; Clapier and Cairns, 2009), and other DNA binding proteins, such as transcription factors (Bell et al., 2011; Zhang et al., 2009). In addition, several types of epigenetic marks, including covalent histone modifications, affect both the structure and stability of nucleosomes, as well as higher-order chromatin structure (Bannister and Kouzarides, 2011; Campos and Reinberg, 2009; Ray-

Gallet and Almouzni, 2010). Thus, a detailed understanding of nucleosome structure and dynamics, as well as the relationship between nucleosomes and relevant regulatory factors, is of great interest to multiple fields and can enhance our knowledge of the basic tenets of biology.

Single-molecule techniques offer the unique ability to both detect the inherent heterogeneities of biomolecules and directly monitor dynamic processes in real time, and are thus important complementary to ensemble studies for understanding various biological systems (Joo et al., 2008; Killian et al., 2011; Moffitt et al., 2008). In particular, DNA stretching experiments utilizing single-molecule manipulation techniques, such as magnetic tweezers or optical tweezers, allow for the direct investigation of the mechanical properties of both single nucleosomes and nucleosome arrays (Brower-Toland et al., 2002; Gemmen et al., 2005; Mihardja et al., 2006; Simon et al., 2011). However, stretching experiments are unable to directly determine the location of a nucleosome on a long DNA template or directly probe the absolute locations of specific histone-DNA interactions in a nucleosome. To overcome these limitations, we developed an optical trapping-based single-molecule unzipping technique as a versatile tool to probe a variety of protein-DNA interactions (Dechassa et al., 2011; Hall et al., 2009; Jiang et al., 2005; Jin et al., 2010; Koch et al., 2002; Koch and Wang, 2003; Shundrovsky et al., 2006). The unzipping technique is a straightforward concept and may be incorporated into different optical trapping configurations (an example is shown in Figure 1.1). Briefly, a single double stranded DNA (dsDNA) is unzipped in the presence of DNA-binding proteins. Mechanical force is applied to separate dsDNA into two single strands (Figure 1.1a). DNA-bound proteins or protein complexes act as barriers to the unzipping fork so that resistance to unzipping provides a measure of the strengths of protein-DNA interactions while the amount of

DNA unzipped reveals the locations of these interactions along the DNA. These locations may be mapped to near base-pair precision and accuracy, making unzipping a powerful high-resolution technique for mapping these interactions.

Unzipping is a unique and extremely powerful single-molecule technique with many advantages: 1) The interaction map of a protein-DNA complex, as characterized by the strengths and locations of interactions, provides important structural information about the complex (Dechassa et al., 2011; Hall et al., 2009; Jin et al., 2010; Shundrovsky et al., 2006); 2) The footprint of a bound complex can be directly measured by unzipping DNA molecules from both directions (Dechassa et al., 2011; Hall et al., 2009; Jiang et al., 2005; Jin et al., 2010); 3) Unzipping directly reveals the presence or absence of a bound protein, making it an ideal method for measuring its equilibrium dissociation constant, even for tight binding in the pM range (Jiang et al., 2005; Koch et al., 2002); 4) Using dynamic force measurements, it is possible to differentiate different bound species which may bind to the same DNA sequence (Koch and Wang, 2003); 5) Unzipping is capable of determining the location of a protein on a very long DNA molecule with near base-pair accuracy, making it ideal for studying the positioning and/or repositioning of proteins and protein complexes along DNA (Shundrovsky et al., 2006).

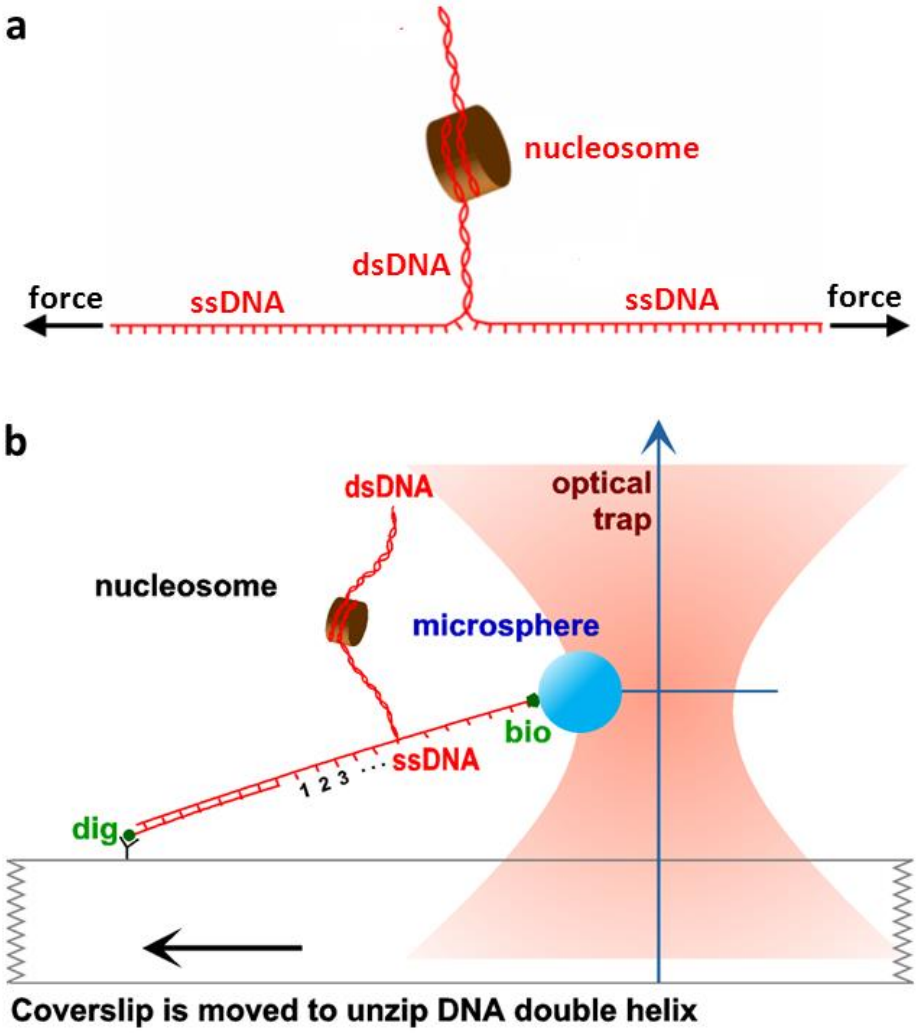


Figure 1.1. Experimental unzipping configuration. (a) A simplified cartoon of the unzipping configuration. A DNA double helix is mechanically unzipped in the presence of DNA-binding proteins, such as a nucleosome, by the application of opposing forces on the two strands. (b) A typical experimental configuration for unzipping. An optical trap is used to apply a force necessary to unzip through the DNA as the coverslip is moved away from the trapped microsphere.

Unzipping has been successfully utilized to study the binding affinity of restriction enzymes (Koch et al., 2002; Koch and Wang, 2003), mismatch detection by DNA repair enzymes (Jiang et al., 2005), the dynamics of nucleosome structure and positioning (Dechassa et al., 2011; Hall et al., 2009; Shundrovsky et al., 2006), and how RNA polymerase overcomes a nucleosome barrier (Jin et al., 2010). For clarity and brevity, we will focus below on the experimental procedures utilizing our particular single-beam optical trapping system to study nucleosome structure and dynamics (Figure 1.1 b).

Sample preparation

2.1 DNA unzipping template design

Here, we detail the construction of unzipping templates that can be used with the optical trapping system shown in Figure 1.1b. Although different experimental configurations may require somewhat different templates, the general template designs share common features (Bockelmann et al., 1998; Koch et al., 2002). The template generally consists of two segments: an anchoring segment and an unzipping segment, separated by a nick (Figure 1.2). At one end of the anchor segment is a tag that will bind to the coverslip and, near the nick is a different tag that will bind to a microsphere. By moving the coverslip away from the trapped microsphere, the unzipping segment can be unzipped.

The anchor segment is generally 1-2 kb long and consists of a dsDNA linker arm with an end-labeled digoxigenin tag. This length will provide sufficient distance between the anchor point

and the unzipping segment to ensure that the trapped microsphere does not contact the coverslip surface. This will facilitate data analysis. However, an anchor segment that is too long will lead to increased Brownian noise of the trapped microsphere and compromise the accuracy of position measurements.

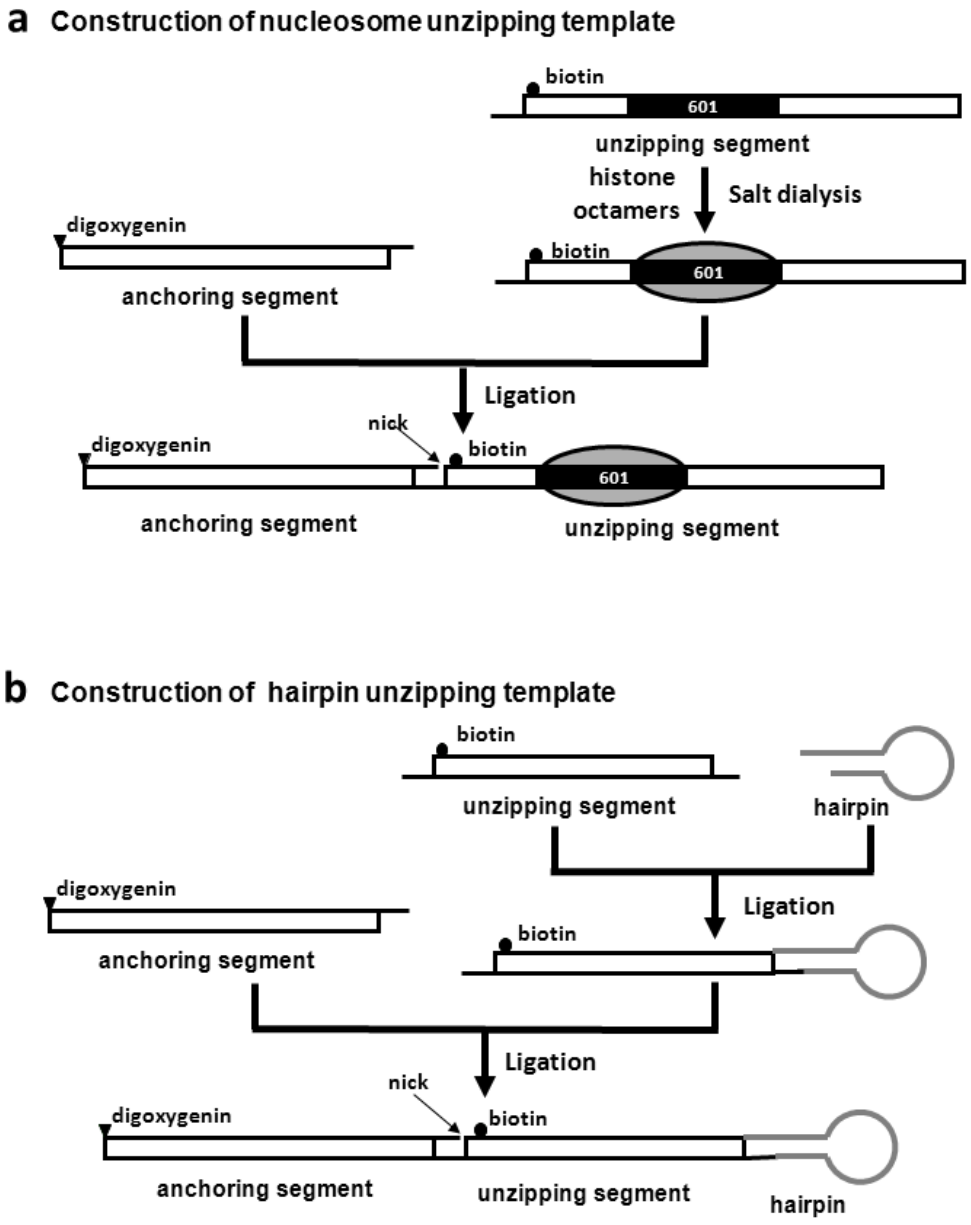


Figure 1.2 Unzipping template construction. (a) Construction of a nucleosome unzipping template. A DNA template for nucleosome unzipping experiments consists of a digoxigenin-labeled anchoring segment and a biotin-labeled nucleosome unzipping segment. As an example, a nucleosome unzipping segment is shown containing a 601 nucleosome positioning sequence. (b) Construction of a hairpin-capped unzipping template. This template consists of a digoxigenin-labeled anchoring segment and a biotin-labeled unzipping segment with a hairpin capped at the distal end.

The unzipping segment consists of an experiment-specific target sequence with an internal biotin tag near the nick. This segment may also contain bound proteins or protein complexes to be studied. The length of the unzipping segment can vary from hundreds to a few thousands of base pairs. Note that a segment that is too short will not allow for optimal data alignment (detailed in a later section). Bound proteins or protein complexes can be located in any region of the unzipping segment, although we typically leave at least 200 bp of flanking DNA on either side of a bound protein to ensure accuracy during data alignment. In some experiments, unzipping should be conducted from both directions (forward and reverse) along the same unzipping segment to investigate possible asymmetric binding. Therefore, both forward and reverse unzipping segments should be prepared.

As an example, we provide below a detailed protocol for constructing a typical forward unzipping template for studying histone-DNA interactions in a single nucleosome (Dechassa et al., 2011; Hall et al., 2009; Shundrovsky et al., 2006). In this template, the anchor segment is 1.1 kb and the unzipping segment is 774 bp. The nucleosome is located near the center of the unzipping template, and is well-positioned on a 147 bp “Widom 601” nucleosome positioning element (601), which has an extremely high affinity for a nucleosome (Figure 1.2 a) (Lowary and Widom, 1998). Labeling and producing the DNA templates is accomplished by standard

enzymatic reactions and purification methods with biotin and digoxigenin labeled nucleic acids. These two labels are especially convenient due to the ease of covalent attachment of streptavidin (Sigma) to carboxylated polystyrene microspheres (Polysciences, Inc.) and the availability of high-affinity anti-digoxigenin antibodies (Roche Applied Sciences, Indianapolis, IN). The reverse nucleosomal unzipping segment is prepared using methods nearly identical to those of the forward unzipping segment, except the entire segment is flipped by the use of different primers, such that the ligatable overhang is located on the opposite end.

Anchor segment preparation:

1) PCR amplify the anchor segment from plasmid pRL574 (Schafer et al., 1991). The forward primer contains a 5' digoxigenin label, designed to be ~1.1kb away from the single BstXI cutting site located on the plasmid.
2) BstXI (New England Biolabs) digest the PCR product to generate a 3' overhang for ligation of the unzipping segment.

Unzipping segment preparation:

1) PCR amplify the unzipping segment from plasmid p601 (Lowary and Widom, 1998). The forward primer was designed to be ~200 bp upstream of the 601 element and to contain a BstXI cutting site. This site will be utilized to generate a 3' overhang complementary to the one produced on the anchor segment. The forward primer contains

an internal biotin label near the 3' overhang. The reverse primer is located ~400 bp downstream of the 601 element.

2) BstXI (New England Biolabs) digest the PCR product. Follow the digestion with the addition of a stoichiometric amount of Calf Intestinal Alkaline Phosphatase (CIP, New England Biolabs) in the same buffer, to remove the phosphate from the 3' overhang. This allows for the generation of a nick after subsequent ligation (discussed below in 2.3).

To characterize the precision and accuracy of the unzipping method in locating a bound protein along the DNA (discussed later), we have also designed multiple unzipping templates of varying lengths, capped with hairpins at distal ends (Figure 1.2b). These hairpins act as strong binding sites by preventing further unzipping at well-defined locations along the DNA and allow for a direction comparison with measured locations. In addition, these unzipping templates are also used to determine the elastic parameters of single stranded DNA (ssDNA) under experimental conditions discussed below. Here, we have outlined a protocol for constructing hairpin-capped unzipping segments of various lengths.

Hairpin-capped unzipping segment preparation:

1) PCR amplify the unzipping segment from p601 using the same forward primer specified above in 'unzipping segment preparation'. The reverse primers are located at various distances downstream from the forward primer to generate templates of different lengths.

The reverse primer is also designed to contain an EarI cutting site that will generate a 5' overhang.

2) EarI (New England Biolabs) digest the PCR products to generate a 5' overhang.
3) The oligonucleotide (Integrated DNA Technologies) utilized to generate the hairpin is designed to form a 3-base hairpin loop and a short dsDNA stem (~12 bp) with a 5' overhang which is complementary to the overhang in the unzipping segment.
4) Ligate the unzipping segment with a hairpin oligonucleotide (1:10 molar ratio) using T4 Ligase (New England Biolabs). Overnight ligation is normally necessary to maximize the ligation yield. Purify the ligated products using agarose gel purification.
5) BstXI (New England Biolabs) digest the gel purified product to generate an overhang near the biotin-label. Following the digestion, add a stoichiometric amount of Calf Intestinal Alkaline Phosphatase (CIP, New England Biolabs) in the same digestion buffer to remove the phosphate from the 3' overhang. This allows for the generation of a nick after ligation with the anchor segment (discussed below in 2.3).

2.2 Nucleosome reconstitution

It has been well established that *in vitro* assembly of nucleosomes and chromatin arrays from highly purified DNA and histone components can be achieved by either salt-gradient dialysis (Luger et al., 1999) or a chaperone mediated approach (Fyodorov and Kadonaga, 2003). We employ a salt-gradient dialysis method for nucleosome reconstitution. Different types of individual histones can be prepared as previously described (Dyer et al., 2004). Purified histone

octamers from several species are also commercially available in forms suitable for reconstitution (Protein Expression/Purification Facility, Colorado State University). In our previous publications (Dechassa et al., 2011; Hall et al., 2009; Jin et al., 2010; Shundrovsky et al., 2006), a well-established salt dialysis method (Dyer et al., 2004; Thastrom et al., 2004) was modified to reconstitute a single nucleosome on a long piece of DNA containing one 601 positioning element. The modified protocol uses a small total volume and requires a low concentration of DNA template and histones (100 nM or even lower) that are suitable for single-molecule studies. The modifications are listed below.

1) The dialysis button is constructed following the procedure from Thastrom et al., 2004. This allows us to work with assembly volumes of ~ 30 µL, which is much smaller than commercially available dialysis chambers.
2) We include 0.2 g/Lsodium azide (Sigma-Aldrich) in both the high and low salt buffers to remove bacteria which may contaminate the buffers and decrease assembly efficiency.
3) We include 0.1 mg/mL acetylated BSA (acBSA) (Ambion), to each dialysis button, as a crowding agent to assist with assembly.
4) The dialysis is performed at 4°C and the dialysis pump is set to 1.2 mL/min for ~18 hours. The flow is then changed to 2.5 mL/min for an additional ~4 hours.

After dialysis, the samples are transferred to zero salt buffer and incubated for 2-3 hours at 4°C. Fine-tuning the molar ratio between the histone octamer and DNA template is critical to achieve a high reconstitution yield and avoid over assembly. In addition, it is also important to remove bubbles from the dialysis button, because bubbles can prevent buffer exchange between the high salt and the sample. After reconstitution, the nucleosome samples may be stored for a few weeks at 4°C.

2.3 Formation of the final unzipping template

The unzipping segment (containing either a nucleosome or a hairpin) is directly ligated to the anchor segment (in a 1:1 molar ratio) immediately prior to use (Figure 1.2). The CIP dephosphorylation of the unzipping segment ensures that only one strand of the DNA is ligated and a nick is generated on the complementary strand during the ligation step. This complete unzipping template is labeled with a single digoxigenin tag at the 5' end of the anchor segment and a biotin tag located near the nick on the unzipping segment. A complete template lacking a nucleosome (naked DNA) is stable for a few days without DNA nicking at 4°C; a template containing a nucleosome should be used within a few hours of ligation.

2.4 Preparation of experimental sample chambers

For single-molecule studies in general, individual DNA tethers need to be immobilized in a single-molecule sample chamber or a flow cell. These allow the user to sequentially flow in different solutions for use with the optical trapping system. In our nucleosome unzipping studies, sample chambers with a ~15 μL volume are prepared at room temperature, and then mounted onto an optical trapping setup. By performing incubations in a humid chamber prior to mounting onto the optical setup, buffer evaporation can be minimized.

Buffer solutions:

1) Sample Buffer (SB): 10 mM Tris HCl (pH 7.5), 1 mM Na_2EDTA, 150 mM NaCl, 3% (v/v) glycerol, 1 mM DTT, 0.1 mg/mL acBSA.
2) Blocking Buffer (BB): SB + 5 mg/mL casein sodium salt from bovine milk.
3) Nucleosome Unzipping Buffer (NUB): 10 mM Tris HCl (pH 8.0), 1 mM Na_2EDTA, 100 mM NaCl, 1.5 mM $MgCl_2$, 3% (v/v) glyercol, 1 mM DTT, 0.02% (v/v) Tween20, 2 mg/mL acetylated BSA (acBSA).

Blocking agents are used to coat the surface of the sample chamber to prevent unwanted protein attachment to the surface. The blocking agent that has been the most successful for us is casein sodium salt from bovine milk (Sigma-Aldrich Co.). AcBSA is thought to mimic conditions of a higher protein concentration in the sample buffer, creating a more 'crowded' environment, and thus prevents protein dissociation from the template (Gansen et al., 2007). We utilize a polyclonal sheep anti-digoxigenin (Roche Applied Science, Indianapolis, IN) for attachment of digoxigenin-labeled DNA samples to sample chamber surfaces. The complete unzipping template is diluted to a desired concentration in SB immediately before introduction to sample chamber. The detailed procedure of creating a tethered DNA sample chamber is given below:

Creating a DNA tethered sample chamber:

1) Apply two thin pieces of double-stick tape (~0.1 mm thick) to a coverslip (24 mm x 40 mm x 0.15 mm). Orient the pieces parallel to one another and separate them by ~5 mm.
2) Place a glass slide on top of the coverslip and perpendicular to it, to create an ~15μL channel down the center.

3) Flow in one volume (~15 µL) of antidigoxigenin solution (20 ng/µL in H_2O).
4) Incubate for five min. Wash with five volumes of BB. Incubate with residual blocker for five min.
5) Wash with five volumes of SB. Immediately flow in one volume of diluted unzipping template in SB. Incubate for 10 min.
6) Wash with five volumes of SB. Flow in one volume of streptavidin-coated beads (5 pM in BB). Incubate for 10 min.
7) Wash with 10 volumes of NUB.

The concentration to which the unzipping sample is diluted prior to being added to the sample chamber is critical for achieving an optimal tether density under single-molecule conditions. A concentration that is too high will lead to multiple tethers (one bead attached to multiple DNA molecules) and a concentration that is too low will make it difficult to locate a suitable unzipping tether. In theory, 10 pM of DNA template are needed to achieve an acceptable tether density. However, since the ligated templates are directly diluted without purification, the appropriate 'flow-in' concentration depends heavily on the ligation efficiency. Therefore, for each new template, several chambers with different template concentrations are often made and evaluated to establish an appropriate flow-in concentration.

Consistency in all aspects of sample preparation and utilization is critical to achieve reproducible results among different sample chambers. In this regard, unzipping experiments are conducted in a temperature and humidity controlled soundproof room. Once prepared, the sample

chambers should be utilized in a timely fashion, typical within 1 hour, to avoid unnecessary complications such as protein dissociation or sample sticking to the surface of the chamber.

Instrumentation and data collection

3.1 Layout of single beam optical trapping apparatus

Since the pioneering work by Arthur Ashkin over 20 years ago (Ashkin et al., 1986), the optical trapping field has grown tremendously due to the unique ability of optical tweezers to monitor and manipulate biological targets with high temporal and spatial resolution. In addition, refinements of established methods and the integration of this tool with other forms of single-molecule manipulation or detection have made this technique of great interest in both physics and biology. The single beam optical trapping instrument that we use has a very straightforward design (Brower-Toland and Wang, 2004; Koch et al., 2002), containing the minimal set of optical components required for the operation of a high precision instrument of its kind (Figure 1.3). A 1064 nm laser (Spectra-Physics Lasers, Inc. Mountain View, CA) is transmitted through a single-mode optical fiber (Oz Optics, Carp, ON), expanded by a telescope lens pair, and focused onto the back focal plane of an 100X, 1.4 NA oil-immersion microscope objective that is mounted in a modified Eclipse TE 200 DIC inverted microscope (Nikon USA, Melville, NY). The focused beam serves as a trap for a 500 nm polystyrene microsphere (Polysciences, Inc.). Forward scattered light is collected by a condenser lens and imaged onto a quadrant photodiode (Hamamatsu, Bridgewater, NJ). A displacement of a trapped sphere imparts a deflection of the

forward scattered light and is captured as a differential voltage signal at the quadrant photodiode. The laser intensity is adjusted by modulating the voltage amplitude applied to an acoustic optical deflector (AOD) (NEOS Technologies, Inc., Melbourne, FL) placed between the laser aperture and the beam expander. Samples are manipulated manually via a micro-stage or via a high precision 3D piezoelectric stage (Mad City Labs, Madison, WI). Analog voltage signals from the position detector and stage position sensor are anti-alias filtered at 5 kHz (Krohn-Hite, Avon, MA) and digitized at 7 to 13 kHz for each channel using a multiplexed analog to digital conversion PCIe board (National Instruments Corporation, Austin, TX).

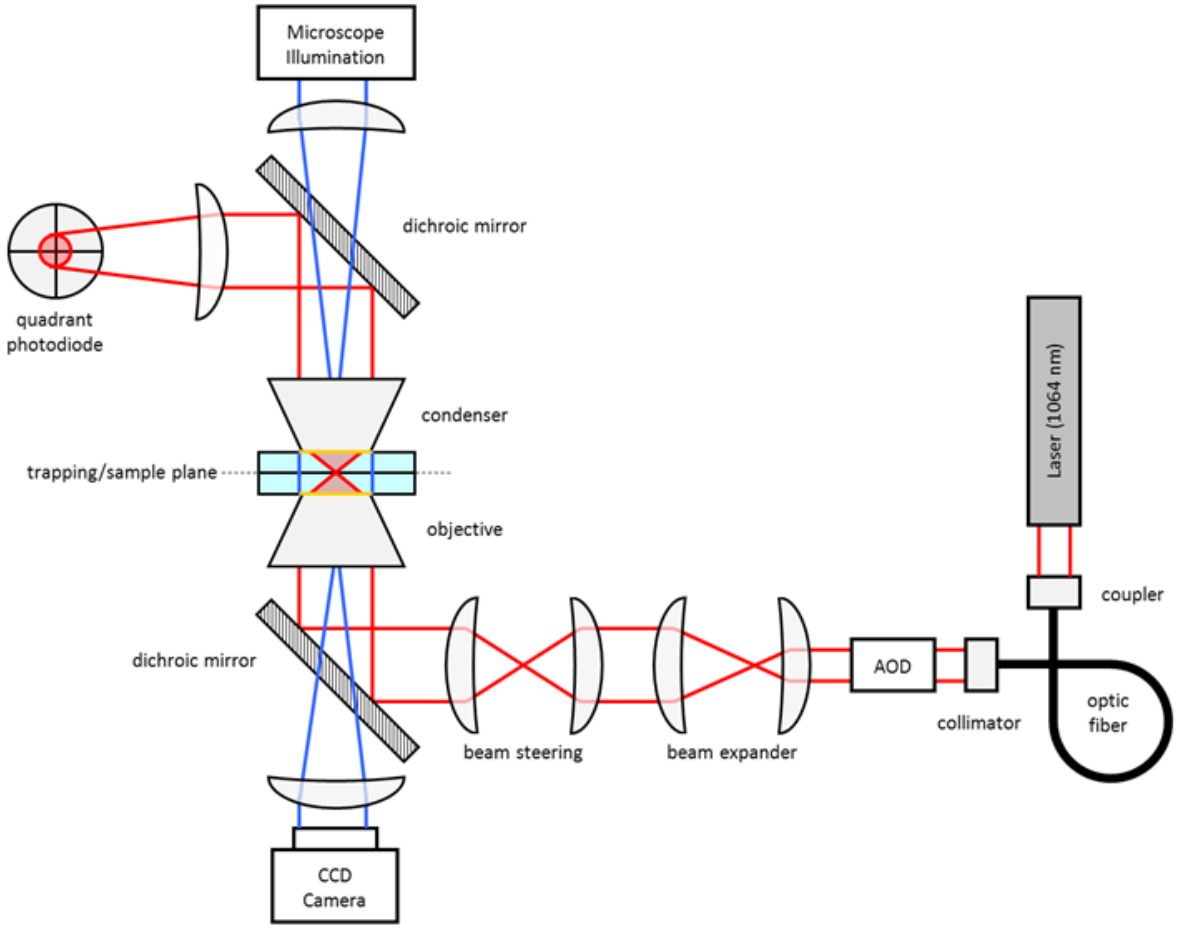


Figure 1.3. Layout of the optical trapping apparatus. See text for a detailed description of the setup.

3.2 Calibration of the optical trapping system

The instrument calibration methods for our optical trapping setup were detailed in a previous publication (Wang et al., 1997). In brief, they include 1) the determination of the position detector sensitivity and the trap stiffness, 2) the determination of the position of the trap center relative to the beam waist and the height of the trap center relative to the coverslip, 3) The location of the anchor position of the unzipping tether on the coverslip, which is determined prior to each measurement by stretching the anchor segment laterally at a low load (< 5 pN). These calibrations are subsequently used to convert raw data into force and extension values.

3.3 Experimental control – loading-rate clamp and force clamp

In nucleosome unzipping experiments, we often use two approaches to disrupt a nucleosome: loading-rate clamp unzipping and force clamp unzipping. The advantages and disadvantages of these two methods are discussed below.

A loading-rate clamp allows the force to increase linearly at a specific rate until the disruption of an interaction. This approach generates distinct force unzipping signatures (unzipping force as a function of the number of base pair unzipped) which can be used to distinguish a nucleosome from other DNA binding proteins. Because the disruption is a thermally activated process (Evans, 2001), the force needed to disrupt a specific interaction in the nucleosome is dependent on the loading rate as well as the starting force. After a disruption, the force naturally drops but is not allowed to recover to the naked DNA unzipping baseline. Consequently the starting force

for a subsequent disruption is higher than for the initial disruption. Thus weak interactions in a nucleosome may be detected if they are first encountered by the unzipping fork. This method may be used to highlight weak histone-DNA interactions near the entrance and exit sites. It will of course detect all strong interactions.

A force clamp allows for the disruption of all interactions in a nucleosome under the same force. This is well-suited to a quantitative analysis of the strength of the detected interactions (Forties et al., 2011). However, more experimentation is normally required to determine an appropriate range of desired unzipping forces. A force that is too small will make the time to disrupt the nucleosome too long to be experimentally accessible, and a force that is too large may overlook specific interactions in the nucleosome. A force clamp is usually implemented with loading rate clamps before and after it to simplify data alignment (discussed further below). A loading rate clamp is most suitable to study the interactions around the periphery of a nucleosome, while the force clamp mode is optimal for determining the interactions around the dyad. More importantly, a loading rate clamp provides a clear force unzipping signature, while a force clamp is more convenient in quantitative analysis of the energy landscape of histone-DNA interactions in a nucleosome.

3.4 Data acquisition

Here we detail the process of data collection using the two aforementioned approaches.

When utilizing the loading rate clamp mode (Dechassa et al., 2011; Hall et al., 2009; Koch and Wang, 2003; Shundrovsky et al., 2006), the microscope coverslip velocity is adjusted to produce a constant force-loading rate by controlling the position of the piezo stage, while the position of the microsphere in the trap is kept constant by modulating the light intensity (trap stiffness) of the trapping laser. Unzipping through a nucleosome is visualized as a group of force peaks up to 30-40 pN ramping up linearly above the naked DNA unzipping baseline (13-16 pN). In the force-clamp mode (Hall et al., 2009), the unzipping begins with a loading rate clamp until the desired force (threshold force) is reached within a nucleosome. The unzipping force is then held constant via feedback control of the coverslip position. The threshold force is carefully selected so that it is much higher than the baseline unzipping force of the naked DNA, but is still low enough to allow sufficient dwell time at most histone-DNA interactions for detection. Upon passing through the nucleosome region, the unzipping reverts to the original loading rate clamp. The distinctive naked DNA unzipping signatures, detected by the loading rate clamp before and after the nucleosome, are important for data alignment (see below).

Apart from the two unzipping modes presented above, we can also modulate the unzipping process to allow the DNA to be unzipped and rezipped multiple times by controlling the unzipping forces and the corresponding position of the piezo stage. This modulation allows unzipping experiments to potentially mimic important biological process, such as a motor protein progressing into a nucleosome (Hall et al., 2009).

Data processing

Data acquired by the optical trapping setup need to be processed according to these steps: 1) determine the elastic parameters of the dsDNA and ssDNA; 2) determine trap height; 3) perform data conversion to force and extension based on geometry; 4) convert data to number of base pairs unzipped; 5) perform data alignment against a theoretical curve. These steps are detailed below.

DNA elastic parameter determination.

Since unzipping experiments involve the extension of dsDNA (the anchoring segment) in series with ssDNA (unzipped DNA) (Figure 1b), elastic parameters of both dsDNA and ssDNA are necessary for data analysis. These parameters are strongly dependent on the buffer conditions used in unzipping experiments. We obtain elastic parameters of dsDNA by stretching dsDNA and fitting the resulting force vs. extension curve to a modified worm-like chain (WLC) model under the same buffer conditions as the actual unzipping experiment following the detailed procedures published previously (Wang et al., 1997). To obtain the elastic parameters of ssDNA, we unzip a template capped with a hairpin at the distal end under the same buffer conditions as the actual unzipping experiment (Koch et al., 2002). Once completely unzipped, the unzipped DNA is then stretched to a high force (up to 50 pN) to obtain the force-extension curve, which reflects elastic contributions from both the dsDNA and ssDNA. Given the elastic parameters of dsDNA under these conditions, this curve allows for the determination of the elastic properties of ssDNA using an extensible freely-jointed chain (FJC) model (Smith et al., 1996).

Trap height determination for individual unzipping curves.

Prior to nucleosome unzipping experiments, we calibrate the height of the trap center relative to the surface of the coverslip when the objective is focused on the coverslip surface (Wang et al., 1997); it is typically found to be ~600 nm. However, the actual trap height during a nucleosome unzipping experiment may differ from the calibrated height by as much as 100 nm due to limited focusing precision. We have therefore implemented a technique to obtain trace-specific trap height of the unzipping data. In this method, we analyze the initial segment of the data prior to strand separation (0-10 pN). Because DNA is not yet unzipped, the expected force-extension curve has been fully characterized as described in section 4.1 and is simply that of the dsDNA anchor segment of known contour length. The trap height is determined when the difference between the converted force-extension curve and the expected curve is minimized.

Data conversion to force and extension based on geometry

Once the trap height is determined, force and extension of DNA as a function of time may be obtained following a method that has been previously described in detail (Wang et al., 1997). At a given time point, a number of parameters must be detected and/or calibrated: DNA anchor point on the coverslip, the position of the trapped bead relative to the trap center, and the stiffness of the trap. The results of the conversion are the force (F) and extension (x) along the direction of the stretched DNA molecule.

Conversion to number of base pairs unzipped (j)

Once the unzipping data are converted to force (F) and extension (x) for a given time point (t), the number of base pairs unzipped (j) at each time point may be obtained. The extension of the DNA (x) contains contributions from both the dsDNA (x_{ds}) and ssDNA (x_{ss}) under the same force:

$$x(F) = x_{\mathrm{ds}}(F) + x_{\mathrm{ss}}(F). \qquad (1)$$

$x_{\mathrm{ds}}(F)$ is determined because the force-extension curve of the anchor segment is fully characterized. $x_{\mathrm{ss}}(F)$ is thus obtained from (1) and is proportional to the number of ssDNA nucleotides. Using the extensible freely-jointed chain (FJC) model, $x_{\mathrm{ss}}(F)$ is converted to the number of base pairs unzipped (j).

As the DNA is extended but prior to strand separation, j = 0 bp, resulting in a vertical rise in the F versus j plot (Figure 4a). Once strand separation starts, a characteristic force signature, determined by the underlying DNA sequence as discussed below, is detected with an increase in j. In the presence of bound proteins this gently varying baseline is interrupted by sharp force rises. When the unzipping fork encounters a strong protein-DNA interaction, force increases linearly while the number of base pair unzipped remains unchanged until the sudden dissociation of the bound protein, leading to a sudden reduction in the force. Therefore, the F versus j plot provides a direct measure of 1) the location of the bound protein on the DNA (j at which the force rise starts) and 2) the strength of the interaction (force magnitude).

Unzipping curve alignment

Although the raw F versus j plot already contains critical information about a bound protein, the precision and accuracy of locating a bound protein are limited to ~ 20-30 bp, due to small but significant uncertainties in a number of parameters (trap height, bead size, and trapped bead position and force). In order to improve on this, we take advantage of the characteristic unzipping force signatures that depend strongly on the DNA sequence and align an experimental unzipping force curve $F_{\exp}(j)$ against a theoretical curve $F_{\text{theo}}(j)$. $F_{\text{theo}}(j)$ is computed based on an equilibrium statistical mechanics model that considers sequence-dependent base pairing energy and DNA elasticity(Bockelmann et al., 1998). During the correlation, the argument of $F_{\exp}(j)$ is both shifted by j_0 number of base pairs and stretched by a factor of a. The best values of j_0 and a are obtained by maximizing the following generalized cross-correlation function:

$$R(a, j_0) = \frac{\int dj \left[F_{\text{theo}}(j) - \overline{F}_{\text{theo}}\right]\left[F_{\exp}(aj + j_0) - \overline{F}_{\exp}\right]}{\sqrt{\int dj \left[F_{\text{theo}}(j) - \overline{F}_{\text{theo}}\right]^2 \int dj \left[F_{\exp}(aj + j_0) - \overline{F}_{\exp}\right]^2}} \qquad (2)$$

where $\overline{F}_{\text{theo}}$ and $\overline{F}_{\exp}$ are the mean values of the $F_{\text{theo}}(j)$ and $F_{\exp}(aj + j_0)$ respectively. The search for optimal j_0 and a may be also facilitated by the use of a SIMPLEX search algorithm.

When using this method to align a trace taken from DNA containing a nucleosome against the known DNA sequence, regions of naked DNA, ~ 100-200 bp, adjacent to the nucleosome should be used for correlation. We found that once unzipping passes a nucleosome, the unzipping curve immediately following the nucleosome did not always show the expected naked DNA pattern (Shundrovsky et al., 2006). Instead, in some traces we observed random high-force peaks that were not present when unzipping naked DNA. We attribute this effect to non-specific

interactions between the end of the DNA and the histone proteins removed from the disrupted nucleosome. For those traces, only the naked DNA preceding a nucleosome can be used for correlation, which may result in somewhat lower precision.

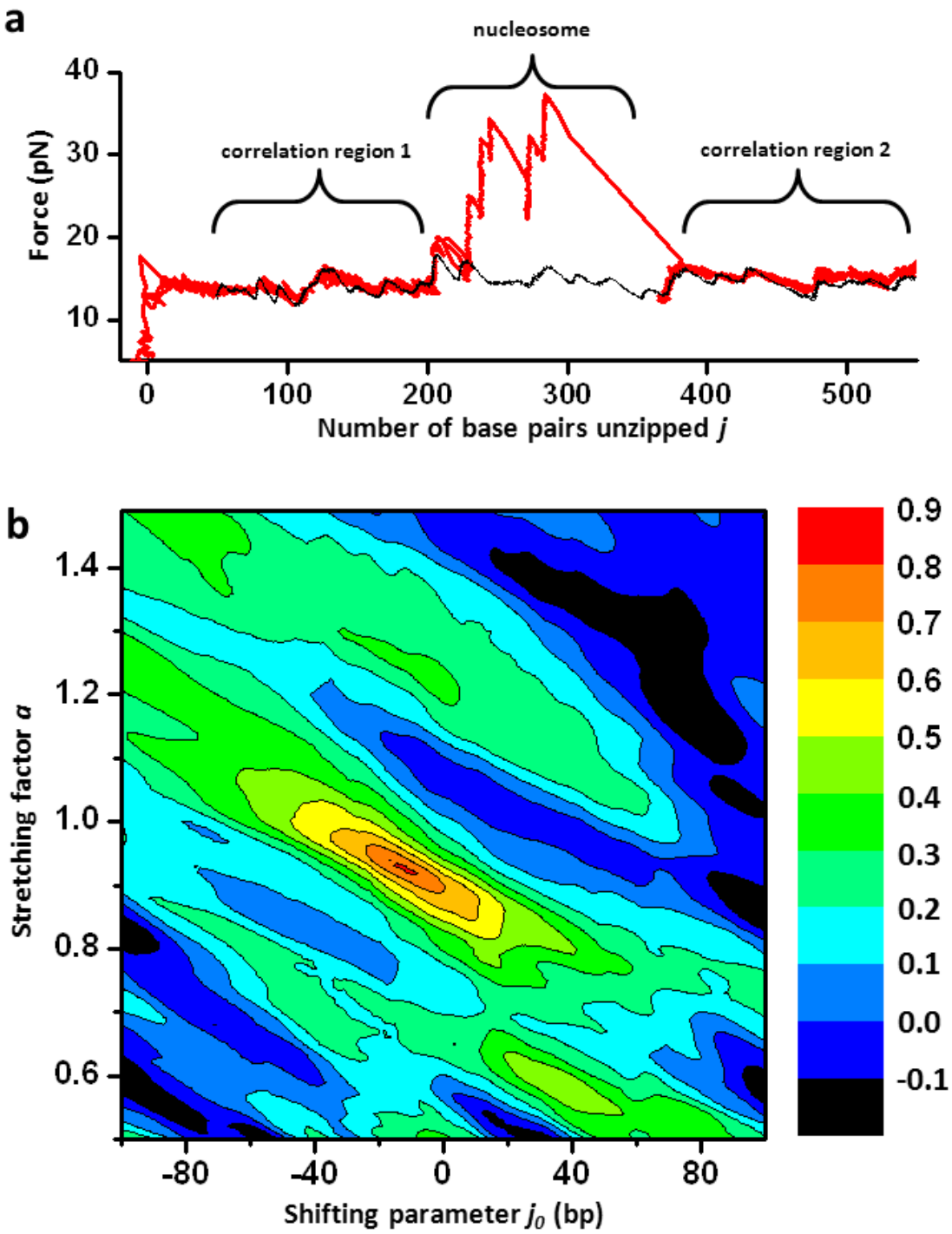


Figure 1.4. Unzipping curve alignment. (a) An example of force versus number of base pairs unzipped plot for a nucleosome unzipping curve (red) after alignment with the corresponding theoretical curve for naked DNA of the same sequence (black). Unzipping was carried out at a loading rate of 8 pN/s. Regions 1 and 2 flanking the nucleosome were used for correlation. Note also that the initial rise of force is located at$j$ = 0 bp, corresponding to stretching of the anchoring segment before strand separation. (b) A two-dimensional intensity graph of the generalized correlation function $R(a, j_0)$ for the trace shown in A. The peak R = 0.80 is located at stretching factor a = 0.93 and shifting parameter j = − 14 bp.

Figure 1.4 is an example of the application of this method to a nucleosome unzipping trace. The correlation was performed using two regions of naked DNA flanking the nucleosome (Figure 1.4a). The generalized correlation function (Figure 1.4b) shows a maximum of 0.80 at $j_0 = -14$ bp and a = 0.93.

Determination of unzipping accuracy and precision

To characterize the ability of the unzipping technique to locate the absolute position of an interaction, we unzip naked DNA templates capped with hairpins at distal ends (Figure 1.5a) and analyze the measured locations of hairpins (Hall et al., 2009). These hairpins mimic strong binding sites at well-defined locations on DNA. As shown in Figure 1.5a, three unzipping templates of varying length, each with a hairpin located near where a nucleosome could be assembled, are unzipped. Each unzipping curve follows that of the naked DNA until it reaches the hairpin where the force rises sharply, providing a clear indication of the hairpin location. These unzipping curves are aligned as described above. Figure 1.5b shows histograms of the detected binding locations for each hairpin template and a comparison with expected locations. Note that accuracy is a measure of the closeness of the measured value with the true value

whereas precision is a measure of the repeatability of measurements. For each template, accuracy is given by the difference between the mean measured location and the expected location while precision is given by the standard deviation of the histogram. For all three templates, the accuracy is within 1 bp and the precision is within 2 bp. Therefore we conclude that the unzipping technique has the capability to determine the absolute sequence position of an interaction with near single base pair accuracy and precision.

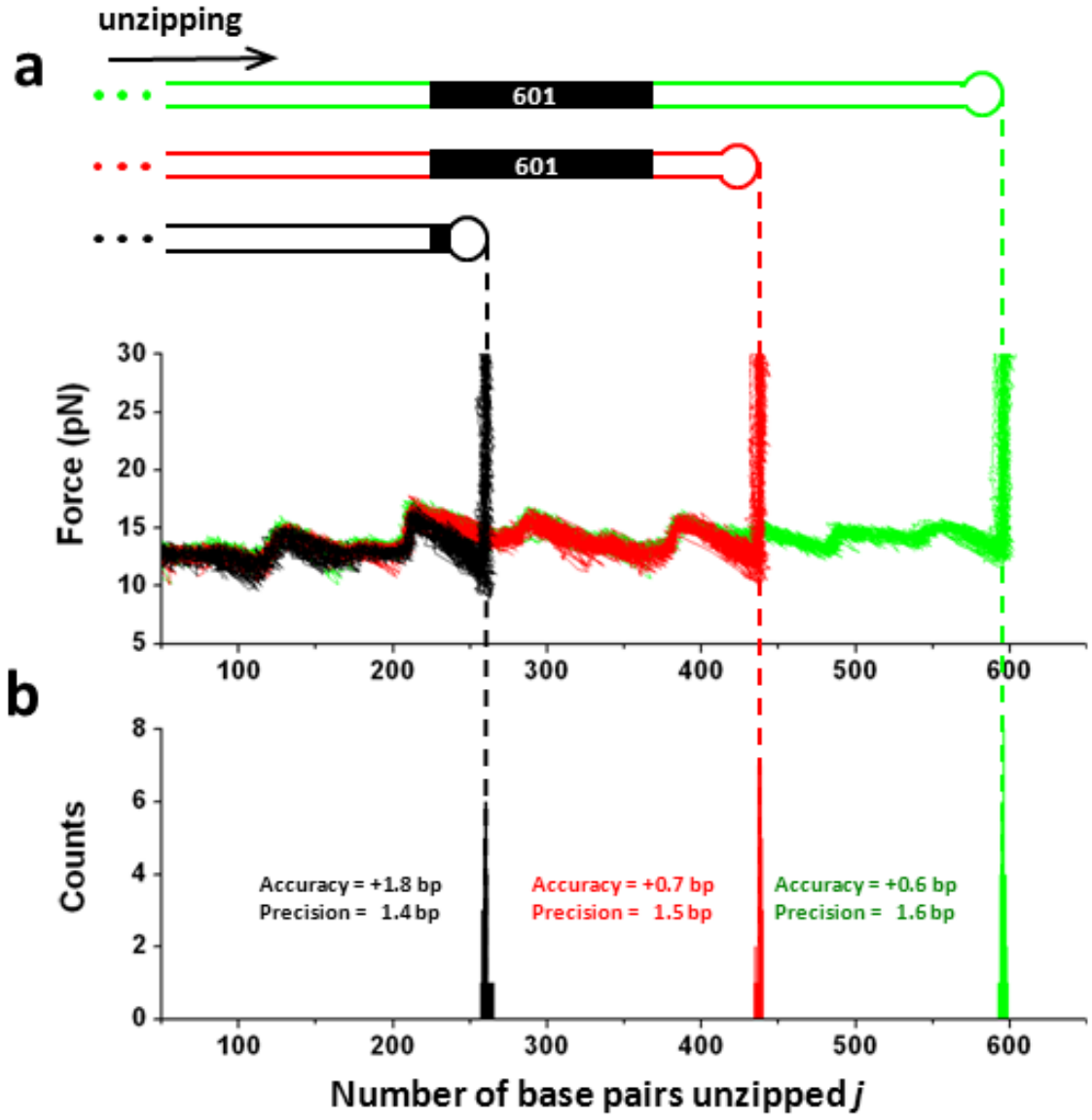


Figure 1.5. Characterization of the accuracy, and precision, of the unzipping method. (a) Three hairpin-capped unzipping templates were unzipped using a loading rate clamp (8 pN/s): 258 bp (black, 21 traces), 437 bp (red, 27 traces), and 595 bp (green, 33 traces). (b) For each template, a histogram was generated from the data points in the vertically rising section only. The measured hairpin location of each template was taken as the mean of the histogram. The accuracy was determined by the difference between the mean of the histogram and the expected value (dashed vertical line). The precision was determined by the standard deviation of the histogram.
(Adapted from Hall et al., 2009, with permission from the publisher.)

Unzipping in nucleosome studies

Unzipping is ideally suited for the manipulation of protein-DNA interactions and the detection of their dynamics at the single-molecule level. Below, we provide a very brief summary of various studies on nucleosome structure, positioning, and remodeling which we have explored using our unzipping technique. For more specific details regarding these experiments or data analysis, we refer the reader to the original publications (Dechassa et al., 2011; Hall et al., 2009; Jin et al., 2010; Shundrovsky et al., 2006)

Nucleosome unzipping signature

A nucleosome has the most distinctive signature when unzipped with a loading-rate clamp. As an example, we have unzipped a DNA containing a positioned nucleosome in both forward and reverse directions (Dechassa et al., 2011; Hall et al., 2009). In either direction, two regions of strong interactions are detected, one preceding the dyad and one near the dyad (Figure 1.6 c). When results from both directions are combined, the unzipping force signatures reveal three distinct regions of interactions, one located around the dyad axis and the other two ~ 40 bp on

either side of the dyad axis. Within each region, interactions are discretely spaced with ~ 5 bp periodicity (Figure 1.6 b). By comparison with the crystal structure of the nucleosome, the dyad region should correspond to contacts from the $(H3/H4)_2$ tetramer at superhelical location (SHL) –2.5 to +2.5, and the two off dyad regions should correspond to contacts from the two H2A/H2B dimers between SHL –3.5 to –6.5 and +3.5 to +6.5 respectively. The absence of the last region for each direction of unzipping also indicates that after the first and second regions are disrupted, the nucleosome structure likely becomes unstable and histone dissociation occurs before the last region can be probed. These features are further discussed below.

We have also verified that the unzipping method could clearly distinguish a nucleosome from a tetrasome consisting only of a $(H3/H4)_2$ tetramer (Figure 1.6 b). Unzipping through a tetrasome exhibits only a single region of strong interactions near the dyad and this region substantially overlaps with the dyad region of interactions for canonical nucleosomes (Dechassa et al., 2011) .

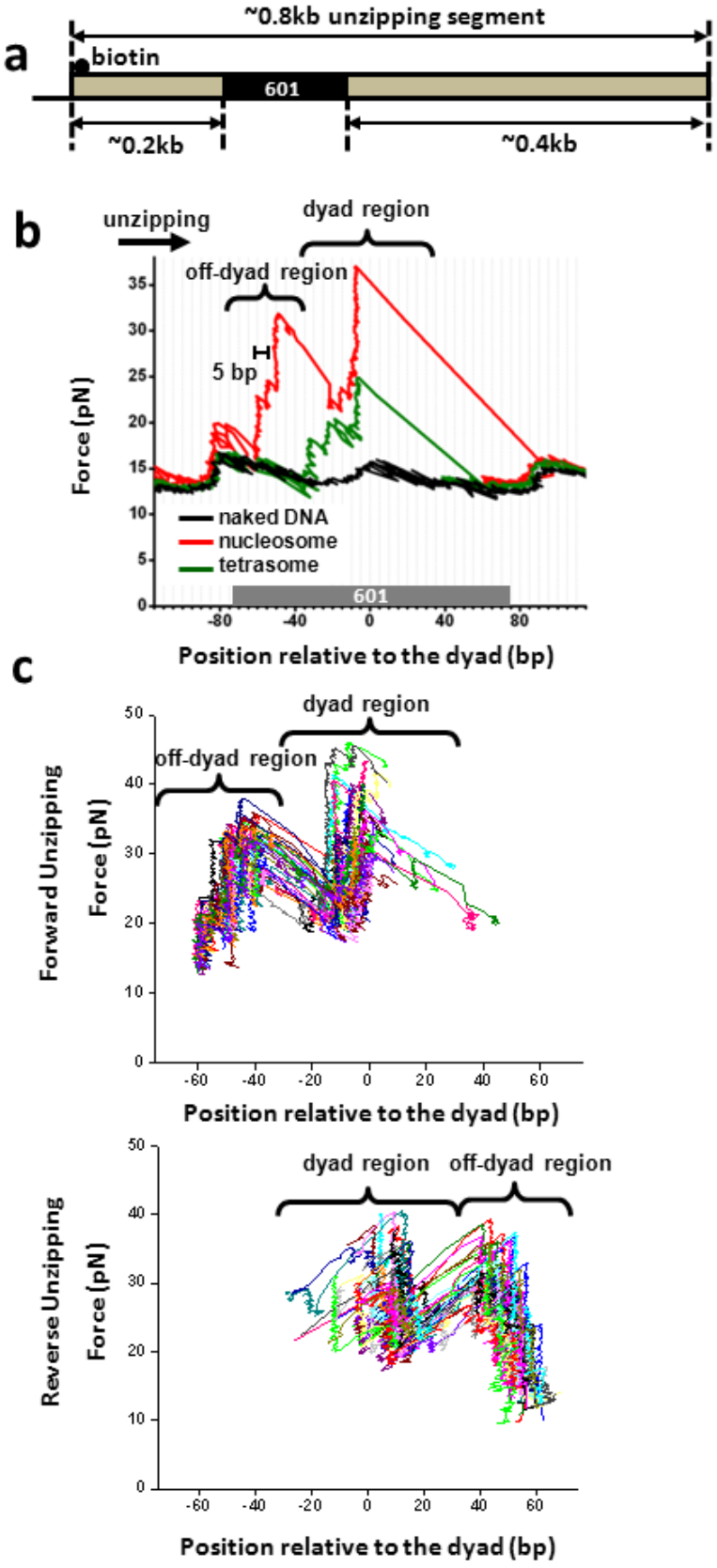

a
~0.8kb unzipping segment
biotin
601
~0.2kb
~0.4kb
b
unzipping
dyad region
off-dyad region
5 bp
Force (pN)
naked DNA
nucleosome
tetrasome
601
Position relative to the dyad (bp)
c
dyad region
off-dyad region
Forward Unzipping
Force (pN)
Position relative to the dyad (bp)
dyad region
off-dyad region
Reverse Unzipping
Force (pN)
Position relative to the dyad (bp)

Figure 1.6. Unzipping through a positioned nucleosome using a loading rate clamp at 8 pN/s. (a) A sketch of the forward nucleosome unzipping segment. (b) Representative force unzipping signatures of naked DNA (black), DNA containing a nucleosome (red), and DNA containing a tetrasome (green). Both the nucleosome and the tetrasome were assembled onto an unzipping segment containing the 601 positioning element. The arrow indicates the unzipping direction. Two distinct regions of interactions, as well as a 5-bp periodicity within each region, were observed for the nucleosome. The tetrasome signature exhibits only a single region of interactions, which substantially overlaps the dyad region identified in the nucleosome. (c) Multiple traces of unzipping through a nucleosome from both forward (upper panel, 31 traces) and reverse (lower panel, 28 traces) directions. Each color represents data obtained from a single nucleosomal DNA molecule. Distinct regions of interactions and a 5-bp periodicity within each region are highly reproducible.

The nucleosome unzipping signature characteristic of a positioned nucleosome is also shared by nucleosomes on arbitrary sequences (Hall et al., 2009). This was demonstrated by assembling nucleosomes onto a DNA segment that does not contain any known positioning elements (Figure 1.7a). The assembly condition was controlled to achieve a relatively low saturation level so that each DNA molecule had at most one nucleosome. When such nucleosomal DNA molecules were unzipped with a loading rate clamp using the same conditions as those of Figure 6, nucleosomes were found at various locations on the template (Figure 1.7b), likely due to a lack of known nucleosome positioning elements on this DNA sequence. Each unzipping trace contains two major regions of strong interaction, with the second region presumably located near the dyad. These nucleosome unzipping signatures possessed essentially identical characteristics to those of the 601 sequence, except that their peak forces within each region were typically smaller by a few pN, reflecting weaker interactions of histone with non-positioning DNA sequences. The key features remained essentially identical: the three regions of strong interactions with the strongest at the dyad, the 5 bp periodicity, and the loss of nucleosome stability upon dyad disruption (Figure 1.7b,c).

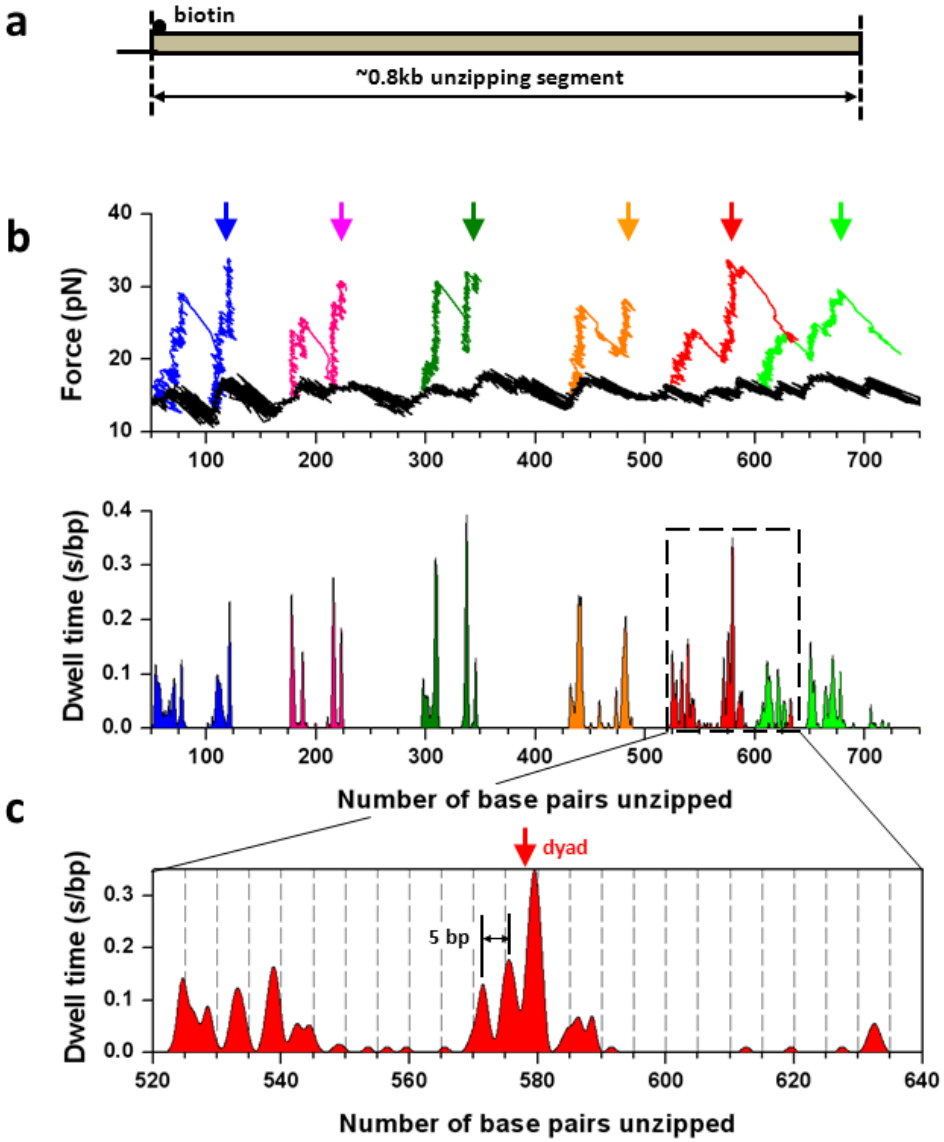


Figure 1.7. Unzipping through a nucleosomes on an arbitrary sequence using a loading rate clamp at 8 pN/s. (a) A sketch of the unzipping segment. (b) Force unzipping signature of a nucleosome at different locations on a DNA template lacking known strong positioning elements. Each color was obtained from a single nucleosome unzipping trace, with the unzipping force shown in the top panel and the corresponding dwell time histogram shown in the bottom panel. The unzipping signature of a naked DNA molecule of the same sequence is also shown (black), as a reference. Vertical arrows indicate the observed dyad locations of these nucleosomes. (c) Close-up of the dwell time histogram for a specific

unzipping trace (red) to emphasize the 5-bp periodicity observed in each interaction region of the unzipping signature.
(Adapted from Hall et al., 2009, with permission from the publisher.)

High resolution mapping of histone-DNA interactions in a nucleosome

To quantitatively assay the strengths of the histone-DNA interactions, we unzipped through individual nucleosomal DNA molecules with a constant unzipping force (Hall et al., 2009). Under a force clamp (Johnson et al., 2007), the dwell times at different sequence positions measure the strengths of interactions at those positions. Thus this method allows direct mapping of the strengths of interactions. Figure 1.8a shows example traces for unzipping DNA through a nucleosome under a constant force. DNA molecules were unzipped from both directions along the DNA. In both cases, the unzipping fork did not move through the nucleosomal DNA at a constant rate but instead dwelled at specific locations within the nucleosome, indicating the presence of strong interactions. In particular, these traces revealed that the fork dwelled with discrete steps spaced by ~5 bp and the longest dwell times tended to occur near the dyad.

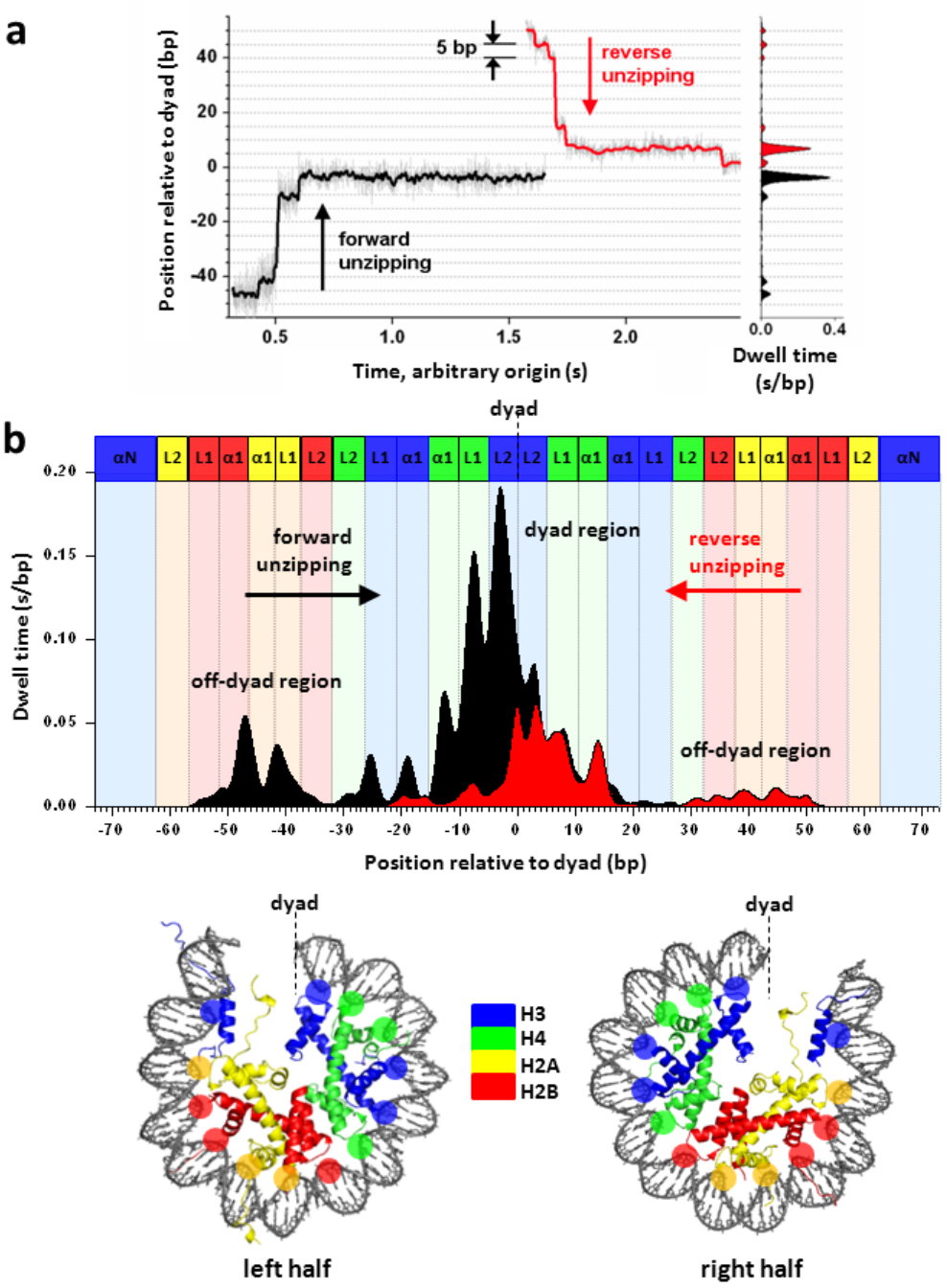


Figure 1.8. Unzipping through a positioned nucleosome using a force clamp at 28 pN. (a) Representative traces of forward (black) and reverse (red) unzipping through a nucleosome under a constant applied force (~ 28 pN). The unzipping fork paused at specific locations when passing through a nucleosome, which are evident from both the traces (left) and their corresponding dwell time histograms (right). (b) A histone–DNA interaction map is constructed by using a total of 27 traces from the forward direction and 30 traces from the reverse direction. Each peak corresponds to an individual histone–DNA interaction and

the heights are indicative of their relative strengths. Three regions of strong interactions are indicated: one located at the dyad and two located off-dyad. The bottom panel is the crystal structure of the nucleosome core particle (Luger et al., 1997), where dots indicate individual histone binding motifs that are expected to interact with DNA. The two halves of the nucleosome are shown separately for clarity. On the top panel, these predicted interactions are shown as colored boxes.
(Adapted from Hall et al., 2009, with permission from the publisher.)

An interaction map was generated by averaging dwell time histogram measurements from many traces from both forward and reverse unzipping, as shown in Figure 1.8b. Several features, consistent with findings using the loading rate clamp, are evident from these plots:

Histone-DNA interactions are highly non-uniform within a nucleosome. There are three broad regions of strong interactions: one located at the dyad and two ~ ±40 bp from the dyad. The locations of all three regions are strongly correlated with those estimated from the crystal structure of the nucleosome (Davey et al., 2002; Luger et al., 1997). The locations of these interactions are also consistent with estimates from our nucleosome stretching experiments (Brower-Toland et al., 2005; Brower-Toland and Wang, 2004; Brower-Toland et al., 2002), although results from those studies are less direct in identifying the absolute locations of strong interactions and are more difficult to interpret.

An ~5 bp periodicity occurs within each region of interaction. According to the crystal structure of the nucleosome, histone core domains are expected to make strong contacts with the DNA minor groove every 10 bp (Davey et al., 2002; Luger et al., 1997). The observed 5 bp periodicity demonstrated that two distinct interactions at each minor groove contact, one from each strand, could be disrupted sequentially rather than simultaneously.

The interactions near the entry and exit DNA are particularly weak. The unzipping fork did not dwell at a 20 bp region of both entry and exit DNA, indicating that the histones are only loosely bound to the DNA. Note that these weaker interactions are detected by the loading rate-clamp described above.

For unzipping in both the forward and reverse directions, the first two regions of interactions encountered were always detected, but not the last region. This indicates that once the dyad region of interactions was disrupted, the nucleosome became unstable and histones dissociated from the 601 sequence.

The total dwell time in the nucleosome was longer in the forward direction compared with that in the reverse direction, indicating nucleosomes were more difficult to disrupt when unzipped in the forward direction, likely reflecting the non-palindromic nature of the 601 sequence.

These mechanical unzipping experiments resemble the action of RNA polymerase which opens up a transcription bubble and unzips the downstream DNA while advancing into a nucleosome. The histone-DNA interaction map has significant implications for how RNA polymerases or other motor proteins may gain access to DNA associated with a nucleosome (Hall et al., 2009; Jin et al., 2010).

Centromeric Nucleosome

The centromere is a specific locus on each chromosome that directs the formation of the kinetochore, a multiprotein complex that interacts with spindle microtubules to promote proper chromosomal alignment and segregation during cell division(Cleveland et al., 2003). The replacement of canonical (major-type) H3 with the centromere-specific histone variant CenH3 (known as CENP-A in humans and Cse4 in budding yeast) is essential for centromere function. Much controversy exists regarding the structural organization of the yeast centromeric nucleosome and the role of the non-histone protein Scm3 in its assembly and architecture. Two rather different types of centromeric nucleosomes (octameric and hexameric) have been proposed for budding yeast (Camahort et al., 2009; Mizuguchi et al., 2007). In the proposed hexameric structure, the two H2A/H2B dimers are replaced by the nonhistone protein Scm3 (Mizuguchi et al., 2007). In fission yeast, H2A/H2B dimers appear to be depleted from the inner centromere regions that contain CenH3 (Mizuguchi et al., 2007). Recent reports on the conformation of DNA in centromeric nucleosomes are also contradictory (Furuyama and Henikoff, 2009; Sekulic et al., 2010). There is evidence from in vitro (Drosophila) and in vivo (S. cerevisiae) experiments that the assembly of centromeric nucleosomes onto plasmid DNA induces positive supercoils (Furuyama and Henikoff, 2009). This suggests that the DNA in centromeric nucleosomes is wrapped around the histone complex in a right-handed, as opposed to the canonical left-handed, superhelix. Because of obvious structural constraints (Dechassa et al., 2011), the composition of these particles might be tetrameric or hexameric(Furuyama and Henikoff, 2009), suggesting a third type of yeast centromeric ‘nucleosome’.

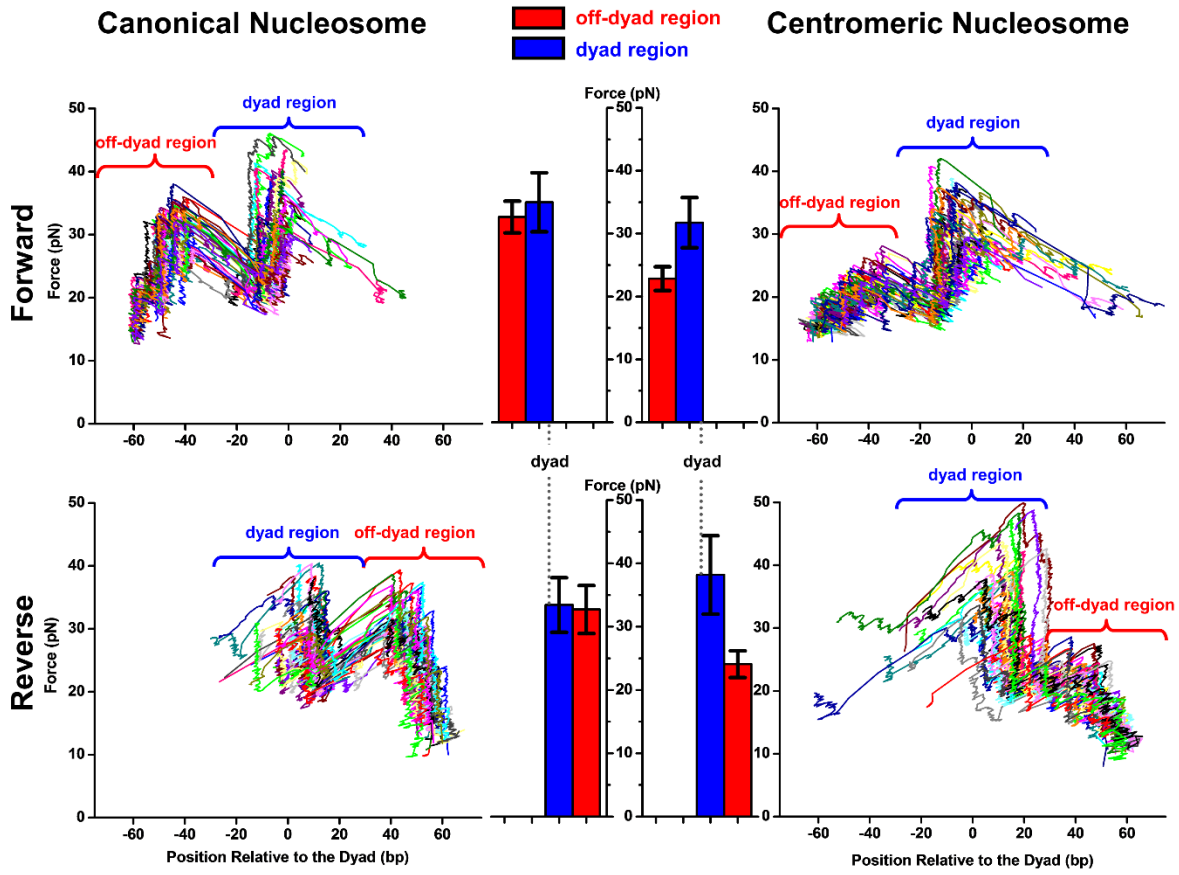


Figure 1.9. Nucleosome disruption signatures under a constant loading rate (8 pN s−1). Representative traces are shown for canonical Sc-nucleosomes (left panel) and Cse4-nucleosomes (right panel), where each colour represents data obtained from a single-nucleosomal DNA molecule. Interaction regions are defined according to the crystal structure27 as super-helical location (SHL) −2.5 to +2.5 (H3/H4 tetramer interaction region) and SHL ±3.5 to ±6.5 (H2A/H2B dimer interaction region). Within the SHL −2.5 to +2.5 region the maximum disruption forces observed are similar between Sc- and Cse4-nucleosomes. Cse4-nucleosomes have significantly weaker interactions throughout the peripheral regions (SHL ±3.5 to ±6.5). A quantitative comparison of the peak forces measured within each interaction region is shown (centre panel). SHL ±3.5 to ±6.5 is shown in red bar and SHL −2.5 to +2.5 shown in blue. Error bars represent s.d. See Supplementary Figures S4–6 for additional data.

To investigate the structural features of *S. cerevisiae* centromeric nucleosomes, we used the unzipping method to provide a quantitative measure of histone-DNA interactions in a centromeric nucleosome. We first examined canonical yeast nucleosomes by unzipping DNA

from either the forward (Figure 1.9, top left) or reverse (Figure 1.9, bottom left) direction along the DNA sequence. When DNA was unzipped from either direction, two clear regions of interactions were detected, one preceding the dyad and one near the dyad. Thus when the results from unzipping from both directions were combined, the unzipping force signatures revealed three distinct regions of interactions, one located around the dyad axis and the other two ~ 40 bp on either side of the dyad axis. Within each region, interactions were discretely spaced with ~ 5 bp periodicity (Figure 1.9, 1.10). These results are indistinguishable from those obtained previously using HeLa histones (Hall et al., 2009) and thus demonstrate that yeast and human canonical nucleosomes have identical histone-DNA interactions as revealed by the unzipping method. Next, we further verified that the unzipping method could clearly distinguish a nucleosome from a tetrasome; unzipping through a DNA molecule containing a tetrasome exhibited only a single region of interactions near the dyad and this region substantially overlapped with the dyad region of interactions for canonical nucleosomes (Figure 1.10).

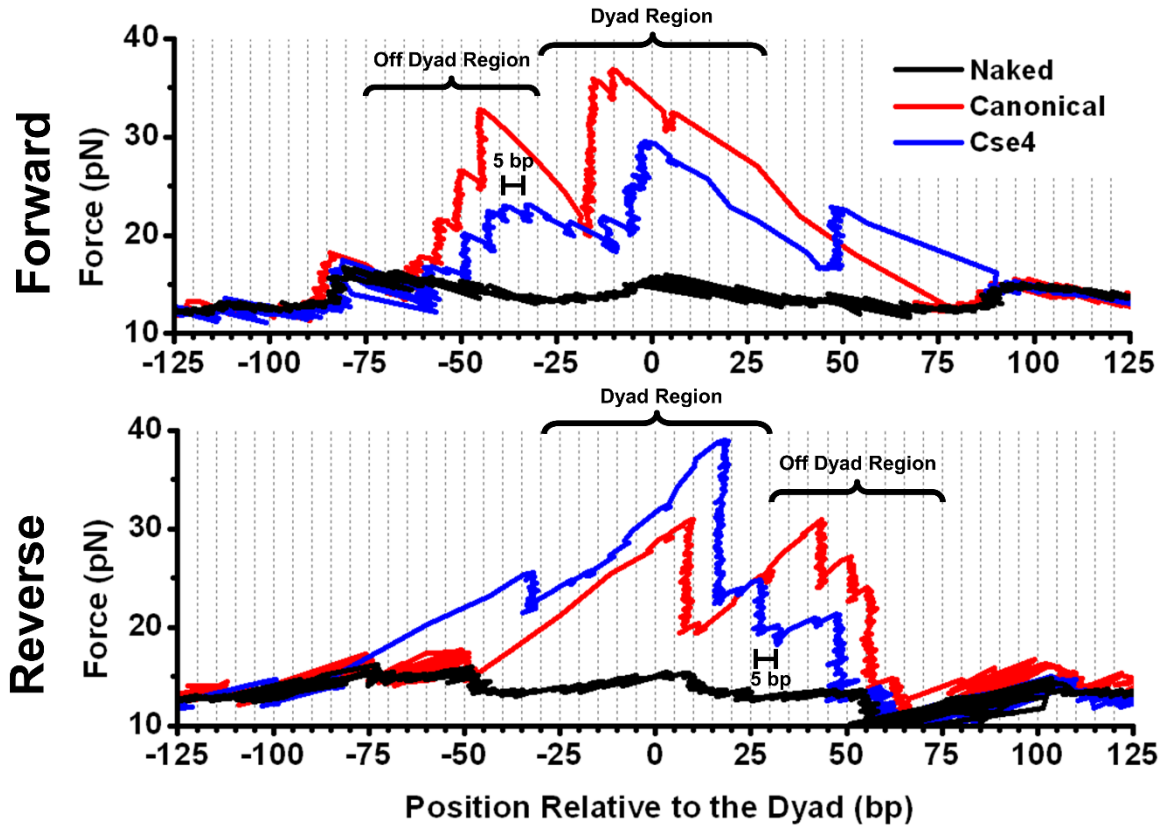


Figure 1.10. The peripheral DNA regions are less well organized in Cse4-nucleosomes Comparison of force signatures for Sc-nucleosome and Cse4-nucleosome. Single representative traces are shown for forward (upper panel) and reverse (lower panel) unzipping of a Sc-nucleosome (red) and Cse4-nucleosome (blue). The baseline unzipping signature of naked DNA (black) is also shown. Data obtained under the same conditions as in Figure 3. Regions are defined as in Figure 1.9. Three distinct regions of interaction were observed for both nucleosomes. The distinct 5 bp periodicity within each region was also apparent, as described previously for HeLa nucleosomes 34. This suggests that, overall, the structure of histone-DNA interactions are conserved between these two types of nucleosomes. However, for a centromeric nucleosome, the outer regions are disrupted at lower forces compared to that of Sc-nucleosome, indicating that its outer turn DNA is more weakly bound.

When Yeast centromeric nucleosomes were similarly examined, and their unzipping signatures resembled their yeast counterparts to a large extent (Figure 1.9, right panels). There existed three regions of interaction and a distinct 5 bp periodicity. These observations suggest that the

centromeric nucleosome contains a full complement of histone octamer. However, the peak forces within the two off dyad regions of strong interactions were significantly lower than those of the canonical nucleosomes. This suggests that the outer turn of DNA in a centromeric nucleosome is significantly more weakly bound to the H2A/B dimers, consistent with an interpretation of the footprinting assay (Dechassa et al., 2011)whereby the outer turn of DNA is transiently accessible to MNase digestion. Interestingly, the dyad region of histone-DNA interactions, which involves direct contributions from Cse4, remains at least as strong as those of canonical nucleosomes. Therefore our data indicate that the overall structure of a centromeric nucleosome resembles that of a canonical nucleosome except for the more loosely bound outer turn DNA.

Conclusions

An increasing interest in the nucleosome as a key regulator of chromatin structure and many cellular processes has inspired the development of a variety of novel techniques, particularly at the single-molecule level (for a recent review, see Killian et al., 2011). The unzipping method detailed here offers high accuracy and precision in locating a nucleosome as well as the ability to elucidate both structural and dynamic features of protein-DNA interactions, complementing more traditional techniques. We anticipate that the unzipping method will continue to play an important role in the study of nucleosome structure, regulation, and remodeling, and is readily extendable to studies of a wide variety of DNA-based activities.

REFERENCES

Andrews, A.J., and Luger, K. (2011). Nucleosome structure(s) and stability: variations on a theme. Annu Rev Biophys *40*, 99-117.

Ashkin, A., Dziedzic, J.M., Bjorkholm, J.E., and Chu, S. (1986). Observation of a single-beam gradient force optical trap for dielectric particles. Opt Lett *11*, 288.

Bannister, A.J., and Kouzarides, T. (2011). Regulation of chromatin by histone modifications. Cell Res *21*, 381-395.

Bell, O., Tiwari, V.K., Thoma, N.H., and Schubeler, D. (2011). Determinants and dynamics of genome accessibility. Nat Rev Genet *12*, 554-564.

Bockelmann, U., Essevaz-Roulet, B., and Heslot, F. (1998). DNA strand separation studied by single molecule force measurements. Phys Rev E *58*, 2386-2394.

Bowman, G.D. (2010). Mechanisms of ATP-dependent nucleosome sliding. Curr Opin Struct Biol *20*, 73-81.

Brower-Toland, B., Wacker, D.A., Fulbright, R.M., Lis, J.T., Kraus, W.L., and Wang, M.D. (2005). Specific contributions of histone tails and their acetylation to the mechanical stability of nucleosomes. J Mol Biol *346*, 135-146.

Brower-Toland, B., and Wang, M.D. (2004). Use of optical trapping techniques to study single-nucleosome dynamics. Methods Enzymol *376*, 62-72.

Brower-Toland, B.D., Smith, C.L., Yeh, R.C., Lis, J.T., Peterson, C.L., and Wang, M.D. (2002). Mechanical disruption of individual nucleosomes reveals a reversible multistage release of DNA. Proc Natl Acad Sci U S A *99*, 1960-1965.

Camahort, R., Shivaraju, M., Mattingly, M., Li, B., Nakanishi, S., Zhu, D., Shilatifard, A., Workman, J.L., and Gerton, J.L. (2009). Cse4 is part of an octameric nucleosome in budding yeast. Mol Cell *35*, 794-805.

Campos, E.I., and Reinberg, D. (2009). Histones: annotating chromatin. Annu Rev Genet *43*, 559-599.

Clapier, C.R., and Cairns, B.R. (2009). The biology of chromatin remodeling complexes. Annu Rev Biochem *78*, 273-304.

Cleveland, D.W., Mao, Y., and Sullivan, K.F. (2003). Centromeres and kinetochores: from epigenetics to mitotic checkpoint signaling. Cell *112*, 407-421.

Das, C., Tyler, J.K., and Churchill, M.E. (2010). The histone shuffle: histone chaperones in an energetic dance. Trends Biochem Sci *35*, 476-489.

Davey, C.A., Sargent, D.F., Luger, K., Maeder, A.W., and Richmond, T.J. (2002). Solvent mediated interactions in the structure of the nucleosome core particle at 1.9 a resolution. J Mol Biol *319*, 1097-1113.

Dechassa, M.L., Wyns, K., Li, M., Hall, M.A., Wang, M.D., and Luger, K. (2011). Structure and Scm3-mediated assembly of budding yeast centromeric nucleosomes. Nat Commun *2*, 313.

Dyer, P.N., Edayathumangalam, R.S., White, C.L., Bao, Y., Chakravarthy, S., Muthurajan, U.M., and Luger, K. (2004). Reconstitution of nucleosome core particles from recombinant histones and DNA. Methods Enzymol *375*, 23-44.

Evans, E. (2001). Probing the relation between force--lifetime--and chemistry in single molecular bonds. Annu Rev Biophys Biomol Struct *30*, 105-128.

Forties, R.A., North, J.A., Javaid, S., Tabbaa, O.P., Fishel, R., Poirier, M.G., and Bundschuh, R. (2011). A quantitative model of nucleosome dynamics. Nucleic Acids Res *39*, 8306-8313.

Furuyama, T., and Henikoff, S. (2009). Centromeric nucleosomes induce positive DNA supercoils. Cell *138*, 104-113.

Fyodorov, D.V., and Kadonaga, J.T. (2003). Chromatin assembly in vitro with purified recombinant ACF and NAP-1. Methods Enzymol *371*, 499-515.

Gansen, A., Hauger, F., Toth, K., and Langowski, J. (2007). Single-pair fluorescence resonance energy transfer of nucleosomes in free diffusion: optimizing stability and resolution of subpopulations. Anal Biochem *368*, 193-204.

Gemmen, G.J., Sim, R., Haushalter, K.A., Ke, P.C., Kadonaga, J.T., and Smith, D.E. (2005). Forced unraveling of nucleosomes assembled on heterogeneous DNA using core histones, NAP-1, and ACF. J Mol Biol *351*, 89-99.

Hall, M.A., Shundrovsky, A., Bai, L., Fulbright, R.M., Lis, J.T., and Wang, M.D. (2009). High-resolution dynamic mapping of histone-DNA interactions in a nucleosome. Nat Struct Mol Biol *16*, 124-129.

Jiang, J., Bai, L., Surtees, J.A., Gemici, Z., Wang, M.D., and Alani, E. (2005). Detection of high-affinity and sliding clamp modes for MSH2-MSH6 by single-molecule unzipping force analysis. Mol Cell *20*, 771-781.

Jin, J., Bai, L., Johnson, D.S., Fulbright, R.M., Kireeva, M.L., Kashlev, M., and Wang, M.D. (2010). Synergistic action of RNA polymerases in overcoming the nucleosomal barrier. Nat Struct Mol Biol *17*, 745-752.

Johnson, D.S., Bai, L., Smith, B.Y., Patel, S.S., and Wang, M.D. (2007). Single-molecule studies reveal dynamics of DNA unwinding by the ring-shaped T7 helicase. Cell *129*, 1299-1309.

Joo, C., Balci, H., Ishitsuka, Y., Buranachai, C., and Ha, T. (2008). Advances in single-molecule fluorescence methods for molecular biology. Annu Rev Biochem *77*, 51-76.

Kaplan, N., Moore, I.K., Fondufe-Mittendorf, Y., Gossett, A.J., Tillo, D., Field, Y., LeProust, E.M., Hughes, T.R., Lieb, J.D., Widom, J., *et al.* (2009). The DNA-encoded nucleosome organization of a eukaryotic genome. Nature *458*, 362-366.

Killian, J.L., Li, M., Sheinin, M.Y., and Wang, M.D. (2011). Recent advances in single molecule studies of nucleosomes. Curr Opin Struct Biol.

Koch, S.J., Shundrovsky, A., Jantzen, B.C., and Wang, M.D. (2002). Probing protein-DNA interactions by unzipping a single DNA double helix. Biophys J *83*, 1098-1105.

Koch, S.J., and Wang, M.D. (2003). Dynamic force spectroscopy of protein-DNA interactions by unzipping DNA. Phys Rev Lett *91*, 028103.

Korber, P., and Becker, P.B. (2010). Nucleosome dynamics and epigenetic stability. Essays Biochem *48*, 63-74.

Lowary, P.T., and Widom, J. (1998). New DNA sequence rules for high affinity binding to histone octamer and sequence-directed nucleosome positioning. J Mol Biol *276*, 19-42.

Luger, K., Mader, A.W., Richmond, R.K., Sargent, D.F., and Richmond, T.J. (1997). Crystal structure of the nucleosome core particle at 2.8 A resolution. Nature *389*, 251-260.

Luger, K., Rechsteiner, T.J., and Richmond, T.J. (1999). Preparation of nucleosome core particle from recombinant histones. Methods Enzymol *304*, 3-19.

Mihardja, S., Spakowitz, A.J., Zhang, Y., and Bustamante, C. (2006). Effect of force on mononucleosomal dynamics. Proc Natl Acad Sci U S A *103*, 15871-15876.

Mizuguchi, G., Xiao, H., Wisniewski, J., Smith, M.M., and Wu, C. (2007). Nonhistone Scm3 and histones CenH3-H4 assemble the core of centromere-specific nucleosomes. Cell *129*, 1153-1164.

Moffitt, J.R., Chemla, Y.R., Smith, S.B., and Bustamante, C. (2008). Recent advances in optical tweezers. Annu Rev Biochem *77*, 205-228.

Park, Y.J., and Luger, K. (2008). Histone chaperones in nucleosome eviction and histone exchange. Curr Opin Struct Biol *18*, 282-289.

Ransom, M., Dennehey, B.K., and Tyler, J.K. (2010). Chaperoning histones during DNA replication and repair. Cell *140*, 183-195.

Ray-Gallet, D., and Almouzni, G. (2010). Nucleosome dynamics and histone variants. Essays Biochem *48*, 75-87.

Schafer, D.A., Gelles, J., Sheetz, M.P., and Landick, R. (1991). Transcription by single molecules of RNA polymerase observed by light microscopy. Nature *352*, 444-448.

Sekulic, N., Bassett, E.A., Rogers, D.J., and Black, B.E. (2010). The structure of (CENP-A-H4)(2) reveals physical features that mark centromeres. Nature *467*, 347-351.

Shundrovsky, A., Smith, C.L., Lis, J.T., Peterson, C.L., and Wang, M.D. (2006). Probing SWI/SNF remodeling of the nucleosome by unzipping single DNA molecules. Nat Struct Mol Biol *13*, 549-554.

Simon, M., North, J.A., Shimko, J.C., Forties, R.A., Ferdinand, M.B., Manohar, M., Zhang, M., Fishel, R., Ottesen, J.J., and Poirier, M.G. (2011). Histone fold modifications control nucleosome unwrapping and disassembly. Proc Natl Acad Sci U S A *108*, 12711-12716.

Thastrom, A., Lowary, P.T., and Widom, J. (2004). Measurement of histone-DNA interaction free energy in nucleosomes. Methods *33*, 33-44.

Wang, M.D., Yin, H., Landick, R., Gelles, J., and Block, S.M. (1997). Stretching DNA with optical tweezers. Biophys J *72*, 1335-1346.

Zhang, Y., Moqtaderi, Z., Rattner, B.P., Euskirchen, G., Snyder, M., Kadonaga, J.T., Liu, X.S., and Struhl, K. (2009). Intrinsic histone-DNA interactions are not the major determinant of nucleosome positions in vivo. Nat Struct Mol Biol *16*, 847-852.

CHAPTER 2

DYNAMIC REGULATION OF TRANSCRIPTION FACTORS BY NUCLEOSOME REMODELING

The chromatin landscape and promoter architecture are dominated by the interplay of nucleosome and transcription factor (TF) binding to crucial DNA sequence elements. However, it remains unclear whether nucleosomes mobilized by chromatin remodelers can influence TFs that are already present on the DNA template. In this chapter, we investigated the interplay between nucleosome remodeling, by either yeast ISW1a or SWI/SNF, and a bound TF. We found that a TF serves as a major barrier to ISW1a remodeling, and acts as a boundary for nucleosome repositioning. In contrast, SWI/SNF was able to slide a nucleosome past a TF, with concurrent eviction of the TF from the DNA, and the TF did not significantly impact the nucleosome positioning. The results provide direct evidence for a novel mechanism for both nucleosome positioning regulation by bound TFs and TF regulation via dynamic repositioning of nucleosomes.

INTRODUCTION

Dynamic access to specific genetic information is critical for cellular development and response to the environment. Thus, processes such as transcription must be mediated by mechanisms that regulate gene function rapidly and reliably (Barrera and Ren, 2006; Kornberg, 2007). In eukaryotic cells, proper transcriptional regulation depends upon transcription factors (TFs) that bind to specific DNA binding sites (Kadonaga, 2004). Additionally, the repression of transcription has often been correlated with the presence of nucleosomes, the basic units of chromatin structure, in which histone-DNA interactions act as a barrier for RNA polymerase elongation along DNA (Li et al., 2007; Petesch and Lis, 2012; Teves et al., 2014). Therefore,

understanding the relationship between TF binding and nucleosomes is essential in understanding gene expression and regulation (Voss and Hager, 2014).

Chromatin landscape and promoter architecture are dominated by the interplay of nucleosome and TF binding. Nucleosomes and TFs have been shown to compete for binding to DNA (Lickwar et al., 2012; Mirny, 2010; Moyle-Heyrman et al., 2011). This competition is based on the respective affinities of the TF and nucleosome for DNA, and depends upon both DNA sequence and histone modifications and variants. However, a nucleosome may also be repositioned through the action of chromatin remodelers, suggesting additional levels of transcription regulation. Some TFs are known to recruit nucleosome remodelers. Earlier studies focused on how these recruiting TFs affect the outcomes of nucleosome remodeling (Dechassa et al., 2010; Nagaich et al., 2004; Voss et al., 2011). However, most TFs identified so far have not been found to recruit nucleosome remodelers and the interplay between nucleosome remodeling and these TFs remains unclear. We hypothesize that nucleosome remodeling, without remodeler recruitment, may regulate the state of a bound TF. Specifically, a remodeler may attempt to move a nucleosome to or through a site pre-occupied by a TF. During such an encounter, the TF may be displaced, or it may act as a roadblock for nucleosome remodeling. Thus, chromatin remodeling may serve as an alternative mechanism to regulate transcription through its influence on a bound TF, and a bound TF may dictate the location of a remodeled nucleosome.

In this chapter, I studied the influence of nucleosome remodeling on a bound TF in a single molecule assay. I used a DNA unzipping technique (Dechassa et al., 2011; Hall et al., 2009; Inman et al., 2014; Jiang et al., 2005; Jin et al., 2010; Li and Wang, 2012; Shundrovsky et al.,

2006) to characterize the locations of a bound TF and a nucleosome simultaneously on long DNA templates to near base pair accuracy. By examining the remodeling behavior upon encountering a bound TF, I determined that the relationship between TFs and nucleosome remodeling not only plays a critical role in nucleosome positioning, but also reveals a novel mechanism for how a TF can be dynamically recycled by nucleosome remodeling.

MATERIALS and METHODS

Plasmids

The plasmids containing the Gal4 binding site and the 601NPE with varied distances were prepared using standard PCR and cloning methods. The cloning segments were generated by standard PCR from the 601 plasmid (Lowary and Widom, 1998) using special primers, one of which contains one Gal4 binding site. The distance between the primer containing the Gal4 binding site and the 601NPE determines the distance between the Gal4 binding site and 601NPE. Then, the PCR product was cloned into the pDrive vector (Qiagen). The finished constructs were confirmed by DNA sequencing.

Nucleosome unzipping template

Nucleosomal DNA templates were prepared using methods similar to those previously described (Koch et al., 2002; Li and Wang, 2012). Briefly, each DNA construct consisted of two separate segments. A ~1.1 kbp anchoring segment was amplified, by PCR, from plasmid pRL574 using a

digoxigenin-labeled primer and then subsequently digested with *Bst*XI (NEB) to produce an overhang. The unzipping templates were amplified, by PCR, from the plasmids described above and amplified with a biotin-labeled primer, digested with *Bst*XI, and dephosphorylated using CIP (NEB) to introduce a nick into the final DNA template. Nucleosomes were assembled from purified HeLa histones onto the unzipping fragment by a well-established salt dialysis method (Lee and Narlikar, 2001). The two segments were joined by ligation immediately prior to use. This produced a complete template labeled with a single dig tag on one end and a biotin tag located 7 bp after the nick in one DNA strand.

Nucleosome remodeling reaction

yISW1a and ySWI/SNF were purified as previously described (Dechassa et al., 2008; Gangaraju and Bartholomew, 2007). yGal4DBD was purchased from Santa Cruz Biotechnology, Inc. After the ligation of the anchoring segment and unzipping segment containing a nucleosome, we incubated 20 nM of the nucleosomal DNA with 200 nM Gal4DBD at 16°C for 30min. Single molecule sample preparation was performed according to protocols previously described (Li and Wang, 2012). The remodeling experiments were conducted in a sample chamber after tethers are formed. SWI/SNF remodeling reactions contained 1.5 nM purified ySWI/SNF, and 1mM ATP in the SWI/SNF remodeling buffer (10 mM Tris·Cl, pH 8.0, 100 mM NaCl, 7 mM $MgCl_2$, 2 mM DTT, 0.1 mg/mL acBSA). ISW1a remodeling reactions contained 1.5 nM purified yISW1a, and 1mM ATP in the ISW1a remodeling buffer (30mM HEPES, pH 7.6, 3mM $MgCl_2$, 5mM NaCl, 0.1mM EGTA, 0.02mM EDTA, 5% glycerol, 0.2 mg/mL acBSA). Both types of remodeling reactions were incubated at 25 ^oC with duration specified. The reactions were stopped by the

addition of 10 mM EDTA and 0.25 mg/ml Salmon Sperm DNA and incubation for 5 min at 25 ^{o}C. Finally, the sample chamber was rinsed with 100 μL sample buffer (10 mM Tris·Cl pH 7.5, 1 mM EDTA, 100 mM NaCl, 1.5 mM $MgCl_2$, 1 mM DTT, 3 % (v/v) glycerol, 0.02 % (v/v) Tween 20, and 2 mg/mL BSA). All single molecule measurements were performed in this sample buffer.

Data collection and alignment

An optical trapping setup as previously described (Brower-Toland and Wang, 2004) was used to unzip a single DNA molecule by moving the microscope coverslip horizontally away from an optical trap. The unzipping methods have been previously described (Li and Wang, 2012) and briefly summarized here. Whenever the unzipping fork encountered an interaction that prevented the fork progression, the unzipping force was ramped up linearly with time (15 pN/s) until the interaction was disrupted. When two interactions occurred in close vicinity, upon the disruption of the first interaction the force was unable to relax back to the baseline before being ramped up again for the second interaction, subjecting this subsequent interaction to a higher initial force. Therefore, for each region of interactions, the dwell time histogram highlighted the edge of the region first encountered. Another feature of this method was the display of the distinctive force signature for a nucleosome, allowing for robust identification of the nucleosome structure.

Data were low pass filtered to 5 kHz, digitized at ~12 kHz, and later low passed filtered to 60 Hz. The precision and accuracy of the experimental curves were improved to near base pair

level by cross-correlation of regions immediately before the Gal4DBD disruption and after the nucleosome disruption, using methods as previously described (Hall et al., 2009; Li and Wang, 2012). For the experimental curves where the nucleosomes are located at the end of the template, the cross-correlation was carried out for a region immediately before the Gal4DBD disruption or nucleosome disruption. To account for minor instrumental drift, trapping bead size variations, and DNA linker variations, the alignment allowed for a small additive shift (~ 10 bp) and multiplicative linear stretch (< 2%) using algorithms similar to those previously described (Hall et al., 2009).

Determination of locations of a TF and a nucleosome

Gal4DBD showed a distinct unzipping signature with a single force peak at 8 bp from the center of the consensus sequence, indicating the front end of the Gal4DBD footprint on the DNA. The disruption force peak was 18-20 pN, significantly larger than the baseline force of ~15 pN. Therefore we determined the center position of a bound Gal4DBD by first detecting the peak force location and then shifting this location by 8 bp in the direction of unzipping.

The positioned nucleosome displayed a much more complex force signature with multiple force peaks and a significantly greater overall force, reflecting the multiple finer and stronger histone-DNA interactions within a nucleosome (Hall et al., 2009; Shundrovsky et al., 2006). We determine the dyad position of a nucleosome by first measuring mean force location within the first force cluster and then shifting this position by 43 bp in the direction of the unzipping

RESULTS

Precise determination of the position of a transcription factor and a nucleosome

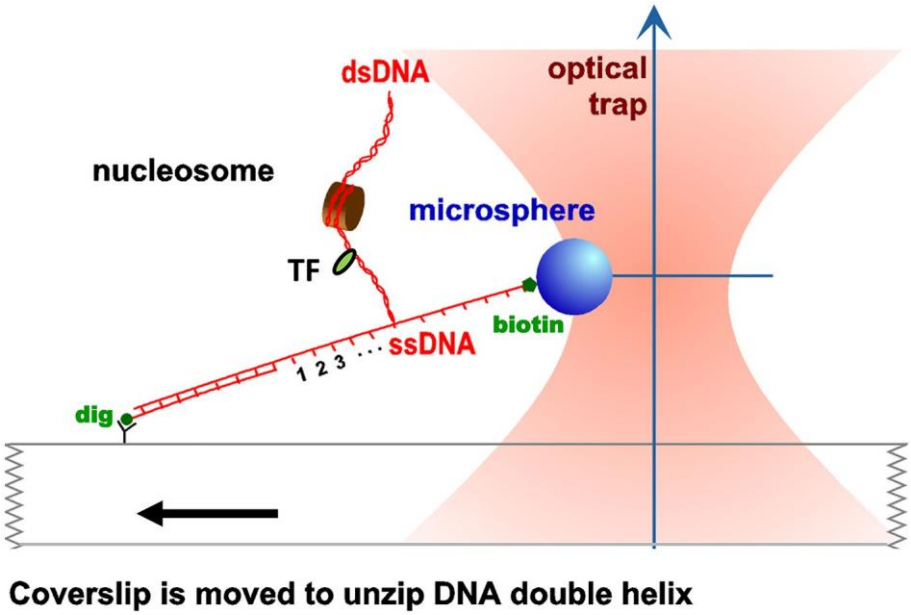


Figure 2.1. Unzipping experimental configuration. The DNA template was attached, at one end, to the surface of a glass coverslip via a digoxigenin–antidigoxigenin linkage, and at its other end to a microsphere via a biotin–streptavidin linkage. As the coverslip was moved away from the trapped microsphere, using a loading-rate clamp, the dsDNA was sequentially converted into ssDNA upon base pair separation. The presence of force peaks above the naked DNA baseline revealed the detected locations of protein–DNA interactions.

In this work, we needed to precisely locate the positions of a nucleosome and a TF before and after nucleosome remodeling. we thus employed the DNA unzipping technique, which has been demonstrated to be a powerful single molecule technique for accurate and precise determination of positions and strengths of DNA-protein interactions (Dechassa et al., 2011; Hall et al., 2009; Inman et al., 2014; Jiang et al., 2005; Jin et al., 2010; Li and Wang, 2012; Shundrovsky et al., 2006). To evaluate the precision of this approach, we constructed a DNA template containing a

single Gal4 sequence for binding to the Gal4 DNA binding domain (Gal4DBD) and a single 601 nucleosome positioning sequence (601NPE) for uniquely positioning a nucleosome (Figure 2.1). Gal4DBD contains only the 147 amino acids of the N terminal domain of the Gal4 protein, and does not have any known remodeler recruitment function.

Figure 2.2 shows representative traces from unzipping DNA molecules without nucleosome remodeling. The top trace of Figure 1 shows the result when the DNA template was unzipped starting from the Gal4 side. Both Gal4DBD (a single smaller peak) and a nucleosome (two clusters of larger peaks) were readily detected above the baseline of the corresponding naked DNA. The bottom trace shows the result when the DNA template was unzipped starting from the nucleosome side. Although the nucleosome unzipping signature was readily detectable, the unzipping signature of Gal4DBD was sometimes masked by that of the nucleosome. Therefore, it was often necessary to carry out unzipping experiments from both directions. Analysis of these unzipping signatures confirmed that unzipping mapped the position of the TF and the nucleosome to near base pair precision (Figure 2.3)

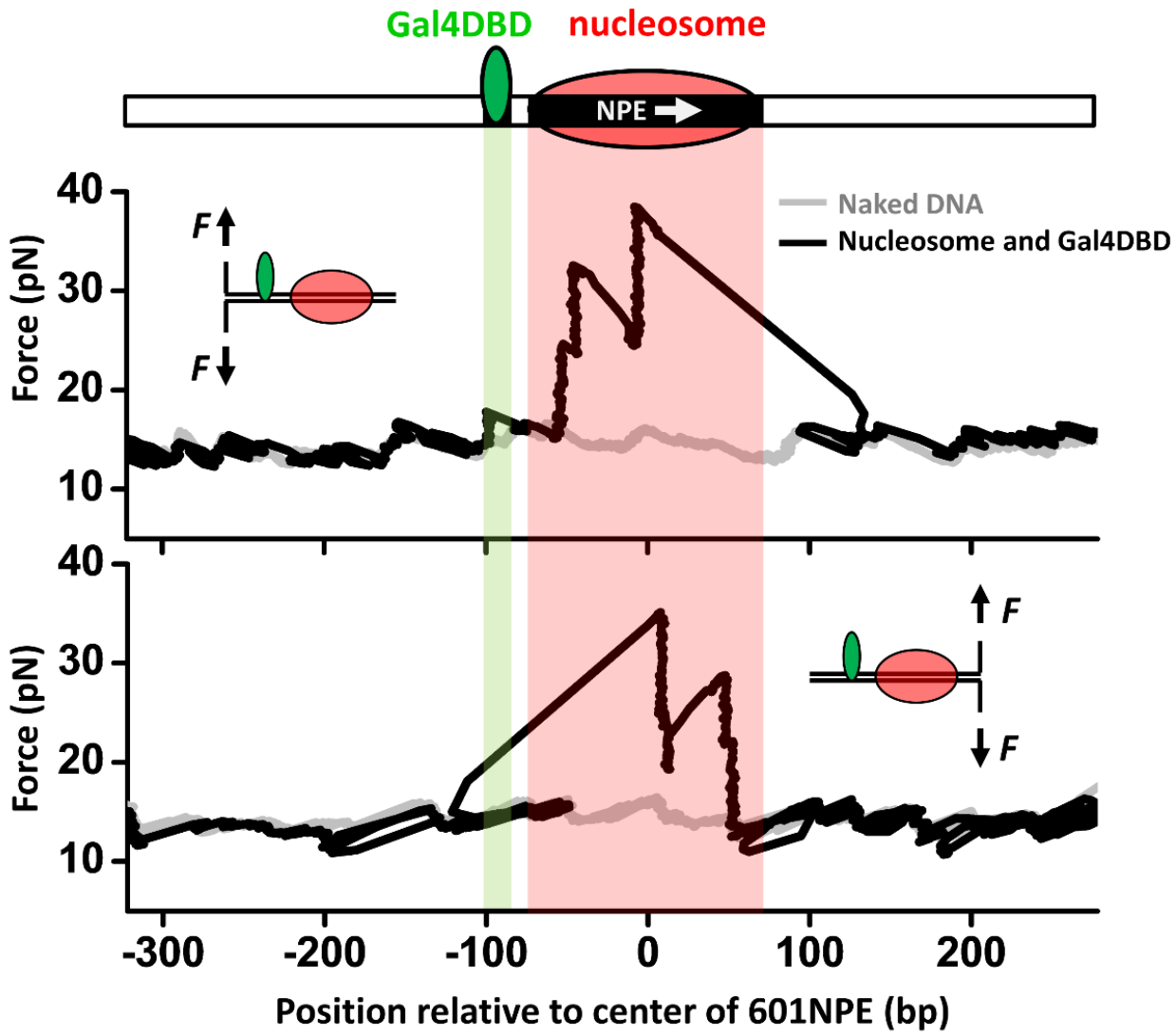


Figure 2.2. Single molecule unzipping technique detects Gal4DBD and nucleosome to near base-pair accuracy. DNA molecules, each containing a nucleosome and a bound Gal4DBD, were unzipped. All unzipped DNA molecules used in this work were in the region of 600bp to 1.2kbp. For clarity, much smaller regions are shown in all figures, with the origin of a template sequence defined as center position (the dyad) of the 601NPE. Shaded regions indicate locations of the Gal4 binding sequence and the 601NPE.

(top panel) Cartoon illustrating the unzipping template design using for this experiment. A Gal4 sequence was separated from a 601NPE by 10 bp. The orientation of the 601NPE sequence is indicated by a white arrow.

(middle panel) Unzipping in the direction in which the bound Gal4DBD was encountered first. (bottom panel) Unzipping in the direction in which the nucleosome was encountered first.

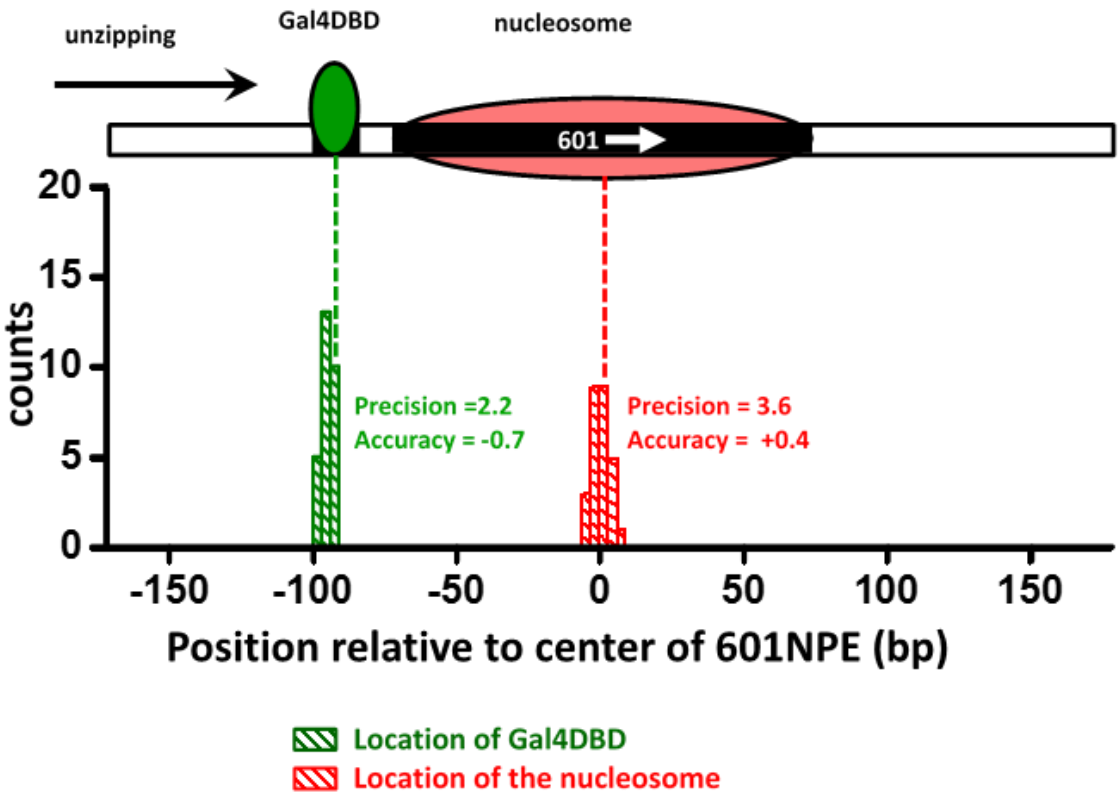


Figure 2.3. Characterization of the precision and accuracy of detection of the locations of Gal4DBD and nucleosome. Single molecule unzipping detected Gal4DBD and a nucleosome simultaneously. The histograms for detected locations of Gal4DBD (green) and nucleosome (red) were obtained by pooling data from multiple single molecule traces, with the expected bound locations represented by their respective dashed lines. For each histogram, the precision was determined by the standard deviation of each histogram, and the accuracy by the difference between the mean of the histogram and the expected value (the vertical dashed line). These data demonstrate both the precision and accuracy to be near base-pair.

These unzipping experiments also revealed tight binding of Gal4DBD to its recognition sequence and slow dissociation. Under our experimental conditions, the equilibrium dissociation constant of Gal4DBD was determined to be 3.4 nM (Figure 2.4). Our experiments were carried out with 95% of Gal4 sites bound to Gal4DBD. In addition, the bound Gal4DBD's lifetime was much

longer than 1 hour (the typical duration of a single molecule experiment) (Figure 2.4). For all experiments involving Gal4DBD, including those in Figure 1, Gal4DBD was allowed to equilibrate with the DNA, and remaining free Gal4DBD was then flushed from the sample chamber. Thus subsequent remodeling reactions were carried out without free Gal4DBD in solution.

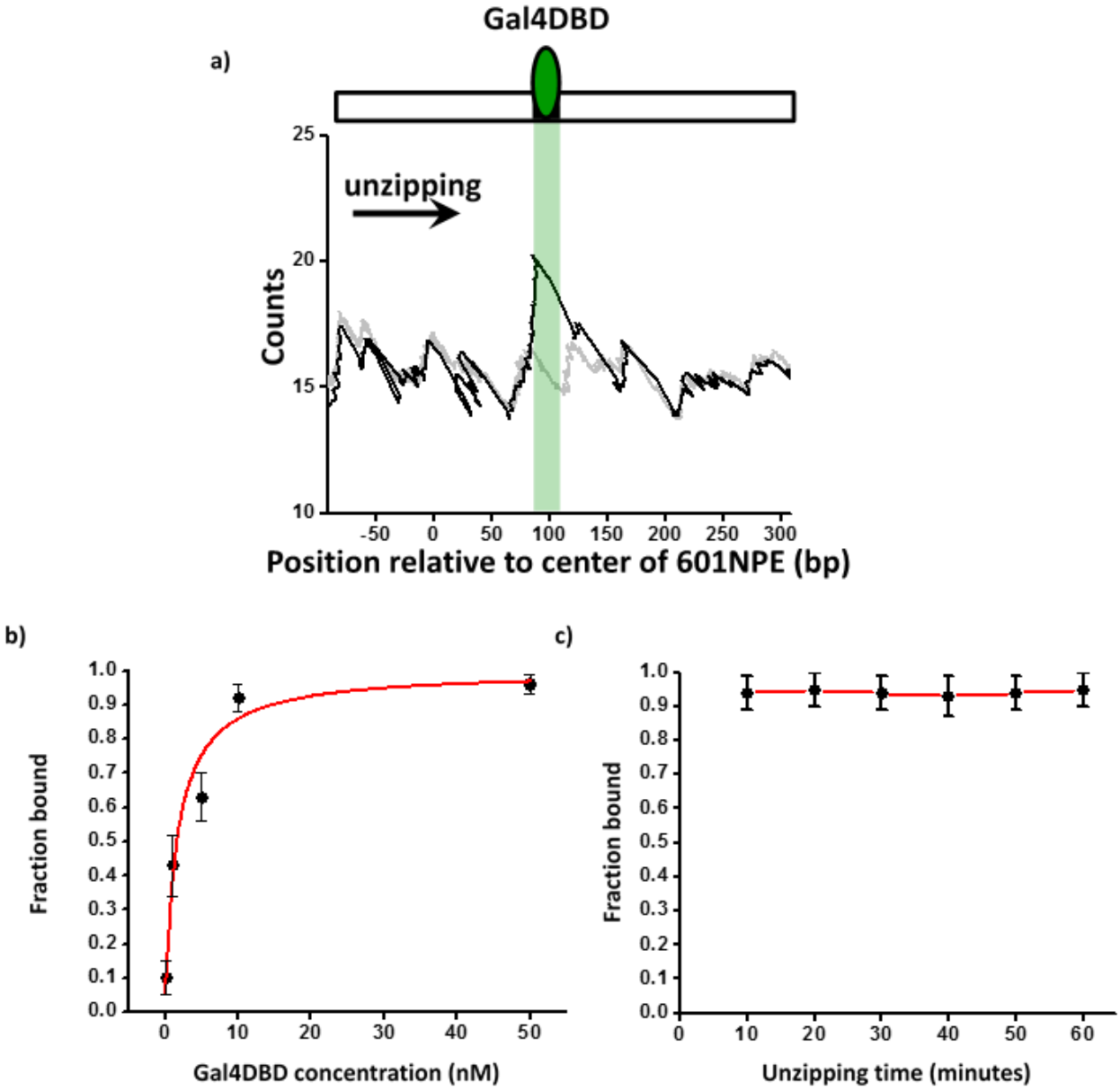


Figure 2.4. Characterization of Gal4DBD binding. To characterize Gal4DBD binding to its binding sequence, DNA unzipping was carried out in the presence of known Gal4DBD concentrations.

a) A representative unzipping trace of a bound Gal4DBD. The location of the binding sequence is shown as a shaded region. Naked DNA unzipping baseline is shown in grey.
b) Fraction of bound Gal4DBD versus the concentration of Gal4DBD. For a given concentration of Gal4DBD, measurements were on multiple DNA molecules to obtain the fraction of Gal4DBD. Data points are represented as (mean ± s.e.m.). The relation for the fraction bound versus [Gal4DBD] was fit to: $\frac{[\text{Gal4DBD}]}{[\text{Gal4DBD}]+K_d}$ (red smooth curve, which yielded the dissociation equilibrium constant K_d = 3.4 nM.
c) Fraction of bound Gal4DBD versus time. This relation shows no significant Gal4DBD dissociation from its binding sequence over a course of one hour. Data were fit to a straight line to guide the eye.

Gal4DBD affects the directionality of ISW1a and SWI/SNF remodeling differently

Chromatin remodelers utilize ATP hydrolysis to move nucleosomes by altering histone-DNA interactions, with the two major families of chromatin remodelers, ISWI and SWI/SNF, differing in their outcomes of the remodeling reaction (Clapier and Cairns, 2009). SWI/SNF family remodelers are known to be associated with nucleosome disruption(Aoyagi et al., 2002; Imbalzano et al., 1996; Logie and Peterson, 1997) and transcriptional activation (Gkikopoulos et al., 2011; Hassan et al., 2001; Kwon et al., 1994); while ISWI family remodelers have been shown to contribute to the formation of evenly spaced nucleosome arrays (Fyodorov and Kadonaga, 2002; Lusser et al., 2005; Torigoe et al., 2011; Tsukiyama et al., 1999) and transcriptional repression(Whitehouse et al., 2007; Yadon et al., 2010). On mononucleosome substrates, many ISWI remodelers have been shown to be sensitive to naked DNA segments flanking the nucleosome, preferentially sliding the nucleosome towards the longer segment of DNA (Blosser et al., 2009; Deindl et al., 2013; Yang et al., 2006). This sensitivity to linker DNA

is believed to underlie their ability to generate evenly-spaced nucleosomal arrays (Gelbart et al., 2001; Stockdale et al., 2006). SWI/SNF remodelers, on the other hand, can shift a histone octamer up to 50 bp off the end of a short DNA fragment (Kassabov et al., 2003). On dinucleosomal templates, SWI/SNF remodelers have been found to shift one nucleosome onto another, indicating nucleosome disruption and eviction characteristics of these remodelers (Dechassa et al., 2010; Engeholm et al., 2009).

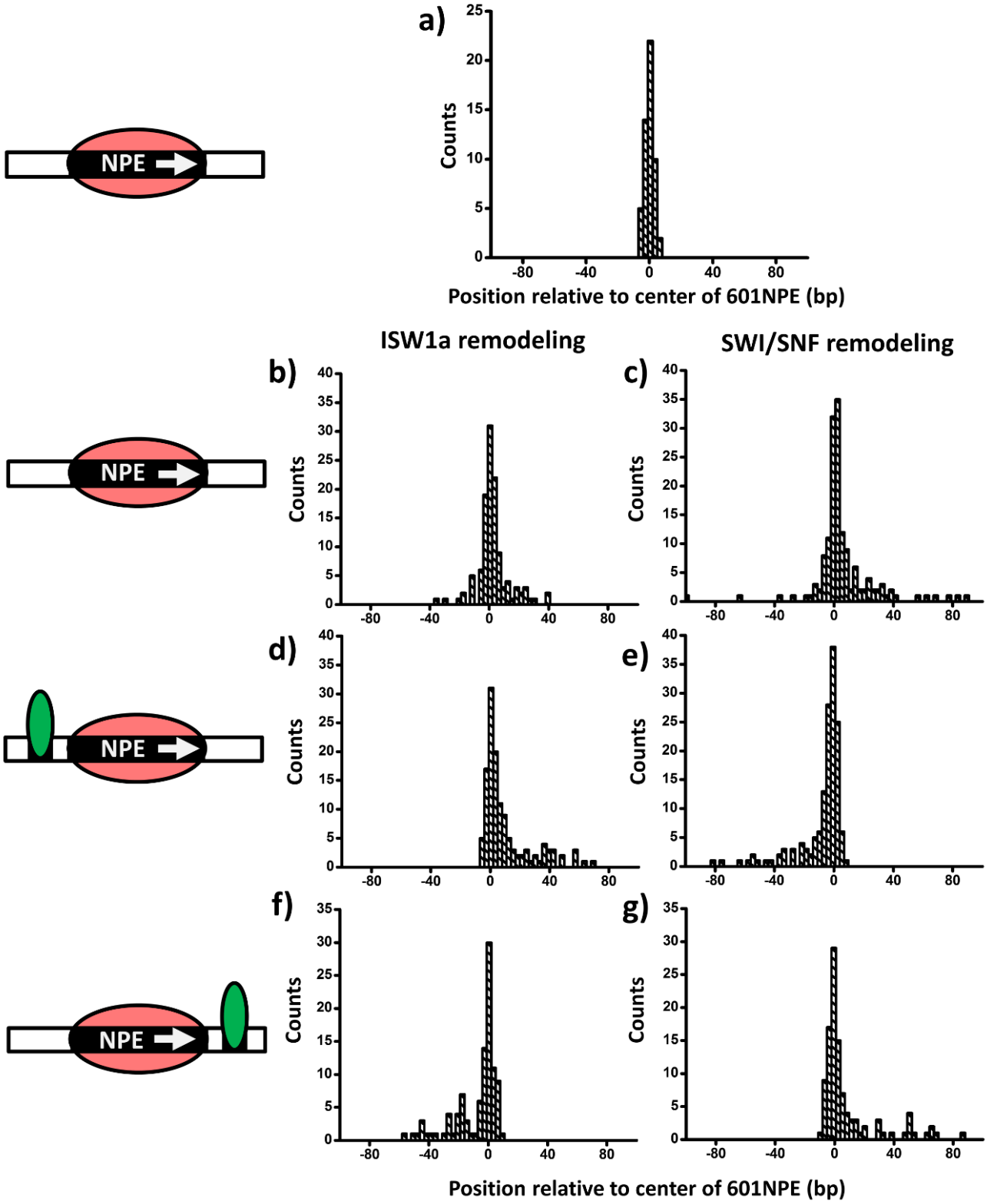

a)
NPE
Counts
Position relative to center of 601NPE (bp)
ISW1a remodeling
SWI/SNF remodeling
b)
c)
d)
e)
f)
g)

Figure 2.5. A bound Gal4DBD affects the directionality of SWI/SNF remodeling and ISW1a remodeling differently. Nucleosomes were remodeled by either 1 nM ISW1a or 1 nM SWI/SNF with 1 mM ATP for 1 min, a time sufficiently short that the majority of nucleosomes were not remodeled (Supplementary Figure 5). Each DNA template was subsequently unzipped. For templates used in d)-g), the 601NPE was separated from the Gal4 binding sequence by 10 bp.

a) Distribution of the location of a nucleosome before remodeling. Data were pooled from measurements on multiple nucleosomal DNA molecules.

b) Distribution of the location of a nucleosome remodeled by ISW1a in the absence of Gal4DBD.

c) Distribution of the location of a nucleosome remodeled by SWI/SNF in the absence of Gal4DBD.

d) Distribution of the location of a nucleosome remodeled by ISW1a with a bound Gal4DBD initially located upstream of the 601NPE.

e) Distribution of the location of a nucleosome remodeled by SWI/SNF with a bound Gal4DBD initially located upstream of the 601NPE.

f) Distribution of the location of the nucleosome remodeled by ISW1a with a bound Gal4DBD initially located downstream of the 601NPE.

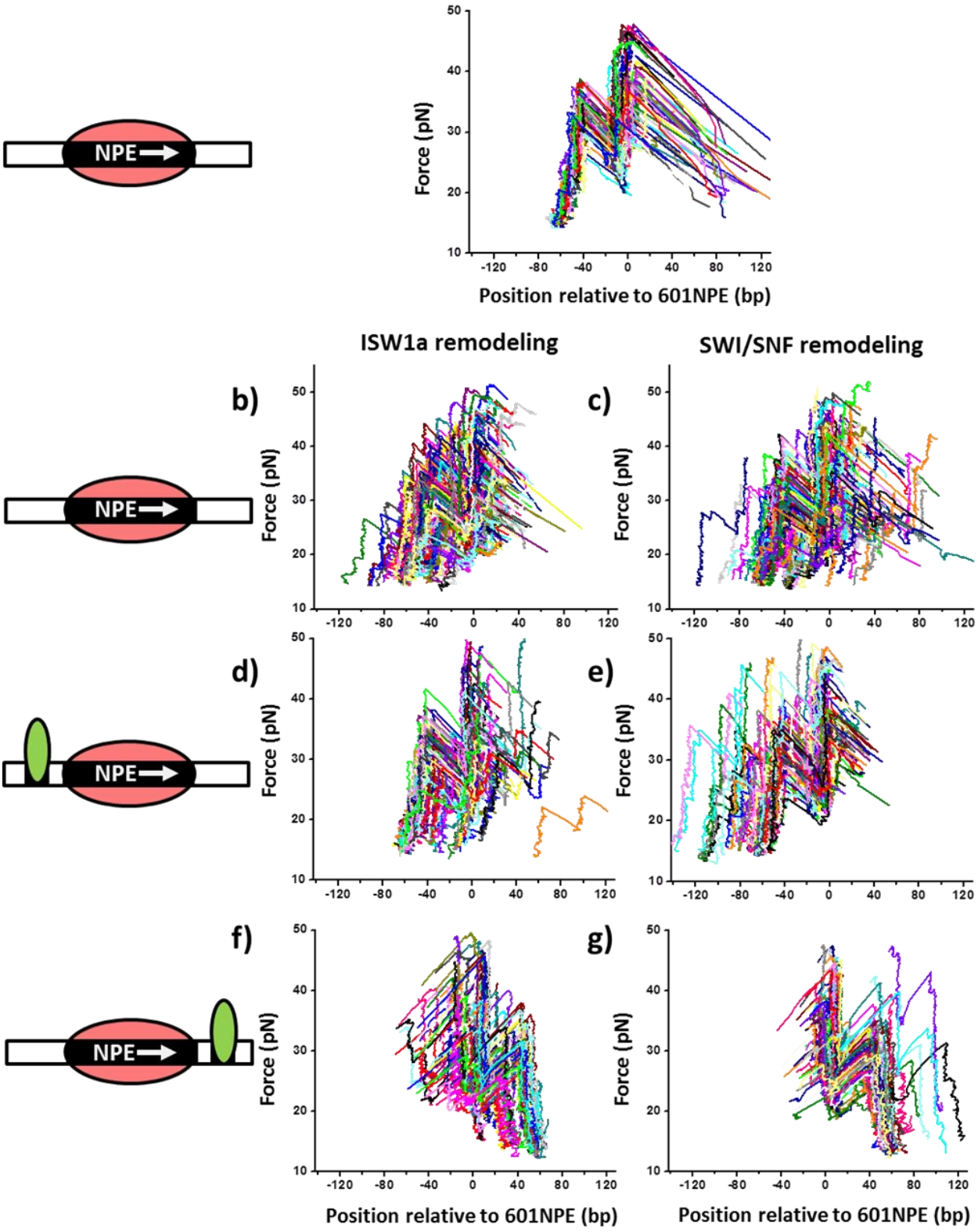
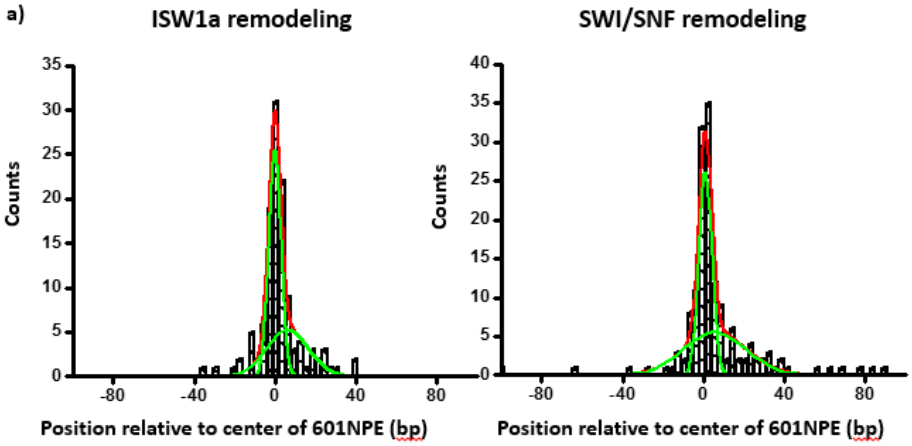

NPE
Force (pN)
Position relative to 601NPE (bp)
ISW1a remodeling
SWI/SNF remodeling
b)
c)
d)
e)
f)
g)

Figure 2.6. Directionality of ISW1a and SWI/SNF remodeling of a nucleosome in close proximity to a bound Gal4DBD. This figure shows the corresponding raw traces for Figure 2.5.

Here we employed ySWI/SNF and yISW1a as model systems to study how a bound Gal4DBD may affect the remodeling of an adjacent nucleosome. First, we investigated the initial direction of nucleosome remodeling in the presence of a bound Gal4DBD in close proximity. This was achieved by limiting the remodeling reaction to the first remodeling event which I define as a single round of remodeler binding, nucleosome remodeling, and remodeler detachment from the nucleosomal DNA (Shundrovsky et al., 2006). I engineered a DNA template in which a Gal4 binding sequence and a 601NPE were separated by 10 bp. The DNA template containing a positioned nucleosome was then remodeled, by either SWI/SNF or ISW1a, for a short period of time (~1 min), with or without the addition of Gal4DBD (Figure 2.5; Figure 2.6). During such a short remodeling time, ~ 56% of nucleosomes pooled from measurements of multiple single molecules were found to remain at the original location, suggesting a lack of remodeling (Figure 2.7). Of the remaining ~ 45% of the nucleosomes that were remodeled, we estimate that ~73% were remodeled only once and ~27% were remodeled more than once, using a method we previously established (Shundrovsky et al., 2006).

a)



b)

remodeler	ISW1a	SWI/SNF
remodeled fraction	41%	47%
Un-remodeled fraction	59%	53%

Figure 2.7. Determination of fractions of remodeled nucleosome.
a) For each nucleosome distribution, remodeled by either ISW1a or SWI/SNF as shown in Figures 2b and 2c, was fit to a double Gaussian function. One Gaussian (narrow green curve) corresponds to the distribution of unremodeled nucleosomes and the other (broader green curve) to the distribution of remodeled nucleosomes. The sum of the two Gaussians is shown as the red curve.
b) A table summarizing fractions of unremodeled percentage and remodeled percentage for both ISW1a and SWI/SNF.

In the absence of Gal4DBD, although the nucleosome unzipping signature did not appear to be altered after remodeling by either ISW1a or SWI/SNF (Figure 2.5), the positions of the nucleosomes were spread out from the original location. Both ISW1a and SWI/SNF were able to

move a nucleosome bi-directionally (Figures 2.5b and 2.5c) without inducing significant changes in the nucleosome structure (Table 2.1). The slight asymmetric distribution of the remodeled nucleosome was likely due to the non-palindromic feature of the 601 sequence (Lowary and Widom, 1998) which leads to some asymmetry in the protein-DNA interactions at the two halves of a nucleosome (Hall et al., 2009). The results from SWI/SNF remodeling were also consistent with those from an earlier study (Shundrovsky et al., 2006).

Force (pN)

	1st Cluster		2nd Cluster	
	mean	s.d.	mean	s.d.
Before remodeling	33.12	2.9	38.02	5.3
ISW1a remodeling without Gal4DBD	31.99	4.2	36.63	7.2
ISW1a remodeling with Gal4DBD upstream	31.67	3.8	35.44	6.4
ISW1a remodeling with Gal4DBD downstream	30.22	3.4	37.24	6.2
SWI/SNF remodeling without Gal4DBD	31.64	3.9	36.06	6.7
SWI/SNF remodeling with Gal4DBD upstream	32.35	3.5	37.56	5.5
SWI/SNF remodeling with Gal4DBD downstream	30.54	4.0	35.67	6.4

The width and distance of interactions (bp)

	1st Cluster		2nd Cluster		Distance of two clusters	
	mean	s.d.	mean	s.d.	mean	s.d.
Before remodeling	38.02	5.3	19.36	7.2	46.00	4.8
ISW1a remodeling with out Gal4DBD	36.64	7.4	18.08	8.5	46.21	7.5
ISW1a remodeling with Gal4DBD upstream	32.68	5.0	25.08	6.4	45.19	5.5
ISW1a remodeling with Gal4DBD downstream	36.66	7.9	21.02	10.5	44.09	9.8
SWI/SNF remodeling without Gal4DBD	35.96	7.3	16.68	7.5	45.67	7.9
SWI/SNF remodeling with Gal4DBD upstream	38.51	7.3	21.71	9.1	45.50	7.7
SWI/SNF remodeling with Gal4DBD downstream	36.75	7.8	20.79	9.0	44.77	8.5

Table 2.1. Comparison of unzipping force signatures of a nucleosome before and after remodeling. We used unzipping to characterize the structure of a nucleosome before or after remodeling by either ISW1a or SWI/SNF, in the presence or absence of Gal4DBD. The structural features include the maximum force in the 1st force cluster, the maximum force in the 2nd force cluster, the width of each cluster, and the distance between the two clusters. Errors show s.d.

Interestingly, in the presence of Gal4DBD, ISW1a moved the nucleosome away from Gal4DBD (Figure 2.5d), whereas SWI/SNF moved the nucleosome towards Gal4DBD (Figure 2.5e). To determine whether such a differential behavior was a result of the DNA sequence used, we engineered another DNA template that was identical to this one, except that the Gal4 binding site was located on the other side of the 601NPE. After adding Gal4DBD, ISW1a again moved the nucleosome away from the Gal4DBD (Figure 2.5f), while SWI/SNF again moved the nucleosome towards the Gal4DBD (Figure 2.5g). These data rule out the possibility of a DNA sequence effect on the directionality of nucleosome movement by the two remodelers. Therefore, we conclude that the bound Gal4DBD affects the directionality of nucleosome movement by the two types of remodelers differently: away from the TF for ISW1a and toward the TF for SWI/SNF.

Our findings on the TF directed SWI/SNF nucleosome remodeling are entirely novel; while our findings on the TF directed ISW1a nucleosome remodeling are in agreement with a previous study that used NURF (a homolog of ISWI complexes in Drosophila) in the presence of Gal4DBD(Kang et al., 2002).

Gal4DBD is a barrier for nucleosome sliding by ISW1a

Since ISW1a moved a nucleosome away from an adjacent Gal4DBD, Gal4DBD may act as a physical barrier, or roadblock, to ISW1a nucleosome remodeling. To specifically investigate whether a bound Gal4DBD may be a barrier to ISW1a remodeling, we designed an unzipping template with a 601NPE at the end of the template and a Gal4 binding at a greater spacing (75 bp) from the 601NPE (Figure 2.8). The use of an end-positioned nucleosome should dictate that the nucleosome movement could only take place towards a bound Gal4DBD.

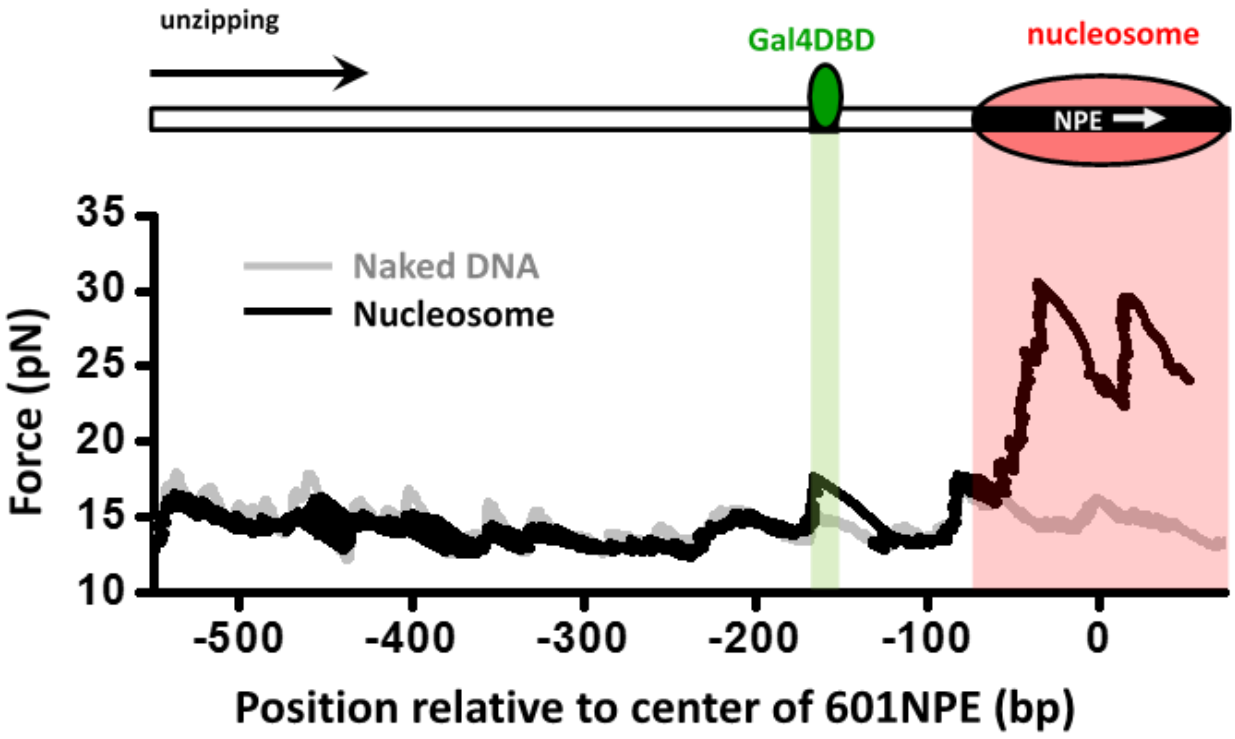


Figure 2.8. Single molecule unzipping simultaneously detects Gal4DBD and an end-positioned nucleosome. Shown is a representative unzipping trace of a DNA molecule containing an end-positioned nucleosome and a bound Gal4DBD.

After ISW1a remodeling for 10 min, which was sufficiently long to allow for multiple rounds of remodeling, the distributions of the nucleosome location showed a significant difference between

the absence and presence of Gal4DBD (Figure 2.9). In the absence of Gal4DBD, the nucleosome was moved away from the template end by several hundred base pairs towards the center region of the template, generating a rather broad distribution. In contrast, in the presence of Gal4DBD, although nucleosomes were still moved away from the end of the template, they were not able to pass the location of the Gal4DBD (Figure 2.9). Instead, the distribution peaked at the midpoint between the Gal4 binding sequence and the 601NPE.

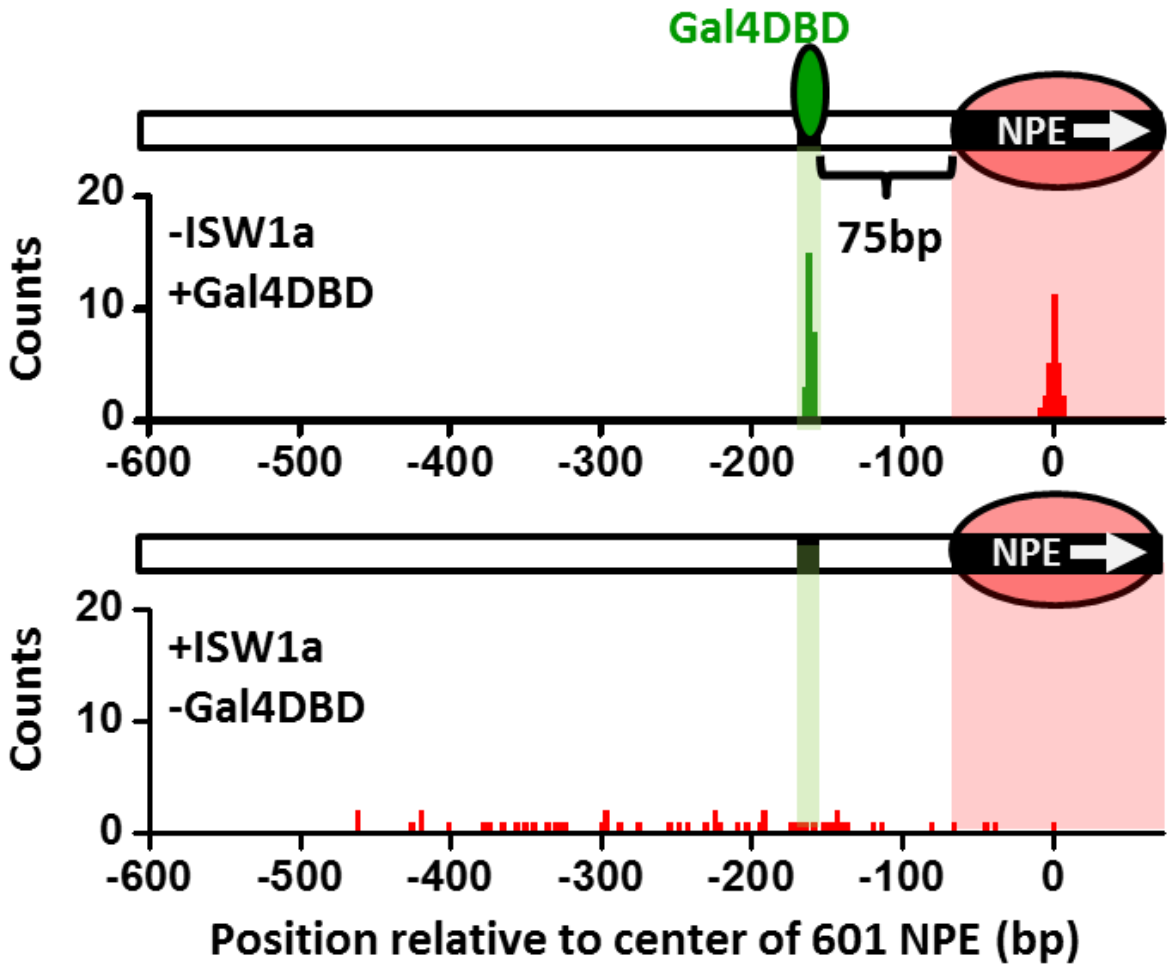


Figure 2.9. ISW1a remodeling is blocked by a bound Gal4DBD. Nucleosomes were remodeled by 1 nM ISW1a with 1 mM ATP for 10 min with or without Gal4DBD. Shaded regions indicate locations of Gal4

binding sequence and 601NPE. Distributions of the locations of the nucleosome and bound Gal4DBD, either before remodeling or without Gal4DBD, as controls.
c) Relationship between the distance the remodeled nucleosome moved and the separation between the Gal4 binding sequence and the 601 NPE. Error bars are SEM.

To further examine the relationship between the location of Gal4DBD and the ISW1a remodeled nucleosome, we used two additional templates with shorter (24 bp and 50 bp) distances between the Gal4 binding site and the 601NPE (Figure 2.10). After ISW1a remodeling in the presence of Gal4DBD, we found that the dyad locations of the remodeled nucleosomes always peaked at the center distance between the Gal4DBD and the DNA end (Figure 2.11). These results demonstrate that Gal4DBD is a physical barrier for ISW1a mediated nucleosome remodeling and that ISW1a is able to use Gal4DBD as a reference point to reposition a nucleosome.

This novel finding is of particular relevance to *in vivo* nucleosome spacing, especially near transcription start and termination sites. Although previous studies have suggested a possible role for ISWI remodelers to space nucleosomes using bound TFs near these sites (Gkikopoulos et al., 2011; Yen et al., 2012), our results provide direct evidence that ISWI remodelers can indeed sense and respond to the presence of a DNA-bound protein such as a TF, which acts as barrier to dictate the placement of nucleosomes.

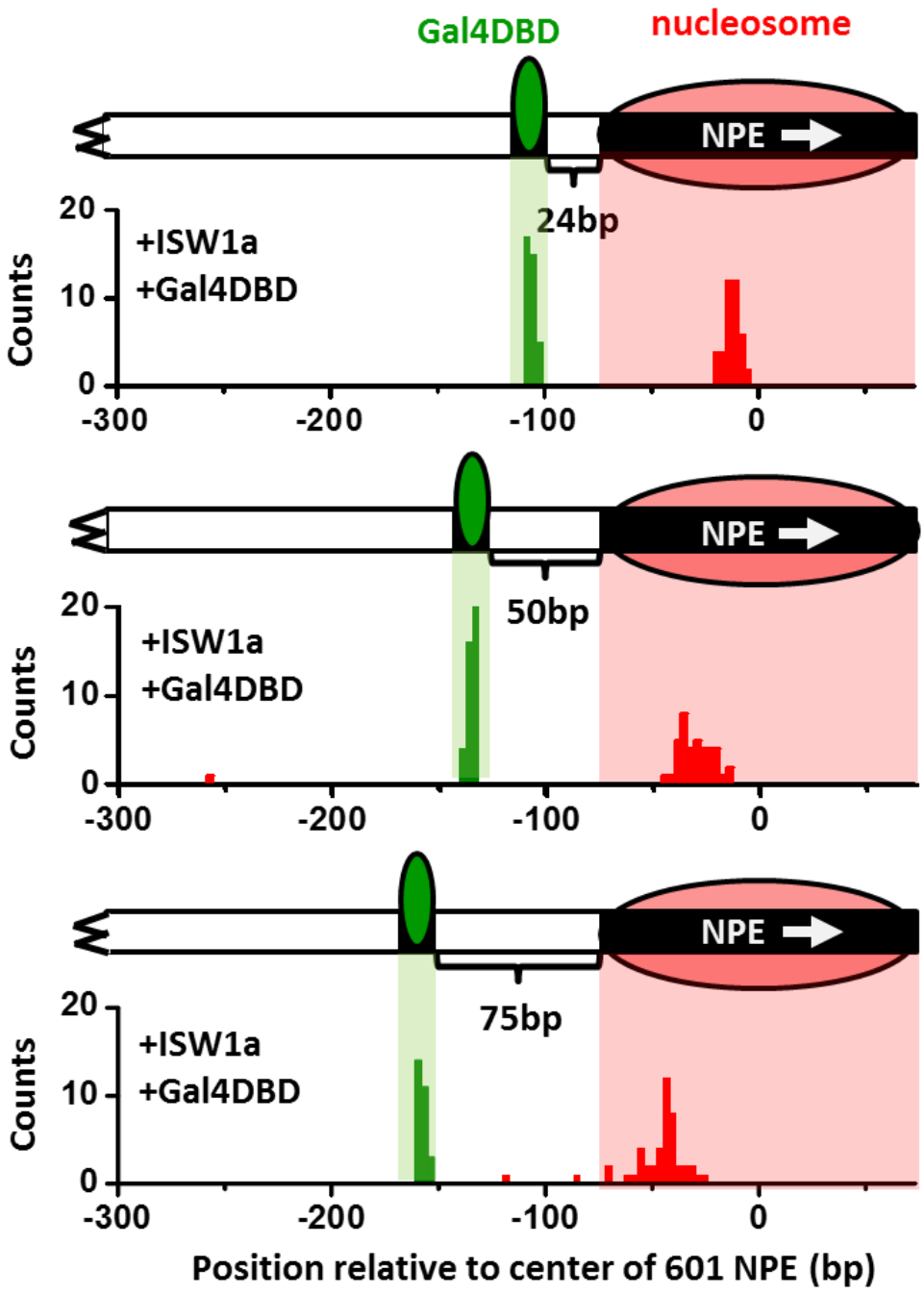
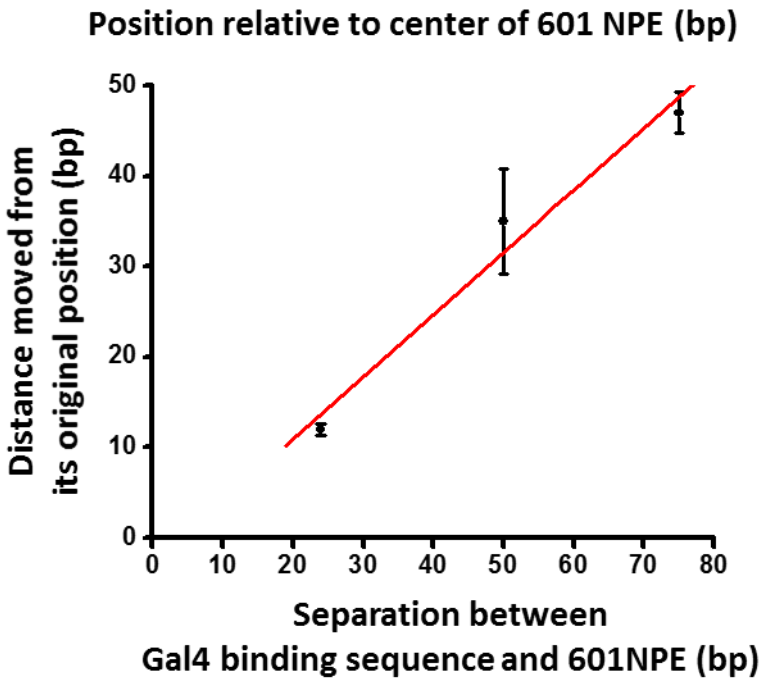


Figure 2.10. Distributions of the locations of the nucleosome after ISW1a remodeling in the presence of Gal4DBD on three different templates of increasing separation between the Gal4 binding sequence and the 601NPE. For each template, the nucleosome position distribution is dominated by a narrow population, but has a few outliers which were moved a much greater distance and some of which even passed the Gal4 binding sequence. These outliers (~5%) were likely a result of templates that did not have a bound Gal4DBD initially (~5%; see main text). This is further supported by the observation that

none of these outlier traces revealed a bound Gal4DBD. Nonetheless, in order to avoid possible bias, these nucleosome positions were still used for further analysis in c) and thus contributed to the error bars in c).



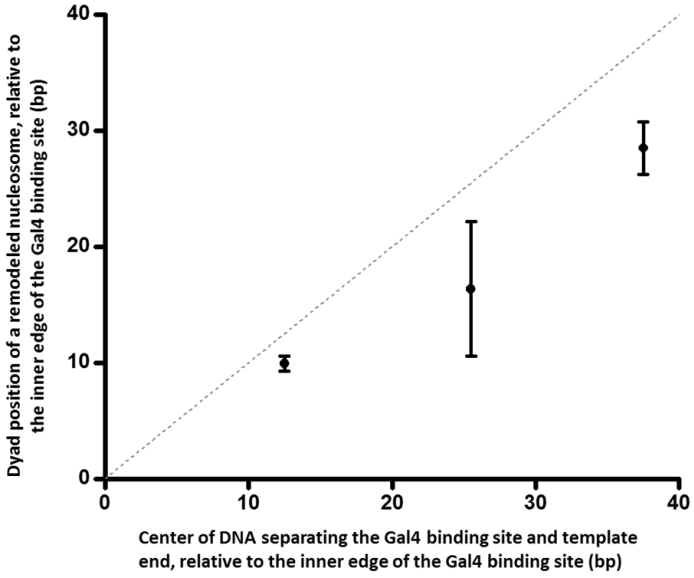


Figure 2.11. ISW1a remodeling is blocked by a bound Gal4DBD. Top Panel) Relationship between the distance the remodeled nucleosome moved and the separation between the Gal4 binding sequence and the 601 NPE. Error bars are SEM. Bottom Panel) this relation shows that ISW1a remodeler tends to center a nucleosome between the Gal4 binding site and the DNA end. Error bars are SEM.

SWI/SNF remodeling evicts Gal4DBD from DNA

In order to test whether a bound Gal4DBD is a physical barrier for SWI/SNF remodeling, we designed a long template (~1200 bp) with a 601NPE separated from a Gal4 binding sequence by 11 bp. The 601NPE was located near the center of a long DNA template to allow ample distance for possible bidirectional sliding of the nucleosome via multiple rounds of remodeling, such that the remodeled nucleosome and possible presence of Gal4DBD could be independently detected. After a 10 min remodeling by SWI/SNF, nucleosomes were repositioned from the center of the template to random locations along the entire sequence. Both in the absence and presence of Gal4DBD, remodeled nucleosomes were detected on both sides of the original Gal4 binding position (Figure 2.12). This indicates that Gal4DBD is not a physical barrier for SWI/SNF mediated nucleosome remodeling.

To determine whether the distributions of the locations of remodeled nucleosomes were similar for measurements made by unzipping the DNA in one direction versus in the other direction, we unzipped multiple DNA molecules, each containing either an unremodeled or a remodeled nucleosome, from both directions. Nucleosome remodeling was carried out in 1 nM SWI/SNF with 1 mM ATP for 5 min. Our data show similar distributions for data obtained in both directions (Figure 2.13).

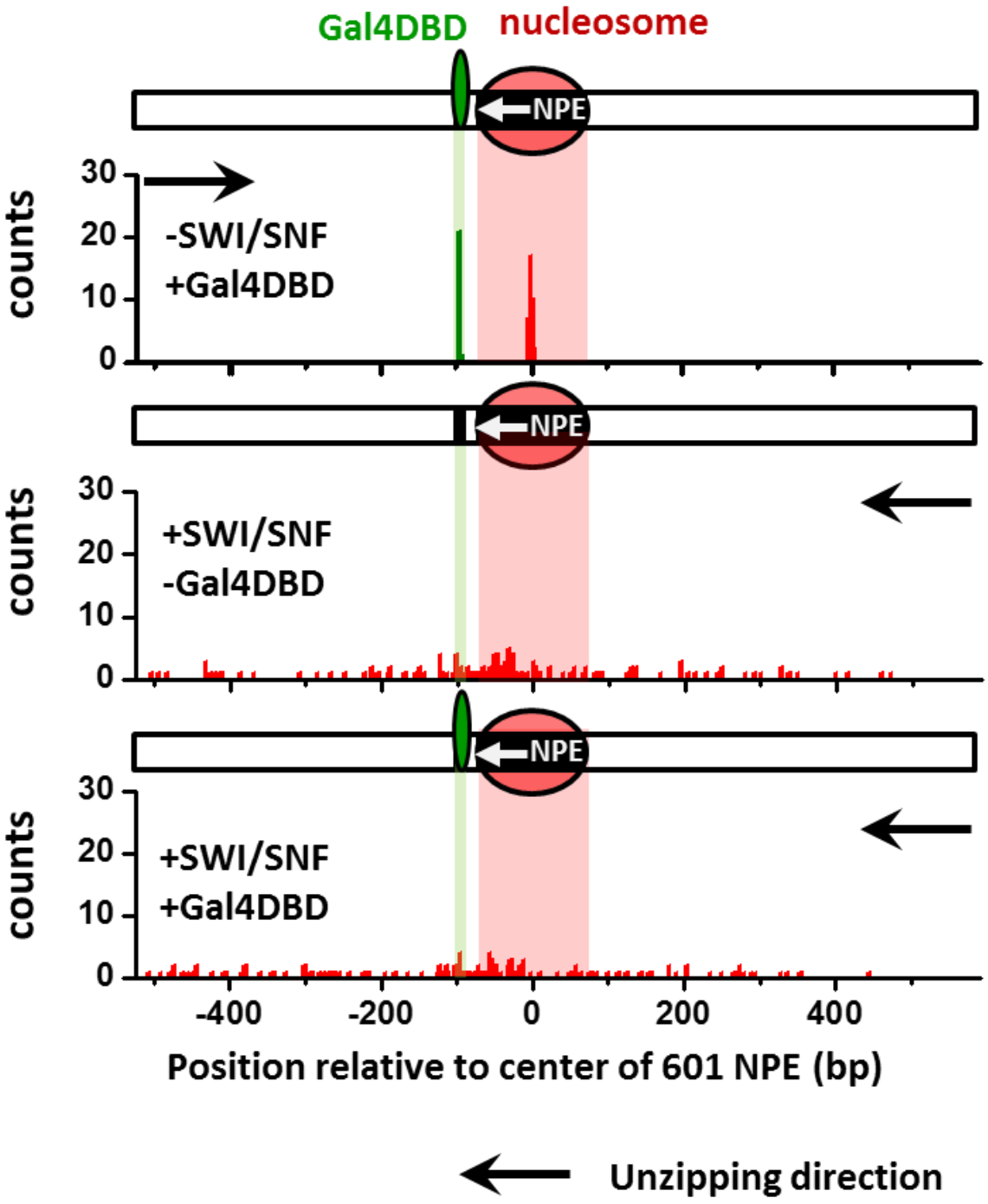


Figure 2.12. SWI/SNF remodeling evicts a bound Gal4DBD from its DNA template. Nucleosomes were remodeled by 1.5 nM SWI/SNF with 1 mM ATP for 10 min with or without Gal4DBD. Shaded regions indicate locations of Gal4 binding sequence and 601NPE. Distributions of the locations of the

nucleosome and bound Gal4DBD before remodeling (upper plot), after remodeling without Gal4DBD (middle plot), and after remodeling with Gal4DBD (lower plot).

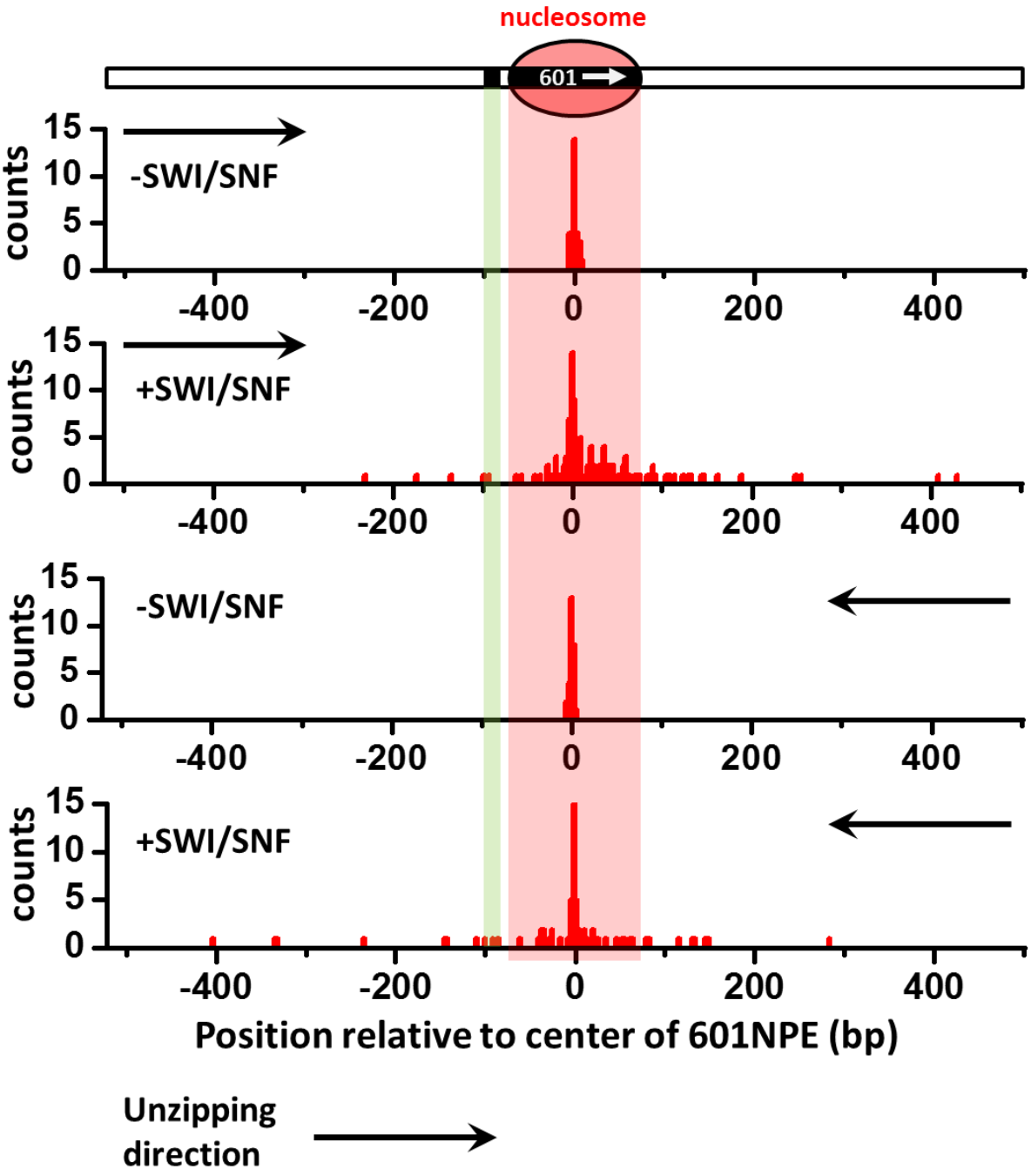


Figure 2.13. Distributions of the locations of SWI/SNF remodeled nucleosomes as determined by unzipping from both directions.

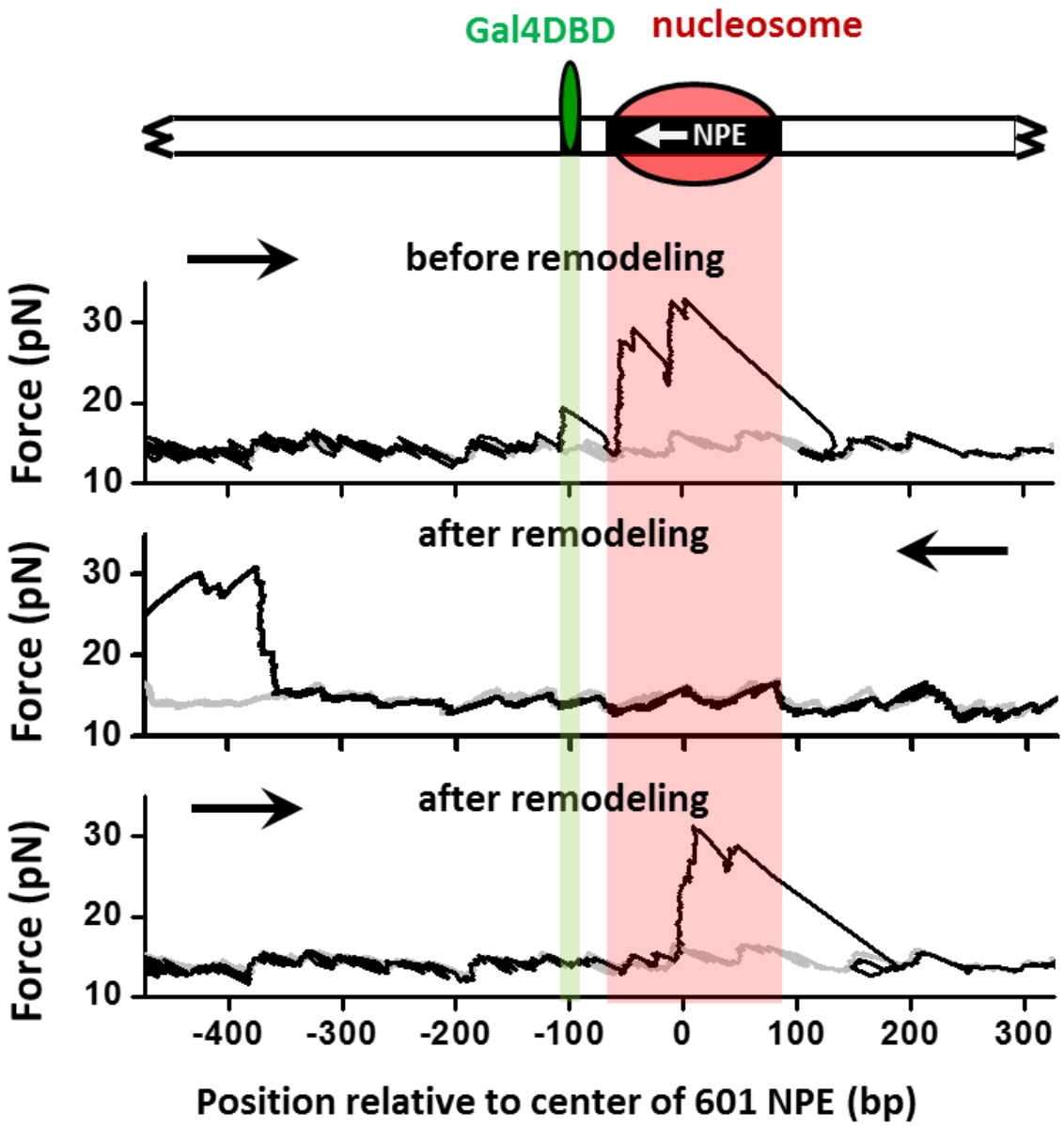


Figure 2.14. SWI/SNF remodeling evicts a bound Gal4DBD from its DNA template. Nucleosomes were remodeled by 1.5 nM SWI/SNF with 1 mM ATP for 10 min with or without Gal4DBD. Shaded regions indicate locations of Gal4 binding sequence and 601NPE. Representative traces in the case of before remodeling (top plot; $N = 55$) and after remodeling (middle and bottom plots; $N = 50$). The middle plot shows an example trace where a nucleosome was remodeled to the opposite side of Gal4DBD relative to its original position; while the bottom plot shows an example trace where a nucleosome was remodeled to the same side of Gal4DBD relative to its original position. Grey traces were taken from the corresponding naked DNA.

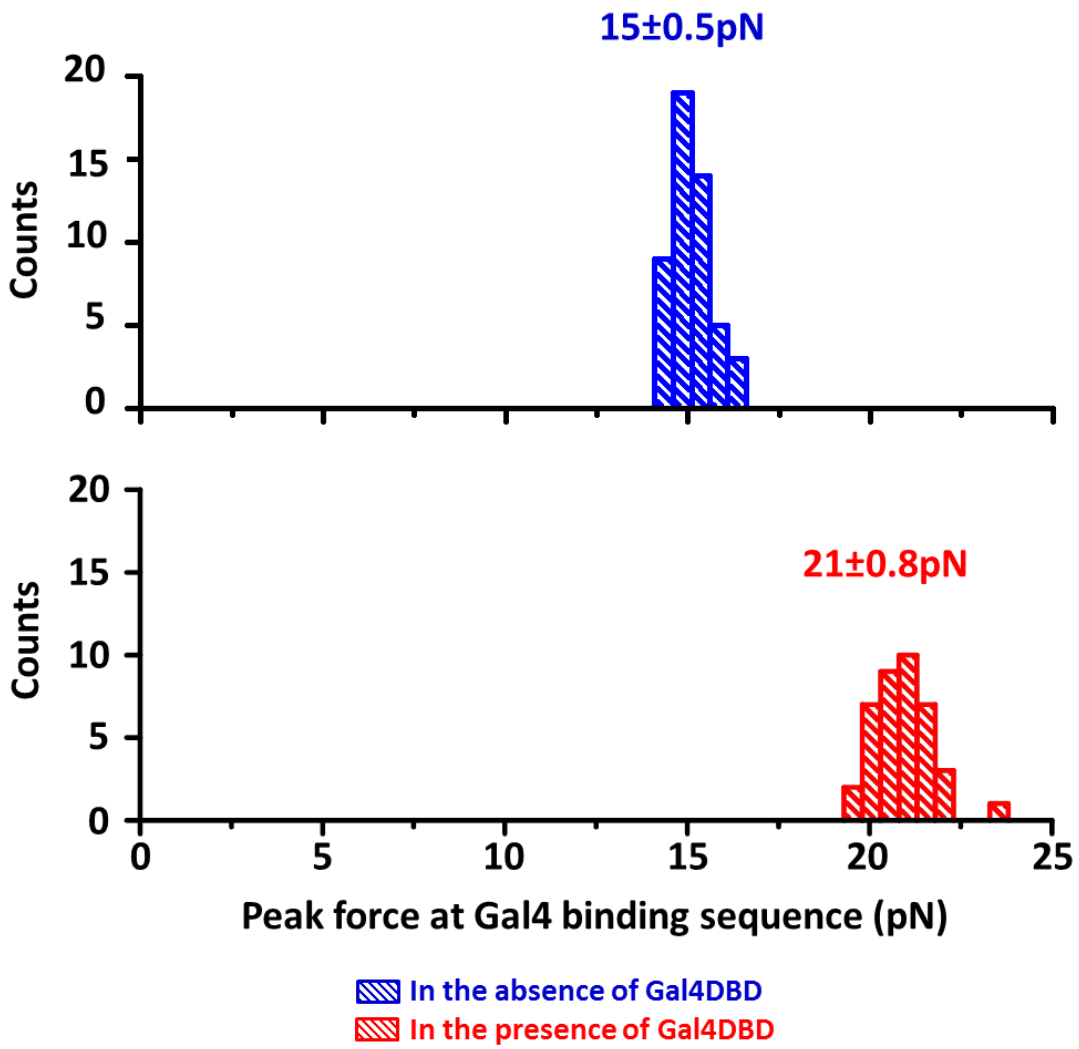


Figure 2.15. Detection of Gal4DBD binding. The presence of a bound Gal4DBD was determined by the magnitude of the force peak at the Gal4 binding sequence. In the presence of a bound Gal4DBD, the peak force increased substantially and was readily differentiable from the baseline DNA force.

What is the fate of the Gal4DBD after a nucleosome has been remodeled? To answer this question, we allowed a nucleosome to be remodeled by SWI/SNF in the presence of Gal4DBD. We then analyzed each trace to determine whether a nucleosome was remodeled to the opposite side of the Gal4 sequence or to the same side of the Gal4 sequence, relative to the 601NPE. For the traces where nucleosomes were remodeled to the opposite side of the Gal4 sequence, we did

not detect any Gal4DBD unzipping signature (Figure 2.14; Figure 2.15). This indicates that SWI/SNF was able to move the nucleosome in such a way that the Gal4DBD was evicted from its binding sequence. For traces where a nucleosome was remodeled to the same side of the Gal4 sequence relative to the 601NPE (Figure 2.15), we also did not detect any Gal4DBD unzipping signature. This implies that these nucleosomes were likely first remodeled towards the bound Gal4DBD, as indicated by Figure 2.5, resulting in the eviction of Gal4DBD. This was followed by subsequent remodeling events that moved nucleosomes to other positions.

Several Control experience were conducted to validate our results on SWI/SNF remodeling (Table 2.2)

To determine whether SWI/SNF with ATP alone is able to displace a bound Gal4DBD in the absence of a nucleosome, we carried out experiments on a DNA template preloaded with Gal4DBD but without a nucleosome in 1.5 nM SWI/SNF with 1 mM ATP for 10 min. DNA molecules were subsequently unzipped to determine the presence of Gal4DBD. The fraction of templates containing a bound Gal4DBD remained the same before and after the remodeling reaction, indicating that SWI/SNF with ATP alone is not able to displace a bound Gal4DBD.

To rule out the possibility that Gal4DBD disruption was due to binding of SWI/SNF to DNA or the nucleosome and not due to nucleosome remodeling, we carried out a control experiment on a DNA template containing a bound Gal4DBD and a nucleosome by incubating the sample with 1.5 nM SWI/SNF for 10 min in the absence of ATP. We subsequently unzipped the DNA template to determine if Gal4DBD was still bound. The fraction of templates containing a bound

Gal4DBD was comparable to that of a template without a nucleosome and without SWI/SNF and ATP added, indicating that in the absence of ATP, SWI/SNF is unable to evict a bound Gal4DBD even in the presence of a nucleosome adjacent to a bound Gal4DBD.

In addition, to test whether SWI/SNF is able to evict Gal4DBD via nucleosome remodeling when a bound Gal4DBD is located farther away from a nucleosome, we used a template where the Gal4 binding site was separated from the 601NPS by 24 bp and carried out unzipping experiments under identical conditions as those shown in Figure 2.12 (Figure 2.16). Out of all traces where the nucleosome was repositioned to the opposite side of the Gal4 binding site by SWI/SNF (N = 13), we did not detect any Gal4DBD binding signature on the template, indicating eviction of Gal4DBD. Shown are example traces, with arrows indicating the unzipping directions.

	Without ATP	**With ATP**	**Without ATP**
	Without SWI/SNF	**With SWI/SNF**	**With SWI/SNF**
	No nucleosome	**No nucleosome**	**Nucleosome**
Number of traces with bound Gal4DBD	**31**	**20**	**30**
Total number of traces	**33**	**21**	**31**
Percentage of Traces with bound Gal4DBD	**94%**	**95%**	**97%**

Table 2.2. SWI/SNF is unable to evict a bound Gal4DBD in the absence of a nucleosome or in the absence of ATP.

Our findings provide the first direct evidence that SWI/SNF nucleosome remodeling is capable of evicting a bound TF. Previously, SWI/SNF was shown to move one nucleosome to invade and eventually disrupt an adjacent nucleosome (Dechassa et al., 2010). Taken together with our findings, SWI/SNF nucleosome remodeling appears to be powerful machinery capable of actively overcoming and removing a variety of obstacles in its vicinity.

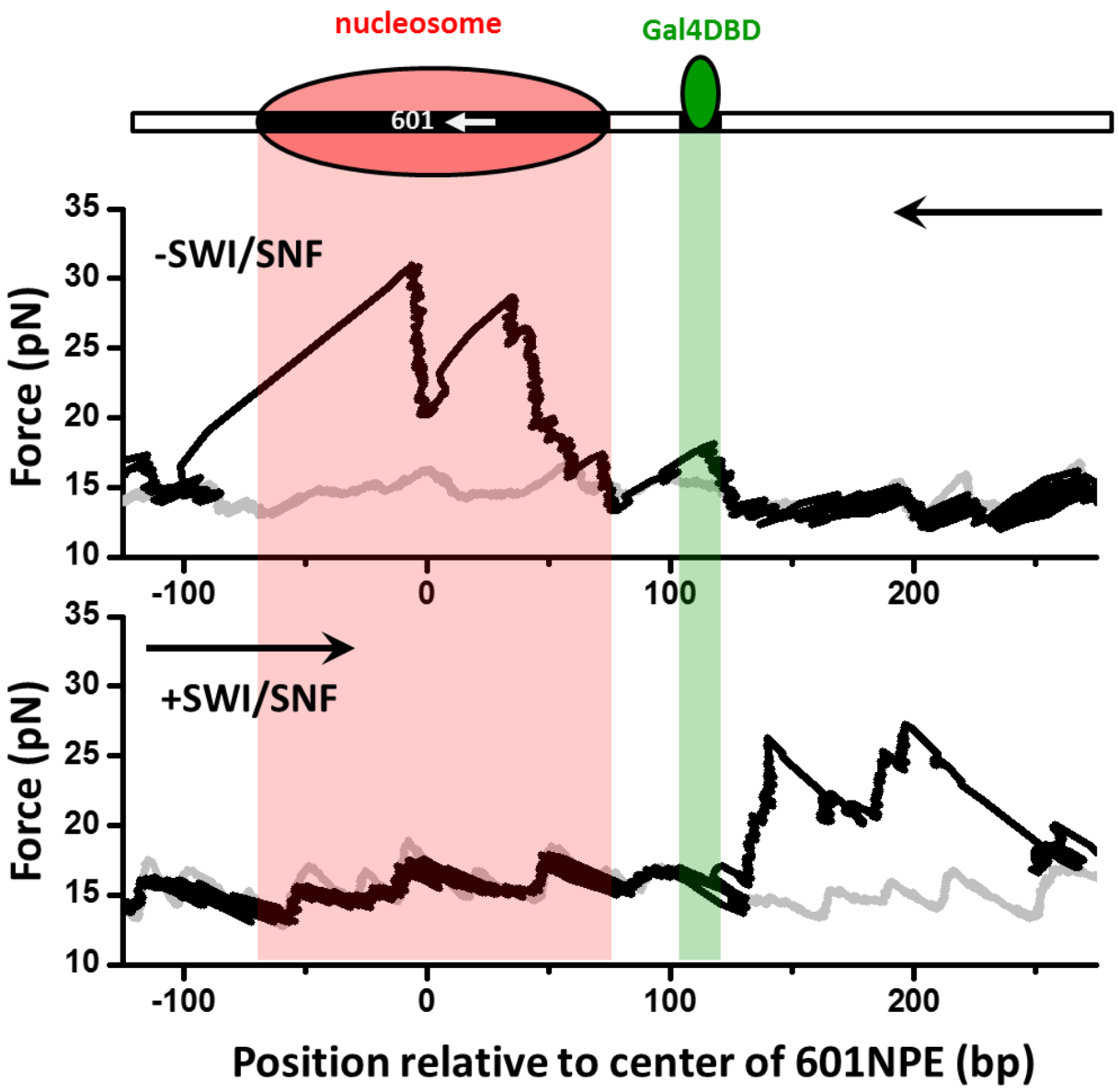


Figure 2.16. Nucleosome remodeling by SWI/SNF on a template with the Gal4 binding site separated from the 601NPE by 24 bp.

Nucleosome remodeling shows similar effects in the presence of a Lac repressor

To rule out the possibility that the interaction between a TF and a nucleosome demonstrated above is specific to Gal4DBD, we replaced Gal4 binding sequence with a Lac repressor binding sequence and repeated the above experiments using the Lac repressor. Because the Lac repressor is only found in prokaryotic cells and has no known relationship with any chromatin remodeler in eukaryotic cells, it can act as a biologically neutral bound protein. When ISW1a remodeled an end-positioned nucleosome on a template also containing a bound Lac repressor, the Lac repressor effectively dictated the position of the remodeled nucleosome, with ISW1a centering the nucleosome on the DNA with the Lac repressor acting as a barrier (Figure 2.17). In contrast, SWI/SNF was able to slide a nucleosome in either direction, displacing the bound Lac repressor (Figure 2.18). Therefore, we conclude that the mechanism of TF regulation by nucleosome remodeling is likely general without any specificity to a particular TF.

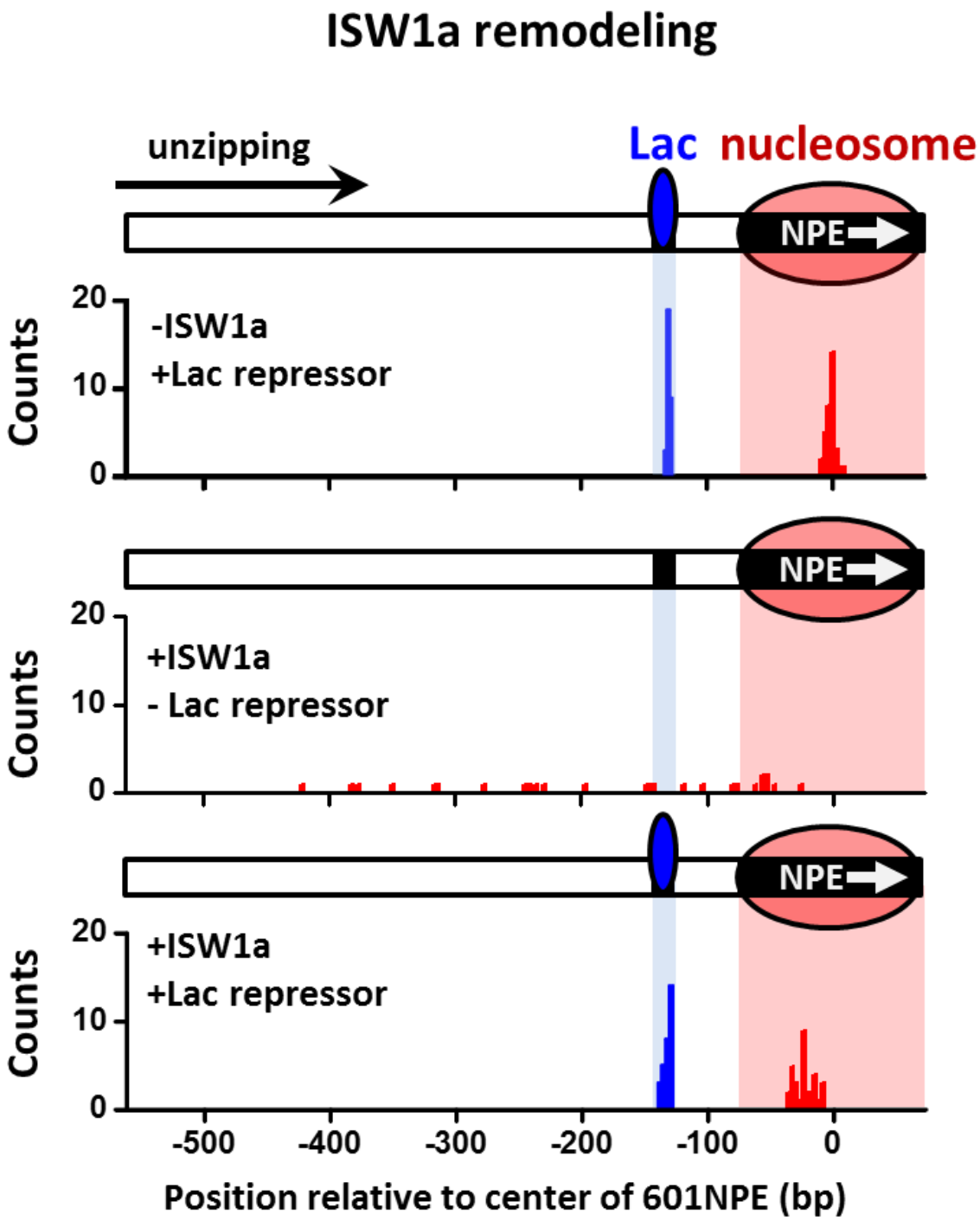


Figure 2.17. ISW1a remodeling is blocked by Lac repressor. Shaded regions indicate locations of Lac repressor binding sequence and 601NPE. Distributions of locations of nucleosomes before remodeling (upper plot), after remodeling by ISW1a without Lac repressor (middle plot), and after remodeling by ISW1a with Lac repressor (lower plot). Lac repressor binding sequence was separated from the 601NPE

by 50 bp. Nucleosomes remodeling was carried out in 1 nM ISW1a with 1 mM ATP for 10 min with or without Lac repressor.

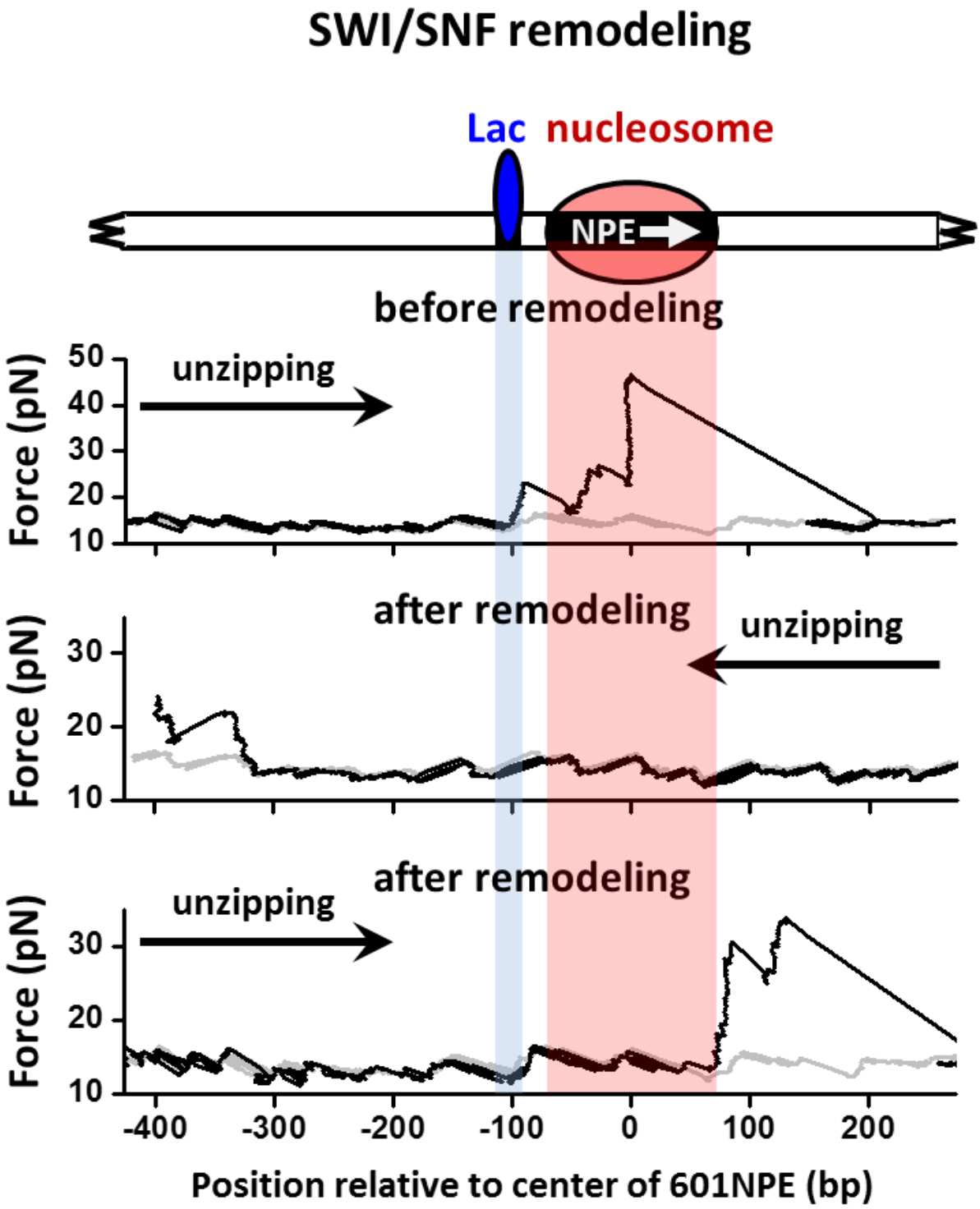


Figure 2.18. SWI/SNF remodeling evicts Lac repressor from the template. Shaded regions indicate locations of Lac repressor binding sequence and 601NPE. Representative traces in the case of before SWI/SNF remodeling (top plot; $N = 25$) and after remodeling (middle and bottom plots; $N = 27$). Lac repressor binding site was separated from the 601NPE by 10 bp. Nucleosomes were remodeled by 1.5 nM SWI/SNF with 1 mM ATP for 10 min. The middle plot shows an example trace where a nucleosome was remodeled to the other side of the Lac repressor and the bottom plot shows an example trace where a nucleosome was remodeled to the same side of Lac repressor.

DISCUSSION

It has been widely acknowledged that SWI/SNF remodelers move nucleosomes toward the ends of a DNA template and ISWI remodelers move nucleosomes away from the ends of a DNA template (Stockdale et al., 2006; Yang et al., 2006; Zofall et al., 2006). However, rather than DNA ends, *in vivo* DNA nucleosome remodeling will most likely encounter DNA binding proteins such as TFs. Here we show that ISW1a senses a bound TF as a boundary during nucleosome repositioning, while SWI/SNF remodeling is unimpeded by the presence of a bound TF and is able to slide a nucleosome and evict the TF (Figure 2.19). In contrast to bulk biochemical studies that typically use DNA ends as boundaries on short DNA templates, our studies using bound TFs as potential barriers better mimic a situation that will more likely occur in vivo.

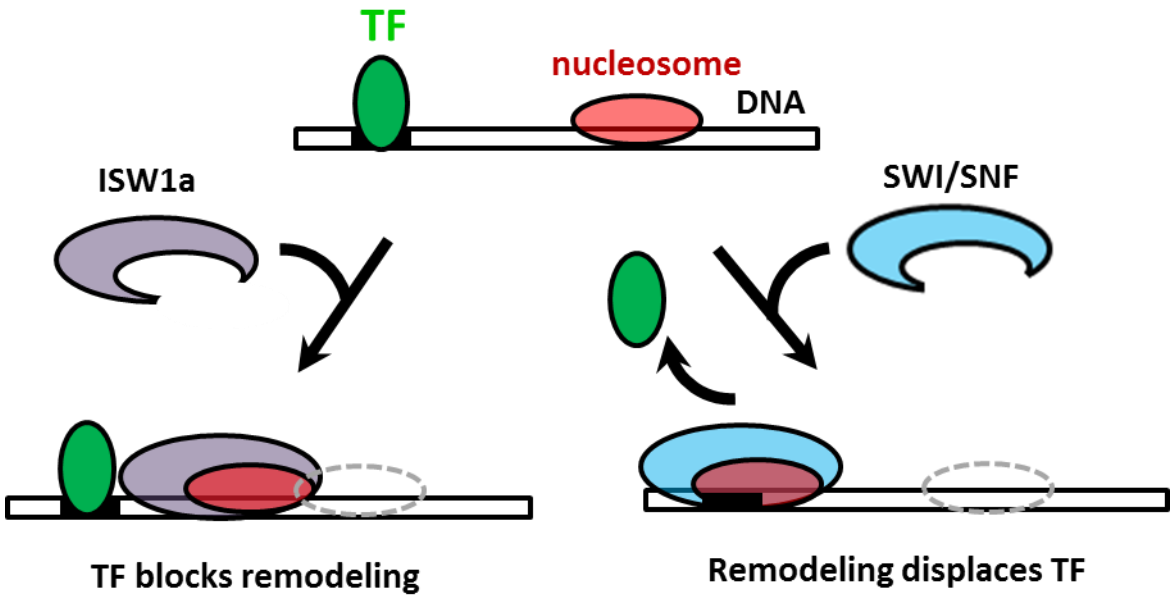


Figure 2.19. The interplay between nucleosome remodeling and a bound TF. When a nucleosome is remodeled by ISW1a (left), nucleosome repositioning is blocked by a bound TF, the TF remains intact, and the remodeled nucleosome is positioned with the TF acting as a boundary and reference point. On the other hand, when a nucleosome is remodeled by SWI/SNF (right), nucleosome positioning is unaffected by a bound TF and the TF is evicted.

Differential directionality of nucleosome remodeling

Our results showed opposite directionality for nucleosome positioning when a nucleosome in close proximity to a TF was remodeled by ISW1a and SWI/SNF (Figure 2). Biochemical studies and the crystal structure of ISW1a indicate that the DNA binding domain of ISW1a binds to ~29 bp of the extranucleosomal DNA, which has been proposed to act as an anchor to pull the nucleosome towards ISW1a (Gangaraju and Bartholomew, 2007; Hauk and Bowman, 2011; Hota et al., 2013; Yamada et al., 2011). Our finding that ISW1a moves nucleosome away from a TF adjacent to the nucleosome is consistent with an important role of the DNA-binding domain

in engaging fully accessible DNA immediately flanking the nucleosome. In contrast, while it has been widely acknowledged that SWI/SNF does not require extra-nucleosomal DNA binding to remodel a nucleosome, cryo-EM and DNA-crosslinking experiments have shown that the SNF6 subunit of SWI/SNF binds to ~15 bp of the extra-nucleosomal DNA and the rest of the SWI/SNF slides the nucleosome away from where the SNF6 subunit binds (Dechassa et al., 2008). Although the SNF6 subunit has not been shown to be essential for remodeling, it has DNA binding affinity(Dechassa et al., 2008; Sengupta et al., 2001). We speculate that it may help to orient SWI/SNF binding on the nucleosome. In the presence of a barrier adjacent to a nucleosome, SNF6 may prefer to bind to the side of the nucleosome with more linker DNA and thus orient SWI/SNF to slide a nucleosome towards the TF.

Nucleosome spacing by ISW1a in vivo

Our study of ISW1a remodeling demonstrates that Gal4DBD is an effective barrier for ISW1a mediated nucleosome remodeling and the ISW1a is able to use Gal4DBD as a reference point to reposition nucleosomes (Figure 3). These results have significant implications for the mechanism of nucleosome spacing in vivo. Genome-wide nucleosome mapping in budding yeast revealed that deletion of Isw1 in yeast disrupts nucleosome spacing (Gkikopoulos et al., 2011), suggesting that ISW1 remodelers are key players in generating evenly distributed nucleosomal arrays. In addition, several recent studies have shown that certain DNA-binding factors located at the promoter region are also responsible for nucleosome positioning (Bai et al., 2011; Hughes et al., 2012; Struhl and Segal, 2013; Whitehouse et al., 2007; Yadon et al., 2010; Yen et al., 2012; Zhang et al., 2010). Our finding that a bound Gal4DBD is a barrier to ISW1a now provides

direct evidence to illustrate that ISW1a can potentially use a TF around the promoter region as a reference point to evenly position nucleosomes into the gene body.

In vivo implications of SWI/SNF action in TF regulation

Our study shows that TF eviction is an intrinsic property of SWI/SNF remodeling and is independent of SWI/SNF recruitment (Figure 4 and 5). It has been previously shown that SWI/SNF recruitment by the glucocorticoid receptor (GR) induced histone loss in nucleosomes and this was immediately followed by GR and SWI/SNF eviction from the template (Nagaich et al., 2004). Our current work demonstrates that TF eviction via nucleosome remodeling can take place without nucleosome loss in the absence of remodeler recruitment. Although SWI/SNF can translocate along naked DNA (Lia et al., 2006; Sirinakis et al., 2011; Zhang et al., 2006), raising the possibility for TF eviction solely by SWI/SNF, we found that in the absence of a nucleosome, SWI/SNF did not displace Gal4DBD from its binding site (Figure 4 – figure supplement 3). Thus SWI/SNF translocation alone is insufficient to displace a bound Gal4DBD and TF eviction requires nucleosome remodeling. We speculate that TF removal may be accelerated once a nucleosome is repositioned over the bound TF. Indeed, a recent single molecule fluorescence study of Gal4 binding kinetics on nucleosomal DNA showed that a nucleosome regulates Gal4 binding not only by preventing Gal4 binding, but also by dramatically increasing the Gal 4 dissociation rate from the DNA (Luo et al., 2014).

SWI/SNF family remodelers are known to be involved in transcriptional activation. Genome-wide mapping of yeast indicates that, apart from localizing to nucleosomes around transcription

start sites, SWI/SNF family remodelers are also enriched upstream of the promoter regions (Yen et al., 2012). Genome wide analysis of the locations of human chromatin remodelers also found that Brg1, Chd4 and Snf2h are highly enriched at the promoter and distal upstream regions (Morris et al., 2014). Since many relevant transcriptional modulators, such as enhancers (Ren, 2010) and insulators (Bell et al., 2001), are located further upstream of promoters, SWI/SNF family remodelers could move promoter nucleosomes to dynamically regulate these factors. Thus, although SWI/SNF alone does not possess any ability to remove TFs on its own, our work shows that SWI/SNF can slide nucleosomes to displace neighboring TFs around the promoter region, providing a mechanistic basis for dynamically clearing both nucleosomes and other bound factors upon SWI/SNF recruitment (Nagaich et al., 2004).

REFERENCES

Aoyagi, S., Narlikar, G., Zheng, C., Sif, S., Kingston, R.E., and Hayes, J.J. (2002). Nucleosome remodeling by the human SWI/SNF complex requires transient global disruption of histone-DNA interactions. Mol Cell Biol *22*, 3653-3662.

Bai, L., Ondracka, A., and Cross, F.R. (2011). Multiple sequence-specific factors generate the nucleosome-depleted region on CLN2 promoter. Mol Cell *42*, 465-476.

Barrera, L.O., and Ren, B. (2006). The transcriptional regulatory code of eukaryotic cells--insights from genome-wide analysis of chromatin organization and transcription factor binding. Curr Opin Cell Biol *18*, 291-298.

Bell, A.C., West, A.G., and Felsenfeld, G. (2001). Insulators and boundaries: versatile regulatory elements in the eukaryotic genome. Science *291*, 447-450.

Blosser, T.R., Yang, J.G., Stone, M.D., Narlikar, G.J., and Zhuang, X. (2009). Dynamics of nucleosome remodelling by individual ACF complexes. Nature *462*, 1022-1027.

Brower-Toland, B., and Wang, M.D. (2004). Use of optical trapping techniques to study single-nucleosome dynamics. Methods Enzymol *376*, 62-72.

Clapier, C.R., and Cairns, B.R. (2009). The biology of chromatin remodeling complexes. Annu Rev Biochem *78*, 273-304.

Dechassa, M.L., Sabri, A., Pondugula, S., Kassabov, S.R., Chatterjee, N., Kladde, M.P., and Bartholomew, B. (2010). SWI/SNF has intrinsic nucleosome disassembly activity that is dependent on adjacent nucleosomes. Mol Cell *38*, 590-602.

Dechassa, M.L., Wyns, K., Li, M., Hall, M.A., Wang, M.D., and Luger, K. (2011). Structure and Scm3-mediated assembly of budding yeast centromeric nucleosomes. Nat Commun *2*, 313.

Dechassa, M.L., Zhang, B., Horowitz-Scherer, R., Persinger, J., Woodcock, C.L., Peterson, C.L., and Bartholomew, B. (2008). Architecture of the SWI/SNF-nucleosome complex. Mol Cell Biol *28*, 6010-6021.

Deindl, S., Hwang, W.L., Hota, S.K., Blosser, T.R., Prasad, P., Bartholomew, B., and Zhuang, X. (2013). ISWI remodelers slide nucleosomes with coordinated multi-base-pair entry steps and single-base-pair exit steps. Cell *152*, 442-452.

Engeholm, M., de Jager, M., Flaus, A., Brenk, R., van Noort, J., and Owen-Hughes, T. (2009). Nucleosomes can invade DNA territories occupied by their neighbors. Nat Struct Mol Biol *16*, 151-158.

Fyodorov, D.V., and Kadonaga, J.T. (2002). Dynamics of ATP-dependent chromatin assembly by ACF. Nature *418*, 897-900.

Gangaraju, V.K., and Bartholomew, B. (2007). Dependency of ISW1a chromatin remodeling on extranucleosomal DNA. Mol Cell Biol *27*, 3217-3225.

Gelbart, M.E., Rechsteiner, T., Richmond, T.J., and Tsukiyama, T. (2001). Interactions of Isw2 chromatin remodeling complex with nucleosomal arrays: analyses using recombinant yeast histones and immobilized templates. Mol Cell Biol *21*, 2098-2106.

Gkikopoulos, T., Schofield, P., Singh, V., Pinskaya, M., Mellor, J., Smolle, M., Workman, J.L., Barton, G.J., and Owen-Hughes, T. (2011). A role for Snf2-related nucleosome-spacing enzymes in genome-wide nucleosome organization. Science *333*, 1758-1760.

Hall, M.A., Shundrovsky, A., Bai, L., Fulbright, R.M., Lis, J.T., and Wang, M.D. (2009). High-resolution dynamic mapping of histone-DNA interactions in a nucleosome. Nat Struct Mol Biol *16*, 124-129.

Hassan, A.H., Neely, K.E., and Workman, J.L. (2001). Histone acetyltransferase complexes stabilize swi/snf binding to promoter nucleosomes. Cell *104*, 817-827.

Hauk, G., and Bowman, G.D. (2011). Structural insights into regulation and action of SWI2/SNF2 ATPases. Curr Opin Struct Biol *21*, 719-727.

Hota, S.K., Bhardwaj, S.K., Deindl, S., Lin, Y.C., Zhuang, X., and Bartholomew, B. (2013). Nucleosome mobilization by ISW2 requires the concerted action of the ATPase and SLIDE domains. Nat Struct Mol Biol *20*, 222-229.

Hughes, A.L., Jin, Y., Rando, O.J., and Struhl, K. (2012). A functional evolutionary approach to identify determinants of nucleosome positioning: a unifying model for establishing the genome-wide pattern. Mol Cell *48*, 5-15.

Imbalzano, A.N., Schnitzler, G.R., and Kingston, R.E. (1996). Nucleosome disruption by human SWI/SNF is maintained in the absence of continued ATP hydrolysis. Journal of Biological Chemistry *271*, 20726-20733.

Inman, J.T., Smith, B.Y., Hall, M.A., Forties, R.A., Jin, J., Sethna, J.P., and Wang, M.D. (2014). DNA y structure: a versatile, multidimensional single molecule assay. Nano Lett *14*, 6475-6480.

Jiang, J., Bai, L., Surtees, J.A., Gemici, Z., Wang, M.D., and Alani, E. (2005). Detection of high-affinity and sliding clamp modes for MSH2-MSH6 by single-molecule unzipping force analysis. Mol Cell *20*, 771-781.

Jin, J., Bai, L., Johnson, D.S., Fulbright, R.M., Kireeva, M.L., Kashlev, M., and Wang, M.D. (2010). Synergistic action of RNA polymerases in overcoming the nucleosomal barrier. Nat Struct Mol Biol *17*, 745-752.

Kadonaga, J.T. (2004). Regulation of RNA polymerase II transcription by sequence-specific DNA binding factors. Cell *116*, 247-257.

Kang, J.G., Hamiche, A., and Wu, C. (2002). GAL4 directs nucleosome sliding induced by NURF. EMBO J *21*, 1406-1413.

Kassabov, S.R., Zhang, B., Persinger, J., and Bartholomew, B. (2003). SWI/SNF unwraps, slides, and rewraps the nucleosome. Mol Cell *11*, 391-403.

Koch, S.J., Shundrovsky, A., Jantzen, B.C., and Wang, M.D. (2002). Probing protein-DNA interactions by unzipping a single DNA double helix. Biophys J *83*, 1098-1105.

Kornberg, R.D. (2007). The molecular basis of eukaryotic transcription. Proc Natl Acad Sci U S A *104*, 12955-12961.

Kwon, H., Imbalzano, A.N., Khavari, P.A., Kingston, R.E., and Green, M.R. (1994). Nucleosome disruption and enhancement of activator binding by a human SW1/SNF complex. Nature *370*, 477-481.

Lee, K.M., and Narlikar, G. (2001). Assembly of nucleosomal templates by salt dialysis. Curr Protoc Mol Biol *Chapter 21*, Unit 21 26.

Li, B., Carey, M., and Workman, J.L. (2007). The role of chromatin during transcription. Cell *128*, 707-719.

Li, M., and Wang, M.D. (2012). Unzipping single DNA molecules to study nucleosome structure and dynamics. Methods Enzymol *513*, 29-58.

Lia, G., Praly, E., Ferreira, H., Stockdale, C., Tse-Dinh, Y.C., Dunlap, D., Croquette, V., Bensimon, D., and Owen-Hughes, T. (2006). Direct observation of DNA distortion by the RSC complex. Molecular Cell *21*, 417-425.

Lickwar, C.R., Mueller, F., Hanlon, S.E., McNally, J.G., and Lieb, J.D. (2012). Genome-wide protein-DNA binding dynamics suggest a molecular clutch for transcription factor function. Nature *484*, 251-255.

Logie, C., and Peterson, C.L. (1997). Catalytic activity of the yeast SWI/SNF complex on reconstituted nucleosome arrays. EMBO J *16*, 6772-6782.

Lowary, P.T., and Widom, J. (1998). New DNA sequence rules for high affinity binding to histone octamer and sequence-directed nucleosome positioning. J Mol Biol *276*, 19-42.

Luo, Y., North, J.A., Rose, S.D., and Poirier, M.G. (2014). Nucleosomes accelerate transcription factor dissociation. Nucleic Acids Res *42*, 3017-3027.

Lusser, A., Urwin, D.L., and Kadonaga, J.T. (2005). Distinct activities of CHD1 and ACF in ATP-dependent chromatin assembly. Nature Structural & Molecular Biology *12*, 160-166.

Mirny, L.A. (2010). Nucleosome-mediated cooperativity between transcription factors. P Natl Acad Sci USA *107*, 22534-22539.

Morris, S.A., Baek, S., Sung, M.H., John, S., Wiench, M., Johnson, T.A., Schiltz, R.L., and Hager, G.L. (2014). Overlapping chromatin-remodeling systems collaborate genome wide at dynamic chromatin transitions. Nature Structural & Molecular Biology *21*, 73-+.

Moyle-Heyrman, G., Tims, H.S., and Widom, J. (2011). Structural constraints in collaborative competition of transcription factors against the nucleosome. J Mol Biol *412*, 634-646.

Nagaich, A.K., Walker, D.A., Wolford, R., and Hager, G.L. (2004). Rapid periodic binding and displacement of the glucocorticoid receptor during chromatin remodeling. Mol Cell *14*, 163-174.

Petesch, S.J., and Lis, J.T. (2012). Overcoming the nucleosome barrier during transcript elongation. Trends Genet *28*, 285-294.

Ren, B. (2010). Transcription: Enhancers make non-coding RNA. Nature *465*, 173-174.

Sengupta, S.M., VanKanegan, M., Persinger, J., Logie, C., Cairns, B.R., Peterson, C.L., and Bartholomew, B. (2001). The interactions of yeast SWI/SNF and RSC with the nucleosome before and after chromatin remodeling. J Biol Chem *276*, 12636-12644.

Shundrovsky, A., Smith, C.L., Lis, J.T., Peterson, C.L., and Wang, M.D. (2006). Probing SWI/SNF remodeling of the nucleosome by unzipping single DNA molecules. Nat Struct Mol Biol *13*, 549-554.

Sirinakis, G., Clapier, C.R., Gao, Y., Viswanathan, R., Cairns, B.R., and Zhang, Y.L. (2011). The RSC chromatin remodelling ATPase translocates DNA with high force and small step size. Embo Journal *30*, 2364-2372.

Stockdale, C., Flaus, A., Ferreira, H., and Owen-Hughes, T. (2006). Analysis of nucleosome repositioning by yeast ISWI and Chd1 chromatin remodeling complexes. J Biol Chem *281*, 16279-16288.

Struhl, K., and Segal, E. (2013). Determinants of nucleosome positioning. Nat Struct Mol Biol *20*, 267-273.

Teves, S.S., Weber, C.M., and Henikoff, S. (2014). Transcribing through the nucleosome. Trends Biochem Sci *39*, 577-586.

Torigoe, S.E., Urwin, D.L., Ishii, H., Smith, D.E., and Kadonaga, J.T. (2011). Identification of a rapidly formed nonnucleosomal histone-DNA intermediate that is converted into chromatin by ACF. Mol Cell *43*, 638-648.

Tsukiyama, T., Palmer, J., Landel, C.C., Shiloach, J., and Wu, C. (1999). Characterization of the imitation switch subfamily of ATP-dependent chromatin-remodeling factors in Saccharomyces cerevisiae. Genes Dev *13*, 686-697.

Voss, T.C., and Hager, G.L. (2014). Dynamic regulation of transcriptional states by chromatin and transcription factors. Nat Rev Genet *15*, 69-81.

Voss, T.C., Schiltz, R.L., Sung, M.H., Yen, P.M., Stamatoyannopoulos, J.A., Biddie, S.C., Johnson, T.A., Miranda, T.B., John, S., and Hager, G.L. (2011). Dynamic exchange at regulatory elements during chromatin remodeling underlies assisted loading mechanism. Cell *146*, 544-554.

Whitehouse, I., Rando, O.J., Delrow, J., and Tsukiyama, T. (2007). Chromatin remodelling at promoters suppresses antisense transcription. Nature *450*, 1031-1035.

Yadon, A.N., Van de Mark, D., Basom, R., Delrow, J., Whitehouse, I., and Tsukiyama, T. (2010). Chromatin remodeling around nucleosome-free regions leads to repression of noncoding RNA transcription. Mol Cell Biol *30*, 5110-5122.

Yamada, K., Frouws, T.D., Angst, B., Fitzgerald, D.J., DeLuca, C., Schimmele, K., Sargent, D.F., and Richmond, T.J. (2011). Structure and mechanism of the chromatin remodelling factor ISW1a. Nature *472*, 448-453.

Yang, J.G., Madrid, T.S., Sevastopoulos, E., and Narlikar, G.J. (2006). The chromatin-remodeling enzyme ACF is an ATP-dependent DNA length sensor that regulates nucleosome spacing. Nat Struct Mol Biol *13*, 1078-1083.

Yen, K., Vinayachandran, V., Batta, K., Koerber, R.T., and Pugh, B.F. (2012). Genome-wide nucleosome specificity and directionality of chromatin remodelers. Cell *149*, 1461-1473.

Zhang, Y., Moqtaderi, Z., Rattner, B.P., Euskirchen, G., Snyder, M., Kadonaga, J.T., Liu, X.S., and Struhl, K. (2010). Evidence against a genomic code for nucleosome positioning Reply. Nature Structural & Molecular Biology *17*, 920-923.

Zhang, Y., Smith, C.L., Saha, A., Grill, S.W., Mihardja, S., Smith, S.B., Cairns, B.R., Peterson, C.L., and Bustamante, C. (2006). DNA translocation and loop formation mechanism of chromatin remodeling by SWI/SNF and RSC. Mol Cell *24*, 559-568.

Zofall, M., Persinger, J., Kassabov, S.R., and Bartholomew, B. (2006). Chromatin remodeling by ISW2 and SWI/SNF requires DNA translocation inside the nucleosome. Nat Struct Mol Biol *13*, 339-346.

CHAPTER 3

REPAIR FACTOR MFD AND ITS FUNCTION IN TRANSCRIPTION ELONGATION

Transcription-coupled repair (TCR) is a sub-pathway of nucleotide excision repair (NER) that acts specifically on lesions in the transcribed strand of expressed genes. It distinguishes itself from global excision repair at the lesion recognition step of the repair process. In bacteria, the transcription repair coupling factor, Mfd, displaces stalled RNA polymerase (RNAP) at a lesion prior to the recruitment of the Uvr repair complex and the ensuing repair. However, many studies implied, but not confirmed, that Mfd could function beyond the scope of transcription coupled repair. Using a combination of biochemical and newly developed single molecule assays, we systematically studied the motion of Mfd and its relationship with RNAP. We demonstrated that Mfd is able to translocate independently on double strand DNA with speed of ~5 bp/s and processivity of several hundred base pairs. This translocation processitivty is highly increased by interacting with RNAP. Additionally, we discovered that Mfd stabilizes transcription elongation complexes and enhances the elongation rate by reducing RNAP pausing frequency and duration. When an RNAP does pause, depending upon the pause duration, the bound Mfd could either pushes the RNAP forward to resume effective elongation or disrupt the RNAP from the template to terminate the transcription. Our data provide important insight into the role of Mfd beyond the scope of transcription coupled repair and significantly contribute to the understanding of Mfd function in a larger context within transcription.

INTRODUCTION

As a sub pathway of nucleotide excision repair (NER), transcription-coupled repair (TCR) acts specifically on lesions in the transcribed strand of expressed genes and distinguishes itself at the lesion recognition step of the repair process (Reardon and Sancar, 2005). In bacteria, when an elongating RNA polymerase encounters the lesion, it becomes arrested by backtracking at a

lesion and is displaced by the transcription repair coupling factor, Mfd. Mfd then recruits the NER lesion-recognition factor UvrA, and dissociates from the DNA. UvrA binds UvrB, and the assembled UvrAB complex initiates repair, a process similar to the global repair (Selby and Sancar, 1993, 1995a, b; Selby et al., 1991). Efficient TCR in E.coli requires Mfd. Mfd's human homolog CSB protein is essential to neurological functions and its defection will cause diseases such as Cockayne syndrome(McMurray, 2010).

The E. coli Mfd is a monomeric 130 kDa protein containing eight structural domains (Figure 3.1). The N-terminal portion, consisting of D1a, D2, D1b, and D3, is joined to the C-terminal portion (D4, D5, D6, and D7) by a flexible linker of 23 amino acids (Deaconescu et al., 2006). Structure and biochemical studies have shown that Mfd can interact with the β subunit of the RNA polymerase when the elongation complex (EC) is arrested by either a DNA lesion, a DNA-bound protein, or by nucleotide depletion(Chambers et al., 2003; Park et al., 2002; Westblade et al., 2010). Using ATP hydrolysis as the energy source, the bound Mfd can displace the stalled RNA polymerase and the nascent transcript from the transcribing DNA (Park and Roberts, 2006). During TCR, this disruption process helps expose the lesion so that the repair enzyme can get access readily once the lesion is recognized.

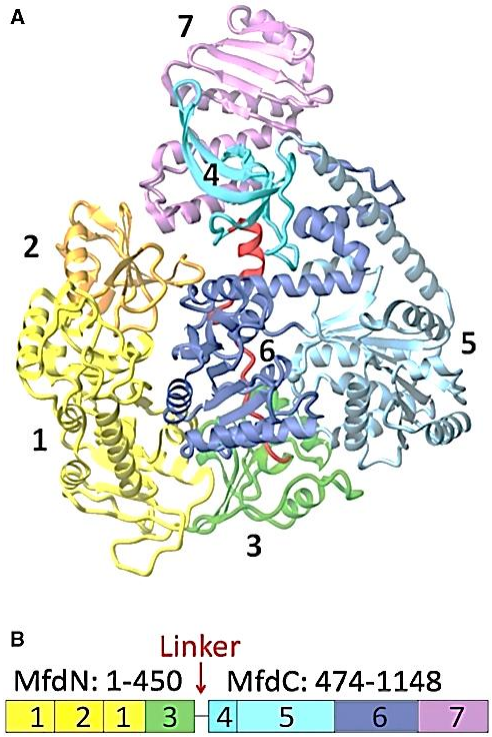


Figure 3.1. Domains of Mfd and structure of the enzyme in its apo-form. (A) Structure of Mfd [PDB-ID 2EYQ, (6)]. Domains are color-coded and the backbone is shown as a ribbon indicating secondary structure. In the orientation shown, the ATP-binding site and the putative DNA-binding site are facing the viewer. (B) Domain organization and definition of MfdN, MfdC and linker regions.

However, TCR is a complex process that requires multiple proteins to function around the lesion sequentially. How Mfd recognizes RNAP and accomplishes the RNAP displacement is not clear (Deaconescu et al., 2012a; Manelyte et al., 2010). Since the crystal structures of RNAP have shown that an elongating RNAP and a stalled RNAP have no difference structurally (Brueckner et al., 2007), it is likely that Mfd can not only bind onto a lesion induced stalled RNAP, but also any type of stalled RNAP. Since Mfd is a DNA translocase and pushes a temporarily stalled RNAP forward (Deaconescu et al., 2012b; Murphy et al., 2009; Smith et al., 2007), it is also likely that Mfd travels with an elongating RNAP and constantly governs the elongating rate of the RNAP by preventing it from pausing induced by multiple reasons. Last but not least,

pervious results within other types of bacterial indicated the possible contribution of Mfd to translesion synthesis (Robleto et al., 2012). Therefore, it is highly possible that Mfd is able to determine the fate of the lesion stalled RNAP based on specific types of lesion. These hypotheses broaden Mfd's role within transcription, from a bridge during transcription coupled repair, to a general factor that associates with elongation to ensure the healthy transcription in bacteria.

In this study, we used a combination of biochemical and single molecule assays to systematically study the relationship between the function of Mfd and an elongating RNAP. In particularly, we developed a new single molecule unzipping tracker methods to robustly monitor the motion of Mfd alone and Mfd-RNAP complexes in real time. We demonstrated that Mfd is able to translocate independently on double strand DNA with low speed (5-10 bp/s) and low processitivity. Additionally, Binding of Mfd to RNAP stabilized the elongation complex by clamping the binding of the RNAP upstream to DNA and upstream DNA, indicating that Mfd may prevent RNAP from backtracking during the elongation process. When an RNAP does pause, depending upon the pause duration and frequency, Mfd either pushes the RNAP forward to resume effective elongation or disrupts the RNAP from the template to terminate the transcription. Our data provides important insight into the role of Mfd beyond the scope of transcription coupled repair and significantly contributes to the understanding of Mfd function in the larger context of transcription.

MATERIALS and METHODS

Plasmids

The plasmids PRL574 containing the T7A1 promoter were obtained originally from Robert Landick Lab and were prepared using standard PCR and cloning methods.

Transcription Template Preparation

Both the transcription templates used in biochemistry and single molecule studies were prepared using the standard PCR methods.

Template construction for transcription gel experiment

To study the anti-pausing effect of Mfd during transcription elongation, we constructed a magnetic bead attached template with T7A1 promoter and A20 downstream to stall the RNAP. As is shown Figure 3.2, The template was PCRed with an end-biotin labeled primer 350bp upstream of the A20 position and a reverse primer ~550bp downstream of the A20 position.

Figure 3.2. Template design for the transcription gel experiment

After the template was constructed, the bead attachment procedure is listed as follows:

1. Vigorously shake the stock bead (NEB). Then pipette out 200µl beads in a 1.5ml centrifuge tube.
2. Put the tube on the separate rack, wait for 10 seconds. The bead should be sucked onto the wall of the tube. Pipette out the supernatant and discard.

3. Add 200ul wash buffer (Transcription buffer in this case: 25 mM Tris-Cl, pH 8.0, 100 mM KCl, 4 mM $MgCl_2$, 1 mM DTT, 3% (v/v) glycerol, 0.15 mg ml^{-1} acetylated BSA,). Take the tube off the rack (NEB), vigorously shake the bead, and briefly centrifuge the tube.
4. Repeat the step 2 and 3 another two times. The bead will be properly washed. This step ensures there is no free streptavidin, which can attach to biotin-labeled DNA and affect the attachment efficiency.
5. After the washing, put the tube back on the rack, wait for 10 seconds, pipette out and discard the supernatant.
6. Add 100µl DNA (100nM) into the tube. Vigorously shake for 15min-30min and briefly centrifuge. The sample is ready to be used.
7. If not immediately used, the DNA sample should be stored at 4 °C and can last for months.

The magnetic bead separation procedure is listed as follows:

a. After the DNA is attached to the beads, place the tube in the magnetic separation rack. Wait for 10 second. Pipette out the supernatant and transfer it to a separate tube (label it as supernatant 1).
b. After separating the supernatant from the pallet, QUICKLY add same volume (as the PTC volume) of the 1xTB100 into the tube, and start tapping the tube to mix the solution. You do not want to vortex the solution right now because protein is bound on the DNA. Then put the tube back on the rack and wait for 10 seconds. Pipet out the supernatant and transfer it to a separate tube (label it as supernatant 2).

c. Repeat step b a couple of times. Add same volume of transcription buffer into the tube. Gently mix the solution.

Transcription Gel to study Mfd's ant-pausing activity

The Paused Transcription Complex (PTC) was formed by incubation of 20 nM *E. coli* RNAP, 4 nM magnetic bead attached transcription DNA templates, 250 μM ApU initiating dinucleotide, 50 μM ATP/GTP, and α-[^{32}P] CTP [5 μCi (1 μCi = 37 GBq) at 3,000 Ci/mmol] in transcription buffer (25 mM Tris•Cl, pH 8.0, 100 mM KCl, 4 mM $MgCl_2$, 1 mM DTT, 3% (v/v) glycerol, 0.15 mg/mL acetylated BSA) for 20 min at 37ºC to form paused transcription complexes (PTCs) which contain DNA, RNAP and 20 nt RNA transcript(Jin et al., 2010).

After PTCs were formed, we performed the magnetic bead separation to eliminate the free excess NTPs in the solution. We then added 0.5μM Mfd and incubated it with PTCs for 5min. After that we started to add 1μM dATP and NTPs with various concentrations. Transcription reactions were quenched at predetermined time points by addition of 10 mM EDTA. Transcripts were analyzed on polyacrylamide sequencing gels and imaged with PhosphorImager (Molecular Dynamics) (Jin et al., 2010; Shundrovsky et al., 2004). The different NTPs concentrations will generate the different types of pausing patterns during the transcription elongation. We compared the pausing patterns in the presence and in the absence of Mfd under same NTPs concentration and same time point.

Template construction for single molecule experiment

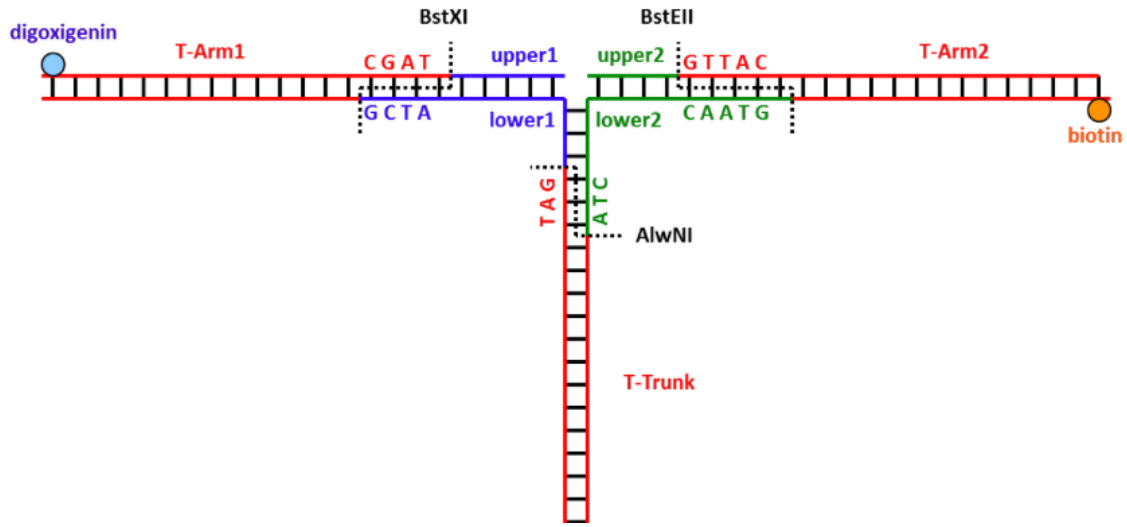


Figure 3.3. Construction of the Y structure. (1) Arm 1 DNA was cut from plasmid pMDW38 (sequence available upon request) and its 5' end was labeled with digoxigenin with a Klenow reaction. (2) Upper 1 (5'-/phos/GCA GTA CCG AGC TCA TCC AAT TCT ACA TGC CGC) and lower 1(5'-/phos/GCC TTG CAC GTG ATT ACG AGA TAT CGA TGA TTG CG GCG GCA TGT AGA ATT GGA TGA GCT CGG TAC TGC ATC G) were annealed to form adapter 1. (3) Adapter 1 was ligated to arm 1 and the product was gel purified to remove un-ligated adapters. (4) Steps 1-3 were repeated for arm 2: upper 1 (5'-CGT TAC GTC ATT CTA TAC ACT GTA CAG) and lower 2 (5'-/phos/GTAAC CTG TAC AGT GTA TAG AAT GAC GTA ACG CGC AAT CAT CGA TAT CTC GTA ATC ACG TGC AAG GC CTA). (5) Arm 1 and arm 2 were annealed via lower 1 and lower 2. (6) Trunk DNA was prepared by PCR from pRL574 (7) Arms were ligated to trunk DNA.

In the singe molecule experiments, we constructed a Y structure template to study the dynamics of Mfd and Mfd-RNAP complex in real time (Inman et al., 2014). The Y structure template was newly developed in our lab based upon the traditional unzipping technique. The Y arms further separate the distance between the trapping microsphere and the unzipping surface, minimizing the surface effect to the dynamics of the targeted protein. Briefly, the Y structure DNA was constructed from three distinct dsDNA segments: two arms and the trunk (Figure 3.3). The arms were made by restriction enzyme cuts from plasmid pMDW38 (sequence available upon request) for symmetric arms or from plasmid pRL574 for asymmetric arms. A single restriction cut (XhoI

or SphI) in this plasmid created an overhang that was subsequently filled in with either dig-dUTP or bio-dATP by Klenow polymerase (NEB) to provide specific attachment to anti-digoxigenin or streptavidin coated microspheres respectively. A second restriction cut (BstXI or BstEII) created an overhang for ligation to an annealed trunk adapter oligos to generate a long (>30 bp) overhang on each arm. The two trunk adaptor oligos from the two arms were complementary to each other and were subsequently annealed to form Y arms with a short trunk (~30 bp). The annealed adaptor oligos were designed to create an overhang for ligation to the full-length trunk. Such a design is modular so that the trunk is interchangeable. Trunk DNA was made via PCR with pRL574 and then cut with a restriction enzyme (AlwNI) to provide the proper overhang for ligation to the Y arms. The trunk PCR primers were designed so that the after AlwNI cutting and ligation with the adaptor, the T7A1 promoter is ~340bp downstream of the unzipping origin.

Single Molecule Unzipping Tracking

The single molecule unzipping tracking experiment was developed based on the traditional unzipping experiment. The sample preparation process was similar to previously described (Koch and Wang, 2003; Li and Wang, 2012) with the following modifications: Y Structure DNA tethers were formed by first adding antidigoxigenin to the sample chamber, which binds nonspecifically to the coverglass surface. Next the digoxigenin end-labeled DNA construct (with PTC complex in Mfd/RNAP complex tracking experiment) was added to chamber, which was followed by the addition of streptavidin-coated 0.48 nm polystyrene microspheres. Finally the transcription solution was flowed into the sample chamber just prior to data acquisition.

Data were collected using a single beam optical trap as described previously(Koch et al., 2002). After the single molecule PTC chamber was formed, Mfd (2uM) was flown in to the chamber to incubate with PTC of the template for 5min. Then dATP solution was added as we started to mechanically unzipping the Y Structure DNA template. Once the unzipping fork encounters a bound protein, the force will rise above the DNA unzipping baseline (15pN). To successfully track the encountered protein or protein complexes, we set the threshold force at 18pN (higher than DNA baseline unzipping, but lower than the force required to unzipping through the bound protein) and kept it constant so that the fork could follow the motion of the encountered protein without disrupting the protein from the template.

The constant force tracking experiments were conducted while the fork position, or velocity were achieved via computer-con- trolled feedback. Data were low-pass filtered to 5 kHz and digitized at ~12 kHz, then were further averaged to 110 Hz. Experiments were conducted at a room temperature. The acquired data signals were converted into force and number of base pairs unzipped as previously described (Koch et al., 2002), with the exception of using a more refined force-extension curve for ssDNA.

RESULTS

Mfd translocation is detected by Single Molecule Unzipping Tracker

In E. coli, Mfd is a monomeric 130 kDa DNA translocase containing eight structural domains. Although structure studies showed that Mfd belongs to a translocase family, the pervious triplex

displacement essay did not detect any independent translocation activity (Smith et al., 2007). We felt that this seemingly conflict might be explained by Mfd's intrinsic property as a weak translocase and would like to revisit this idea using the single molecule real time tracking technique.

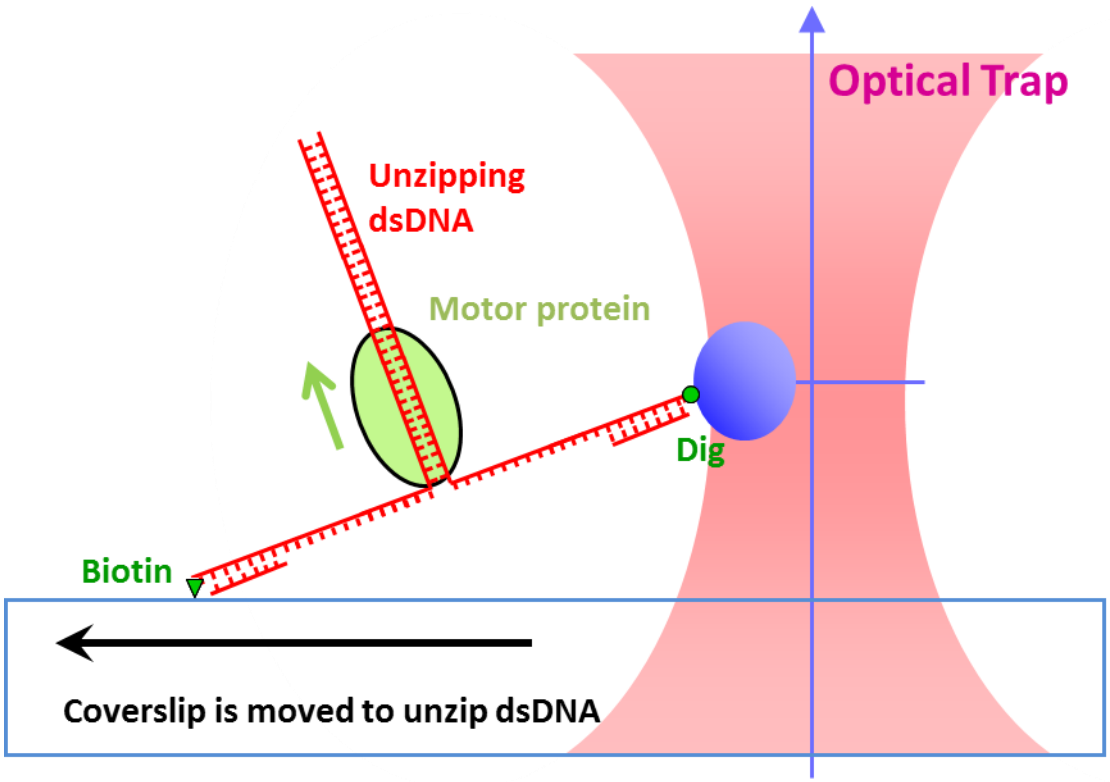


Figure 3.4. Experiment configuration of single molecule unzipping tracker. The DNA template was first mechanically unzipped until the fork encountered the motor protein. Then the force will set as a constant number (above the DNA unzipping baseline but below the force needed to disrupt the DNA) so that the unzipping fork can follow the motion of the motor protein in real time.

Based on the single molecule unzipping technique (Li and Wang, 2012), we developed a single molecule unzipping tracking assay (unzipping tracker), a robust technique to monitor the real-

time motion of motor proteins with minimum perturbation (Figure 3.4). When the unzipping force reaches the bound motor protein, the force will be tuned so that it will unzip the DNA but won't disrupt the motor protein. In this case, the unzipping fork can be used to track the motion of the motor by following the motor or letting the motor push again it. Compared to the pervious optical trapping based real-time method, unzipping tracker frees the targeting protein from being labeled to the surface or the trapping bead, both of which could affect the activity and the function of the target. It is a technique specifically suitable for monitoring the behavior of weak motor proteins with low speed and low processivity.

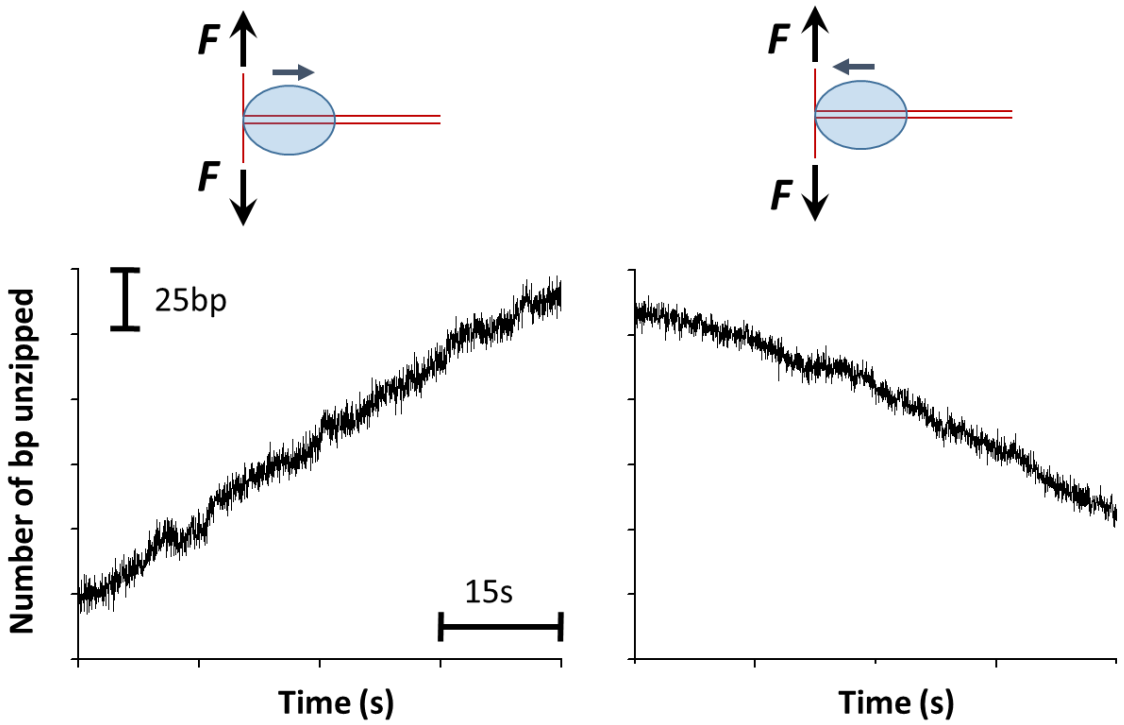


Figure 3.5. Real time tracking Mfd translocating on double stranded DNA with 2mM ATP. The left panel show the representative tracking trace when unzipping fork follows the direction of the Mfd motion. The right panel shows the representative tracking trace when unzipping fork moving against Mfd motion. (Studies were conducted collaboratively by me and Dr. Yi Yang. The data were collected by Dr. Yi Yang)

With unzipping tracker in hand, we monitored the behavior of a single Mfd on a long piece of DNA. To our surprise, we successfully detected a robust behavior of Mfd moving along the DNA by itself (Figure 3.5). Unzipping tracker monitored Mfd's movement in real time by either using the unzipping fork to follow the Mfd or letting Mfd push against the unzipping fork. Statistically speaking, Mfd translocates on DNA alone with an average speed of 5-10bp/s (Figure 3.6). The similar result obtained from the double direction tracking ruled out the possible effect of unzipping fork on the function of Mfd protein. The processivity histogram shows a single population around the zero, corresponding to the Mfd population that is not moved. The distribution difference between the speed statistics and the processitivity statistics indicates that of all the Mfd molecules that we worked on, a significant portion of the Mfd moved at a distinct speed, however, the processivity of these moved Mfd molecules is extremely low of 187bp (Figure 3.7). To our knowledge, this is the first time that Mfd's intrinsic translocation was detected. As a DNA translocase, Mfd indeed demonstrates the translocation capability by performing ATP hydrolysis. In addition, our results show that Mfd is a weak motor compared to other tranlocase in the same category, reconciling the seemingly conflict between the results from structure study and triples displacement assays.

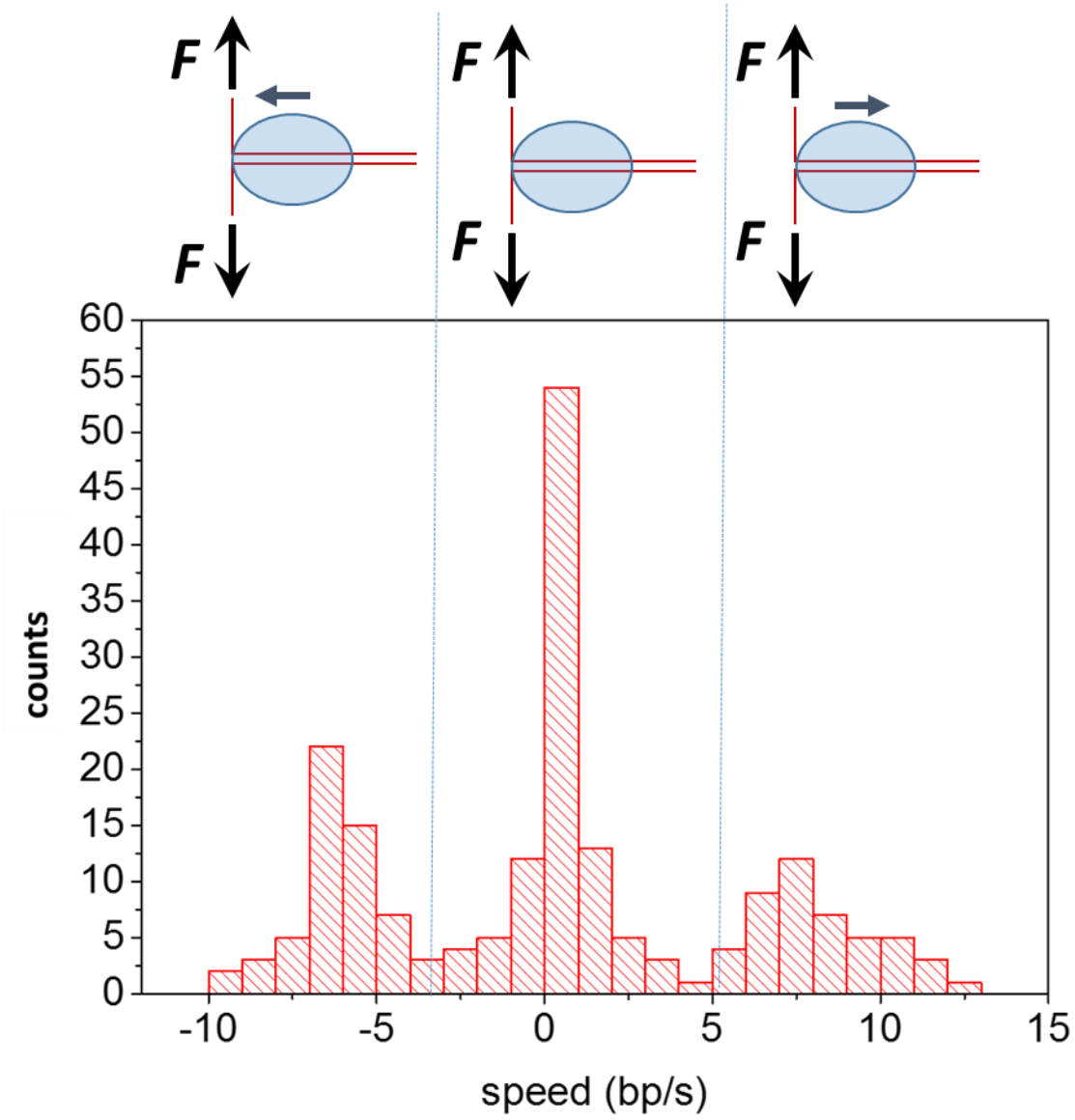


Figure 3.6: Statistical Analysis of Mfd translocation speed on double stranded DNA with 2mM ATP. The entire figure can be divided into three panels. The left panel show the translocation speed when mfd moves against the unzipping fork. The right panel shows the translocation speed when mfd moves in the same direction. The middle panel shows the portion of Mfd molecules that moved very slow or don't move at all. (Studies were conducted collaboratively by me and Dr. Yi Yang. The data were collected by Dr. Yi Yang)

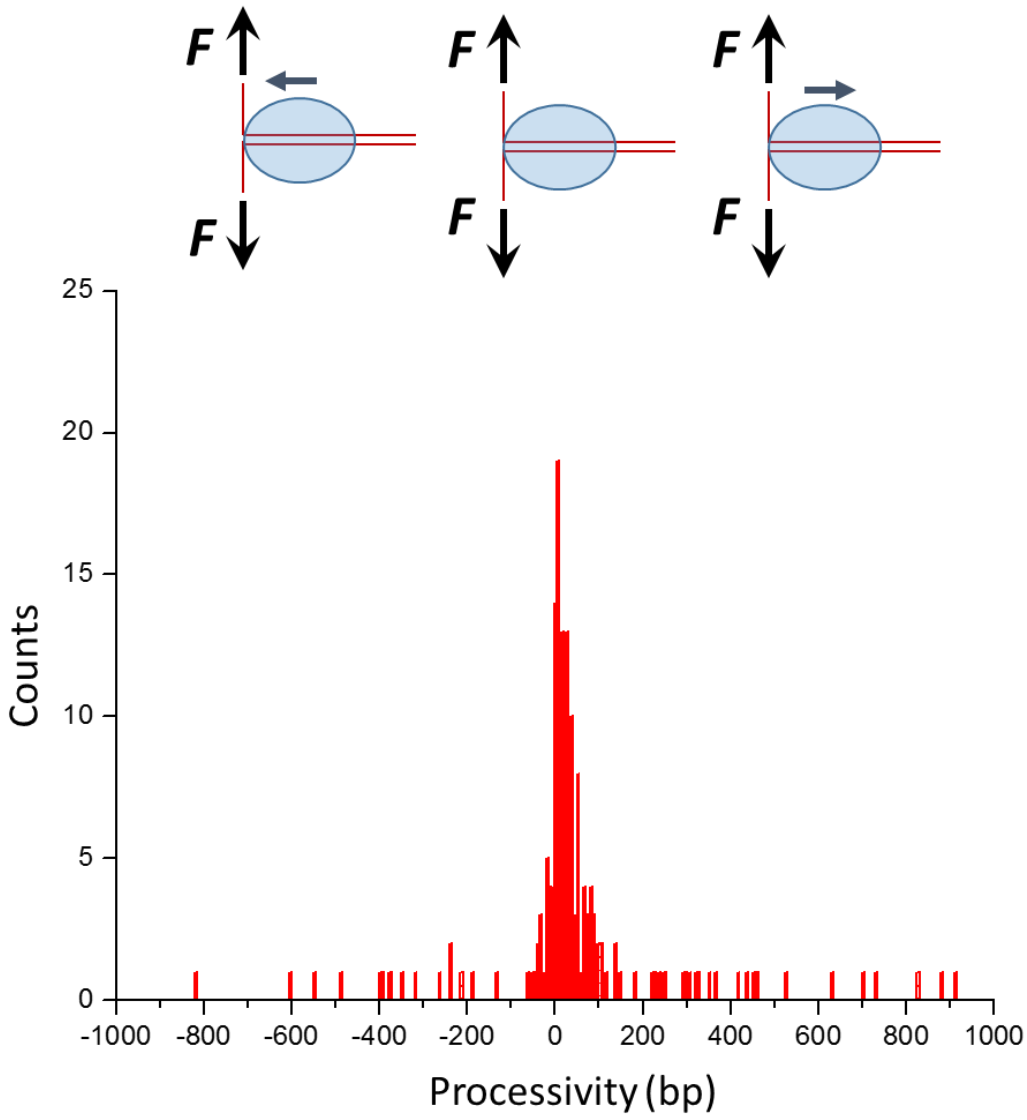


Figure 3.7: Statistical Analysis of Mfd translocation processitivity on double stranded DNA with 2mM ATP. (Studies were conducted collaboratively by me and Dr. Yi Yang. The data were collected by Dr. Yi Yang)

Our finding of Mfd as a weak motor not only provides important insights on Mfd's function itself, but also indicates that Mfd's function can be transient and might be facilitated by other protein or protein complexes. Indeed, the recent biochemistry (Haines et al., 2014) and single molecule experiments(Graves et al., 2015) have shown that after disrupting RNAP, Mfd is capable of traveling on the DNA up to 1000bp with a 3bp/s speed. In combination of pervious

results, our data could imply that Mfd needs to be stabilized by RNAP in vivo in order to fully function as a translocase.

Mfd stabilizes RNAP with two layers of binding to elongation complex.

Since RNAP stabilizes the Mfd as a DNA translocase, we used traditional unzipping technique to further map out the major interactions between Mfd and RNAP when both of them were on the DNA, with the purpose of study the detailed interactions between RNAP and Mfd (Jin et al., 2010). After a paused transcription complex (PTC) at +20 nt position was formed as previously described, we unzipped through the PTC to map out the unzipping signature of the PTC alone on the template as the baseline. After that, we added Mfd with or without ATPrS, hopefully to see the unzipping signature changes, which would be attributed to the interaction between Mfd and RNAP or the interaction between Mfd and DNA.

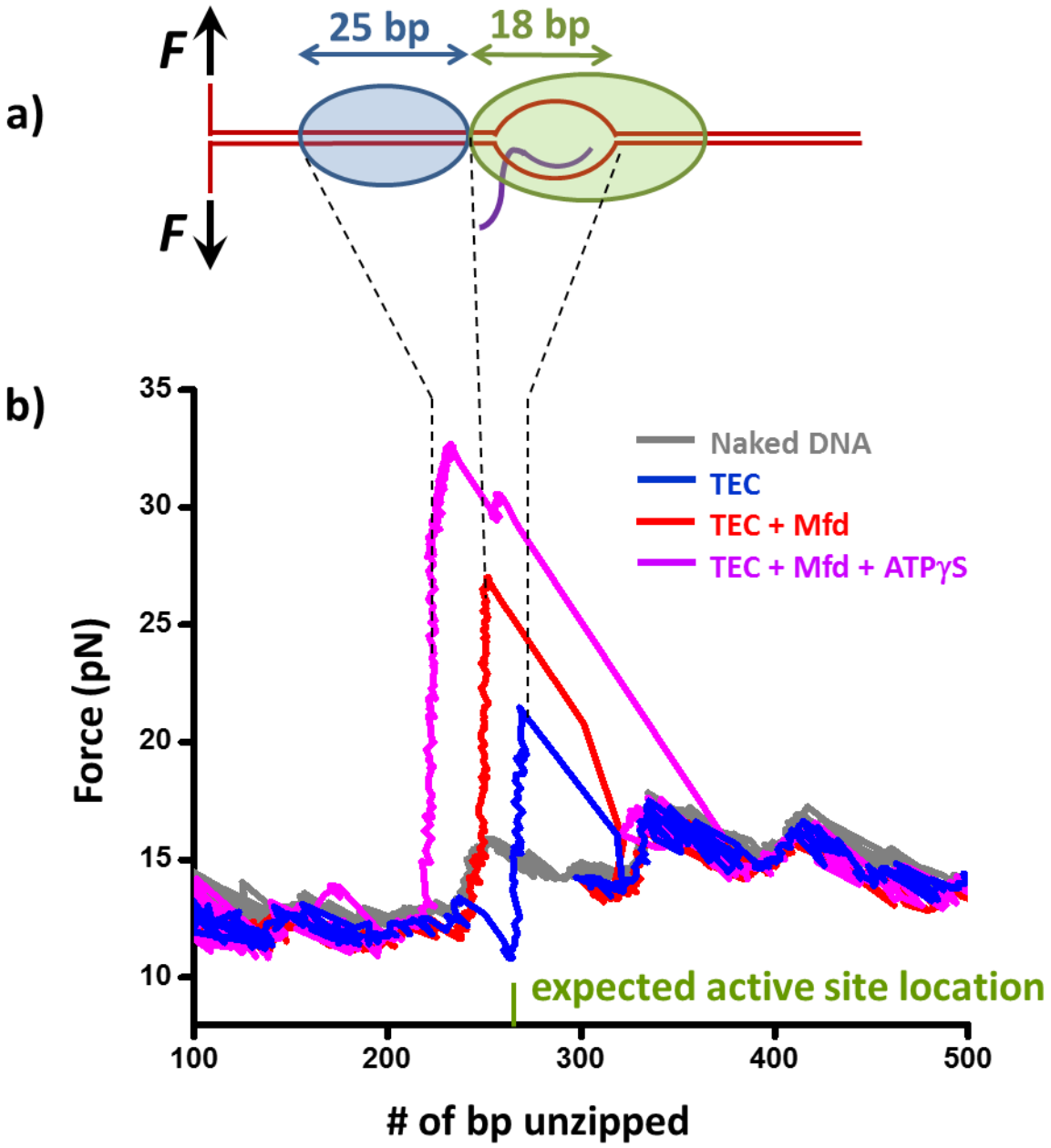


Figure 3.8: Representative traces of the unzipping signatures of RNAP and RNAP-Mfd conmplexes. The grey line indicates the DNA baseline. The blue line indicate the unzipping signature of RNAP alone, The red line indicates the unzipping signature of Mfd and RNAP in the absence of ATP. The pink line indicates the unzipping signature of Mfd and RNAP in the presence of ATPrS. (Studies were conducted collaboratively by me and Dr. Yi Yang)

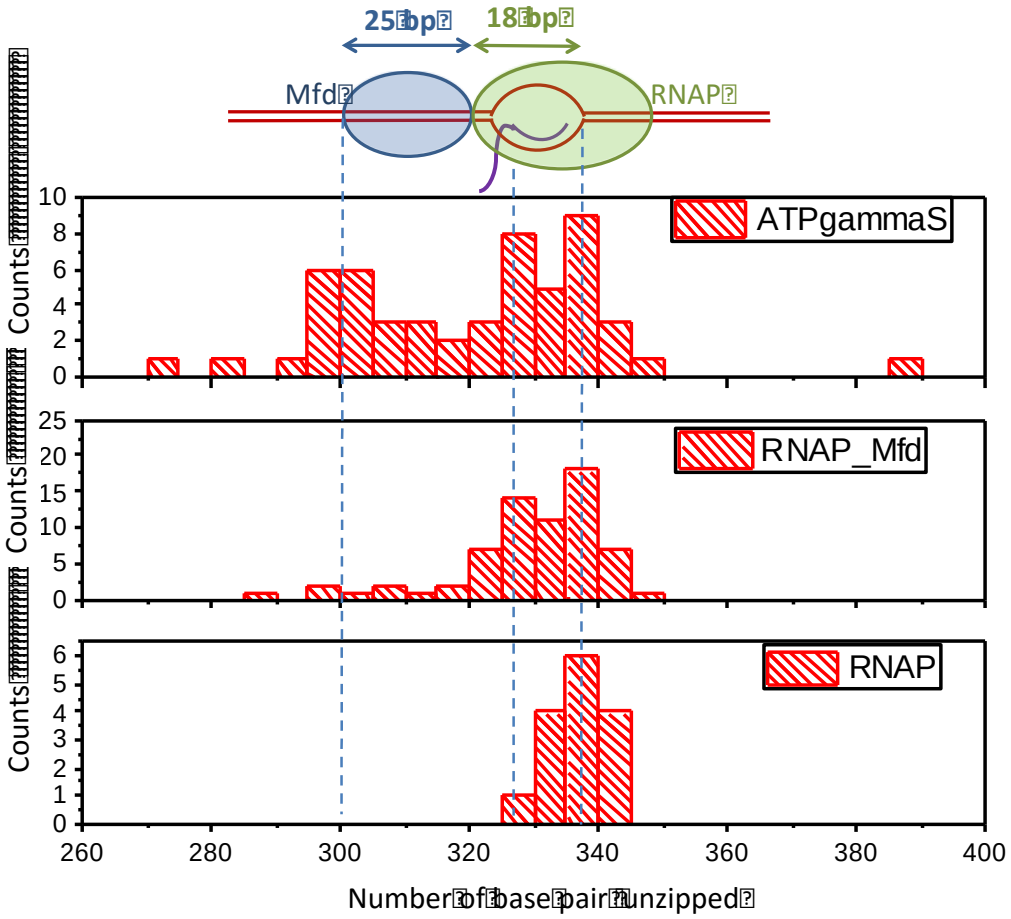


Figure 3.9: Statistical analysis of the positions of the unzipping signature of RNAP, RNAP and Mfd without ATP as well as RNAP and Mfd in the presence of ATPrS. (Studies were conducted collaboratively by me and Dr. Yi Yang)

As is shown in Figure 3.8, the three scenarios showed drastically different unzipping signatures. Without adding Mfd or ATPrS, We detected the onset of the force drop when the unzipping fork encountered the transcription bubble formed by RNAP while the location of the subsequent force rise indicated the end of the transcription bubble and the beginning of the dsDNA that was clamped downstream by the RNAP. The magnitudes of the force drop and the rise varied from trace to trace as would be expected from a thermally activated, off-equilibrium process.

When we added the Mfd to the PTC without ATPrS, we detected a smaller force drop (sometimes no force drop) at the upstream of the DNA (Figure 3.9). Instead the force rise started at the upstream region of the transcription bubble (where the force drop started in the case of PTC alone). The shrinkage of the detected bubble size indicates that Mfd is not only able to interact with the upstream of RNAP, but also helps RNAP to clamp the upstream of the transcription bubble to stabilize the PTC complex.

After we added ATPrS, we detected a second force rise ~25bp upstream of the RNAP (Figure 3.9 top panel). The location of new peak is consistent with Mfd's binding signature to upstream of the PTC predicted by the pervious biochemistry study and crystal structure, indicating that with the help of ATP, Mfd can exert a second layer of stabilization to RNAP for more robust elongation.

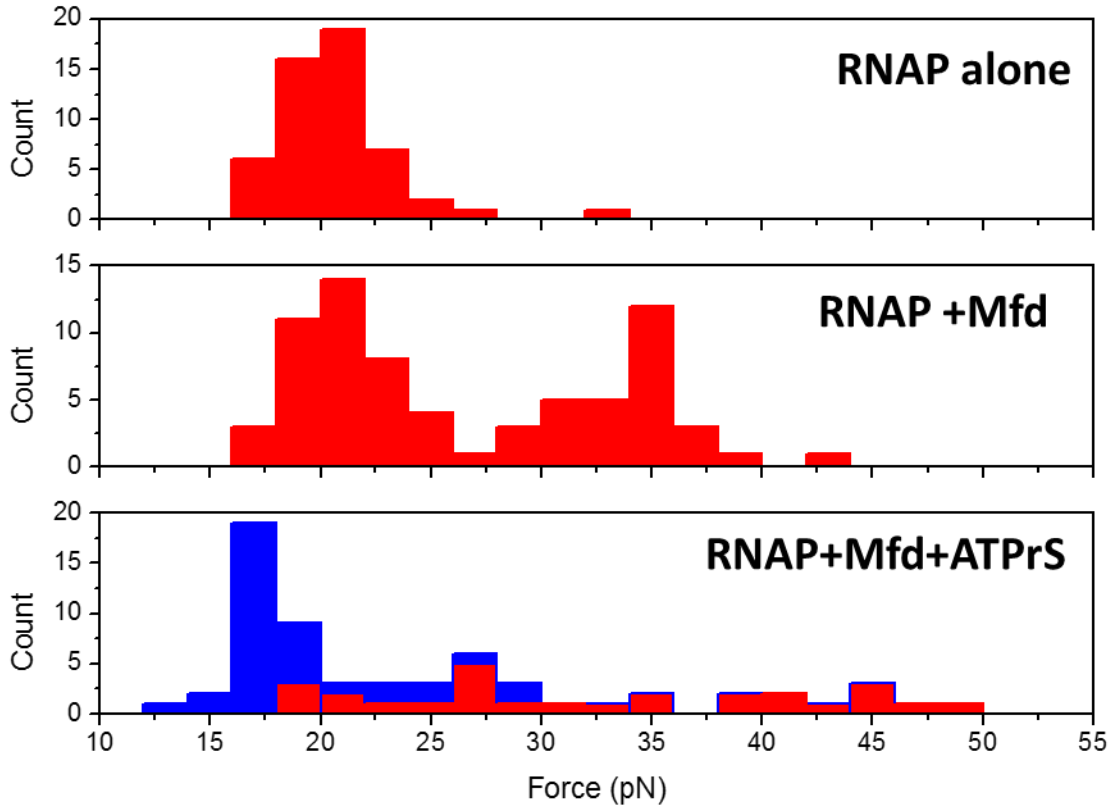


Figure 3.10. The statistical force analysis of unzipping signature of RNAP alone, RNAP-Mfd complexes without ATP, and RNAP-Mfd with ATPrS. (Studies were conducted collaboratively by me and Dr. Yi Yang)

The statistical force analysis (Figure 3.10) of the unzipping peak also showed that RNAP alone's unzipping force peaked at ~20pN. However, with the addition of Mfd, the statistical force analysis demonstrated a second peak around 35pN. With the addition of Mfd and ATPrS, the RNAP's force statistics became more spread-out, reaching as high as 50pN. The unzipping force increase indicates that the addition of Mfd stabilizes the entire complex (The unzipping force histogram is spread towards the high force region after Mfd binds on RNAP).

The two layers of binding could imply that Mfd plays a unique role during transcription elongation, with binding to RNAP as a co-factor, and binding to DNA to ensure the healthy elongation of RNAP. The two layers of Mfd interactions with PTC, along with the results shown in pervious section indicate an interesting relationship between Mfd and RNAP. Both motors relies on each other to maximize their own functions.

Mfd determines the fate of a paused RNAP by its pausing degree

The upstream DNA binding of Mfd in the presence of ATPrS made us speculate that Mfd might use it to prevent RNAP from pausing to ensure robust elongation. Additionally, since Mfd has been known to disrupt a lesion stalled RNAP, what determines the outcome of a paused RNAP after being treated by Mfd? To answer this question, we used the biochemical transcription assay to symmetrically study the pausing pattern difference with and without the involvement of Mfd.

In order to generate elongation complexes with different speed and pausing likelihoods, we varied the NTPs concentration from 1μM to 1mM during the RNAP elongations. In addition, we use dATP as Mfd's energy source to separate the nucleotide allocation between RNAP and Mfd. Interestingly, when we compare the pausing pattern of RNAP at the same elongation time frame, but under different NTPs concentrations, we found dramatically different results. When the NTPs concentration was higher than 200μM, Mfd did not seem to affect the elongation pausing pattern and the run-off efficiency of the RNAP. When the NTPs concentration was tuned between 50μM to 200uM, we started to detect the pausing pattern moving to the downstream of the template and the run off efficiency improvement in the presence of Mfd (Figure 3.11e).

Interestingly, when NTPs concentration was tuned below 50μM, we started to detect the opposite effect. In another word, the elongation pausing pattern moved to the upstream of the template in the presence of Mfd. This indicates that under this condition, Mfd would preferably terminate the elongation when RNAP pauses during the elongation (Figure 3.11e).

The pervious study on RNAP elongation dynamics and pausing(Bai et al., 2007) has predicted that during the main elongation pathway, the RNAP rapidly translocates between the pre- and post translocation states under thermal activation. The binding of NTP stabilizes the post-translocation state, and subsequent NTP incorporation biases the polymerase forward by one base pair, with the binding and incorporation kinetics dependent on the NTP type. Additionally, There exist two alternative pausing mechanisms: pausing could be caused either by a slow rate within the main pathway (pre-translocation pause) or by a relatively fast rate of entry into a nonproductive branch pathway with a slow rate of returning to the main pathway (backtracking pause). A large number of DNA sequences may produce back-tracked or pre-translocation pause sites, depending on the stability of the TEC at those sites. Compared with a pre-translocation pause, a backtracking pause tends to occur less frequently but for a longer duration.

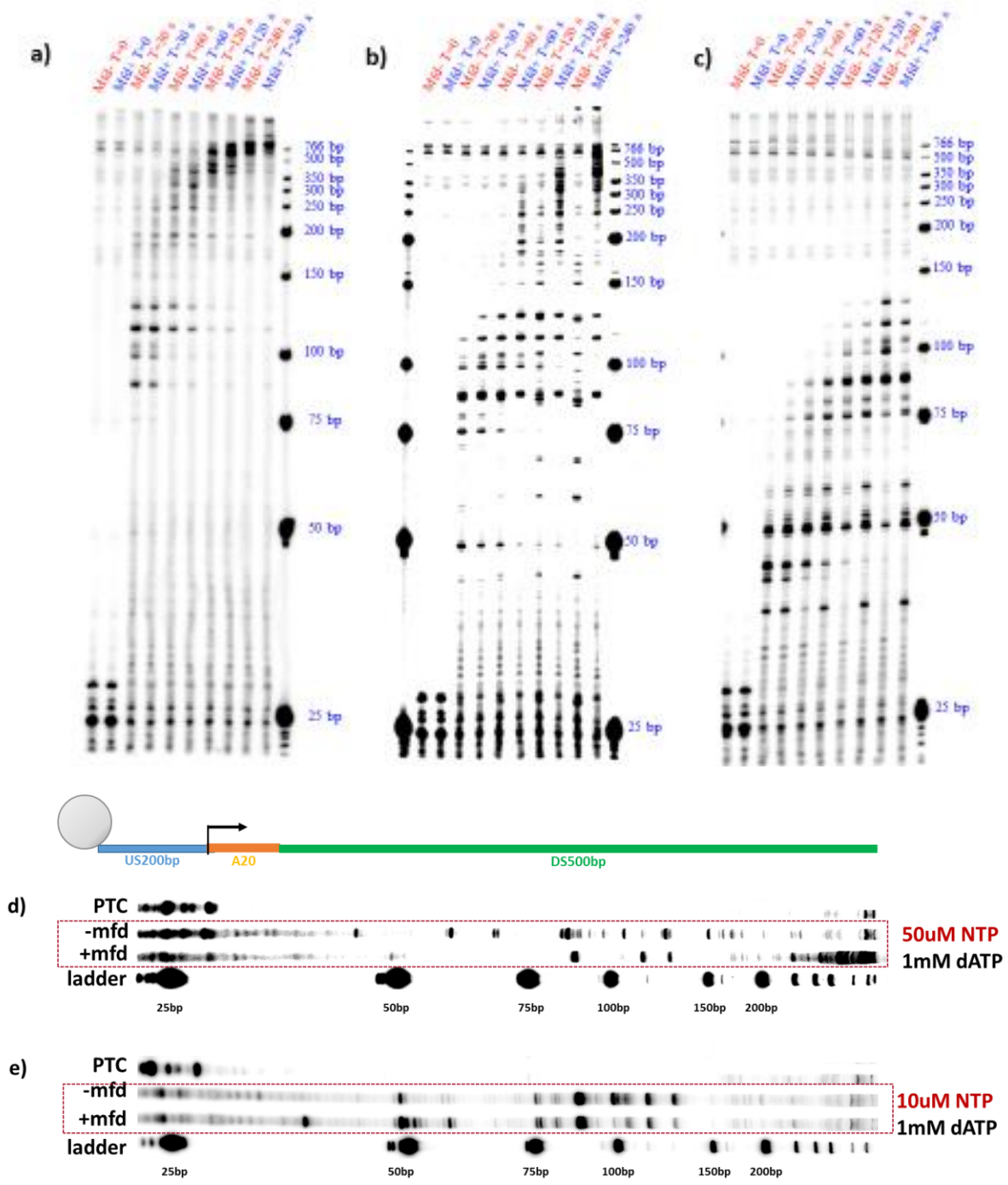


Figure 3.11. The transcription gel analysis to show the dichotomy behavior of Mfd on a paused RNAP. The ATP concentration is varied: a) 100 μM; b) 50 μM; c) 20 μM.; d) direct comparison of elongation patterns at T=240 second with 50μM NTP and 1mM dATP; direct comparison of the elongation pattern at T=240 second with 10μM NTP and 1mM dATP (Preliminary studies were conducted by me, and the gels shown here were obtained by Dr. Chuang Tan)

The dichotomy behavior of Mfd on a paused RNAP has demonstrated that Mfd determines the fate of a pausing RNAP based on its pausing degree or even its pausing mechanism. If RNAP is going through a very transient pausing, it may primarily dwell near the pre- and post-transclocated states even without the help of Mfd. When RNAP is going through a medium pausing with slow speed, Mfd will constantly push the RNAP forward to reduce pausing and increase its elongation speed. However, when RNAP is extremely slow with significantly long pausing, Mfd will most likely treat the RNAP as a stalled one and clean it up from the template so that another round of efficient transcription can be restarted. Additionally, as we decreased the NTP concentration, pauses happened more frequently and their duration became longer, indicating that more backtracked pauses and fewer pre-translocation pauses occurs. Therefore, Mfd could be more likely to help RNAP resume elongation from a pre-translocation pause, and be more likely to terminate the elongation when RNAP goes through a backtracked pause. This function of Mfd is believed to be important in terms of ensuring efficient elongation and wisely allocated the RNAP resources.

DISCUSSION

It has been widely acknowledged that Mfd plays an important role in transcription coupled repair by connecting the lesion stalled RNAP and the repair system. Here we show Mfd and RNAP stabilize with each other to maximize their own function. RNAP increases the processivity of Mfd translocation to make Mfd a strong motor, Mfd in turn stabilizes the RNAP elongation complex by monitoring its elongation and regulate RNAP pausing. Our study goes above and

beyond of the role of Mfd in TCR and provides unique insight of Mfd's role not only in transcription coupled repair but also in a bigger transcription elongation context.

Mfd translocation

Using the newly developed single molecule unzipping tracker, we showed, for the first time, that Mfd translocates on double strand DNA with 5-10bp/s speed and small processitivity. Previous studies used the triplex displacement assay to conclude that Mfd alone could not translocate on DNA and attributed this result to the auto inhibiting effect of D7 domain with in Mfd (Smith et al., 2007). More specifically, An N-terminal clamp formed a cover with D7 domain to restrain the motor domain's activity as a translocase (Murphy et al., 2009). While these explanation was consistent with Mfd's crystal structure (Deaconescu et al., 2006), we feel that the close and open of the cover that inhibits the motor function could be interpreted as a dynamic process rather than a static picture. Therefore, it is very likely that the autoihibitory effect is transiently lifted when Mfd interacts with DNA. In addition, the biochemistry study also indicates that binding to RNAP triggers a conformational change of Mfd to keep the auto inhibitory domain permanently open(Murphy et al., 2009). However, we believe that the function of RNAP is not to open a completely closed auto inhibitory domain, instead it is to lock a dynamic close-open process at the open state. This explanation is confirmed by us and others(Graves et al., 2015; Haines et al., 2014) which showed that Mfd has a large processitivity after RNAP disruption.

The stalled RNAP seems to make Mfd a more reliable translocase on DNA, by increasing the processitivity and increase the unzipping force of Mfd. Our results, in combination of the other

people's studies (Graves et al., 2015; Haines et al., 2014) implies that RNAP is likely to facilitate Mfd during transcription elongation or transcription coupled repair or both. However, we are still not sure, at this point, where the disrupted RNAP is traveling with Mfd or fall off the template. Given the fact that Mfd travels robustly on the template after disrupting the RNAP, we suspect that the disrupted RNAP still attaches to the mfd to lock it at its open conformation to ensure a robust Mfd's translocation (Graves et al., 2015).

Mfd/RNAP relationship and its implication in vivo

The interesting relationship between Mfd and RNAP might indicate that Mfd can be part of RNAP during the active transcription elongation (Figure 3.12). First of all, the self-translocation of Mfd on DNA provides an alternative pathway for Mfd to search for a paused or elongating RNAP. (Figure 3.12 A, B, A'). After Mfd finds the elongating RNAP, Mfd follows the elongating RNAP by binding to the upstream of the RNAP. This layer of binding stabilizes RNAP on DNA by increasing RNAP's upstream clamping (Figure 3.12 C). When RNAP pauses (Figure 3.12 D). Mfd utilizes the energy sources (ATP or dATP) to grab the upstream DNA. This second layer of binding triggers Mfd to push the paused RNAP forward. Depending on the pausing degree or mechanism, Mfd either pushes the RNAP forward to resume the elongation (Figure 3.12F) or pushes the RNAP off the template to terminate the elongation (Figure 3.12E).

Many early studies on Mfd have been presented to imply that Mfd and RNAP could act together not only in transcription coupled repair, but also in transcription elongation. 1) The crystal structure of RNAP showed that elongating RNAP and stalled RNAP had no structural difference (Brueckner et al., 2007). If Mfd could bind to a stalled RNAP, it is equally likely that Mfd would

bind to an elongating RNAP; 2) in vivo studies showed that the copy number of Mfd was comparable to the copy number of active elongation RNAP (Kad and Van Houten, 2012; Klumpp and Hwa, 2008). This implied that Mfd might act as an indispensable part of the RNAP during active elongation; 3) Mfd's yeast counterpart Rpb9 is a party of RNAP II and has been shown to have important function in TCR.(Li and Smerdon, 2002)

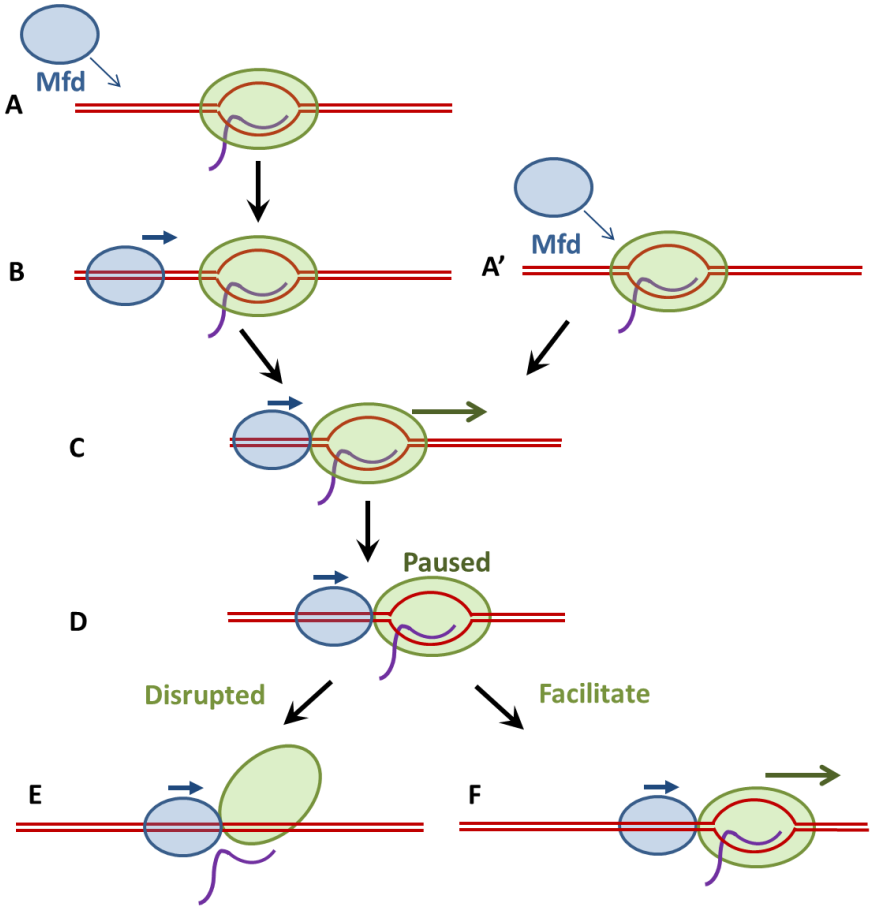


Figure 3.12. The function of Mfd in transcription elongation: Mfd locates a elongating RNAP through 1D or 3D search (A-B or A). Mfd travels with elongating RNAP (C). When RNAP pauses (D), Mfd either disrupt the paused RNAP (E) or push the paused RNAP to resume elongation (F)

In our work, Mfd and RNAP stabilizes with each other to maximize their own functions. This makes Mfd a potential co-factor of RNAP during transcription elongation. In bacterial, multiple

factors, such as GreA, GreB, NusG, have been known to act with RNAP to guarantee the healthy elongation (Burova et al., 1995; Marr and Roberts, 2000). However, none of these factors relies on ATP hydrolysis to perform their functions. As an ATPase, Mfd could provide a regulation at the mechanical level to regulate RNAP, especially when elongation is going through extremely conditions. Specifically, the two layer interaction between Mfd and RNAP provides an interesting mechanism of how Mfd regulate transcription elongation.

Mfd in Transcription coupled repair

The fact that Mfd can translocate after disrupting RNAP may provide new understanding of Mfd's role in transcription couples repair. Historically, RNAP was considered as the primary lesion recognition entity during TCR. Mfd was treated as a bridge to connect RNAP and repair enzyme(Deaconescu et al., 2012a). In particular, RNAP disruption and repair enzyme recruitment were believed to happen in a concerted fashion. However, small-angle X-ray diffraction experiment shows that after Mfd binds to RNAP, the UvrA binding domain was still not completely open, opening up the possibility that repair enzyme recruitment would after RNAP disruption.(Deaconescu et al., 2012b). Our finding complements well with other studies on Mfd tranlocation, showing that after RNAP disruption, Mfd switches from a weak motor to a robust translocase, with large stability on DNA template and long processitivity. In combination with a recent biochemistry work that shows that Mfd has an intrinsic affinity to the lesion (Haines et al., 2014), our results supports a recently proposed model that Mfd might serve a secondary lesion recognition entity when RNAP cannot recognize the lesion effectively for some reasons. Indeed, during transcription elongation, RNAP can stall for different reasons before it

reaches the lesion. In the scenarios when RNAP cannot reach the lesion, the associated Mfd can fulfill the lesion recognition job by itself, as a secondary guarantee during lesion recognition and repair.

REFERENCES

Bai, L., Fulbright, R.M., and Wang, M.D. (2007). Mechanochemical kinetics of transcription elongation. Phys Rev Lett *98*, 068103.

Brueckner, F., Hennecke, U., Carell, T., and Cramer, P. (2007). CPD damage recognition by transcribing RNA polymerase II. Science *315*, 859-862.

Burova, E., Hung, S.C., Sagitov, V., Stitt, B.L., and Gottesman, M.E. (1995). Escherichia coli NusG protein stimulates transcription elongation rates in vivo and in vitro. J Bacteriol *177*, 1388-1392.

Chambers, A.L., Smith, A.J., and Savery, N.J. (2003). A DNA translocation motif in the bacterial transcription--repair coupling factor, Mfd. Nucleic Acids Res *31*, 6409-6418.

Deaconescu, A.M., Artsimovitch, I., and Grigorieff, N. (2012a). Interplay of DNA repair with transcription: from structures to mechanisms. Trends Biochem Sci *37*, 543-552.

Deaconescu, A.M., Chambers, A.L., Smith, A.J., Nickels, B.E., Hochschild, A., Savery, N.J., and Darst, S.A. (2006). Structural basis for bacterial transcription-coupled DNA repair. Cell *124*, 507-520.

Deaconescu, A.M., Sevostyanova, A., Artsimovitch, I., and Grigorieff, N. (2012b). Nucleotide excision repair (NER) machinery recruitment by the transcription-repair coupling factor involves unmasking of a conserved intramolecular interface. Proc Natl Acad Sci U S A *109*, 3353-3358.

Graves, E.T., Duboc, C., Fan, J., Stransky, F., Leroux-Coyau, M., and Strick, T.R. (2015). A dynamic DNA-repair complex observed by correlative single-molecule nanomanipulation and fluorescence. Nat Struct Mol Biol *22*, 452-457.

Haines, N.M., Kim, Y.I., Smith, A.J., and Savery, N.J. (2014). Stalled transcription complexes promote DNA repair at a distance. Proc Natl Acad Sci U S A *111*, 4037-4042.

Inman, J.T., Smith, B.Y., Hall, M.A., Forties, R.A., Jin, J., Sethna, J.P., and Wang, M.D. (2014). DNA Y Structure: A Versatile, Multidimensional Single Molecule Assay. Nano Letters *14*, 6475-6480.

Jin, J., Bai, L., Johnson, D.S., Fulbright, R.M., Kireeva, M.L., Kashlev, M., and Wang, M.D. (2010). Synergistic action of RNA polymerases in overcoming the nucleosomal barrier. Nat Struct Mol Biol *17*, 745-752.

Kad, N.M., and Van Houten, B. (2012). Dynamics of lesion processing by bacterial nucleotide excision repair proteins. Prog Mol Biol Transl Sci *110*, 1-24.

Klumpp, S., and Hwa, T. (2008). Growth-rate-dependent partitioning of RNA polymerases in bacteria. Proc Natl Acad Sci U S A *105*, 20245-20250.

Koch, S.J., Shundrovsky, A., Jantzen, B.C., and Wang, M.D. (2002). Probing protein-DNA interactions by unzipping a single DNA double helix. Biophys J *83*, 1098-1105.

Koch, S.J., and Wang, M.D. (2003). Dynamic force spectroscopy of protein-DNA interactions by unzipping DNA. Phys Rev Lett *91*, 028103.

Li, M., and Wang, M.D. (2012). Unzipping single DNA molecules to study nucleosome structure and dynamics. Methods Enzymol *513*, 29-58.

Li, S., and Smerdon, M.J. (2002). Rpb4 and Rpb9 mediate subpathways of transcription-coupled DNA repair in Saccharomyces cerevisiae. EMBO J *21*, 5921-5929.

Manelyte, L., Kim, Y.I., Smith, A.J., Smith, R.M., and Savery, N.J. (2010). Regulation and rate enhancement during transcription-coupled DNA repair. Mol Cell *40*, 714-724.

Marr, M.T., and Roberts, J.W. (2000). Function of transcription cleavage factors GreA and GreB at a regulatory pause site. Mol Cell *6*, 1275-1285.

McMurray, C.T. (2010). Mechanisms of trinucleotide repeat instability during human development. Nat Rev Genet *11*, 786-799.

Murphy, M.N., Gong, P., Ralto, K., Manelyte, L., Savery, N.J., and Theis, K. (2009). An N-terminal clamp restrains the motor domains of the bacterial transcription-repair coupling factor Mfd. Nucleic Acids Res *37*, 6042-6053.

Park, J.S., Marr, M.T., and Roberts, J.W. (2002). E. coli Transcription repair coupling factor (Mfd protein) rescues arrested complexes by promoting forward translocation. Cell *109*, 757-767.

Park, J.S., and Roberts, J.W. (2006). Role of DNA bubble rewinding in enzymatic transcription termination. Proc Natl Acad Sci U S A *103*, 4870-4875.

Reardon, J.T., and Sancar, A. (2005). Nucleotide excision repair. Prog Nucleic Acid Res Mol Biol *79*, 183-235.

Robleto, E.A., Martin, H.A., and Pedraza-Reyes, M. (2012). Mfd and transcriptional derepression cause genetic diversity in Bacillus subtilis. Front Biosci (Elite Ed) *4*, 1246-1254.

Selby, C.P., and Sancar, A. (1993). Molecular mechanism of transcription-repair coupling. Science *260*, 53-58.

Selby, C.P., and Sancar, A. (1995a). Structure and function of transcription-repair coupling factor. I. Structural domains and binding properties. J Biol Chem *270*, 4882-4889.

Selby, C.P., and Sancar, A. (1995b). Structure and function of transcription-repair coupling factor. II. Catalytic properties. J Biol Chem *270*, 4890-4895.

Selby, C.P., Witkin, E.M., and Sancar, A. (1991). Escherichia coli mfd mutant deficient in "mutation frequency decline" lacks strand-specific repair: in vitro complementation with purified coupling factor. Proc Natl Acad Sci U S A *88*, 11574-11578.

Shundrovsky, A., Santangelo, T.J., Roberts, J.W., and Wang, M.D. (2004). A single-molecule technique to study sequence-dependent transcription pausing. Biophys J *87*, 3945-3953.

Smith, A.J., Szczelkun, M.D., and Savery, N.J. (2007). Controlling the motor activity of a transcription-repair coupling factor: autoinhibition and the role of RNA polymerase. Nucleic Acids Res *35*, 1802-1811.

Westblade, L.F., Campbell, E.A., Pukhrambam, C., Padovan, J.C., Nickels, B.E., Lamour, V., and Darst, S.A. (2010). Structural basis for the bacterial transcription-repair coupling factor/RNA polymerase interaction. Nucleic Acids Res *38*, 8357-8369.