Access to the Genome:
A Study of Transcription Factor Binding Within Nucleosomes

Dissertation

Presented in Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy in the Graduate School of The Ohio State University

By

Matthew Brehove

Graduate Program in Physics

The Ohio State University

2016

Dissertation Committee:
Professor Michael G. Poirier, Advisor
Professor Ralf Bundschuh
Professor Comert Kural
Professor Louis DiMauro
Abstract

All the DNA in a cell’s nucleus is packaged into a material called chromatin consisting of DNA and DNA-associated proteins. The basic unit of chromatin is the nucleosome which consists of ~147 bases of DNA wrapped around a protein core composed of two copies each of histones H2A, H2B, H3, and H4. DNA wrapped into a nucleosome is inaccessible to most DNA processing machinery. This machinery needs access to the DNA to perform processes such as transcription, replication, and repair. The cell uses many mechanisms to modulate this protection including nucleosome unwrapping, sliding, remodeling, disassembly, and chemical modification.

This work used transcription factors along with fluorescent methods to probe the accessibility of DNA near the ends of the nucleosome. The DNA in this entry-exit region can spontaneously unwrap to provide transient access to DNA binding factors such as transcription factors. We studied the effect of histone post-translational modifications on the accessibility of DNA in the entry-exit region. We found that the phosphorylation at H3Y41ph increased DNA accessibility in vitro by 3-fold, which is similar in size to the effect of a previously studied modification H3K56ac. Since both of these modifications are associated with active genes and could occur together, we studied nucleosomes with both modifications and they showed a 17±5-fold increase in accessibility. This indicates that the
cell could use multiple modifications in the entry-exit region to adjust the accessibility of nucleosomal DNA by over an order of magnitude.

We also studied the effects of the removal of one H2A/H2B dimer. These partially formed nucleosomes, called hexasomes, are a component of chromatin that has been largely unstudied. Our results have shown that hexasomes have DNA accessibility similar to that of fully-formed nucleosomes on the side with the remaining H2A/H2B dimer. However, on the side without the dimer, our results showed the DNA remains permanently unwrapped by approximately 40 bases. The unwrapping caused this dimer-distal side to have far greater DNA accessibility amounting to that of naked DNA. These results show that there is a dramatic alteration of DNA accessibility caused by the formation of hexasomes.
Acknowledgements

I would first like to thank my parents Richard and Theresa Brehove for always saying “we support you no matter what” when I told them I wanted to fly to the other side of the continent for six years. Their encouragement and habit of watching Star Trek are what made me interested in science from the beginning. I also want to thank my sisters Lisa and Christina for their support and encouragement. I was also encouraged and held accountable by my friend Stephanie Mate as we shared our frustrations in graduate education.

I want to especially thank those people who allowed me to fail enough to earn this degree. Justin North, who taught me at least 70% of the things I know about biochemistry, taught me more about patience. After I had mistakenly allowed an ice box melt at room temperature overnight, ruining several of Justin's samples, his reaction on realizing it the next morning was one of unfazed forgiveness and understanding.

I want to thank Michael Poirier for all of the helpful advice and guidance he has given me over the years, but I especially want to thank him for the time after I had ruined an octamer preparation eight times in a row. Because when I came to him and explained that I obviously wasn’t capable of doing biochemistry and suggested finding another lab or leaving graduate school, he simply told me to work on the optical setup until I felt better. Later, the preparation worked on the 9th try.
When a first-year student chooses a lab, it is important that they choose one with good lab-mates. I made an excellent choice when I joined the Poirier Lab. All of my lab-mates were willing to drop what they were doing and show me where reagents were, give advice, change dialysis buffers while I was away, and have a good time hanging out in the office.

- To Yi Luo who answered all my constant appeals for help with the ever- idiosyncratic single molecule microscope and helped collect data for the *Journal of Biological Chemistry* paper.
- To Matt Gibson for making the triclops microscope which I used extensively, and fixing most of the equipment that broke in the lab.
- To Marek Simon for being the first one to show me around the Poirier Lab.
- To Morgan Bernier thank you for keeping the lab running and teaching me protein expression.
- To Alex Moony for introducing me to many fantastic board games.
- To Michael Darcy and Ben Donovan for taking charge of the lab and training new students as Matt Gibson and I moved to working on our dissertations. The Poirier Lab is in good hands.

I would also like to thank my co-authors John C. Shimko for making our modified proteins and Tao Wang for performing the MNase digestion and SAXS experiments as well as their advisors Jennifer Ottesen and Karolin Luger. Lastly, I would like to thank my committee members Comert Kural, Ralf Bundschuh, and Louis DiMauro for agreeing to be part of my committee and follow me through this process.
Vita

2010........................................B.S. Physics, University of California Santa Barbara
2012........................................M.S. Physics, Ohio State University
2010-2013 .................................Graduate Teaching Associate, Department of Physics, The Ohio State University
2013-2016 .................................Graduate Research Associate, Department of Physics, The Ohio State University

Publications


Fields of Study

Major Field: Physics
Table of Contents

Abstract .............................................................................................................................. ii

Acknowledgements ........................................................................................................ iv

Vita ................................................................................................................................ vi

Publications .................................................................................................................... vi

Fields of Study ............................................................................................................... vi

Table of Contents .......................................................................................................... vii

List of Tables .................................................................................................................. xiii

List of Figures ................................................................................................................. xiv

1. Introduction ................................................................................................................ 1

   1.1 Epigenetics ......................................................................................................... 1

   1.2 DNA Packaging in Cells .................................................................................. 2

   1.3 Post-Translational Modifications ..................................................................... 4

       1.3.1 PTMs on Histones ................................................................................... 5

   1.4 Partially Formed Nucleosomes ....................................................................... 6

   1.5 Transcription Factors ....................................................................................... 7
1.6 Our Biophysical Approach ................................................................. 8

2. Fluorescent Methods ........................................................................... 10

2.1 Introduction to Fluorescence Resonance Energy Transfer .................. 10

2.2 Protein Induced Fluorescence Enhancement ..................................... 13

2.3 Measuring Transcription Factor Binding with FRET ....................... 15

2.4 Ensemble Fluorescence Measurements ............................................. 18

3. Single Molecule Total Internal Reflection Microscopy ......................... 21

3.1 Introduction to smTIRF ................................................................. 21

3.2 Single Molecule TIRF Setup .......................................................... 23

3.3 Single Molecule TIRF Data Acquisition ........................................... 25

3.4 Single Molecule Data Analysis ....................................................... 28

4. Sample Preparation Methods ............................................................. 33

4.1 Labeled DNA Preparation ............................................................ 33

4.1.1 Polymerase Chain Reaction (PCR) ........................................... 33

4.1.2 Primer Labeling ......................................................................... 35

4.2 Histone Octamer Preparation ......................................................... 36

4.3 Histone Labeling and Purification .................................................. 40

4.3.1 Histone Labeling ....................................................................... 40

4.2.2 Histone Octamer Purification ................................................... 42
4.4 Nucleosome Reconstitution and Purification .............................................. 42

5. Histone Core Phosphorylation Regulates DNA Accessibility ........................... 46

5.1 Introduction ................................................................................................. 46

5.2 Experimental Methods ............................................................................... 49

5.2.1 Preparation of DNA Constructs .............................................................. 50

5.2.2 Preparation of H3Y41ph and H3K56ac ..................................................... 51

5.2.3 Nucleosome Preparation ....................................................................... 52

5.2.4 LexA Preparation .................................................................................... 53

5.2.5 Micrococcal Nuclease Measurements of Nucleosome DNA Accessibility .... 53

5.2.6 Size Exclusion Chromatography with Multi-angle Light Scattering (SEC-
MALS) ........................................................................................................... 55

5.2.7 Small Angle X-ray Scattering SAXS ....................................................... 56

5.2.8 FRET Measurements of LexA Binding within the Nucleosome ............... 57

5.3 Results ......................................................................................................... 58

5.3.1 H3Y41E and H3T45E Increase Nucleosome Unwrapping ....................... 58

5.3.2 H3Y41 Phosphorylation and H3Y41E and H3T45E Each Increase DNA
Accessibility within the DNA Entry-Exit Region of the Nucleosome .................. 61

5.4 Discussion ................................................................................................... 64

6. Histone Modification Combinations ............................................................. 68
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.1</td>
<td>Introduction</td>
<td>68</td>
</tr>
<tr>
<td>6.2</td>
<td>Results</td>
<td>69</td>
</tr>
<tr>
<td>6.2.1</td>
<td>The Combination of H3Y41ph and H3K56ac Multiplicatively Increases DNA Accessibility</td>
<td>69</td>
</tr>
<tr>
<td>6.2.2</td>
<td><em>Combining H3Y41ph and H3K56ac Amplifies the Differences between Their Mimics</em></td>
<td>71</td>
</tr>
<tr>
<td>6.2.3</td>
<td>H3Y41E and H3T45E Mimics Show Redundant Effect</td>
<td>72</td>
</tr>
<tr>
<td>6.3</td>
<td>Discussion</td>
<td>75</td>
</tr>
<tr>
<td>7.0</td>
<td>Hexasomes Show Dramatically Increased Accessibility for Dimer-Distal DNA</td>
<td>77</td>
</tr>
<tr>
<td>7.1</td>
<td>Introduction</td>
<td>77</td>
</tr>
<tr>
<td>7.2</td>
<td>Experimental Methods</td>
<td>80</td>
</tr>
<tr>
<td>7.2.1</td>
<td>Reconstitution of Hexasomes</td>
<td>80</td>
</tr>
<tr>
<td>7.2.2</td>
<td>Exonuclease III Digestions</td>
<td>82</td>
</tr>
<tr>
<td>7.2.3</td>
<td>Ensemble FRET Experiments</td>
<td>84</td>
</tr>
<tr>
<td>7.2.4</td>
<td>Ensemble PIFE Experiments</td>
<td>85</td>
</tr>
<tr>
<td>7.3.5</td>
<td>Single Molecule FRET Experiments</td>
<td>86</td>
</tr>
<tr>
<td>7.3.6</td>
<td>Single Molecule PIFE Experiments</td>
<td>87</td>
</tr>
<tr>
<td>7.3.1</td>
<td>Exonuclease Mapping Shows Asymmetric Protection in Hexasomes</td>
<td>87</td>
</tr>
</tbody>
</table>
E.1 General PCR Primer Design .................................................................................. 128

E.2 Design Considerations for Ligated Constructs ...................................................... 129
List of Tables

Table 1 Summary of Nucleosomal DNA Accessibility as Measured by LexA Binding

................................................................................................................................................. 74
List of Figures

Figure 1 Crystal Structure of a Nucleosome Core Particle ........................................ 3
Figure 2 Tyrosine Phosphorylation .............................................................................. 5
Figure 3 Fluorescence Resonance Energy Transfer (FRET) .......................................... 11
Figure 4 Cis and Trans Isoforms of the Cy3 Fluorophore ............................................. 15
Figure 5 Measuring Transcription Factor Binding with FRET ...................................... 16
Figure 6 Schematic Diagram of the Prism-Based smTIRF Microscope ......................... 23
Figure 7 Single Molecule FRET Nucleosome Design .................................................. 25
Figure 8 Single Molecule Time Traces .......................................................................... 28
Figure 9 Lifetime Analysis of Idealized Trace Data ....................................................... 32
Figure 10 Nucleosomal DNA Preparation Schematic ................................................... 35
Figure 11 Histone Protein Expression and Purification .................................................. 38
Figure 12 Histone Octamer Refolding and Reconstitution ............................................. 39
Figure 13 Nucleosome Reconstitution Sucrose Gradient Fractions ................................. 43
Figure 14 Design and Analysis of Nucleosomes Used for Ensemble FRET Experiments .......................................................... 49
Figure 15 Analysis of Purified Fully Synthetic Proteins H3Y41ph and H3Y41ph/K56ac .......................................................... 52
Figure 16 Micrococcal Nuclease (MNase) Digestion of WT, H3Y41E, H3T45E Nucleosomes

Figure 17 SEC-MALS Nucleosome Quality Control

Figure 18 SAXS Shows H3Y41E and H3T45E Mutations Result in Extended Nucleosomes

Figure 19 FRET Measurements of LexA Binding Within Nucleosomes

Figure 20 Relative Reduction of the S1/2 (S1/2 unmod / S1/2 PTM) for PTM Combinations

Figure 21 Close Up of the Nucleosome Entry-Exit Region

Figure 22 Hexasome Structure

Figure 23 Hexasome Reconstitution Dimer Titration

Figure 24 Sucrose Gradient Purification Fractions

Figure 25 DNA Constructs Used in Hexasome Studies

Figure 26 ExoIII Mapping in Nucleosomes and Hexasomes

Figure 27 Swapping the 601 Dimer Side Preference

Figure 28 Single Molecule FRET Data Analysis

Figure 29 Ensemble and Single Molecule TF Binding on the Dimer-Proximal Side

Figure 30 FRET Efficiencies on the Weak and Strong Side of Nucleosomes and Hexasomes

Figure 31 Ensemble and Single Molecule TF Binding on the Dimer-Distal Side

Figure 32 Diagram Showing the Assembly of Flow Cells
1. Introduction

This chapter first gives an introduction to epigenetics in general and describes how DNA is packaged in the cell. It will give the reader background on post-translational modifications (PTM’s), partially formed nucleosomes, and transcription factors which all serve to regulate gene expression. The next three chapters will detail the experimental methods we used in these studies. Chapters 6-8 show the specific strategies and results from our study of transcription factor binding inside of PTM modified nucleosomes and partially formed nucleosomes.

1.1 Epigenetics

An analogy can be made between a cell's genome and a cookbook. Like a cookbook, the genome contains genes with exact instructions for making every kind of protein in the cell. These proteins, by their sequence of amino-acid subunits and folded structure, perform all manner of cellular functions from digesting nutrients to replicating the genome for reproduction. In the same way, a cookbook will tell a cook how to prepare every dish for every occasion. But a cookbook doesn't tell the cook when to make what dish or how much to make. It is true that all of the information is about the cell's behavior is encoded in its DNA because DNA is what persists through generations and so must contain all relevant information. But it is also true that skin cells, blood cells, and brain cells all share the same DNA, yet behave very differently. Cells also adjust their protein
production due to all manner of environmental factors such as stress, temperature, or the presence of foods or toxins. This regulation of the genome comes from the varying concentration of many kinds of micro-RNAs and transcription factors, but there is also a code written on top of the genetic code. Chemical modifications to DNA and DNA-associated proteins along the genome serve as temporary markers to control the output of the genes associated with them and tell the cook what to cook. The study of these marks (which can sometimes be inherited) is called epigenetics.

1.2 DNA Packaging in Cells

The human genome has about 2 meters of DNA divided into 23 chromosomes and is stored in a nucleus that is only a few microns across. This packing is quite an energetic feat considering that DNA is a stiff polymer that has an energetic cost associated with bending. An unconstrained polymer like DNA that obeys the worm-like-chain model has an average end-to-end distance given by

\[ \langle R^2 \rangle = 2Pl \left( 1 - \frac{P}{l} \left( 1 - e^{-\frac{l}{P}} \right) \right) \]  

where \( R \) is the end-to-end distance; \( l \) is the contour length of the polymer; and \( P \) is the persistence length of the polymer or the characteristic length over which the DNA can be thought of as stiff. Thus, if it was allowed to expand without compaction the average chromosome would have a root-mean-squared (RMS) end-to-end distance of over 93\( \mu \)m given that the DNA persistence length is 50nm. In the cell the DNA is compacted because it is wrapped around a protein complex called histone octamer which is composed of two
copies each of histones H2A, H2B, H3, and H4. 147 base pairs of DNA wrapped around a histone octamer core form a nucleosome as shown in Figure 1.

**Figure 1 Crystal Structure of a Nucleosome Core Particle.** This figure shows x-ray crystallography data from a nucleosome shown in cartoon diagram. It is approximately 10nm across with a core of two copies each of the four histone proteins and 147 base pairs of DNA wrapped 1.6 times around the core. (Protein Data Bank (PDB): 1AOI)

DNA has a 166 base pair persistence length, but in a nucleosome, just 147 bases are wrapped in 1.6 turns\(^1\) bending much more than it would be due to thermal energy. With only thermal bending, this DNA would have an RMS end-to-end distance equal to 87% of its contour length. Given the persistence length of DNA, one can estimate\(^3\) that the whole energy of bending the DNA into a nucleosome is \(52 \, k_B T\) and given that the free energy of nucleosome formation is \(42 \, k_B T\) \(^4\) this implies \(94 \, k_B T\) of energy is involved in the charge-charge interactions and hydrogen bonding between the histone core and the DNA. These are largely due to the many positively charged lysines and argenines on the histone surface.
These nucleosomes form all along the genomic DNA with about 20-30 bases of linker DNA in between them. These strings of nucleosomes are referred to as chromatin. Chromatin can be compact and inaccessible as in heterochromatin or open and more accessible as in euchromatin.

The function of the nucleosome is not merely compaction. When DNA is wrapped in a nucleosome, it is inaccessible to other DNA-binding proteins such as transcription factors, DNA repair complexes, and DNA and RNA polymerases. It must first be disassembled, moved, or partially unwrapped in order for transcription, replication, or repair factors to operate. It thus serves as a gatekeeper for the DNA. It also serves as a template for epigenetic modifications which can regulate its ability to be disassembled, moved or unwrapped\textsuperscript{5-7}.

1.3 Post-Translational Modifications

After mRNA is translated into protein using the 20 standard amino acids, some residues are covalently modified by other enzymes with the addition of different functional groups. The most common modifications in the proteome are phosphorylations and acetylations\textsuperscript{8}, and an estimated 30\% of human proteins are phosphorylated\textsuperscript{9}. A diagram of phosphorylated tyrosine is shown in Figure 2. These PTM's can affect the performance and function of proteins and are used heavily by the cell in signaling pathways\textsuperscript{10,11}. To give one example unrelated to this work: the 5th amino acid on the RNA Polymerase II protein (PolII) is a Serine which is phosphorylated by TFIIH when PolII starts transcription. This modification causes an RNA capping enzyme to be recruited to PolII and cap the end of
the newly synthesized RNA\textsuperscript{11}. This gives a generic example that shows how PTM’s can dynamically modify functional properties of active proteins.

![Tyrosine Phosphorylation](image)

**Figure 2 Tyrosine Phosphorylation.** Diagram of the amino acid Tyrosine (left) and its phosphorylated form (right).

Mass spectrometry has been a key tool for discovering and mapping new PTM sites and thousands have been discovered in the proteome\textsuperscript{10,12,13}. In mass spectrometry, the molecule of interest is given a charge and accelerated through a voltage into a region of the perpendicular magnetic field. The field deflects them according to mass and charge giving a readout of the mass/charge ratio\textsuperscript{10}. A mass change of a known protein can point to the addition or subtraction of atoms on one of the residues and reveal the modification.

### 1.3.1 PTMs on Histones

The first evidence that histones were modified was found by Allfrey *et al.* in 1964\textsuperscript{14} who found that acetyl and methyl groups were added to histones after translation. Histones are heavily modified proteins with a total of 130 covalent modifications having been discovered by mass spectrometry\textsuperscript{15}. The most common modifications are acetylations, phosphorylations, and methylation,\textsuperscript{16} but there are over 30 different types\textsuperscript{17}. Different histone PTMs have been shown to be associated with DNA replication, DNA damage repair, and transcriptional activation and repression\textsuperscript{7,16}. They are also used to mark different chromatin states such as euchromatin (open) and heterochromatin (closed).\textsuperscript{18} Changes in
histone modification patterns throughout the genome have been linked to many forms of cancer\textsuperscript{19–23} and heart disease\textsuperscript{24,25} and enzymes that act on epigenetic histone modifications have proven to be useful drug targets\textsuperscript{24–26}. Some PTM’s are gene activating marks such as acetylation while others are repressive\textsuperscript{27} and still others can be either activating or repressive depending on the other cellular factors\textsuperscript{17,28}.

Many of the above studies determined PTM location via Chromatin Immuno-Precipitation followed by sequencing (ChIP-seq)\textsuperscript{29}. In this method, antibodies are raised to histone peptides with the desired modification. The cells being analyzed are lysed, and their DNA cut by sonication. Then the prepared antibodies are added, bind to the modified histones, and are pulled out of solution. The nucleosomal DNA is then purified, sequenced, and the sequences aligned to the cell’s genome giving a map of where in the genome the target modifications are located.

1.4 Partially Formed Nucleosomes

The structure of all histone proteins contains a conserved helix-loop-helix motif that allows H2A to dimerize with H2B and H3 to dimerize with H4\textsuperscript{30,31}. The H3/H4 dimers also bind to form an H3/H4 tetramer\textsuperscript{30}. Thus, the nucleosome can be thought of as being made of one H3/H4 tetramer and two H2A/H2B dimers. While standard nucleosomes contain two copies each of all four histones, sub-nucleosomal structures with missing H2A/H2B dimers have been detected\textsuperscript{32–36}. These are referred to as hexasomes (missing one dimer) or tetrasomes (missing both dimers). Hexasomes and tetrasomes are thought to be intermediates in the formation of full nucleosomes because the current model of nucleosome formation involves the deposition of tetramer on DNA followed by dimer\textsuperscript{37,38}. 
Although there is some evidence that free H3/H4 exist as dimers complexed with a histone chaperone and not tetramers. This deposition occurs primarily during the DNA synthesis phase of the cell cycle when histones are divided up between daughter DNA strands. Hexasomes are also formed when RNA polymerases transcribe through a nucleosome. RNA PolII has been shown to transcribe through the nucleosome, leaving the tetramer in place but expelling one of the dimers. Hexasomes have also been shown to form in vitro as a result of the action of the remodeler Remodels Structure of Chromatin (RSC) in the presence of the histone chaperone Nucleosome Assembly Protein 1 (NAP1). Rhee et al. confirmed the existence of hexasomes near transcription start sites using ChIP-exo to determine the correlation between dimer occupancy on one side of the nucleosome and the other. Formation of hexasomes during transcription may allow other polymerases easier access to transcribed DNA.

### 1.5 Transcription Factors

Generally speaking, transcription factors are proteins that facilitate or inhibit transcription. Typically, the term is used either in reference to “general transcription factors” or “specific transcription factors.” General transcription factors complex with RNA Polymerase II (PolII) and are required for all gene transcription in eukaryotes. Specific transcription factors (which are used in this study) bind to specific, short, gene regulatory sequences and recruit PolII and other factors to initiate transcription of the gene or genes with which they are associated. Specific transcription factors are typically composed of a sequence-specific DNA-binding domain and an activation domain which
recruits PolII. Experiments have shown that these domains can even be artificially swapped to give certain activators the ability to bind to other sites\textsuperscript{42}.

TF binding sites are often located inside of nucleosomes where they are inaccessible to TF binding due to being blocked by the octamer\textsuperscript{43}. However, nucleosomes and chromatin structures are dynamic\textsuperscript{44–48} providing transient exposure of buried DNA target sites to DNA binding proteins. Nucleosomes, for instance, have been shown to undergo spontaneous partial unwrapping,\textsuperscript{49–51} sliding,\textsuperscript{45} and gapping\textsuperscript{52}. Under the influence of ATP-driven chromatin remodelers, they can also be partially or completely disassembled\textsuperscript{53}. The particular dynamic investigated in this work is the partial unwrapping or “breathing” that can expose about 30 bases near either end of the nucleosome (termed the “entry-exit region”). The prevailing model for transcription factor binding in the entry-exit region involves three states: i) a fully wrapped state, ii) the unbound partially unwrapped state, and iii) the bound partially unwrapped state. These states are shown schematically in Figure 5B. It is estimated that approximately 30\% of TF sites fall within this region of the nucleosome\textsuperscript{43}.

\textbf{1.6 Our Biophysical Approach}

We study nucleosomes as physical systems with attention to the energies, forces, and kinetics involved. Our experiments all involve purified protein and DNA where reactions take place \textit{in vitro} so that we can control all variables involved and narrow in on the particular phenomena of interest. The theme in this approach is characteristic of the physical sciences and hence is termed “biophysics” to contrast with the more traditional biological “top-down” approach were whole organisms or populations are studied.
The primary measurement tools in these studies are small-molecule fluorescent probes attached to the DNA or protein of interest. The emission from these can be used as a reporter of protein binding in the case of Protein-Induced Fluorescence Enhancement (PIFE) and separation distance in the case of Fluorescence Resonance Energy Transfer (FRET). Fluorescent techniques can report on molecular changes both in solution and at the single molecule level. Advantages to small molecule fluorescence include smaller steric bulk than Green Fluorescent Protein (GFP) techniques, no need to apply forces to probe the sample as with force spectroscopy techniques, and the ability to measure reactions in equilibrium. These fluorescent methods will be described in chapter 2 and their application to single molecule transcription factor binding will be described in chapter 3.
2. Fluorescent Methods

This chapter discusses the fluorescence methods used in this work including Fluorescence Resonance Energy Transfer and Protein-Induced Fluorescence Enhancement. It then discusses how these techniques are used to measure transcription factor (TF) binding inside of nucleosomes.

2.1 Introduction to Fluorescence Resonance Energy Transfer

Fluorescence Resonance Energy Transfer (FRET) is a technique for optically detecting the distance between two fluorescent molecular labels. It has been thoroughly reviewed as a physical system and as a tool for biophysical investigation. One small-molecule fluorescent label, called the donor, is optically excited and then transfers its energy to the other label, called the acceptor. This acceptor then emits a photon of longer wavelength than the original excitation photon. The efficiency of this energy transfer is dependent on the distance between donor and acceptor labels. The system makes for a good molecular ruler allowing an experimenter to attach these donor and acceptor labels to molecular components of interest, illuminate in the donor's absorption range, and watch for emission from the donor and acceptor. Acceptor emission indicates energy transfer efficiency and close proximity of the labels while donor emission indicates lower energy transfer efficiency and larger distances between the labels. FRET has been a popular technique in biochemistry and biophysics because it can be used to report the formation of protein complexes, conformational changes within proteins or protein complexes, and DNA-protein interactions.
The ability to absorb photons is primarily found in molecules that have delocalized bonds such as Cyanine 3 (Cy3) and Cyanine 5 (Cy5), the fluorophores used in this work. In order to absorb a photon, spacing between electrical energy levels (shown as $S_0$ and $S_1$ in Figure 3) must be close to the photon energy. When a molecule absorbs a photon, vibrational energy levels above the first electronic excited state can also be excited allowing the molecule to absorb a range of wavelengths. A molecule in these higher vibrational states typically loses energy to the environment in a process called vibrational relaxation which causes it to transition to the lowest vibrational state in the same electrical state. This is shown schematically in Figure 3A. Vibrational relaxation occurs much faster than other timescales in this system.
Fluorophores can interact with the electromagnetic field through dipole coupling. For a classical analogy of this quantum effect, one can imagine an excited fluorophore as an oscillating dipole. This dipole can absorb energy in the presence of an oscillating electric field as is the case for absorption described above. Classical oscillating dipoles have an electric field that can be divided into a far-field which falls off as 1/r and a near field which falls off as 1/r^3. Energy leaving the dipole through the far field as radiation is analogous to the fluorophore releasing a photon. The energy can also leave if it is absorbed by another dipole residing in the near field. This dipole-dipole energy transfer is analogous to FRET.

The rate of FRET energy transfer depends on the relative orientation of the fluorophores because the molecular structure gives the dipole a preferred orientation. The rate also varies with distance as 1/R^6. More explicitly the rate of energy transfer\(^5\) is:

\[
k_T = \left( \frac{9 \ln 10 \Phi^D \kappa^2 f(v)}{128 \pi^5 N_A n^4 \tau_D R^6} \right)
\]

Where \(\Phi^D\) is the donor emission quantum yield \(\tau_D\) is the donor lifetime in the absence of the acceptor, \(f(v)\) is the spectral overlap integral, \(N_A\) is Avagadro’s number, \(n\) is the index of refraction, and \(\kappa\) is a constant that depends on the relative orientation of the donor and acceptor molecules (\(\kappa^2\) averages to 2/3 in the freely rotating case).

The overall FRET efficiency\(^5\) is defined as

\[
E = \frac{\text{number of quanta transferred from D to A}}{\text{number of quanta absorbed by D}}
\]

This efficiency depends on both the rate of energy transfer and the combined rate of all deexcitation pathways.
\[
E = \frac{k_T}{\tau_D^{-1} + k_T} = \frac{1}{1 + \left(\frac{R}{R_0}\right)^6}
\]

The efficiency also depends on the distance between the fluorophores with respect to \(R_0\) which is a combination of constants from equation 2. This gives FRET efficiency a strong dependence on fluorophore separation at separation distances near \(R_0\).

Acceptor/donor label pairs must be designed so that the emission spectrum of the donor overlaps with the absorption spectrum of the acceptor in order for the transfer to conserve energy. It also requires that the acceptor does not absorb at the excitation wavelength or else acceptor fluorescence will not be indicative of FRET. Lastly, in attaching fluorophores to our biomolecules we use 6-carbon linkers so that the fluorophores are free to rotate and any orientation dependence is averaged out.

**2.2 Protein Induced Fluorescence Enhancement**

The other fluorescent technique used in this work is Protein Induced Fluorescence Enhancement (PIFE). PIFE only requires one fluorophore which is attached to DNA. For fluorophores such as Cyanine 3 (Cy3), a protein in the vicinity (<4nm) of the fluorophore results in an increase in the fluorophore’s quantum efficiency (the fraction of absorbed photons that a fluorophore emits) up to 2-fold. This effect has been shown to be a robust, linear reporter of fluorophore-protein distance in several dyes and proteins\textsuperscript{61–65}. Thus PIFE can, like FRET, be used as an optical ruler or proximity indicator. As such, it has been used to determine the binding constants, substrate specificities, binding positions and kinetics of DNA binding proteins as well as observations of DNA translocases such as RIG I\textsuperscript{63,64}. 
PIFE has several advantages over FRET. Firstly, PIFE can report on shorter length scales than typical FRET probes. Secondly and perhaps more importantly, PIFE experiments don’t require the studied protein to be labeled. Labeling proteins can be difficult and usually requires mutating some amino acid residues. Labeling efficiencies can be low and hard to measure. Additionally, if high concentrations of labeled protein are needed in single molecule experiments, background from fluorophores in solution interferes with the FRET signal.

The study of this fluorescence increase started in 1994 when Aramedia et al. studied the fluorescence of several symmetric carbocyanines. They found that the fluorescence increased with increasing viscosity and decreased with increasing temperature. They hypothesized that after absorption, a trans-to-cis isomerization could sometimes occur which provided a non-radiative decay pathway. This de-excitation pathway competed with the radiative decay pathway and lowered the overall quantum yield. Later studies found that the presence of proteins near the dye prevented this isomerization and thus increased the quantum yield. This hypothesis was recently confirmed by Stennett et al. using time-resolved fluorescence and transient spectroscopy to confirm the presence of the cis isomer was depleted in the presence of protein.
Figure 4 *Cis and Trans Isoforms of the Cy3 Fluorophore*. Being held in the *trans* state causes increased quantum efficiency as a prominent non-radiative decay pathway is removed.

2.3 Measuring Transcription Factor Binding with FRET

There are a number of approaches used to investigate DNA accessibility within nucleosomes. One important approach is the use of DNA endonucleases. Numerous studies have used DNaseI, which cleaves DNA with significant sequence bias\(^\text{67–69}\). Therefore, DNaseI digestions result in a digestion pattern, which changes when the DNA is wrapped within nucleosomes. This provides insight into where nucleosomes restrict DNA accessibility. Alternatively, restriction endonucleases can be used to quantitatively measure accessibility to their specific DNA target sequence located within a nucleosome\(^\text{51,70,71}\). More recent approaches use FRET to quantify nucleosome and chromatin dynamics, including DNA unwrapping, nucleosome repositioning, nucleosome disassembly and DNA-protein binding within the nucleosome\(^\text{43,50,72–75}\). These ensemble measurements have provided important insight into how nucleosomes and chromatin regulate DNA accessibility.
Figure 5 Measuring Transcription Factor Binding with FRET. (A) Diagram of the DNA used to form nucleosomes. The Cy3 fluorophore is attached to the 5’ end of the 147 base 601 nucleosome positioning sequence. The TF binding site is shown in red. (B) Crystal structure of the nucleosome with the H2AK119 amino acid highlighted in red (Cy5 attachment point) and the last base of the DNA highlighted in green (Cy3 attachment point). (C) Three-state model of a TF binding inside a nucleosome. Wrapped nucleosomes spontaneously unwrap allowing the TF to bind to its site and hold the nucleosome in an unwrapped state until the TF dissociates.

In these studies, we have designed a nucleosome FRET system for measuring transcription factor (TF) binding in the entry-exit region of the nucleosome. The primary goal of the nucleosome labeling design is to position the labels on the sample such that there is high FRET efficiency when the nucleosome is fully wrapped and low FRET efficiency when the nucleosome is held in an unwrapped state due to the binding of a TF. This makes a fluorescent reporter of TF occupancy inside the nucleosome. In general, there are two labeling strategies for detecting structural changes within the nucleosome using FRET. One approach is to label the DNA around the nucleosome with both FRET donor
and acceptor\textsuperscript{52,76}, while an alternative is to attach one label to the DNA and the other label to a histone protein\textsuperscript{62,77,78}. We chose this latter approach because it can make use of labeled heterodimer, allowing us to verify the presence of heterodimer when the sample is being analyzed on the single molecule microscope. A diagram of this labeling strategy is shown in Figure 5. Because histones typically yield low labeling efficiencies (70-100\%), we chose to put the acceptor on the histone octamer because the (Ratio)\textsubscript{A} method described in section 2.4 can determine FRET efficiency irrespective of acceptor labeling efficiency.

In order to assure that the nucleosome forms at the correct location on the DNA, we use the 147 base pair 601 nucleosome positioning sequence which has been shown to accurately position nucleosomes in salt dialysis reconstitutions\textsuperscript{79}. This sequence has exceptionally high octamer affinity and has been used in many biophysical experiments that require exact positioning of nucleosomes. The other feature of the nucleosomes used in single molecule experiments is a 75 base pair extension on the unlabeled side of the DNA. This extension has a biotin-functionalized end which is used to attach the nucleosome to the surface of a microscope slide and hold the nucleosome away from the surface to reduce any surface-interaction effects.

Lastly, the nucleosomal DNA has a transcription factor binding site placed 8 base pairs from the end of the nucleosome. This location was chosen based on the crystal structure of the TF in order to place the bulk of the TF protein facing toward the octamer. In this way, TF binding would ensure the nucleosome is held in an unwrapped state. Other TF site positions have also been studied\textsuperscript{80}.
2.4 Ensemble Fluorescence Measurements

The most common and easiest way of measuring FRET or PIFE involves measuring an ensemble of molecules in solution. Here the sample is placed in a quartz cuvette and illuminated in the fluorophore’s absorption spectrum while an emission spectrum is taken. This has advantages over the single molecule fluorescence techniques discussed in section 4 in both being easier and avoiding complications from surface interactions. Using the DNA constructs shown in Figure 5, ensemble measurements can determine the average dissociation constant for a protein to its target site within a nucleosome and nucleosome-free DNA with about 10-20% accuracy. Ensemble measurements also work as controls for single molecule techniques.

The ensemble fluorescence measurements in this work were performed in a Fluoromax 4 (Horiba) fluorometer in 12 μl cuvettes. The experimental buffer conditions were the same as those used in the single molecule imaging buffer but without the oxygen scavenging system and triplet state quenchers (see section 3.3). For FRET experiments of protein binding to recognition sites within nucleosomes, FRET efficiencies were calculated through the \((\text{Ratio})_A\) method\(^5\) at varying binding concentrations (0-100 μM) and 2-10nM nucleosome concentrations. The method requires emission scans separately exciting the donor and acceptor. Nucleosome FRET efficiency is then given by

\[
E = \frac{2 \left( \frac{\varepsilon_A^D F_A^D}{F_A^A} - \varepsilon_A^D \right)}{\varepsilon_A^D d^+}
\]

Fluorophore emission \((F)\) superscripts refer to the donor \((D)\) and acceptor \((A)\) fluorophores and subscripts refer to the excitation frequencies for the donor \((D_{ex})\) and acceptor \((A_{ex})\).
Hence, $F_{Aex}^A$ is the fluorescence emission of the acceptor when directly excited, and $F_{Dex}^A$ is the fluorescence emission of the acceptor when the donor is excited. $\varepsilon_{Aex}^A, \varepsilon_{Dex}^A$ and $\varepsilon_{Dex}^D$ are the molar extinction coefficients of acceptor and donor at the acceptor and donor excitation wavelengths. $d^+$ is the donor labeling efficiency, which for DNA prepared with these methods is nearly 1. A prefactor of 2 reflects the presence of two acceptor labeling sites per nucleosome, only one of which is in range of the donor.

The (Ratio)$_A$ method will yield a FRET efficiency, but this alone will not quantify the accessibility of nucleosomal DNA. What we need to find is the dissociation constant ($K_d$) of the TF to the target site. $K_D = \frac{[DNA_{free}][TF_{free}]}{[TF-DNA]}$. This can be found by titrating the [TF] around the expected $K_d$ while keeping the concentration of DNA far below the expected $K_d$. The resulting titration can be fit to a non-competitive binding curve

$$E = E_F + \frac{E_0 - E_F}{1 + \frac{[TF]}{S_{1/2}}}$$

where $E$ is the FRET efficiency, $S_{1/2}$ is defined as the concentration at which half of the nucleosomes are bound (i.e. the FRET efficiency has decreased halfway to $E_F$), $E_0$ is the FRET efficiencies at zero [TF], and $E_F$ is the FRET efficiency with all the nucleosomes bound. If the [DNA] is significantly lower than the $K_d$ then the $S_{1/2} \approx K_d$.

Protein binding to its recognitions site on nucleosome-free DNA can also be measured in bulk by PIFE by quantifying the enhancement of Cy3 fluorescence emission for increasing protein concentrations around the $K_d$ and with Cy3 labeled DNA concentration at a concentration below the $K_d$. Small variations in DNA concentration can
add noise to a PIFE signal, so it is beneficial to include an additional fluorophore (e.g. Cy5) on the opposite DNA end that doesn’t undergo PIFE and whose fluorescence intensity can be used to control for varying sample concentration.
3. Single Molecule Total Internal Reflection Microscopy

This chapter discusses single molecule total internal reflection fluorescence (smTIRF) microscopy techniques and how they are applied to measuring transcription factor (TF) binding inside of nucleosomes. This chapter was published in part in Methods in Enzymology.81

3.1 Introduction to smTIRF

An advantage of using fluorescence to detect nucleosome structural changes is that fluorescence can be detected at the single molecule level. The technique described above in chapter 2 for observing TF binding inside of a nucleosome with FRET can be adapted for use at the single molecule scale as well. To extend the TF binding experiment to the single molecule level we used total internal reflection microscopy (TIRF or smTIRF). In smTIRF, fluorescently labeled samples are attached to a functionalized microscope slide inside a microfluidic channel (flow cell) filled with the appropriate buffer. A laser beam is incident on the functionalized slide surface at a supercritical angle causing the beam to reflect off the slide-buffer interface. This creates an evanescent wave that illuminates the first ~100 nanometers of the flow cell near the slide surface where the fluorescent sample is attached (Figure 6). Illuminating the sample this way reduces background because the illuminated volume is smaller than if the beam were to illuminate the whole channel. Hundreds of sample molecules can be imaged at once using this technique as long as they are spaced far enough apart for the individual molecules to be resolved. In our case, the attachments are made with a biotin-streptavidin linkage between
the biotin functionalized microscope slide and the biotin attached at the end of the 75 base pair DNA extending from the nucleosome (Figure 7).

This smTIRF microscopy technique has been used to investigate nucleosome dynamics including chromatin remodeling\textsuperscript{86,87}, nucleosome unwrapping\textsuperscript{88,89}, transcription factor binding within nucleosomes\textsuperscript{62}, nucleosome assembly\textsuperscript{90}, nucleosome gapping\textsuperscript{52}, and heterochromatin protein 1 interactions with nucleosomes\textsuperscript{91}. These single molecule studies provide important mechanistic insight into the regulation of DNA accessibility by determining the distribution of nucleosome structural states and detecting transient states that are difficult to detect by ensemble measurements. Furthermore, smTIRF measurements simultaneously determine each equilibrium transition rate between states with minimal sample. This is in contrast to ensemble rate measurement techniques like stopped-flow\textsuperscript{74}. This type of single molecule data can then be modeled to provide further insight into the structural dynamics that regulates DNA accessibility\textsuperscript{92}.

The smTIRF microscopy and flow cell preparation used in this work is largely based on previously published protocols\textsuperscript{93} with modifications for the study of TF binding to nucleosomes and nucleosome-free DNA. Detailed methods for smTIRF and ensemble data analysis of TF binding can be found in the literature\textsuperscript{94}. Section 3.2 outlines the components of the TIRF optical system, section 3.3 describes how the smTIRF data was collected, and section 3.4 describes how the data was analyzed.
3.2 Single Molecule TIRF Setup

![Figure 6 Schematic Diagram of the Prism-Based smTIRF Microscope](image)

The laser paths were filtered through bandpass filters F1, F2. Beam paths converge on dichroic D1 and were focused through lens L1 into the prism at a supercritical angle. This illuminates the samples attached to the quartz slide. Fluorescence from the samples is collected through the objective and split into emission channels by the DualView optical assembly and projected onto different halves of the EMCCD camera.

We used a quartz prism-based smTIRF microscope for our single molecule fluorescence measurements as described in Roy et al. which we outline below. The experimental equipment used is shown in Figure 6, but conceptually it can be divided into three groups: the excitation path, the microscope body, and the emission path. The excitation path spatially aligns lasers toward the microscope body in order to excite the fluorophores. An arbitrary number of lasers may be used to excite various fluorophores, but for a two-fluorophore FRET system (and for PIFE measurements), we used two 50 mW CW diode pumped lasers (CrystaLaser) at 532 and 638 nm. The power output is variable.
and was adjusted to balance signal to noise ratio (SNR) and fluorescence lifetime. Typical powers were about 50 mW for 532 nm and 13 mW for 638 nm. Laser output was cleaned using Semrock Brightline filters 531/22 (532 nm laser) and 640/14 (638 nm laser) (Figure 6; F1 and F2). Lasers were aligned with a Semrock Brightline 580 nm dichroic long pass filter (Figure 6; D1) and focused using a 250mm focal length plano-convex lens (ThorLabs) (Figure 6; L1) into a Pellin-Broca prism (Melles Griot) atop a 1”x3”x1mm quartz slide (G. Finkenbeiner) to produce a TIRF field. The prism is held in place by a custom-made mount.

The microscope body was a IX71 inverted microscope (Olympus). The emission signal was acquired through a 60x water immersion objective (UPlanSApo 60x/1.20w, Olympus). The long working distance allowed for imaging through the thick flow cell to the surface of the quartz slide on which an evanescent TIRF field excited immobilized, fluorophore-labeled samples.

Fluorescence emission was directed unfiltered through the microscope body’s camera port. Once light exited the microscope body, it moved to a DualView two-channel simultaneous-imaging system (Optical Insights) which housed a dichroic mirror (Chroma Technology, T635lpxr) (Figure 6; D2) that spectrally split the fluorescence light into donor and acceptor channels. The donor and acceptor channels were filtered through bandpass filters, Chroma Technology D585/30 and D680/35 respectively (Figure 6; F3 and F4), and steered towards two halves of a PhotonMax electron multiplying (EM) CCD camera (Princeton Instruments). The DualView was adjusted to ensure sub-pixel resolution alignment of the donor emission and acceptor emission images on each half of the CCD.
Images were collected with WinView (Roper Scientific) using 2x2 pixel binning and variable exposure times, depending on the specifics of the rates being studied.

3.3 Single Molecule TIRF Data Acquisition

This section details the flow cell setup and acquisition steps. For step-by-step protocols see Appendix C. Single Molecules experiments were performed at room temperature. Samples were kept separately on ice until just before use and imaging buffers were prepared just before use for each flow cell.

**Figure 7 Single Molecule FRET Nucleosome Design** (A) Diagram of DNA constructs used for single molecule experiments measuring transcription factor association and dissociation rates while binding to sites within nucleosomes (top) and to naked DNA (bottom). A biotin/streptavidin linkage at one end affixes the molecule to the surface of a quartz slide and a donor Cy3 molecule on the other end is excited through a TIR evanescent field. (B) Crystal structure of the nucleosome core particle (PDB ID: 1KX5). The TF binding site is highlighted at bases 8-27 in red (continued... )
(...Figure 7 continued) red with the positions of the Cy3/Cy5 FRET pair as green and red spheres, respectively. (C) Three-state model of TF binding within a nucleosome. A normally wrapped nucleosome in a high FRET state undergoes thermal fluctuations to transition from a closed to open state at which point the TF binding site is exposed and a TF can bind and lock the nucleosome in an unwrapped, low FRET state. This also shows the method of surface attachment used in the single molecule experiments. (D) Two-state model of a TF binding to free DNA. The protein binding changes the local environment of the attached Cy3 molecule and increases its quantum efficiency, detectable as an increase in fluorescence intensity from the molecule.

Prior to binding nucleosomes or DNA to the surface, flow cells were first hydrated with wash buffer and pacified with blocking buffer containing BSA (Appendix A). The wash buffer was designed to match the conditions of the imaging buffer including pH, ionic strength, viscosity, etc. The surfactant Tween-20 was also included in the wash buffer to help prevent nonspecific sticking. The wash buffer described in Appendix A was used for experiments that measure transcription factor binding inside of nucleosomes and to nucleosome-free DNA.

After the flow cell was further pacified with blocking buffer, it was prepared to specifically attach nucleosomes. This was done by first depositing streptavidin on the biotin-PEG molecules present on the flow cell surfaces. Streptavidin then provides an attachment point for biotinylated DNA. There was a much higher density of biotin-PEG attachment points on the surface than required for single molecule imaging. Therefore, sample density was controlled directly through the concentration of sample applied to the flow cell. This allowed for increasing or decreasing density as appropriate to the experiment in question rather than requiring preparation of slides with varying biotin-PEG density. A sample concentration of 10pM typically yielded a good molecule density (roughly 500 molecules per field of view), but this concentration was varied as needed.
A glucose oxidase/catalase-based oxygen scavenging system (Gloxy) was used in the imaging buffer to increase fluorescence lifetime before photobleaching. In addition to this system, triplet state quenchers were used to improve fluorophore stability and inhibit fluorophore blinking. These quenchers were Cyclooctatetraene and 3-Nitrobenzyl alcohol (COT/NBA), Trolox (both used for Cy3/Cy5 FRET experiments), and BME (only used for Cy3 PIFE experiments). We used BME instead of Trolox and COT/NBA in the PIFE experiments because BME is an extremely effective triplet state quencher for Cy3 and easier to prepare. However, BME promotes a long-lived Cy5 dark state and so COT/NBA and Trolox were used for FRET measurements.

After the flow cell was prepared it was mounted on the microscope, a TIRF field was established, and signal was confirmed before data acquisition began. Active flow cells lasted on the order of 30 minutes to 1 hour before the glucose was consumed by the glucose-oxidase.
3.4 Single Molecule Data Analysis

After the videos were collected they were saved in a tiff stack as it is a universal, compressible, and extensible format. The stored data was then analyzed by custom software which identified the molecules, and assembled fluorescence vs time traces. The molecules were identified by the presence of an acceptor fluorophore. During acquisition, the acceptor was directly excited for a number of frames. These frames were averaged
together to produce a low background image which was used to detect local maxima that were above a noise threshold. Identified molecules were then scrutinized for the presence of both fluorophores during the FRET phase of the experiment. To do this, traces were formed as the average of the molecules maxima and the 8 pixels that surround it. A commensurate trace was formed from the same location on the donor half of the image to form two trace channels per molecule. These traces measure the total donor and acceptor emission of each molecule in time.

When a transcription factor binds, the donor fluorescence signal increases while the acceptor decreases. Thus we looked for traces that have a negative cross-correlation coefficient on short timescales. The cross-correlation coefficient at longer timescales tends to be more indicative of a long-term increase or decrease in signal or background and so it was ignored. This was done by splitting the traces into 100 frame chunks (skipping the acceptor excitation frames) and calculating the cross-correlation coefficient between the donor and acceptor for each chunk. Traces were sorted based on the number of 100 frame chunks that showed a cross-correlation coefficient of below -0.25. This caused traces with the most TF binding events to be shown first and traces with no TF binding events to be pushed to the end. This sorting helped speed the manual trace selection.

Traces likely exhibiting fluctuations were semi-automatically selected for further processing, but it was necessary to cut off the acceptor excitation portion of the experiment, any photobleaching events and to truncate traces appropriately to reject spurious event information such as fluorescent debris momentarily increasing one signal without a commensurate change in the other. Summed together, these potential confounding signals
required human input to individually select for sections of traces before moving on to the computationally intensive determination of idealized FRET states (Figure 8).

vbFRET, a MATLAB package provided by Dr. Ruben Gonzalez, was used to generate idealized FRET (or PIFE) traces from selected, truncated traces. In preparation for import into vbFRET, the background of each fluorescent signal was subtracted and a pseudo FRET trace was calculated for each molecule using \( E = I_A / (I_A + I_D) \), or the intensity of the acceptor divided by the total intensity of the molecule at a given time. In order to determine binding and unbinding rates, only dwell times are necessary outputs from vbFRET, in which case the actual FRET value is arbitrary. Example FRET traces and the idealized fits output from vbFRET are shown in Figure 8B. PIFE signals were scaled from 0 to 1 to better align with vbFRET’s expected input range. Molecules which fit to a two state model (one for a fully wrapped nucleosome and one for an unwrapped nucleosome) are selected and dwell time histograms are produced as shown in Figure 9A-B.

In order to use all the information in our dataset and avoid using an arbitrary histogram bin size, cumulative sum distributions were produced and fit to find characteristic dwell times as shown in Figure 9C. The characteristic dwell times correspond to the binding and unbinding rates for each concentration. It is important to show that the high FRET state dwell times are inversely proportional to protein concentration to confirm the fluctuations are in fact a result of protein binding and that low FRET state dwell times do not depend on protein concentration (Figure 9D). Additionally, it is possible to calculate from these idealized traces an excellent analog to bulk fluorescence measurement of the fraction of
nucleosomes unbound. The fraction of time spent in the unbound state (for FRET this is the high FRET state and for PIFE this is the low-intensity state) averaged over all molecules directly measures the fraction of nucleosomes unbound at a given concentration of TF. For comparison to smTIRF data, the ensemble FRET efficiency and PIFE intensities were normalized from 0 to 1, which represents the average fraction unbound or unbound at a given protein concentration. The results are shown in Figure 9.
Figure 9 Lifetime Analysis of Idealized Trace Data. (A) Example binned high state dwell time histogram (cyan) with single exponential decay fit (black). (B) Same as (A) but with low state dwell times. (C) Cumulative sum lifetime in the high (red) and low (cyan) states along with fits to a cumulative sum single exponential distribution (black) to find characteristic dwell times. (D) Characteristic dwell times are plotted as a function of TF concentration. Low state dwell times (cyan) show no concentration dependence whereas high state dwell times (red) show an inverse concentration dependence. (E) Ensemble and single molecule results compared.
4. Sample Preparation Methods

The preparation of nucleosomes is a many-step biochemical process. It varies slightly in the experiments included here and any deviations will be noted where appropriate. The goal of this section is to describe the preparation methods used for fluorescently labeled nucleosomes as they relate to transcription factor binding experiments. The biochemical process followed two parallel paths: one for the preparation of histone octamer and one for the preparation of nucleosomal DNA. After the was prepared it was reconstituted onto nucleosomal DNA and the resulting nucleosomes were purified. Schematic diagrams of this process are shown in Figures 10-13. The preparation and labeling strategy of nucleosomes for fluorescent measurements is designed to help ensure that a majority of the nucleosome population is homogenous. This section was published in part in Methods in Enzymology81.

4.1 Labeled DNA Preparation

To detect nucleosomes trapped in partially unwrapped states by FRET, the DNA needs to be labeled so that the fluorophore is positioned within the Förster radius of the histone label in the fully wrapped state. The DNA label then needs to shift further away from the histone label in the partially wrapped state.

4.1.1 Polymerase Chain Reaction (PCR)
Labeled DNA constructs used for transcription factor (TF) binding experiments are shown in Figure 7. The label was attached to the oligonucleotides that were then incorporated into duplex DNA via polymerase chain reaction (PCR). PCR is an in vitro method of DNA replication that can be used to copy segments from double-stranded template DNA\textsuperscript{96}. In PCR, the template DNA is heated until the complementary strands dissociate and then cooled to allow short, single-stranded DNA primers to specifically bind to the template. DNA polymerases add complementary bases to the primers, extending the primers to the full length of the template. This process can be repeated many times by cycling the temperature. With one primer PCR will generate single stranded DNA. If a reverse primer is added that is complementary to another site on the other template strand, the complementary template strand is also replicated. A PCR reaction with forward and reverse primers creates more template DNA each cycle, thus creating a chain reaction that exponentially amplifies the template strand between the forward and reverse primer locations. The resulting DNA incorporates the forward and reverse primers at its ends and in our case included the desired label. A summarized diagram of this process is shown in Figure 10.

Here, a bacterial plasmid that contained the Widom 601 nucleosome positioning sequence\textsuperscript{97} was used as the template for PCR. The 601 sequence helped ensure that the DNA label was correctly positioned relative to the fluorophore on the histone octamer. The fluorophore was attached via an amine-functionalized six carbon linker on the 5’ end of the DNA primer. Because the labeled DNA primer is purified away from the unlabeled primer by reverse-phase HPLC, the final labeling efficiency is nearly 100%.
Figure 10 Nucleosomal DNA Preparation Schematic. Purchased custom primers (orange) (Sigma-Aldrich) with an amino-modified 5’ end were labeled with Cy3 NHS-ester (green stars). The unlabeled primers are purified away to assure all primers are labeled. Labeled forward primers (orange) and unlabeled reverse primers (green) are used in PCR to make the final DNA construct. Synthesized DNA strands and free nucleotides are shown in blue. The construct was then purified by phenol-chloroform extraction and anion exchange chromatography before being used to form nucleosomes.

4.1.2 Primer Labeling

Before labeling, primers were ethanol precipitated three times to remove any free amine which may be present after synthesis. Excess ethanol was removed from the DNA pellet after the third precipitation by vacuum before resuspension with 0.1M sodium tetraborate buffer pH8.5 (Sigma B9876) to a DNA concentration of 400 µM. The N-hydroxy-succinimide (NHS) functionalized fluorophore was then dissolved in anhydrous dimethylformamide (DMF) (Sigma D4551) to 20mM of the NHS fluorophore. 5-fold molar excess of NHS-functionalized fluorophore was added gradually in 5µl additions to
the primers with thorough mixing. The reaction was allowed to proceed overnight on a lab rotisserie at room temperature and then quenched by ethanol precipitation and stored in 0.5xTE.

After labeling, the primers were purified by reverse-phase HPLC using a C18 column (Vydac C18 218TP54). This column separates samples based on hydrophobicity allowing it to discriminate between unlabeled and labeled primer because the cyanine dyes are highly hydrophobic. The sample was eluted at 1 ml/min flow rate with a gradient from 90% Buffer A (0.1M Triethylamine (Sigma T0086), 1mM EDTA pH 7) and 10% Acetonitrile to 70% Buffer A and 30% Acetonitrile over 20 minutes. Peaks that showed both 260nm and 550nm absorbance (DNA and Cyanine 3) corresponded to labeled primer and were collected. The acetonitrile was removed by vacuum concentration and finally ethanol precipitation was used to change the buffer to 0.5xTE.

The labeled primers were then used for PCR as described above. The resulting PCR products were purified via anion exchange chromatography on a MonoQ (GE Lifescience) column following manufacturers’ protocols. This column separates samples by the amount of negative charge on each molecule, making the longer PCR product elute after the shorter primers and free DNA bases.

4.2 Histone Octamer Preparation

Labeled histone octamer preparation takes place in 5 steps: histone expression, histone purification, octamer refolding, octamer labeling, and octamer purification. The histone protein expression and purification has been detailed previously\textsuperscript{98,99}, but I will give a short outline which is summarized by Figure 11. Schematic representation of the
refolding steps can be seen in Figure 12. The histone octamer preparation began with the transformation of plasmid DNA containing the histone gene into the *E. coli* based BL21(DE3)pLysS cells. The cells were grown in liquid cultures to an optical density of \( \text{OD}_{600\text{nm},1\text{cm}} = 0.5 \). Then the cells were induced with the Isopropyl β-D-1-thiogalactopyranoside (IPTG) to start the protein expression. After the cells had expressed the histone protein, they were lysed (broken up) with freeze-thaw cycles and then the resulting cellular lysate was centrifuged to separate out the inclusion bodies which the cells use to contain the histones. Then the inclusion bodies were lysed and purified by size exclusion chromatography to remove the genomic DNA. Lastly, cation exchange chromatography was used to remove any remaining proteins. The purified histone protein was lyophilized into a dry pellet and stored at -80°C.
Histone Protein Expression and Purification. Plasmid DNA is shown as blue circles. Histone protein is shown in red. A plasmid containing the histone gene is transformed into bacterial cells which are grown and then induced with IPTG to begin histone production. After 4 hours of production the cells are lysed, centrifuged, and the precipitated histones purified by size exclusion chromatography and ion exchange chromatography before being stored in lyophilized pellets.

For general histone octamer preparation, the whole octamer was refolded and labeled. In the case of the hexasome experiments, H2A/H2B dimer was refolded, labeled and purified separately from H3/H4 tetramer. The dimer and tetramer were then combined during the nucleosome reconstitution step. It should be noted that since the labeling chemistry used labels cysteine residues, the natural cysteine at H3(C110) needed to be mutated to an alanine H3(C110A) in order to avoid extraneous labels on the octamer. Mutation of H3(C110) to an alanine minimally impacts nucleosome structure and dynamics\textsuperscript{100,101} but does slightly impact nucleosome unwrapping\textsuperscript{102}. Thus nucleosomes with the H3(C110A) mutation were never compared directly to nucleosomes without the mutation. The experiments in this work primarily use H2A(K119C) as the labeling site.
**Figure 12 Histone Octamer Refolding and Reconstitution.** Purified histones (red, yellow, dark blue and green) are unfolded and combined in equimolar ratios. The guanidine is then removed by double dialysis to refold the octamer. The mutated cysteine is then labeled with Cy5 maleimide. The octamer is then purified by size exclusion chromatography. The purified octamer is combined with nucleosomal DNA and slowly dialyzed into low-salt conditions forming nucleosomes. Excess DNA is finally purified away by sucrose gradient purification.

Refolding begins by resuspending lyophilized histone proteins. Aliquots of H3(C110A), H2A(K119C), H2B, and H4 were resuspended in unfolding buffer (Appendix A) to no more than 5 mg/ml to avoid aggregation. Samples were incubated on ice for 1 hr, thoroughly mixed, and centrifuged at 20,000g to pellet aggregates before determining the concentration of each histone on a UV-VIS spectrophotometer. Based on the concentration of each histone, equal moles of H3 and H4 were combined with 10% molar excess H2A and H2B to give a 1/1/1.1/1.1 molar ratio of H3/H4/H2A/H2B. The excess H2A and H2B helped ensure complete incorporation of H3-H4 tetramer into full octamer because the separation of tetramer from octamer is difficult in subsequent purification steps.
Unfolded histones were refolded through double dialysis. We have found that double dialysis results in a sufficiently slow rate of change of buffer concentration that maximizes efficient octamer refolding. A schematic diagram showing the method of making dialysis chambers is shown in Appendix B. These microdialysis chambers were placed within a 25cm x 50mm piece of 6000-8000 MWCO dialysis tubing filled with about 50mL of the unfolding buffer. The 50mL dialysis bag containing the microdialysis chambers was dialyzed against three changes of 2L refolding buffer for at least 6 hours per change with the last change going overnight. The refolded histones were reclaimed using a pipette to pierce the membrane and remove the sample within.

4.3 Histone Labeling and Purification

There are two general labeling strategies for detecting structural changes within the nucleosome via FRET. One approach is to label the DNA around a nucleosome with both FRET donor and acceptor\textsuperscript{52,76}, while the alternative is to attach one label to the DNA and the other to a histone protein (Figure 7). We find it beneficial to place the acceptor fluorophore on the H2A-H2B heterodimer (e.g. H2A(K119C)) because this allows for direct determination of which molecules contain two heterodimers and hence a full octamer. This is because heterodimers dissociate from nucleosomes before the H3-H4 tetramer\textsuperscript{38,103}. Moreover, the (Ratio)\textsubscript{A} method used for determining FRET efficiencies is not impacted by the labeling efficiency of the acceptor fluorophore\textsuperscript{57}, which can vary between 70-95%.

4.3.1 Histone Labeling
The refolded (but not yet purified) histone octamer was reduced by adding TCEP pH 7.1 (Appendix A) to 10mM final concentration and incubating on ice for 30 minutes. Then the TCEP was dialyzed away in three changes of 300ml PIPES buffer (Appendix A) for at least 2 hours per change with the last change going overnight. This promotes the free thiol state of the cysteine and removes any reducing agents which can lower final labeling efficiency. The octamer was reclaimed and its absorbance measured by UV-VIS spectrophotometer.

The reduced octamer was placed in a sealed tube and purged of oxygen using a slow and steady stream of argon flowing over the surface of the sample for 30 minutes at 4°C. An aliquot of 2M HEPES pH 7.1 (Sigma H3375) was purged as above, however, in this case, the argon was bubbled through the buffer for more efficient oxygen removal.

The maleimide functionalized fluorophore (GE Healthcare PA25031 for Cy5 labeled histones) was brought to room temperature and 0.25mg was dissolved in 14µl of anhydrous N,N-Dimethylformamide (Sigma 227056) to 18mg/ml final concentration. Just before labeling, the samples and HEPES buffer were removed from the argon stream and equilibrated to room temperature for 5-10 minutes.

2M HEPES pH 7.1 was added to the sample to 100mM final concentration to rapidly raise the pH. The fluorophore was then immediately added in 2µl additions mixing well between additions to limit micelle formation of the hydrophobic label until a 7.5-fold molar excess of label was reached. The reaction was allowed to proceed at room temperature for 1hr on a lab rotisserie, and then at 4°C overnight. Finally, the reaction was quenched with
10 mM final concentration dithiothreitol (DTT) and stored at 4°C on ice until purification by size exclusion chromatography.

4.2.2 Histone Octamer Purification

Purification of refolded octamer was necessary to remove aggregates and free histones as well as the excess label. Purification was performed by size-exclusion chromatography on a GE Superdex 200 10/300 GL column pre-equilibrated with 2 column volumes of refolding buffer (Appendix A). The sample was eluted at 0.5 ml/min for 1 column volume in refolding buffer and 0.5 ml fractions were collected. Full octamer eluted at about 15 ml followed by dimer at roughly 18 ml. Purified octamer was stored in 40% glycerol stocks at -20°C.

In the calculation of the final concentration of labeled octamer, the absorptivity of the dye at 276 nm was not taken into account because the local environment and rotational freedom of the fluorophore impacts its absorbance spectrum. Therefore, the precise contribution to the 276 nm absorbance peak was not known and so the true concentration could not be determined accurately. In a reconstitution, histone octamer beyond the given ratio results in poorly formed nucleosomes or aggregation of histones onto DNA and loss of sample. Therefore, we knowingly overestimated histone concentration by ignoring the dye contribution rather than underestimate it. We used an absorbance coefficient of $\varepsilon=48825(\text{cm}^{-1}\text{M}^{-1})$ for measuring octamer concentration and 13035(\text{cm}^{-1}\text{M}^{-1}) and 22371(\text{cm}^{-1}\text{M}^{-1}) for measuring dimer or tetramer respectively.

4.4 Nucleosome Reconstitution and Purification
Nucleosomes are reconstituted with DNA and histone octamer by salt dialysis. There are multiple methods for salt dialysis including the stepwise method\(^5\), the pump method\(^9\), and the double dialysis method\(^7\). We used the double dialysis method as described for histone octamer refolding above. This produces a gradual change in salt concentrations similar to the pump method, but with the ease of making just a few buffer changes.

**Figure 13 Nucleosome Reconstitution Sucrose Gradient Fractions.** 5% native polyacrylamide gel imaging Cy3 (top) and Cy5 (bottom) of the fractions taken from a nucleosome sucrose gradient purification (lower fractions show faster sucrose mobility). The DNA construct was 222 base pairs as shown in Figure 7. Fractions 9-11 were pooled and concentrated for future experiments. The lower band had a lower Cy5/Cy3 intensity ratio indicating either lower FRET from a mispositioned nucleosome or, more likely, the absence of one of the Cy5 labeled heterodimers. (see Figure 24)

To begin the reconstitution, we first combined purified histone octamer and DNA in a 1:1.25 octamer:DNA molar ratio. Our reconstitution typically contained 50-100pmol of octamer. Additionally, the reconstitution mix contained 2M NaCl, 0.5xTE, and 1mM benzamidine in a 50ul volume. This reconstitution mix was placed in a microdialysis
chamber using 6000-8000MWCO dialysis tubing as a dialysis membrane. The microdialysis chamber was placed in a 50mm dialysis bag with 75ml of high-salt buffer (Appendix A) and dialyzed against 2 changes of 4L low salt buffer (Appendix A) for at least 6 hours per change with the last change going overnight. The reconstitution mixture was reclaimed by piercing the microdialysis chamber with a pipette as in refolding.

At this stage, the sample was a combination of well-positioned nucleosomes, aggregates, de-positioned nucleosomes, free DNA, and possibly hexasomes. Purification helped ensure a well-defined homogenous sample for further single molecule work. Purification was performed by sucrose gradient. Gradients were made using a BioComp Gradient Master from 5% to 30% sucrose in 0.5xTE using Beckman-Coulter thin-wall polypropylene ultracentrifuge tubes with 13.2ml total volume. Gradients were then balanced and nucleosomes are loaded onto the top of the prepared gradients. The gradients were then centrifuged for 22 hours at 41k rpm, 4°C in a Beckman-Coulter Ti-41 swinging bucket rotor.

To fractionate the gradients, a 1mm glass capillary tube was attached to 1mm I.D tube and inserted into the bottom of the gradients. The tube was connected to a peristaltic pump and pulled at 1ml/min fractionating into 0.4ml fractions. In this way, the earliest fractions corresponded to molecules with the highest sedimentation velocities. Aggregates were embedded in the bottom of the centrifuge tube, nucleosomes sedimented to fractions ~7-14, and free DNA to fractions ~17-24. The fractions were analyzed on a 5% native polyacrylamide gel with 0.3xTBE (Appendix A) to determine the location of well-formed nucleosomes. Fractions containing correctly positioned nucleosomes were pooled (Figure
13). Using an Amicon 30k MWCO centrifugal filter the pooled samples were concentrated and washed three times with cold 0.5xTE which served as their final storage buffer.
5. Histone Core Phosphorylation Regulates DNA Accessibility

This chapter describes experiments that aim to determine the effect of phosphorylation modifications on the core of the histone octamer in the entry-exit region. Here we discuss how MNase digestions, small angle x-ray scattering (SAXS), and ensemble transcription factor (TF) binding experiments point to increased DNA unwrapping and accessibility in the entry-exit region caused by phosphorylations. This chapter was published in part in Journal of Biological Chemistry105.

5.1 Introduction

Nucleosomes spontaneously partially unwrap due to thermal fluctuations, providing DNA binding complexes transient access to sites within the nucleosome50,51. This occurs most frequently near the DNA entry-exit region of the nucleosome and can be influenced by histone Post-Translational Modifications (PTMs)7. Histone PTMs function by two general mechanisms. One is a signaling function whereby single or combinations of histone PTMs provide binding sites for recruiting specific regulatory complexes16,17,106. PTMs located within the accessible histone tail regions of the nucleosome appear to primarily function by this signaling mechanism. The second mechanism is the direct alteration of nucleosome stability and dynamics by single or combinations of histone
PTMs,\textsuperscript{7,107,108} which in turn regulates DNA accessibility to transcription, replication, and repair complexes. H3K56ac functions by this second mechanism. This modification is located at the DNA-histone interface, about 10 bases into the nucleosome. This modification is located within promoters\textsuperscript{109,110} and at sites of DNA repair\textsuperscript{111}; it is involved in nucleosome assembly during replication\textsuperscript{112} and enhances transcription\textsuperscript{88,113–115}. H3K56ac significantly enhances the probability of DNA to be partially unwrapped\textsuperscript{88} by increasing the unwrapping rate\textsuperscript{77}, which enhances DNA accessibility to proteins within the nucleosome\textsuperscript{5,102}.

A number of recently identified histone PTMs are also located within the entry-exit region of the nucleosome\textsuperscript{116}, including phosphorylation of H3Y41\textsuperscript{117} and H3T45\textsuperscript{118} (see Figure 14). H3Y41 is phosphorylated by JAK2 in human cells and inhibits binding of HP1\textsuperscript{α}\textsuperscript{117,119}, which is involved in heterochromatin formation. ChIP sequencing data indicate that H3Y41ph is present at transcriptional start sites and correlates closely with H3K4 trimethylation, a mark of active genes\textsuperscript{120}.

H3T45 is phosphorylated in human cells by PKC\textsuperscript{118} and DYRK1A\textsuperscript{121}. PKC phosphorylation is associated with apoptosis, whereas DYRK1A phosphorylation represses HP1 binding similarly to H3Y41ph. In budding yeast, H3T45 is phosphorylated by the S-phase kinase Cdc7-Ddf4\textsuperscript{122}. The level of this modification peaks during DNA replication, and loss of H3T45ph causes replicative defects. The addition of a phosphate group at H3Y41 or H3T45 introduces a negative charge and steric bulk near the DNA phosphate backbone (see Figure 14), potentially disrupting DNA-histone interactions. These observations, combined with previous results that H3K56ac and H3R42
trimethylation increase DNA unwrapping\textsuperscript{88,123}, suggest that H3Y41ph and H3T45ph may function to directly increase nucleosome unwrapping.

Here we report the influence of H3Y41 and H3T45 phosphorylation and phosphorylation mimics on nucleosome unwrapping and DNA accessibility. We find that the phosphorylation mimics H3Y41E and H3T45E, and the chemically correct PTM, H3Y41ph, significantly increase nucleosome unwrapping and increase DNA accessibility to transcription factor (TF) binding by about 3-fold. These studies show that phosphorylation in the nucleosome entry-exit region increases nucleosome unwrapping and DNA accessibility similarly to H3K56 acetylation.
Figure 14 Design and Analysis of Nucleosomes Used for Ensemble FRET Experiments. (A) Nucleosome crystal structure (PDB: 1AOI) showing the modified amino acids in blue: H3Y41, H3T45, and H3K56. For FRET measurements, Cy3 is attached at the 5’ end of the DNA molecule (green), and Cy5 is attached at H2AK119C (orange). The LexA binding site is located from the 8th to the 27th bp of the nucleosome (dark red). The region of the nucleosome containing the three PTMs is enlarged to indicate the residues’ orientation relative to the DNA. (B) Schematic diagrams of the DNA molecules used in this study. (C) Electrophoretic Mobility Shift Assay (EMSA) of the nucleosomes used in FRET measurements.

5.2 Experimental Methods

This section explains the particular details of the experiments probing the effects of histone core phosphorylations. The small angle x-ray scattering (SAXS) and micrococcal
nuclease (MNase) experiments were performed by Tao Wang in the Luger lab while the synthesis of the H3Y41ph and H3K56ac proteins was performed by John C. Shimko in the Ottesen lab. The fluorescence experiments were performed by Matthew Brehove in the Poirier lab. Background and further detail for the fluorescent methods can be found in chapter 2 and further detail on fluorescent nucleosome sample preparation can be found in chapter 4. The TF used in the fluorescence experiments is the bacterial TF LexA, which does not bind in the context of nucleosomes in vivo but nevertheless has been used as a model TF for nucleosome experiments50,62.

5.2.1 Preparation of DNA Constructs

The nucleosomal DNA used for SAXS (Figure 14B 601-147) and MNase digestion assays (Figure 14B 601-207) was prepared as described by Dyer et al.124. The nucleosomal DNA used for FRET measurements (Figure 14 DNA-LexA) was prepared by PCR as described in section 4.1. The template used for PCR was a plasmid containing the 601 nucleosome positioning sequence with the 8th through 27th base pairs replaced with the LexA recognition sequence (TACTGTATGAGCATACAGTA)50,102. The oligonucleotides used as primers (Sigma-Aldrich) in the PCR reaction were (Cy3-CTGGAGATACTGTATGAGCATAACAGTACATCGTGGTAC) and (ACAGGATGTATATACACGTGCCTGGAGACTA). The oligonucleotide containing the LexA site contained an amine attached to the 5’ end and was labeled with Cy3-NHS (GE Healthcare) and then purified by reverse phase HPLC. The labeled primers were then used for PCR resulting in a PCR product incorporating Cy3. The product was then phenol-extracted and purified by anion exchange chromatography (see section 4.1).
The 601 sequence is asymmetric in its propensity to unwrap\textsuperscript{125,126}. Here the LexA site has been introduced on the side of the nucleosome that has a reduced probability of unwrapping.

5.2.2 Preparation of H3Y41ph and H3K56ac

Fully synthetic H3Y41ph and K56ac were prepared by sequential native chemical ligation as described previously\textsuperscript{102,127} with the following changes. Peptides were synthesized with standard N-(9-fluorenly)methoxycarbonyl (Fmoc)-Nα protection strategies using (2-(6-chloro-1H-benzotriazole-1-yl)-1,1,3,3-tetramethylaminium hexafluorophosphate) (HCTU) activation on an AAPPTec Apex 396 automated peptide synthesizer. All Met residues were substituted with norleucine to eliminate oxidative side products. C-terminal peptide H3(91–135)-A91C, C110A was synthesized on Rink Amide MBHA LL resin (Novabiochem). Peptides H3(47–90) A47Thz with or without K56ac and H3(1–46) with or without Tyr(P)-41 were prepared on Fmoc-Dbz(Alloc)-derivatized resin (where Dbz is 3,4-di-aminobenzoic acid and Alloc is allyloxy carbonyl) as described in Ref.\textsuperscript{128} and cleaved and purified as the C-terminal N-acylurea derivatives. Synthetic histones were analyzed by reverse phase-HPLC and MALDI-TOF MS (Figure 15).
Figure 15 Analysis of Purified Fully Synthetic Proteins H3Y41ph and H3Y41ph/K56ac. (A). Analytical reverse phase-HPLC of H3Y41ph with a gradient of 32-63% acetonitrile/0.1%TFA (B). MALDI-TOF MS of H3Y41ph: ([MH⁺] m/z expected 15280, observed 15283; [MH²⁺] m/z expected 7640, observed 7640). (C) Analytical reverse phase-HPLC of H3Y41ph/K56ac with a gradient of 32-63% acetonitrile/0.1%TFA. (D). MALDI-TOF MS of H3Y41ph/K56ac: ([MH⁺] m/z expected 15322, observed 15322; [MH²⁺] m/z expected 7661, observed 7660).

5.2.3 Nucleosome Preparation

Recombinant histone octamers were refolded and purified as described in section 4.2. For experiments that used PTM mimics, the mutations for H3Y41E, H3T45E, and H3K56Q were introduced into the plasmid expressing H3C110A by site-directed mutagenesis (see section 5.1 for discussion on these mimics/modifications). Each histone was expressed and refolded into histone octamer as described in section 4.2-4.3. Refolded histone octamer was purified by gel filtration chromatography. Histone octamer that was to be fluorophore-labeled was refolded with H2AK119C and was labeled before gel filtration purification. Histone octamer was then reconstituted onto DNA by salt dialysis and purified by sucrose gradient as detailed in section 4.4.
5.2.4 LexA Preparation

LexA was expressed in *Escherichia coli* BL21(DE3)pLysS cells (Invitrogen) from pJWL288 plasmid and purified as described previously. Briefly, the cells were lysed with one freeze-thaw cycle using external lysozyme and centrifuged to remove cell debris. Polyethyleneimine was added to 0.35% to precipitate and remove the DNA. Then ammonium sulfate was added to 0.4g/ml to precipitate and isolate the LexA. The LexA was further purified with a cellulose phosphate column with an NaCl gradient of 200mM to 800mM at pH 7. The fractions which contained LexA were combined and further purified on a ceramic hydroxyapatite column (BioRad 158-2200) on a gradient of 50mM to 400mM potassium phosphate.

5.2.5 Micrococcal Nuclease Measurements of Nucleosome DNA Accessibility

Nucleosomes reconstituted with 207-bp DNA that contained a centrally located 147-bp 601-nucleosome positioning sequence were subject to MNase digestion studies. MNase reactions were performed by combining 30μl of nucleosome (20 ng/μl) or DNA (20 ng/μl) into 2.5μl of BSA (10mg/ml), 25μl of 10× MNase buffer (New England Biolabs), 2μl (200units/μl) of MNase, and double-distilled H2O to bring up the total reaction volume to 250 μl. 60 μl of reaction mixture was collected at different time points. The reaction was quenched by 5μl of 0.5M EDTA and stored on ice. 4.2 μl of 10% SDS and 1μl of proteinase K (20 mg/ml) were added to each reaction mixture and incubated at 55 °C for 30 min. DNA was isolated by phenol-chloroform extraction. The DNA quantity and length following an MNase digestion of nucleosomes with canonical H3, H3Y41E, or H3T45E were analyzed by 6% native PAGE (Figure 16). The relative mobility ($R_f$) of each
DNA ladder band and MNase digestion product was measured with ImageQuant TL. The $R_f$ and length of each DNA ladder band were correlated and used to calculate the length of each band observed in the MNase reactions.

![Image](image.png)

**Figure 16** Micrococcal Nuclease (MNase) Digestion of WT, H3Y41E, H3T45E Nucleosomes. (A) Nucleosomes reconstituted with 601-207 DNA and histone octamer containing either WT H3, H3Y41E or H3T45E were digested with MNase. The digestions were quenched at 0, 1 and 2 minutes. The length of protected DNA was analyzed on a 6% polyacrylamide gel, and visualized by SYBR Gold staining. (B) The length of protected DNA as a function of digestion time. The migration distance of each band (A) was measured with the 1D gel analysis function of ImageQuant TL software. Compared to WT nucleosome, mutant nucleosomes are less resistant to MNase.
5.2.6 Size Exclusion Chromatography with Multi-angle Light Scattering (SEC-MALS)

Superdex 200 HR 10/30 column (24 ml total volume, GE Healthcare) was run in-line with the MALS instrument. The flow rate was 0.3 ml/min. 100μl of nucleosomes at 0.3 mg/ml was injected in a buffer containing 20 mM Tris, pH 7.5, 1mM EDTA, pH 8.0, 1mM TCEP, and 0mM KCl. The same samples were used in SAXS (at a different KCl concentration). The molecular weight of each sample was calculated using the ASTRA software (Wyatt technologies).

Figure 17 SEC-MALS Nucleosome Quality Control. WT, H3Y41E, and H3T45E nucleosomes have molecular weights consistent with a full complement of histones. To determine the molecular weight (MW) and quality of nucleosomes prior to SAXS experiments, size-exclusion chromatography combined with multi-angle light scattering (SEC-MALS) was performed. The molecular weight of each sample was calculated by ASTRA software (Wyatt technologies). The calculated MW for each nucleosome is within error of their theoretical MW (200 kDa), and samples are mono-disperse. Because the light scattering signal from the peak eluted at 7-8 ml in each nucleosome sample is also present in the buffer control (data not shown), it is not due to free DNA or other contaminants.
5.2.7 Small Angle X-ray Scattering SAXS

Nucleosomes containing the 147-bp 601-nucleosome positioning DNA sequence and WT, H3Y41E, or H3T45E octamers were used in the SAXS measurements. SEC-MALS (see Figure 17) was used to check the purity and monodispersity of each nucleosome sample. All SAXS data collection was done at the SIBYLS beam line (12.3.1) at the Advanced Light Source (Berkeley, CA). Nucleosomes were measured either in the reference buffer (20 mM Tris-HCl, pH 7.5, 1 mM EDTA, 1 mM DTT) or in the reference buffer with 50mM KCl to investigate the influence of ionic strength on nucleosome structure. To optimize the data quality and minimize radiation damage, exposure series of 0.5, 1, 2, and 6 s were performed. Data was processed by PRIMUS\textsuperscript{130}. The dimension of nucleosomes was estimated by GNOM\textsuperscript{131}. Ten random molecular envelopes were constructed for each nucleosome by DAMMIN\textsuperscript{132}. They were superimposed by DAMSUP\textsuperscript{133}. The average molecular envelopes of these 10 random models were calculated by DAMAVER\textsuperscript{134}. The averaged model was filtered by DAMFILT. Convex shells of all models were built and visualized (Figure 18)\textsuperscript{135}

For the radius of gyration ($R_g$) determination, data from 1/16\textsuperscript{th} dilutions, exposed for 1s, were used. This provided strong data at low angle with minimal inter-particle repulsion due to the charged nature of the nucleosomes. A typical example for WT was collected at ~0.1 mg/ml. Data from replicate experiments were evaluated simultaneously for each individual mutant. Radius of gyration values were calculated by Guinier analysis,\textsuperscript{136} with a new algorithm applied to the triplicates of experimental scattering data. The algorithm optimizes a bias-variance tradeoff criterion and allows us to determine $R_g$
values at higher precision than previously possible, while at the same time providing statistically well-founded uncertainties.

5.2.8 FRET Measurements of LexA Binding within the Nucleosome

FRET efficiencies were determined from fluorescence spectra as described previously\textsuperscript{43} and detailed in chapter 2. Briefly, fluorescence spectra were measured with a Fluoromax-4 (Horiba) photon-counting steady-state fluorometer at room temperature (Figure 19). The Cy3 donor fluorophore was excited at 510 nm, and the fluorescence emission was measured from 550 nm to 750 nm. The Cy5 acceptor fluorophore was directly excited at 610 nm, and the fluorescence emission was taken from 650 nm to 750 nm. From these fluorescence spectra, the acceptor emissions due to donor and acceptor excitations were determined. Fluorescence emissions were measured by integrating the fluorescence spectrum from 656 to 674 nm after subtracting out the fluorescence spectra of the sample buffer without nucleosomes. The FRET efficiency was then calculated via the (Ratio)A method as described in section 2.4.

LexA binding was detected by measuring a reduction in the FRET efficiency caused by LexA trapping the nucleosome in a partially unwrapped state. To quantify changes in LexA binding, we performed LexA titrations from 0 to 10 μM with 5 nM Cy3-Cy5-labeled nucleosomes in 0.5×Tris-EDTA with 75mM NaCl. The emission spectra of each 20μl sample was taken in the fluorometer at room temperature. For each data point, a corresponding blank spectrum was taken with the same LexA concentration to correct for background fluorescence. The FRET efficiency as a function of LexA concentration was fit to a non-competitive binding curve. For each modification and mimic studied, we
determined the relative $S_{1/2} = S_{1/2\ unmod}/S_{1/2\ mod}$, which is inversely proportional to the relative change in the probability that the LexA target site is accessible for binding. Each titration was taken in triplicate, and the standard deviation of the three measurements was used as an estimate of the measurement uncertainty.

5.3 Results

5.3.1 H3Y41E and H3T45E Increase Nucleosome Unwrapping

To investigate whether phosphorylation of residues at the DNA entry/exit site affects DNA unwrapping and enhances accessibility within nucleosomes, we prepared histones with individual H3Y41E and H3T45E substitutions. The glutamic acid introduces a negative charge that mimics certain aspects of phosphorylation. Nucleosomes were reconstituted with histone octamers containing these mutations, and 207-bp DNA molecules where the central 147 bp contain the 601-nucleosome positioning sequence$^{97}$. This allows for reconstitution of homogeneously positioned nucleosomes (Figure 7). To determine whether H3Y41E and H3T45E increase DNA unwrapping, we first assessed their impact on the rate of MNase digestion of nucleosomal DNA. MNase cleaves linker DNA but pauses when encountering histone-bound DNA, resulting in the protection of a $\sim$150-bp DNA fragment. Transient unwrapping of DNA from the histone octamer surface allows MNase to proceed further into the nucleosomes, resulting in increased rates of cleavage and smaller products.

We analyzed the MNase digestion time course of naked DNA, unmodified nucleosomes, and nucleosomes containing either H3Y41E or H3T45E by polyacrylamide gel electrophoresis (PAGE, Figure 16). In unmodified nucleosomes, we observe a single
~140-bp fragment, corresponding to the nucleosome boundaries\textsuperscript{50,51} after 2 minutes of digestion. In contrast, MNase digestion of nucleosomes containing H3Y41E or H3T45E did not result in a well-defined 150-bp fragment. Instead, a distribution of shorter DNA lengths was produced following a 60s MNase digestion, whereas a 120s digestion converged to an ~120-bp fragment (Figure 16B). This reduction in the length of DNA protected from MNase digestion by the histone octamer is a strong indication of increased DNA unwrapping. By comparison, crystallographic analysis of nucleosomes containing the H3 histone variant CenpA shows no changes in structure apart from the disorder observed for the last 10 bp on either end\textsuperscript{137}. Single molecule measurements have shown that CenpA-containing nucleosomes have increased exposure of terminal DNA and protect ~120 bp of DNA from MNase digestions\textsuperscript{138}. In combination, this indicates that MNase has increased access to DNA within nucleosomes containing either H3Y41E or H3T45E, suggesting that these phosphorylation mimics increase DNA unwrapping from the histone octamer. To further investigate the influence of H3Y41E and H3T45E on overall nucleosome structure, we carried out SEC-MALS and SAXS measurements. Nucleosomes were reconstituted with 601-147 DNA (Figure 14B) and unmodified histones or histones containing either H3Y41E or H3T45E. SEC-MALS confirmed that nucleosomes reconstituted both with and without these modification mimics were monodisperse and had the expected molecular mass of 200kDa (Figure 17). This verifies that nucleosomes with these amino acid mutations form canonical nucleosomes and not altered DNA-histone complexes such as the altosome\textsuperscript{139,140}, or nucleosomes lacking histones.
SAXS measurements allow us to determine the molecular envelope and $R_g$, which is defined as the average distance from the center of mass for the ensemble of molecules (Figure 18A). At low ionic strength, unmodified nucleosomes have an $R_g$ of 43.1 ± 0.15 Å, consistent with published data\textsuperscript{135}, whereas nucleosomes with H3T45E and H3Y41E both have statistically significant increased $R_g$ values. This effect is more pronounced for H3T45E at 50mM salt. This is also apparent in the molecular envelopes calculated from SAXS data collected at low ionic strength, which are consistent with increased DNA unwrapping from the entry-exit region of the nucleosome (Figure 18B). These results provide additional evidence that phosphorylation of the H3 residues around the nucleosome entry-exit region increases DNA unwrapping to enhance the accessibility of DNA sites within the nucleosome to proteins.
Figure 18 SAXS Shows H3Y41E and H3T45E Mutations Result in Extended Nucleosomes. (A) Experimental radii of gyration ($R_g$ in Å) for nucleosomes shown in (A), at 0 mM KCl (black, left-hand side) and 50mM KCl (blue, right-hand side). The values shown are $R_g$ (Å) and two standard deviations, giving approximately a 95% confidence interval. (B) The molecular envelope of nucleosomes containing 601-147 and either WT H3, H3Y41E or H3T45E, calculated ab initio from SAXS data taken at 0 mM KCl. The shell was superimposed onto the crystal structure of the nucleosome (PDB: 1AOI) without histone tails.

5.3.2 H3Y41 Phosphorylation and H3Y41E and H3T45E Each Increase DNA Accessibility within the DNA Entry-Exit Region of the Nucleosome

We next considered whether the increase in nucleosome unwrapping detected by MNase and SAXS measurements might enhance the interaction of proteins with DNA within the entry-exit region of the nucleosome. To investigate this, we used a FRET-based assay developed by Li and Widom\textsuperscript{50}, modified to assess the impact of histone PTMs on
DNA unwrapping. In this assay, the recognition sequence of a model transcription factor, LexA, is introduced near the DNA entry-exit region of the nucleosome such that LexA binding is occluded in the wrapped nucleosome state but accessible in the partially unwrapped state. The nucleosome is reconstituted with the 601-LexA DNA molecule, in which the 8th through 27th base pairs of the 147-bp 601-nucleosome positioning sequence are substituted with the LexA recognition sequence (Figure 14B). The DNA molecule is labeled with the Cy3 donor fluorophore at the 5’ end near the LexA site, whereas the histone octamer is labeled at H2AK119C with the acceptor fluorophore, Cy5. In the fully wrapped nucleosome state, the Cy3-Cy5 pair exhibits energy transfer, whereas a partially unwrapped nucleosome that is trapped by LexA binding generates a significantly lower energy transfer efficiency. As LexA is titrated from 30 to 3000 nM with a fixed nucleosome concentration of 5 nM, LexA binds its recognition site and traps nucleosomes in a partially unwrapped state. A LexA titration with these Cy3-Cy5-labeled nucleosomes results in a decrease in FRET efficiency. This can be fit to a non-cooperative binding isotherm to determine the $S_{1/2}$, which is the concentration of LexA where 50% of the nucleosomes are bound (Figure 19). A change in the $S_{1/2}$ is a quantitative measure of a change in the accessibility of the LexA target site, such that a decrease in the $S_{1/2}$ implies an equal increase in the LexA site accessibility. We carried out LexA titrations with nucleosomes containing the phosphorylation mimic H3Y41E and found that it decreased the $S_{1/2}$ by 2.8±0.4-fold. Because glutamate is chemically distinct from phosphotyrosine, we prepared the modified protein H3Y41ph using sequential native chemical ligation. We found that H3Y41ph decreases the LexA $S_{1/2}$ by 3.1±0.4-fold relative to unmodified
nucleosomes. This suggests that within error H3Y41E accurately mimics the influence of H3Y41ph on the probability of LexA binding to its DNA target site (Figure 19, or Table 1). This change in probability is related to the change in the free energy difference between the exposed and unexposed state by $\Delta \Delta G_{PTM} = -k_B T \ln(S_{1/2 PTM}/S_{1/2 unmod})$ and implies that $\Delta \Delta G_{Y41ph} = 1.1 \pm 0.1 k_B T = 0.7 \pm 0.1 \text{ kcal/mol.}$

We next investigated the impact of the phosphorylation mimic H3T45E, which contacts the same DNA minor groove as H3Y41. We found that the glutamate substitution for H3T45 decreased the $S_{1/2}$ by a factor of $2.2 \pm 0.5$, which is similar to the reduction observed for H3Y41E, H3Y41ph, H3K56Q, and H3K56ac (Figure 19, Table 1). We conclude that phosphorylation at H3T45 is likely to increase site exposure for DNA-protein binding within the entry-exit region of the nucleosome similarly to phosphorylation at H3Y41. Furthermore, the similarity between histone modifications in the entry-exit region (Figure 19, Table 1) suggests that single histone PTMs in this region tend to increase DNA accessibility by about a factor of 3, irrespective of the precise location and nature of the modification.
Figure 19 FRET Measurements of LexA Binding Within Nucleosomes. (A) Kinetic model of a transcription factor (blue oval) binding to a partially unwrapped nucleosome and trapping the nucleosome in this partially unwrapped state. In this state the Cy3 (green star) and Cy5 (red stars) are separated causing low FRET. (B) Example emission spectra with Cy3 excitation for increasing [LexA]. Cy3 emission increases and Cy5 emission decreases as [LexA] increases, which is due to LexA binding and trapping the nucleosome in a partially unwrapped state with lower FRET. (C) Change in FRET efficiency for increasing concentrations of LexA with unmodified (black), H3Y41ph (green) and H3Y41ph/K56ac (blue) nucleosomes. FRET efficiency changes are normalized to change from 1 to 0 and are fit with a noncooperative binding curve. (D) Relative $S_{1/2}$ reduction ($S_{1/2 \text{ unmod}} / S_{1/2 \text{ PTM}}$) for each single PTM and PTM mimic. Error bars reflect the uncertainty of the $S_{1/2}$ mean over three measurements.

5.4 Discussion

We find that H3Y41ph, H3Y41E, and H3T45E increase DNA accessibility by increasing unwrapping. A similar effect has also been reported for H3R42me2 and H3K56ac. Together, these results indicate that histone PTMs within the DNA entry-exit region of the nucleosome generally increase DNA accessibility by about 3-fold. These PTMs are involved in transcription activation, and H3Y41ph and...
H3K56ac\textsuperscript{109,110} have both been identified by ChIP sequencing to occur within nucleosomes in the regions around the transcription start site. Therefore, this unwrapping mechanism is consistent with a biological function where these PTMs increase DNA accessibility to transcription regulatory complexes facilitating gene expression.

Even a modest 2-fold change in gene expression is biologically significant as is observed for dosage compensation and haplo-insufficiency diseases. H3Y41, H3T45, and H3K56 all coordinate the same 5-bp minor groove section of DNA (Figure 14A). The side chain of H3Y41 extends into the minor groove of the DNA double helix at the first (and last) contact between histones and DNA backbone. It does not make any direct interactions with the DNA but rather appears to prevent further compression of the minor groove, thereby determining the entry and exit angle of linker DNA. The addition of a phosphate to this amino acid generates a steric clash in addition to a repulsive charge, thereby shifting the site exposure equilibrium toward partially unwrapped states. Although phosphotyrosine is significantly larger than glutamate, H3Y41ph and H3Y41E induce a similar increase in DNA accessibility. This suggests that electrostatic effects, rather than steric effects, determine DNA site accessibility at this site.

H3K56 is oriented near the phosphate of the 9th base pair of the DNA and stabilizes the same DNA minor groove as H3Y41 (Figure 14). Acetylation of H3K56 eliminates this interaction, reducing the binding free energy and increasing the probability for the DNA to partially unwrap. The hydroxyl group of H3T45 forms a hydrogen bond with the DNA phosphate between the 3rd and 4th base of the 3’ DNA strand. Either phosphorylation of threonine or substitution with a glutamate mimic as in our H3T45E nucleosomes would
disrupt this hydrogen bond. We, therefore, cannot rule out the reduction of a favorable DNA-histone interaction in the context of the H3T45E nucleosomes, in addition to the likely electrostatic repulsion resulting from the introduction of negative charge through glutamate substitution or phosphorylation. However, not only are glutamate and phosphothreonine relatively similar in size, but there appears to be space for phosphothreonine to extend between the DNA and the C-terminal tail of H2A in the structure of the wrapped nucleosome, suggesting that glutamate may accurately mimic H3T45ph. The question of whether mimics can replicate the effect of a naturally occurring chemical modification is highly debated. Glutamine can be an adequate mimic of acetylated lysine, for example in the effect of histone tail acetylation on chromatin self-association. However, we previously found that the influence of acetylylsine and phosphothreonine on nucleosome stability is not fully replicated by glutamine and glutamate, respectively, in the nucleosome dyad region. The effect of mimics in the entry-exit region is less clear.

The kinases that phosphorylate H3Y41 and H3T45 are expected to be sterically occluded from these sites within fully wrapped nucleosomes. This raises a question of how these modifications are introduced. H3K56, which is also inaccessible within fully wrapped nucleosomes, is acetylated before H3 is assembled into nucleosomes. In budding yeast, H3T45 phosphorylation occurs within newly assembled chromatin, suggesting that H3T45 might be phosphorylated before nucleosome assembly. It is also possible that nucleosome unwrapping would provide kinases transient access to both H3Y41 and H3T45. A large number of additional histone PTMs have been identified that are located
within the DNA-histone interface\textsuperscript{116}. Our studies suggest that many of these histone PTMs may function to increase DNA accessibility via enhanced DNA unwrapping.
6. Histone Modification Combinations

This chapter details fluorescent experiments identical to those described in Chapter 5, but involving multiple modifications or modification mimics. These experiments show how the effects of different modifications on the same histone combine to increase accessibility in the entry-exit region. Parts of this chapter were published in *Journal of Biological Chemistry*\textsuperscript{105}.

6.1 Introduction

There are 128 amino acid locations on the 4 core histones which have been found to be modified\textsuperscript{144}. If we grant that it is possible for each of these modifications to occur independently of each other, then the number of possible nucleosome modification states is huge. In fact, it is possible that every nucleosome is unique. That is, every nucleosome in every one of the cells in all the humans who have ever lived could theoretically be in a unique modification state. While this is unlikely to be the case, histones do typically carry multiple modifications\textsuperscript{145–147}.

To build off the work of the previous chapter, we wished to study how histone PTM’s in the entry-exit region of the nucleosome work together to disrupt DNA-histone contacts and increase DNA accessibility. We used fluorescent transcription factor (TF) binding measurements and preparation methods identical to those described in chapters 2.3, 4, and 5, but used H3 proteins with combinations of the modifications or modification
mimics used in chapter 5. We considered the possibility that H3Y41ph and H3K56ac could occur within the same nucleosome because both are involved in transcriptional regulation and have been identified by ChIP-seq to occur within nucleosomes around transcription start sites\(^{109-112,117,119,120}\). Thus, we decided to investigate the combination of H3Y41ph and H3K56ac and, for completeness, studied all 3 pairs of modification mimics (H3(Y41E, K56Q), H3(T45E,K56Q), and H3(Y41E,T45E)) as well.

6.2 Results

6.2.1 The Combination of H3Y41ph and H3K56ac Multiplicatively Increases DNA Accessibility

We constructed H3Y41ph/K56ac histones by sequential native chemical ligation as discussed in section 5.2.2. We then prepared Cy3-Cy5-labeled nucleosomes containing the 601-LexA DNA sequence (see Figure 14) with histone octamer containing H3Y41ph/K56ac as discussed in chapter 4. We assessed via FRET the \(S_{1/2}\) of LexA binding to its site in partially unwrapped nucleosomes. We find that the \(S_{1/2}\) of LexA binding to nucleosomes containing H3Y41ph and H3K56ac was reduced by a factor of 17±5 relative to unmodified nucleosomes (Figure 20, Table 1). This implies that these two modifications in combination dramatically increase the probability that a DNA target site within the entry-exit region is exposed for DNA-protein binding relative to either modification alone.

We compared this \(S_{1/2}\) for LexA binding to nucleosomes containing both H3Y41ph and K56ac to the LexA \(S_{1/2}\) with either individual PTM. If the two PTMs independently influence LexA binding, each individual \(\Delta\Delta G\) should combine additively, i.e. \(\Delta\Delta G_{Y41ph} + \Delta\Delta G_{K56ac} = \Delta\Delta G_{combined}\).
$\Delta\Delta G_{K56ac} = \Delta\Delta G_{Y41ph/K56ac}$, and each individual relative $S_{1/2}$ should combine multiplicatively, i.e.,

$$\left( \frac{S_{Y41ph}}{S_{unmod}} \right) \left( \frac{S_{K56ac}}{S_{unmod}} \right) = \left( \frac{S_{Y41ph/K56ac}}{S_{unmod}} \right)$$

This is because the $S_{1/2}$ of LexA binding to a modified nucleosome relative to an unmodified nucleosome is related to the $\Delta\Delta G$ by a Boltzmann weight,

$$\frac{S_{1/2}^{mod}}{S_{1/2}^{unmod}} = e^{-\Delta\Delta G_{mod}/k_B T}$$

Therefore, the expected multiplicative change in the $S_{1/2}$ is $10 \pm 2$ (Figure 20, Table 1). The observation that the reduction in $S_{1/2}$ of Y41ph, K56ac is greater than this product implies that at minimum H3Y41ph and H3K56ac combine to multiplicatively increase DNA accessibility. This 17-fold decrease in the $S_{1/2}$ implies a $\Delta\Delta G_{Y41ph,K56ac}$ of $2.8 \pm 0.2 k_B T$ or $1.7 \pm 0.1$ kcal/mol, which is about 5–10% of the free energy for nucleosome formation. The sum of the individually measured free energy changes induced by H3Y41ph and H3K56ac is $\Delta\Delta G_{Y41ph,K56ac} = 2.3 \pm 0.4 k_B T = 1.4 \pm 0.2$ kcal/mol. This indicates that there is about $0.5 k_B T$ of additional $\Delta\Delta G$ that is introduced by combining these PTMs and results in the measured $S_{1/2}^{Y41ph,K56ac}$ being about a factor of 2 higher than predicted from the product of each individual $S_{1/2}$. 
Figure 20 Relative Reduction of the $S_{1/2}$ ($S_{1/2\, \text{unmod}} / S_{1/2\, \text{PTM}}$) for PTM Combinations. (A) H3Y41ph and H3K56ac, (B) H3Y41E and H3K56Q, (C) H3Y41E and H3T45E, (D) H3T45E and H3K56Q. Single mimics shown in green are from Figure 19D. Double modifications and modification mimics are shown in blue. Red bars show the product of the LexA $S_{1/2}$ shifts with the nucleosomes containing the single modifications ($S_{1/2\, \text{mod1}} * S_{1/2\, \text{mod2}}$). If the modifications change the unwrapping free energy additively, the $S_{1/2}$ shifts should combine multiplicatively.

6.2.2 Combining H3Y41ph and H3K56ac Amplifies the Differences between Their Mimics

Given the wide use of mutations to mimic PTMs, we investigated the combined influence of the histone PTM mimics H3Y41E and H3K56Q. We find that the combined amino acid substitution H3Y41E/K56Q reduces the $S_{1/2}$ of LexA binding to nucleosomes by 4 fold (Figure 20B, Table1). This reduction of the $S_{1/2}$ is over three times less than the reduction caused by H3Y41ph/K56ac. These results indicate that small differences
between individual histone PTMs and their mimics can be magnified when combined together and demonstrates a limitation to using mutations to mimic PTMs.

We wanted to look at additional pairs of PTM’s but were unable to obtain the H3T45 phosphorylation and thus performed TF binding experiments on H3(T45E, K56Q), and H3(Y41E, T45E) using glutamate as a phosphorylation mimic. The results from H3(T45E, K56Q) showed a 2.8 ± 0.6 fold decrease in $S_{1/2}$ which is larger than the decrease due to H3T45E alone (2.2 ± .5), but still within error. The $S_{1/2}$ shift from H3(T45E, K56Q) was also not significantly different from the shift predicted by the model outlined above.

### 6.2.3 H3Y41E and H3T45E Mimics Show Redundant Effect

Since the $S_{1/2}$ shifts from H3Y41E and H3T45E individually were higher than for H3K56Q, we anticipated the double PTM mimic H3(Y41E, T45E) would serve as a stronger test of the thermodynamic model presented above which predicted a 6.2 fold shift. However, H3(Y41E, T45E) showed roughly the same $S_{1/2}$ shift as the PTM mimics individually and a significantly lower shift than the thermodynamic model would suggest (see Figure 20C). In this case the $\Delta \Delta G$’s do not appear to combine additively.

If one assumes the $\Delta \Delta G$ from H3Y41ph or H3Y41E is caused by the addition of a negative charge near the DNA, it seems that a second negative charge should approximately double the $\Delta \Delta G$. However, this assumes the there are no structural changes caused by the first negative charge. If the DNA were to spend more time in a partially unwrapped state as a result of the first charge, the second could have a greatly reduced effect on $\Delta \Delta G$. Unfortunately, if this is the case it would seem to contradict our finding that the $\Delta \Delta G$’s of H3K56ac and H3Y41ph do combine additively.
All 3 residues considered here coordinate one DNA gyre facing the octamer (Figure 21). However, the Lysine 56 residue is likely in range of the phosphate groups on the next DNA gyre further in. The alpha carbon of Lysine 56 is 1.0 nm away from the closest phosphate group on the outer gyre and 1.2 nm away from the closest phosphate group on the inner gyre. It is possible that when either Tyrosine 41 or Threonine 45 is phosphorylated, the outer gyre spends most of its time detached from the octamer and the Lysine 56 rotates down to associate with the lower gyre. This would explain why the second phosphorylation does not add to the $\Delta \Delta G$, while Lysine 56 acetylation still contributes to the $\Delta \Delta G$ of unwrapping even when Tyrosine 41 is phosphorylated. This is a tentative hypothesis, however, and we can’t rule out the possibility that this is specific to the use of glutamate as a phosphorylation mimic.
<table>
<thead>
<tr>
<th>Sample</th>
<th>( \left( \frac{S_{1/2 \text{mod}}}{S_{1/2 \text{Unmod}}} \right)^{-1} )</th>
<th>( \Delta \Delta G ) (kcal/mol)</th>
<th>( \left( \frac{S_{1/2 \text{mod1}}}{S_{1/2 \text{Unmod}}} \right)^{-1} \left( \frac{S_{1/2 \text{mod2}}}{S_{1/2 \text{Unmod}}} \right)^{-1} )</th>
<th>( \Delta \Delta G_{\text{mod1}} + \Delta \Delta G_{\text{mod2}} ) (kcal/mol)</th>
</tr>
</thead>
<tbody>
<tr>
<td>H3Y41ph</td>
<td>3.1 ± 0.4</td>
<td>1.1 ± 0.1</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>H3Y41E</td>
<td>2.8 ± 0.4</td>
<td>1.0 ± 0.1</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>H3T45E</td>
<td>2.2 ± 0.5</td>
<td>0.8 ± 0.2</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>H3K56ac</td>
<td>3.3 ± 0.4</td>
<td>1.2 ± 0.1</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>H3K56Q</td>
<td>1.8 ± 0.3</td>
<td>0.6 ± 0.2</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>H3Y41ph/K56ac</td>
<td>17 ± 5</td>
<td>2.8 ± 0.3</td>
<td>10 ± 2</td>
<td>2.3 ± 0.2</td>
</tr>
<tr>
<td>H3Y41E/K56Q</td>
<td>4 ± 1</td>
<td>1.4 ± 0.3</td>
<td>5 ± 1</td>
<td>1.6 ± 0.3</td>
</tr>
<tr>
<td>H3Y41E/T45E</td>
<td>2.5 ± 0.8</td>
<td>0.54 ± 0.4</td>
<td>6.2 ± 1.5</td>
<td>1.8 ± 0.3</td>
</tr>
<tr>
<td>H3T45E/K56Q</td>
<td>2.8 ± 0.6</td>
<td>0.61 ± 0.2</td>
<td>4.0 ± 1</td>
<td>1.2 ± 0.4</td>
</tr>
</tbody>
</table>

**Table 1** Summary of Nucleosomal DNA Accessibility as Measured by LexA Binding. If the modifications’ changes in unwrapping free energy are additive, the \( S_{1/2} \)'s are expected to be multiplicative.
6.3 Discussion

Our observation that H3Y41ph and H3K56ac in combination increase DNA accessibility by over an order of magnitude, whereas individual modifications increase accessibility by 2-3 fold, suggests that different combinations of PTMs can be used to tune DNA accessibility. This could be used to precisely control the level of transcription in a cell.

Our observation that H3Y41E in combination with H3K56Q did not fully replicate the increased unwrapping induced by H3Y41ph in combination with H3K56ac suggests
that steric effects and/or differences in electrostatic interactions become more important when H3K56ac is combined with H3Y41ph. This means that relatively subtle differences between mimics and modifications can be magnified when the residues function synergistically, as for H3Y41ph/K56ac. These findings suggest an additional note of caution regarding interpretation of results acquired using acetylation or phosphorylation mimics in combination in the entry-exit region of the nucleosome.

While the effects of the modifications H3K56ac and H3Y41ph combined multiplicatively, the effects of the modification mimics H3T45E and H3Y41E appeared redundant with regards to unwrapping. We hypothesized that this redundancy resulted from a structural change which involved the unwrapping of the outermost DNA gyre, screening the effects of a second negative charge. We also hypothesized that in the case of a phosphorylated Tyrosine 41 or Threonine 45, the Lysine 56 would make contact with the inner DNA gyre and thus its acetylation would have the effect of further increasing DNA unwrapping. These results suggest that the interactions of histone PTM combinations in the entry-exit region could multiplicatively increase the probability for DNA unwrapping, but that this may not be the case for all modifications, especially ones located near the same DNA gyre. Future studies that investigate these additional histone PTMs and other histone PTM combinations in the DNA entry-exit region of the nucleosome are essential for understanding the function of the numerous PTMs that are located within this region of the nucleosome.
7. Hexasomes Show Dramatically Increased Accessibility for Dimer-Distal DNA

This chapter details the methods and results of ensemble and single molecule transcription factor binding experiments on hexasomes.

7.1 Introduction

Both nucleosomes and other chromatin structures control DNA’s accessibility to complexes that regulate DNA processes, including transcription, repair, and replication\textsuperscript{149–151}. Nucleosome are dynamic\textsuperscript{44–46,48}, providing transient exposure of buried DNA target sites to DNA binding proteins. In addition, numerous factors regulate nucleosome and chromatin structural dynamics including histone Post-Translational Modifications (PTMs)\textsuperscript{152}, proteins that specifically bind to these modifications\textsuperscript{153,154}, and chromatin remodeling complexes\textsuperscript{155}.

The histone octamer that forms the core of the nucleosome can be thought of as being composed of two H2A/H2B dimers and one H3/H4 tetramer. One of the most dramatic structural changes a nucleosome can undergo is the removal of one of the two H2A/H2B dimers. Subnucleosomal structures with one or two missing H2A/H2B dimers have been detected\textsuperscript{32–36}. These are referred to as hexasomes (missing one dimer) or tetrasomes (missing both dimers) (Figure 22). Hexasomes and tetrasomes are thought to be
intermediates in the formation of full nucleosomes because the current model of nucleosome formation involves the deposition of tetramer on DNA followed by dimer\textsuperscript{37,38}. While H3 and H4 are thought to be deposited as tetramer, there is some evidence that free H3/H4 exist as dimers complexed with a histone chaperone and not tetramers\textsuperscript{39}. This deposition occurs primarily during the DNA synthesis phase of the cell cycle when histones are divided up between daughter DNA strands. Hexasomes are also formed when RNA polymerases transcribe through a nucleosome\textsuperscript{34}. RNA PolIII has been shown to transcribe through the nucleosome, leaving the tetramer in place but expelling one of the dimers\textsuperscript{34}. Hexasomes have also been shown to form \textit{in vitro} as a result of the action of the remodeler Remodels Structure of Chromatin (RSC) in the presence of the histone chaperone Nucleosome Assembly Protein 1 (NAP1)\textsuperscript{40}. Rhee \textit{et al.} confirmed the existence of hexasomes near transcription start sites using ChIP-exo to determine the correlation between dimer occupancy on one side of the nucleosome and the other\textsuperscript{36}. Formation of hexasomes during transcription may allow other polymerases easier access to transcribed DNA\textsuperscript{41}.

The properties and biological functions of hexasomes have yet to be studied in depth. Studies have shown through mass spectrometry that hexasomes can be reconstituted by salt dialysis\textsuperscript{35}. The only structural study of reconstituted hexasomes was done by Arimura \textit{et al.}\textsuperscript{156} and used exonuclease digestions, micrococcal nuclease digestions, dynamic light scattering, and small angle x-ray scattering to conclude that hexasomes show increased unwrapping and accessibility to nucleases on one side of the hexasome.
Histone octamer is a symmetric complex and only an asymmetric DNA sequence would make the whole nucleosome asymmetric. The 601 nucleosome positioning sequence was discovered by Lowary & Widom in 1998\textsuperscript{97} to have an exceptionally high propensity to form well-positioned nucleosomes. It was found by iteratively selecting from $5 \times 10^{12}$ random sequences in a process called SELEX. The 601 sequence is an asymmetric (non-palindromic) 147 base sequence used in many biophysical studies\textsuperscript{157} to position reconstituted nucleosomes. An as-of-now unpublished study by Levendosky \textit{et al.}\textsuperscript{158} was the first to report that due to asymmetries in the 601 sequence, hexasomes reconstituted onto DNA with the 601 sequence form with the dimer always on one side of the nucleosome. Here we make use of this property to prepare homogenous hexasomes with fluorescent labels that allow us to measure transcription factor (TF) binding to hexasomal DNA. For clarity, we have chosen one-half of the 601 sequence to be the ‘left’ side and one to be the ‘right’ side as shown in Figure 25. In this case, it is the left side of the 601 sequence that has a stronger affinity for the dimer and the right side that has a weaker affinity. The left side has also been shown to unwrap at greater force than the right\textsuperscript{125,126}.

The experiments in this study use Fluorescence Resonance Energy Transfer (FRET) and Protein-Induced Fluorescence Enhancement (PIFE) to observe TF binding to hexasomal DNA on both the dimer-proximal and dimer-distal sides of the hexasome. We then compared the accessibility and dynamics of hexasomes to that of nucleosomes and naked DNA. The transcription factors used were Gal4 from \textit{Saccharomyces cerevisiae} and LexA from \textit{E. coli}. Both have been used as model TF’s for nucleosome accessibility experiments\textsuperscript{50,62}.
**Figure 22 Hexasome Structure.** Structure is based on the nucleosome crystal structure (PDB: 1AOI) with the missing H2A/H2B dimer shown as sticks and 40 bases of DNA manually unwrapped. Cy3 and Cy5 labeling locations are marked in green and red.

### 7.2 Experimental Methods

#### 7.2.1 Reconstitution of Hexasomes

These experiments required that we reconstitute hexasomes and nucleosomes. Chapter 4 details the refolding of octamer and the reconstitution and purification of nucleosomes. Hexasome reconstitution and purification follows the same steps with one key exception. Instead of refolding and labeling octamer, H3/H4 tetramer, and H2A/H2B dimer are refolded, labeled and purified separately. The purified tetramer and dimer are added directly to the reconstitution mixture along with the DNA at an experimentally determined ratio around 1:1 dimer:tetramer. These reconstituted hexasomes are then purified by sucrose gradient as described in section 4.4.
Figure 23 shows the result of reconstituting with different dimer:tetramer ratios. A dimer:tetramer titration such as shown in Figure 23 was necessary for each preparation of tetramer and dimer because determining exact concentrations of these by UV-spectroscopy was found to be prone to error.

![Dimer/Tetramer Ratio Graph](image)

**Figure 23 Hexasome Reconstitution Dimer Titration.** 5% native acrylamide gel showing the results of 10pmol test reconstitutions with varying dimer:tetramer ratios. The 10pmol test reconstitutions shown here resulted in lower conversion efficiencies than the larger 50pmol reconstitutions used for experiments. Here tetramer is found to have slower mobility than nucleosomes or hexasomes despite its smaller mass, likely due to a less compact structure.

There were two labeling strategies used for these experiments, both of which involved Cy3 end-labeled DNA. The first, described in chapter 4, involves an acceptor Cy5 label on H2AK119C. This is the more common labeling site because the detection of FRET would confirm the presence of the H2A/H2B dimer, which has been shown to dissociate from the nucleosome before the tetramer. The second strategy is to label the tetramer with Cy5 at H3V35C. The H3V35C labeling site places the Cy5 close to the end of the DNA as shown in Figure 22 and shows a FRET efficiency similar to H2AK119C. This was used so
that the FRET efficiency values from nucleosomes could be directly compared to hexasomes with the dimer on the left and right side of the 601 sequence.

Figure 24 Sucrose Gradient Purification Fractions. The upper gel shows sucrose gradient fractions for nucleosomes reconstituted with purified octamer and the lower shows fractions from hexasomes reconstituted with purified dimer and tetramer in a 0.4:1 ratio. Both reconstitutions used S-FRET DNA (Figure 25) and the Cy3 fluorophore was imaged. Hexasomes, being less dense run slower than nucleosomes and faster than tetrasomes in the gradient.

7.2.2 Exonuclease III Digestions

Exonuclease III (ExoIII) is a nuclease that digests one strand of double-stranded DNA one base at a time in the 3’ to 5’ direction. ExoIII has been used in experiments previously\textsuperscript{142} to determine the position of single nucleosomes on DNA longer than 147 base pairs by exposing the nucleosomes to ExoIII. The nucleosome samples are then denatured and the single stranded DNA products are analyzed by denaturing acrylamide gel. We used SW-Exo and WS-Exo DNA constructs for these experiments which are
shown in Figure 25. All ‘Exo’ DNA was made with a Cy3 end labeled forward primer, a Cy5 end labeled reverse primer, and 30 base pairs of flanking DNA on each side.

To perform the exonuclease experiments, between 0.003 and 0.3 units/µl of ExoIII (NEB M0206S) are added to 300fmol of sample in 1xNEB Buffer 1 (1mM Bis-Tris-Propane-HCl, 1mM MgCl₂, 0.1mM Dithiothreitol) for a 30µl reaction. The samples were incubated at 37°C for 5 minutes and the reaction was quenched with equal volume formamide. The resulting digests were heated to 95°C before being run on a 15% denaturing acrylamide gel for 2 hours. The urea in the denaturing gel (Appendix A) assured that the digested DNA ran as single strands. The resulting gel was imaged for Cy3 and Cy5 on a Typhoon fluorescent imager. The bands on the Cy3 image correspond to the protection on the right side and the bands in the Cy5 image correspond to the protection on the left side.
**Figure 25 DNA Constructs Used in Hexasome Studies.** These are the 601-based DNA constructs used in this study. The 147 base 601 sequence shown in light blue has a Gal4-2C site 8 bases from either the left or right side. Bases 27-42 of the original 601 sequence are highlighted green and marked ‘S’ corresponding to the ‘strong’ side of the 601 sequence that dimer preferentially forms on. The corresponding sequence on the right side is highlighted dark blue and labeled ‘W’ for ‘weak’. When these highlighted regions on the right side are swapped, they are replaced by their reverse complements for symmetry (see text). 601 sequence is shown defining the left and right side. See Appendix D for sequences.

### 7.2.3 Ensemble FRET Experiments

An overview of the ensemble FRET measurements is given in Section 2.3-2.4. FRET samples were made with S-FRET or W-FRET DNA constructs shown in Figure 25. S-FRET refers to the ‘strong’ side of the 601 sequence that preferentially binds to the dimer. S-FRET has a Gal4-2C TF binding site on the 8th base of the 601 sequence and a Cy3 fluorophore at the left end which gives us a fluorescent reporter of TF binding as described in Chapter 2. The Gal4-2C site (CCGGAGGGCTGCCCTCCGG) is a variant of the Gal4
site which has been used before to measure TF binding inside of nucleosomes\textsuperscript{62}. This variant is used because it gives a faster off-rate, which allows us to capture more events in our single molecule time traces. S-FRET also has a 75 base pair linker and a biotin label on the right side of the 601 sequence for surface attachment. W-FRET refers to the ‘weak’ side of the 601 sequence that has less ability to bind dimer. W-FRET has the Gal4-2C site and the Cy3 fluorophore on the right side of the 601 sequence with the 75 base linker and biotin on the left side. This allows fluorescent reporting of TF binding on the right side of 601. S-FRET and W-FRET were used for both ensemble and single molecule experiments for nucleosomes and hexasomes.

Ensemble FRET experiments were performed in 12\textmu l quartz cuvettes on a Fluoromax-4 (Horiba) fluorometer at 5nM sample concentration in wash buffer (Appendix A). FRET was measured with the (Ratio)\textsubscript{A} method discussed in section 2.4. Blank scans were taken for each Gal4 concentration to account for any Gal4 fluorescence. Error was estimated by taking all measurements in triplicate, calculating the \( S_{1/2} \) for each titration, and reporting the error in the mean.

7.2.4 Ensemble PIFE Experiments

The single molecule and ensemble PIFE experiments were performed with hexasomes reconstituted on W-PIFE DNA which had a Cy3 fluorophore at the end of a LexA TF binding site. For PIFE experiments we found that LexA produced a stronger PIFE signal than Gal4 and end-labeled DNA gave a stronger signal than internally labeled DNA. Thus we truncated the 601 sequence by 7 bases in order to keep the end-label Cy5 close to
the LexA site. The strong side of the 601 sequence had 75 bases of DNA and a biotin at the end for use in single molecule experiments.

Ensemble PIFE experiments were carried out in 2ml cuvettes in a Fluoromax-4 (Horiba) fluorometer. Emission scans were taken from 530-750nm while illuminating at 510nm and from 630-750nm while illuminating at 610nm to quantify the emission of Cy3 and Cy5. PIFE was quantified by integrating the emission from 560nm to 580nm for Cy3 and dividing this by the emission from Cy5 from 656nm to 674nm. In this way, we controlled for small variations in the concentration of sample.

Emission spectra were taken at 0nM to 10nM \([\text{LexA}]\) in wash buffer with 0.2nM hexasomes or naked DNA. The resulting PIFE signal was fit to a non-competitive binding curve to find the \(S_{1/2}\) of LexA binding. Error was estimated by taking all measurements in triplicate, calculating the \(S_{1/2}\) for each titration, and reporting the error in the mean.

7.3.5 Single Molecule FRET Experiments

The single molecule FRET experiments were done on both nucleosomes and hexasomes made with the S-FRET DNA construct shown in Figure 25 in T130 Trolox imaging buffer (Appendix A). The details are provided in chapter 3. Briefly, the hexasomes or nucleosomes were attached to the surface of a polyethylene glycol (PEG) coated slide with a biotin-streptavidin linkage. Once on the surface, a TIRF field was established with a 638nm laser to confirm the presence of Cy5 labeled dimer. Then the TIRF illumination was switched to 532nm to excite the Cy3 donor and observe FRET changes due to Gal4 binding. At least three flow channels were prepared and imaged for each Gal4 concentration measured in order to estimate the error of the measured binding rates. As
described in section 3.4, Cy3 and Cy5 intensity time traces were prepared from the local maxima found during 638nm illumination. Time traces were selected based on anti-correlation of Cy3/Cy5 intensities and then analyzed by vbFRET to determine idealized FRET state traces for each protein concentration studied. Custom software was then applied to the idealized traces to determine binding and unbinding rates.

7.3.6 Single Molecule PIFE Experiments

The details concerning single molecule PIFE experiments are discussed in chapter 3 and use the same optical setup and flow cell design as used for FRET. These PIFE experiments were performed on hexasomes made with the W-PIFE DNA construct as shown in Figure 25. They were performed in imaging buffer (Appendix A) with 2-betamercaptoethanol at 1% v/v instead of Trolox. A TIRF field was established with a 532nm laser to excite the Cy3 fluorophore. Images were acquired at 1s intervals and time traces were constructed from the local maxima points. At least three channels were imaged for every LexA concentration used. Individual traces were then manually selected for two-state fluctuating signals with one photobleaching event. After the photobleaching event was manually removed from the trace, the trace was normalized and analyzed by the vbFRET software which performed Hidden-Markov-Model analysis to determine the binding state at each point. Custom software used for FRET analysis was then applied to determine binding and unbinding rates.

7.3 Results

7.3.1 Exonuclease Mapping Shows Asymmetric Protection in Hexasomes
We reconstituted hexasomes and nucleosomes on SW-Exo DNA constructs and performed exonuclease digestions on the purified samples. Both hexasomes and nucleosomes showed ExoIII pausing 30 bases from the left side of the SW-Exo DNA corresponding to the beginning of the 601 sequence. The same was true for nucleosomes on the right side of the SW-Exo DNA. However, for hexasomes, the right side showed pausing at 70 base pairs from the end which corresponds to about 40 bases inside the 601 sequence. This shows that while ExoIII pauses at both edges of nucleosomes, it is allowed to digest 40 bases further into hexasomes on one side (Figure 26). This is consistent with the findings of Arimura et al.\textsuperscript{156} which found that hexasomes show an ExoIII stall point about 35 bases into one side of the hexasomes. Our digestions also show that hexasomes formed on 601 DNA show a strong preference for the dimer forming on the left side of the 601 sequence. This can be seen by noting that the right side of the hexasomes showed no stalling where ExoIII stalled for nucleosomes. This indicates that there are few if any dimers residing on the right side of the 601 sequence.
Figure 26 ExoIII Mapping in Nucleosomes and Hexasomes (A) Diagram showing the double stranded Exo-DNA construct with the 601 sequence in blue and the 30 flanking bases in gray and Cy3 and Cy5 fluorophores are shown as green and red stars. ExoIII is shown in yellow digesting the 3’ end of the DNA. (B) Cy3 image showing digestions of hexasomes, nucleosomes, and naked SW-Exo DNA. Digestions were done with 0.003, 0.01, 0.03, 0.1 0.3 Units/μl. Lanes ‘A’ and ‘T’ show sanger sequencing ladders of Exo-III DNA for the ‘A’ and ‘T’ bases respectively. Horizontal lines denote the beginning of the 601 sequence and every 10 bases afterward. (C) The Cy5 image of the gel in (B).

7.3.2 Dimer Sequence Preference

We then wanted to determine if there were specific regions of the 601 sequence which caused the asymmetry in dimer preference. We hypothesized that bases 27-32 (106-121 on the right side), which directly contacted the dimer could be responsible for the preference. We, therefore, used site-directed mutagenesis (SDM) (Agilent 200514) on the SW-Exo plasmid to swap the 27-32 and 106-121 base regions, replacing each side with the reverse complement of the other to make WS-16-Exo. The resulting plasmid was used to make the WS-16-Exo DNA construct with ‘WS’ indicating weak affinity for dimer on the left side
and strong affinity for dimer on the right side. After reconstituting hexasomes onto these DNA constructs we performed ExoIII digestions and the results are shown in Figure 27.

The WS-16-Exo hexasomes showed further ExoIII digestion on the left side and less digestion on the right side indicating that these 16 bases can affect which side of the 601 sequence binds to the heterodimer. Swapping these regions did not completely reverse the dimer side preference, as the digestion on both sides showed multiple pause points even at the highest ExoIII concentration. This indicates that some hexasomes still have the dimer on the left side even after these DNA sequences were swapped. It is possible that bases further into the 601 sequence can affect the propensity to bind dimer even if they do not contact the dimer directly. In order for nucleosomal DNA to contact the dimer, DNA further into the nucleosome that does not contact the dimer needs to bend. If the DNA sequence adjacent to the dimer-contacting region is less flexible, this may prevent dimer from binding. Further studies are needed to elucidate the exact nature of this dimer sequence preference.
Figure 27 Swapping the 601 Dimer Side Preference  The two gel images on the right are taken from Figure 26. The two gel images on the left show Cy3 and Cy5 images of ExoIII digestions of hexasomes reconstituted on WS-16-Exo DNA. The digestions were carried out at 0.003, 0.01, 0.03, 0.1, and 0.3 Units/μl. Lanes ‘A’ and ‘T’ show sanger sequencing ladders of Exo-III DNA for the ‘A’ and ‘T’ bases respectively. Horizontal lines denote the beginning of the 601 sequence and every 10 bases afterward.

7.3.3 Hexasomes Show Accessibility Similar to Nucleosomes on the Dimer-Proximal Side.

In order to determine whether the loss of a dimer changed DNA accessibility on the dimer-proximal side of the nucleosome, we performed ensemble and single molecule FRET experiments on nucleosomes and hexasomes reconstituted on S-FRET DNA constructs. These experiments used dimer with a Cy5 label at H2AK119C in order to confirm the presence of dimer in the individual molecules studied. In these ensemble FRET experiments we titrated the concentration of Gal4 and observed a decrease in FRET that indicated Gal4 binding to its site inside the nucleosome. This gave us a quantitative measure of the accessibility of nucleosomal DNA to TF binding. The results show a relatively small shift in the $S_{1/2}$ from 19.6 ± 0.8 nM in nucleosomes to 46.9 ± 4 nM in
hexasomes as shown in Figure 29A-B. These $S_{1/2}$'s are equivalent to the dissociation constant $K_D$ of Gal4 to nucleosomal or hexasomal DNA because they are well above the 5nM hexasome or nucleosome concentration. Considering the dramatic structural differences between nucleosomes and hexasomes (the removal of ¼ of the component histones), this shift seems modest and is comparable in size to the accessibility shift from a single PTM\textsuperscript{43,105}.

![Figure 28](image)

**Figure 28 Single Molecule FRET Data Analysis** (A) FRET traces at varying Gal4 concentrations for hexasomes reconstituted on S-FRET DNA with Cy5 on H2AK119C. (B-C) dwell times histogram for low FRET events (B) and high FRET events (C). (D-E) Normalized cumulative sum distribution for low (B) and high (C) FRET events used for fitting.

Chromatin is a dynamic material that undergoes changes in response to stimuli and DNA regulatory events. Nucleosomes themselves are dynamic structures that regulate TF occupancy by altering the binding and unbinding rate of TF to DNA\textsuperscript{62}. Since our ensemble FRET measurements only measured overall TF occupancy, we use single molecule TIRF FRET in order to investigate TF binding dynamics on the dimer-proximal side of
hexasomes. This yielded the rate of binding and unbinding of Gal4 to nucleosomal DNA. As the Gal4 concentration was varied the binding rates \( k_{ON} \) increased linearly, while the unbinding rates remained roughly flat. These results, shown in Figure 29, were used to determine the binding rate 5.5 ± 1.2(s\(^{-1}\)nM\(^{-1}\)) and unbinding rate 0.32 ± 0.03(s\(^{-1}\)) of Gal4 to the dimer-proximal side of hexasomes. These were comparable to the rates of binding and unbinding for nucleosomes which were 11.4 ± 3 (s\(^{-1}\)nM\(^{-1}\)) and 0.37 ± 0.07 (s\(^{-1}\)) respectively. The binding rate of hexasomes was roughly half that of nucleosomes and this difference corroborates the 2-fold decrease in \( K_D \) found in the ensemble experiments.

**Figure 29 Ensemble and Single Molecule TF Binding on the Dimer-Proximal Side.** (A) Results from ensemble FRET experiments on hexasomes and nucleosomes reconstituted on S-FRET DNA. The data is fit to a non-competitive binding curve. (B) \( S_{1/2} \) calculated from (A). (C) Fraction of high FRET points in smFRET traces. The points are fit to non-competitive binding curves which happen to overlap in (C). (D) \( S_{1/2} \) calculated from (C). (E) On-rates for each Gal4 concentration and (F) the on rate per pM of Gal4. (G) off rates for each Gal4 concentration and (F) the average off rates. All error bars represent standard error in the mean over three measurements.
7.3.4 Hexasomes are Unwrapped on the Dimer-Distal Side

In order to compare dimer-proximal and dimer-distal hexasomes side by side, we reconstituted nucleosomes and hexasomes on the S-FRET and W-FRET DNA constructs using octamer and tetramer with Cy5 labeled at H3V35C. Labeling the tetramer would allow for direct comparison between FRET values regardless of whether we examined the dimer-proximal or dimer-distal side of the hexasome. The results are shown in Figure 30. Both nucleosomes and hexasomes showed similar FRET values on the S-FRET construct indicating that the dimer-proximal side of hexasomes is fully wrapped on average. With the W-FRET construct, nucleosomes showed high FRET, but hexasomes showed no FRET indicating that the dimer-distal side of the hexasome is unwrapped for the vast majority of the time.

Even if the dimer-distal side of the hexasome shows no FRET in ensemble measurements, the DNA could spend a fraction of its time in a transiently wrapped state. In order to detect these states, we used hexasomes reconstituted on W-FRET DNA with Cy5 labels on the H3V35C. We attached the hexasomes to the surface of quartz flow cells as described in section 3.3. After confirming the presence of tetramer with 638nm illumination, the illumination was switched to 532nm to watch for high FRET events. After analyzing the traces of 261 molecules with both Cy3 and Cy5 we found only 7 showed any Cy5 emission above background and none of those showed any anticorrelated Cy3/Cy5 events that would indicate FRET. This indicates that there are no detectable transiently wrapped states on the dimer-distal end of a hexasome. If there are any transient wrapped states, they last much shorter than 75ms frame rate of our camera.
Figure 30 FRET Efficiencies on the Weak and Strong Side of Nucleosomes and Hexasomes. (A) Schematic diagrams showing the dimer and tetramer in the hexasome (green) with the associated DNA (blue). Cy3 and Cy5 fluorophores shown as green and red stars. The Gal4 binding site is shown in red. Labeling the tetramer allows both sides of the nucleosome to have acceptors regardless of the location of the dimer. (B) FRET efficiencies for nucleosomes and hexasomes. Error bars represent standard error in the mean for triplicate measurements.

7.3.5 Hexasomes Show Accessibility Similar to Naked DNA on the Dimer-Distal Side

It is possible that, despite being unwrapped, the hexasome still reduces accessibility to the unwrapped DNA. Note that the primary mechanism by which nucleosomes reduce the occupancy of TFs in the entry-exit region is increasing the rate of TF dissociation and the details of this mechanism have not been established\(^\text{62}\). Since the hexasomes showed no FRET on the dimer-distal side we used PIFE to measure TF binding. We found that our PIFE signal was strongest with the LexA TF and end-labeled DNA. We thus shortened our W-FRET DNA construct by 7 bases and replaced the Gal4-2C binding site with a LexA binding site. The resulting W-PIFE construct is shown in Figure 25.

We reconstituted hexasomes on W-PIFE DNA and performed ensemble TF binding experiments with varying LexA concentrations to determine the accessibility of unwrapped
hexasomal DNA. The results are shown in Figure 31. We compared the $S_{1/2}$ of binding between hexasomes and naked DNA and found them to be within error. We then attached the same hexasomes and DNA to the quartz flow channel and performed single molecule PIFE experiments. The results showed a LexA binding rate of $0.023 \pm 0.003 \text{ (s}^{-1}\text{nM}^{-1})$ and an unbinding rate of $0.026 \pm 0.001 \text{ (s}^{-1})$ which were both similar to naked DNA. Both of these results show the accessibility and dynamics of dimer-distal hexasomal DNA is similar to naked DNA. It is important not to compare these rates and dissociation constants with those we found on the dimer-proximal side of the hexasome because the Gal4 and LexA transcription factors have vastly different affinities for their target sites. Direct comparisons between the accessibility of nucleosomal DNA and naked DNA have been made previously and have found to differ by a factor of 10,000. Thus we should interpret these findings as a factor of $\sim 10,000$ difference in accessibility between the dimer-distal and dimer-proximal side of a hexasome.
Fig. 31 Ensemble and Single Molecule TF Binding on the Dimer-Distal Side. (A) Results from ensemble PIFE experiments on hexasomes and naked DNA reconstituted on W-PIFE DNA. The data is fit to a non-competitive binding curve. (B) $S_{1/2}$ calculated from (A). (C) Example Cy3 emission traces at 0.3, 1, and 3 nM LexA with Cy3 emission histograms. (D) On-rates for each LexA concentration and (E) the on rate per pM of LexA. (F) Off rates for each LexA concentration and (G) the average off rates. All error bars represent standard error in the mean over three measurements.

7.4 Discussion

Hexasomes have been a largely unstudied component of chromatin and their biological function is not yet known. Here we have studied the dynamics of TF binding to hexasomal DNA and took a brief look into the origin of sequence specificity of hexasome formation. This study took advantage of a newly discovered property of the 601 nucleosome positioning sequence to make homogenous hexasome samples with the dimer forming only on one side of the 601 sequence. In addition to confirming this property by
ExoIII digests, we located a 16 base pair region on each side of the sequence (bases 27-32 and 106-121) which is largely responsible for its preference to form dimer on its left side. In comparing the left and right 16 base pair regions we considered the patterns in nucleosome positioning sequences that give them a high affinity for histone octamer. Lowary and Widom\textsuperscript{97} found that a key pattern in many nucleosome positioning sequences is pairs of T/A bases spaced periodically every 10 base pairs with pairs of C/G bases 5 bases offset. Sequences with this pattern showed increased bendability in native acrylamide gel electrophoresis assays. This 10 base pair periodicity corresponds to the 10 base pair helical periodicity of DNA and is likely to give the DNA a curvature with no stress. The 16 base pair region on the left which has a higher affinity for dimer is TCAATTGGTCGTA\textsubscript{GAC} and the region on the right is ACTCCCTAGTCTCCAG. Here I have bolded the bases in phase with the T/A pattern in the 601 sequence and underlined the bases in phase with the G/C bases. The region on the left is in better keeping with the pattern than the region on the right. This could cause increased bendability on the left side of the 601 sequence which would decrease the energetic barrier to dimer incorporation. Future studies could narrow this region even further by swapping only the bases in phase with the A/T or G/C patterns.

The results of the ensemble and single molecule TF binding experiments showed that the dimer-proximal side of the hexasome (left side of the 601 sequence) behaves much like a fully formed nucleosome in terms of accessibility and binding/unbinding rates of TF’s. In the case of the dimer-distal side, not only does the hexasome remain permanently unwrapped, but the presence of the hexasome did not interfere substantially with the
binding of TF to DNA nor accelerate its dissociation in comparison with naked DNA as fully formed nucleosomes have been shown to do. This paints a picture of the hexasome similar to that shown in Figure 22, which shows a nucleosome crystal structure with 40 base pairs unwrapped and one dimer missing. It is impressive that such a large alteration on one side of the nucleosome could have such a minor impact on the other. The shift in DNA accessibility on the dimer-proximal side between nucleosomes and hexasomes is similar in size to the effects of a single PTM in the entry-exit region. This separation of influence likely implies that PTM’s on one dimer will not cause structural changes on the other side of the nucleosome.
8. Summary and Future Directions

The studies presented here have centered around fluorescent measurements of transcription factor (TF) binding inside of nucleosomes. These measurements are not only relevant for modeling TF occupancy on sites buried in nucleosomes but also give a quantitative measure of the accessibility of nucleosomal DNA in the entry-exit region for any specific DNA binding factor.

In the first experimental chapter (chapter 5) we looked at the effect of two post-translational modifications (PTM’s), H3Y41ph and H3T45ph, on the accessibility of nucleosomal DNA in the entry-exit region. In some experiments, we used glutamate as an amino acid substitute for phosphorylations because these mimic the addition of a negative charge. Micrococcal nuclease digestions on nucleosomes reconstituted with H3Y41E and H3T45E both showed shorter digestion products than nucleosomes with wild-type H3. Small Angle X-Ray Scattering (SAXS) performed on the same samples showed that H3Y41E and H3T45E nucleosomes showed an increased radius of gyration with respect to wild type and produced molecular envelopes that were consistent with nucleosomes that were partially unwrapped. Both of these results point to the phosphorylation mimics increasing the amount of time a nucleosome spends in a partially unwrapped state.

Next, we wanted to know if this increased unwrapping would impact the ability of transcription factors to bind in the entry-exit region of the nucleosome. We were able to
quantify the accessibility of DNA in the entry-exit region with the FRET-based TF binding assay described in chapter 2. This showed that H3Y41E, H3T45E, and the chemically correct modification H3Y41ph were all ~3x more accessible to TF binding than the wild-type nucleosomes. This result implies that in this context glutamate is an acceptable phosphorylation mimic both qualitatively and quantitatively. We can also begin to see a pattern that indicates PTM’s in the entry-exit region of the octamer which reduces charge increase accessibility by about a factor of 3, as was found for H3K56ac77.

Since both H3Y41ph and H3K56ac are modifications associated with actively transcribed genes, it is likely that they could co-occur on nucleosomes and so we prepared H3 with both modifications and performed the same TF binding experiments with them. This showed a $17 \pm 5$ fold increase in accessibility which is consistent with the $\Delta \Delta G$’s from each individual modification being additive. Thus it is possible for combinations of modifications in the entry-exit region to produce a much larger effect than either modification alone. However, when we performed the same assay on the double mimic H3(Y41E, T45E), we found that the effects of these two mimics appeared redundant. Section 5.4 discusses a possible resolution to this discrepancy, but one should be careful in generalizing the results from the H3(Y41ph, K56ac) to other modification combinations.

The 6th chapter discusses ensemble and single molecule experiments performed on hexasomes (nucleosomes with one missing H2A/H2B heterodimer). TF binding experiments on the dimer-proximal side of the hexasome showed similar accessibility to that of a nucleosome with only a 2-fold decrease in accessibility. This is much smaller effect than one might imagine for such a profound structural change. The dimer-distal side
of the hexasome was found to be unwrapped with no long-lived transiently wrapped state. This unwrapped DNA showed accessibility to TF binding very similar to naked DNA with similar binding and unbinding rates.

These hexasome experiments took advantage of a recently discovered property of the 601 positioning sequence which is the propensity to form hexasomes with the dimer preferentially on one side. We first confirmed the dimer’s position by exonuclease mapping. We then investigated this phenomenon further and found a 16-base pair region of the sequence which is primarily responsible for this asymmetry. When this region was swapped for the complementary region on the other side of the 601 sequence, dimers primarily formed on the opposite side in salt dialysis reconstitutions.

8.1 Future Directions

Our studies on the effects of PTM’s were all done with ensemble methods and therefore we could not determine whether the change in accessibility was due to an increase in the binding rate or a decrease in the unbinding rate. One would expect the former because SAXS experiments show that, at least for the PTM mimics, DNA in the entry-exit region was unwrapped more often. Additionally, previous studies have shown that H3K56ac increases the binding rate of TF in the entry-exit region\textsuperscript{43}. Our results showing a 17-fold change between H3(Y41ph, K56ac) and unmodified nucleosomes means that the effect should be easy to capture with our single molecule techniques.

The study of hexasomes briefly looked into the origin of the 601 sequence’s ability to form hexasomes with the dimer reliably on one side. The 16 bases identified were not entirely responsible for the effect. Futures studies will look at the effects of a larger range
of bases and narrow down exactly which bases in the identified regions contribute most to
dimer positioning. It is likely that this will focus on the periodic T/A bases discussed in
section 7.4. These experiments might lay the groundwork for genome-wide studies which
look for patterns in DNA sequence that give preferential direction to hexasomes.

Lastly, it would be interesting to investigate the distribution of dwell times that
occur in our single molecule experiments that observe Gal4 binding inside of nucleosomes.
As can be seen in Figure 28, the dwell-time distributions do not fit well to an exponential
distribution. A possible reason for this could be that there are actually more than two states
for the system (bound and unbound), which are not distinguished by FRET efficiency.
Adding a third fluorophore to the Gal4 TF and imaging all three colors would enable
colocalization to provide a separate fluorescent indicator of binding. This could help us
determine the nature of any third state that the nucleosome might exist in.
References


43. North, J. a *et al.* Regulation of the nucleosome unwrapping rate controls DNA


71. Anderson, J. D. & Widom, J. Sequence and position-dependence of the equilibrium


77. North, J. A. *et al.* Regulation of the Nucleosome Unwrapping Rate Controls DNA Accessibility.


85. Ha, T. *et al.* Probing the interaction between two single molecules: fluorescence resonance energy transfer between a single donor and a single acceptor. *Proc. Natl.*


112. Li, Q. *et al.* Acetylation of histone H3 lysine 56 regulates replication-coupled


155. Narlikar, G. J., Sundaramoorthy, R. & Owen-Hughes, T. Mechanisms and functions


Appendix A: Buffers

- Unfolding Buffer
  - 7M Guanidine-HCL (Biomedicals 820539)
  - 10mM Tris-HCl, pH 7.5 (Sigma T1503)
  - 10mM Dithiothreitol (DTT, Sigma 43816)

- Refolding Buffer
  - 2M NaCl (Sigma S9888)
  - 10mM Tris-HCl, pH 7.5 (Sigma T1504)
  - 1mM ethylenediaminetetraacetic acid (EDTA, Sigma E1644), pH 8.0
  - 2mM 2-mercaptoethanol (BME, Sigma M3145) Add just before use, leave out for size exclusion elution buffer

- TCEP
  - 200mM tris(2-carboxyethyl)phosphine (Sigma C4706)
  - pH to 7.1 with NaOH, store at -80

- PIPES buffer
  - 5mM Sodium PIPES, pH 6.1 (Sigma P6757)
  - 2M NaCl
  - 0.5xTE
  - 5mM Tris-HCl pH 8.0
  - 0.5mM EDTA

- High Salt Buffer
  - 5mM Tris-HCl pH 8.0
  - 0.5mM EDTA
  - 1mM Benzamidine (BZA, Sigma 12072)
  - 2M NaCl

- Low Salt Buffer
  - 5mM Tris-HCl pH 8.0
  - 0.5mM EDTA
  - 1mM BZA

- 0.3x TBE
  - 7mM Tris
  - 27mM Boric Acid
  - 0.6mM EDTA pH 8.0

- Wash buffer
  - 10mM Tris-HCl, pH 8.0
  - 130mM NaCl
- 10% (v/v) glycerol (Sigma G5516)
- 1.5x (0.0075%, v/v) Tween-20 (Sigma P1379)

**Blocking buffer**
- 1mg/mL Ultrapure BSA (Ambion AM2616) in wash buffer

**1x Trolox Buffer (For Nucleosome FRET)**
- Dissolve 10mg Trolox (Sigma 238813) in 10ml of 50mM Tris-HCl pH 8.0 for 1 hour on a shaker table at 37°C. Check pH and adjust to 8.0 with 1M NaOH if necessary. Always prepare fresh.

**1x Tris Buffer (For DNA PIFE)**
- 50mM Tris-HCl pH 8.0

**1x Pre-Imaging buffer**
- 0.75x Trolox buffer (for FRET) or 0.75x Tris Buffer (for PIFE)
- 10% glycerol
- 130mM NaCl
- 0.075% Tween 20
- 1% (v/v) BME (for PIFE only)

**40% glucose (Sigma G8270)**

**25x Gloxy**
- 37.5 mg/ml Glucose Oxidase (Sigma G2133)
- 1830 μg/ml Catalase (Sigma C3155)
- 10mM Tris-HCl pH 8.0
- 50mM NaCl
- Centrifuge for 5 minutes at 15,000 g and filter by 0.2μm filter; store at -20°C. Make dilutions of Gloxy to 8x in 10mM Tris-HCl pH 8.0, 50mM NaCl and keep on wet ice on the day of the experiment.

**1000x COT/NBA**
- 11.5% Cyclooctatetraene (COT, Sigma 138924)
- 13.5% Nitrobenzyl alcohol (NBA, Sigma 146056)
- 75% Ethanol
- Store at -20°C; make dilutions to 25x in dH2O and keep on wet ice on the day of the experiment.

**15% Denaturing Gel**
- 12.62g Urea (Sigma U1250)
- 11.25ml Acrylamide (National Diagnostics EC-852)
- 3.75ml Water
- 6ml 5xTBE
- 300ul Ammonium persulfate (APS) (Amresco K8330100TABS)
- 10ul N,N,N’,N’-Tetramethylethlenediamine (TEMED Sigma-Aldrich T9281-25ml)
- Add mixed bed resin to activate the Urea while dissolving. Filter and then add APS and TEMED before pouring. Run at 26W for 2h
Appendix B: Flow Cell Preparation Protocol

This protocol was published in Methods in Enzymology\textsuperscript{81}. Flow cells are formed from a three-layer stack as shown in Figure 32. Eight channels are cut in Parafilm (Bemis PM966) which is then sandwiched between a passivated 50x24x0.16mm glass coverslip (Fisher 12-544-E) and a 1”x3”x1mm quartz slide (G. Finkenbeiner) with drilled inlet and outlet holes. These holes may be drilled with a CO\textsubscript{2} laser engraver or Starlight 115005 diamond drill bit to a diameter of 0.029 inches. Carefully clean and dry all glassware before use in functionalizing flow cell components.

![Figure 32 Diagram Showing the Assembly of Flow Cells. The quartz slide has 16 (0.029-inch diameter) holes, two for each of the channels. Parafilm forms the horizontal boundaries of the channels and holds the slide and coverslip together to form the completed flow cell. A Clean slides and coverslips](image)

118
Pacification and functionalization of slides and coverslips begins with cleaning any organics from the surfaces. Slides and coverslips are mounted vertically and separate in a custom made Teflon holder throughout the cleaning process.

Required for this protocol:

- Chromasolv Toluene (Sigma 34866)
- Chromasolv Ethanol (Sigma 270741)
- Concentrated H₂SO₄ (Sigma 320501)
- 50% (w/v) H₂O₂ (Sigma 516813)
- NaOH pellets (Sigma 306576)

1. Sonicate the quartz slides and coverslips in toluene and then ethanol for 20 minutes each.
2. Thoroughly rinse in dH₂O 3 times to remove organics.
3. Put on acid-safe personal protection equipment (gown, gloves, and face shield) in preparation for piranha cleaning. In a clean beaker on a heat resistant surface in an acids fume hood, add 2 parts H₂SO₄ followed by 1 part fresh 50% H₂O₂ to cover the slides, typically 450mL total. Vigorous bubbling and heating indicates a successful reaction. When the reaction has completed after about 30 minutes, remove the slides and rinse three times with dH₂O as before.
4. In a new beaker sonicate in 1M Filtered NaOH for 20 minutes.
5. Rinse 3 times with dH₂O as above
6. Dry slides overnight at 75°C.

**B.1 Silanize Slides and Coverslips**

Amine-functionalized silane is covalently bonded to hydroxyl groups on the quartz and glass present after cleaning to serve as functional groups for subsequent PEG derivatization. The slides and coverslips should be placed in the Teflon holders use in 5.1.1 above for this protocol.
Required for this protocol:

- 2% v/v (3-aminopropyl)triethoxysilane (MP Biomedicals 215476680)
  - (lasts 3 months if stored in refrigerated desiccator)
- Acetone (Sigma 650501)

1. In a dry beaker cover the slides with 500mL of water-free acetone. Fresh acetone from an unopened, factory sealed bottle is essential to eliminate water contamination which quenches the silane reaction.
2. Swirl the slides for 10 minutes vigorously without exposing them to air.
3. While swirling, add 8mL silane drop-wise over the course of 1 minute. Allow the reaction to proceed for 2 minutes.
4. Quench the reaction by transferring the slides to a beaker containing 50% dH₂O 50% acetone.
5. Rinse with dH₂O 3 times as in section 5.1.1 above.
6. Incubate at 75°C for 6 hours.

B.2 Coat slides and Coverslips with PEG

Functionalized PEG is covalently bound to the amino-silane derivatized surface and is responsible for both surface pacification and biotin functionalization for sample immobilization. This is accomplished through a mixture of 5000 Dalton average molecular weight monofunctional PEG-Succinimidyl Valerate (mPEG-SVA) and biotin-PEG-SVA (bioPEG-SVA).

Required for this protocol:

- 5000 Da monofunctional-PEG-SVA (Laysan Bio, mPEG-SVA-5000)
  - Stored in 50mg aliquots at -80°C for 1 year
- 5000 Da biotin-PEG-SVA (Laysan Bio, Biotin-PEG-SVA-5000)
  - Stored in 5mg aliquots at -80°C for 1 year
- 0.1M Potassium Tetraborate (Sigma P5754)
  - pH to 8.4 with HCl and store at -20°C

1. Dissolve 2 mPEG-SVA aliquots into 450µL of 0.1M Potassium Tetraborate Buffer each to make 10% (w/w) solutions. Suspend 1 bioPEG-SVA aliquot in 200µL 0.1M Potassium Tetraborate Buffer to make a 2.5% (w/w) solution. Vortex each to fully suspend PEG powder. The solution should be clear or slightly cloudy.
2. Add 20µL of bioPEG-SVA to each 10% mPEG-SVA solution yielding 1:100 bioPEG:mPEG mass ratio. Vortex to mix and centrifuge at 20,000g for 3 minutes to pellet undissolved solids. Transfer supernatant to a clean 1.5mL tube.
3. Remove all amino-silane derivatized quartz slides from the slide holder and place on a clean glass plate.
4. Pipette 50µL of bioPEG/mPEG solution onto the center of each slide. Slowly place one coverslip centered on top of the quartz slide. Ensure no air bubbles are formed.
5. Track slide/coverslip pairs and the sides receiving PEG treatment for later reassembly.
6. Cover against dust and let sit at room temperature for 1 hour.
7. Carefully separate the coverslip from the quartz slide and wash in dH₂O to remove extra PEG solution by washing in a dH₂O bath.
8. Air dry the slides and coverslips vertically in the Teflon slide holder from section 5.1.1 above.
9. Clean and dry the glass plate.
10. Repeat step 4-7 above ensuring the treatment is performed on the same side as before with the same pair.
11. Incubate at 37°C until dry.

**B.3 Assemble Flow Cells with Parafilm**

Cleaned, PEG functionalized slides and coverslips are finally assembled into fully formed flow cells by attaching them with a strip of Parafilm (American National Can) with flow channels cut into it and adhered with heat.

1. Prepare Parafilm by cutting channels in line with previously drilled inlet/outlet holes on the quartz slides.
2. Place one PEGylated quartz slide on a clean, dry glass plate with the PEG side facing up. Place a single sheet of prepared Parafilm with paper backing removed on the quartz slide and align with the inlets/outlets holes on the slide. Cover with a PEGylated coverslip, PEG side down. Flow cell assembly is shown in Figure 32. Press the coverslip slightly to temporarily adhere it to the slide.
3. Heat a glass plate to 90°C on a hot plate.
4. Place one flow cell assembly at a time on the hot glass, quartz slide down. Heat just long enough for the Parafilm to turn partially clear. Gently press the coverslip down to permanently adhere it to the Parafilm. Do not heat long enough to completely melt the Parafilm to avoid channel distortion and slide cracking.
5. Store flow cells at room temperature in low humidity. They are good for at least 1 month. Store under dry nitrogen in a desiccator for longer term storage.
6. Flow cells are dismantled by soaking in acetone for 15 minutes and gently prying the coverslip and slide apart. Quartz slides can be reused repeatedly.
Appendix C: smTIRF Data Acquisition

This protocol was published in *Methods in Enzymology*.

### C.1 Attaching the Sample

Required for this protocol:

- Wash buffer
- Blocking Buffer
- 20µg/mL Streptavidin (Sigma S4762) in wash buffer.
- Nucleosome or DNA sample at 20nM in 0.5xTE

1. Hydrate the flow cell with 100-200µL wash buffer for 2 minutes.
2. Inject 20µL of Blocking buffer into the flow cell.
3. Incubate the flow cell for 2 minutes.
4. Wash blocking buffer from flow cell with 100-200µL of wash buffer.
5. Inject 20µL of Streptavidin in wash buffer and incubate for 2 minutes.
6. Meanwhile dilute biotinylated, fluorescent nucleosome or DNA sample to 1-10 pM in 20µL wash buffer. The optimal sample concentration to yield a final desired molecule distribution on the flow cell surface will need to be empirically determined for each sample. Note nucleosome concentrations of 1-10 pM in solution are not stable and need to be diluted just before surface attachment.
7. Wash Streptavidin from flow cell with 100-200µL of wash buffer.
8. Immediately inject 20µL of biotinylated, fluorescent nucleosomes
9. Incubate for 5 minutes. During incubation prepare 50µL imaging buffer for injection into the flow cell.
   - Final Imaging Buffer
     - 0.75x Pre-Imaging Buffer
     - 1.6% glucose
     - 1x COT/NBA (not used for PIFE)
     - 0.3x Gloxy
     - 1x Transcription Factor or protein of interest
10. Wash unbound sample from the flow cell with 100-200µL of wash buffer.
11. Inject 50µL imaging buffer into the channel.
Imaging the Flow Cell

Required for this protocol:

- Immersion oil (Cargille Type FF, index of refraction matched to the quartz slide)
- Pellin-Broca prism (Melles Griot)

1. Clean the top surface of the quartz slide using a Kimwipe or lens paper.
2. Place a small drop of immersion oil at the center of the flow channel. The diameter of the drop on a clean surface should be a little smaller than the width of the channel. Too much or too little immersion oil will adversely affect the TIRF field through the formation of diffraction patterns.
3. Mount the flow cell with the microscope coverslip down (immersion oil up) to the microscope body. Lower the prism mount and roughly align the immersion oil with the objective so that the prism, once lowered, is wetted in a pool of immersion oil, optically coupling it to the quartz slide. Place the prism in its holder and gently lower it onto the slide, ensuring no air bubbles are trapped between the two. Slide the prism back and forth to disperse immersion oil.
4. Using bright field illumination, raise the objective to focus through the flow cell onto the bottom surface of the quartz slide.
5. Change from bright field to TIRF illumination and change the microscope output from the eyepiece to the DualView two-channel simultaneous imaging system and CCD camera.
6. Turn on each laser successively to ensure the generation of a good TIRF field.
7. Place the camera in a real-time video readout mode and re-focus the image.
8. Adjust the location of the excitation beam to give uniform illumination over the entire field of view.
9. When the acceptor is directly excited, only the acceptor channel should show emission. When the donor channel is excited, emission in the acceptor channel indicates FRET.
10. Specifics of data acquisition are highly dependent on the experiment and experimental setup, but typical settings are a high EM gain, roughly 50-200ms exposure time, a 2000 frame acquisition window and 2x2 pixel binning of a 512x512 CCD pixel image. EM gain is set to maximize signal to noise (SNR), exposure time is minimized to catch fast fluctuations while maximizing SNR, and the total acquisition time should be long enough to allow for significant photobleaching of the field of view.
11. For FRET measurements:
a. Begin data acquisition by first directly exciting the acceptor fluorophore without donor excitation for 25-50 frames. These frames will be used to construct a high SNR average image to identify points of interest during data analysis.

b. Switch laser sources from acceptor exciting to donor exciting and allow the remainder of the acquisition window to elapse.

10. For PIFE measurements:
   a. Simply excite the fluorophore and acquire data for the full acquisition window.

11. Save the data and proceed to a new, un-photobleached location within the flow cell for another video acquisition.

12. Repeat steps 7-13 until enough data is collected or the imaging buffer degrades (roughly 1 hour). After data is collected, this protocol may be repeated on the other channels of the flow cell to allow for up to 8 independent experiments.
Appendix D: DNA Sequences

601 sequence shown in black. Flanking DNA shown in gray. Gal4-2c sites shown in red. LexA site shown in orange. The swapped 16 bases that make contact with the H2A/H2B shown in green (strong) and blue (weak).

**Widom 601 sequence:**
CTGGAGAATCCCGGTGCGGAGGCGCTCAATTGGTCTGAGACGCTCTAGCA
CCGGCTAAACGCACGTACGCCTCGTCGCCCTCCCCCCGTCTTTTAACCGCCAAGGGGAT
TACTCCCTAGTCTCCAGGCAGTCTCGAGATATATACATCCTGT

**S-FRET**
aminoC6-5'CTGGAGACCCGAGGGCTGCCCTCCGGTCAATTGGTCTAGACAGCTCTAGCA
CTGGCGACCCGCTTAAACGCACGTACGCCTCGTCGCCCTCCCCCCGTCTTTTAACCGCCAAGGGGAT
TACTCCCTAGTCTCCAGGCAGTCTCGAGATATATACATCCTGT

**W-FRET**
Biotin-3'-TCTAGAGCTAGCCTACGTGAGAAGCTGCTGACGAGGCTTGGTCTAGACAGCTCTAGCAATTGGTCTAGACAGCTCTAGCA
CTGGCGACCCGCTTAAACGCACGTACGCCTCGTCGCCCTCCCCCCGTCTTTTAACCGCCAAGGGGAT
TACTCCCTAGTCTCCAGGCAGTCTCGAGATATATACATCCTGT

**DNA-PIFE**
aminoC6-5'-ATCTAAGGTCTAGCCTACGTGAGAAGCTGCTGACGAGGCTTGGTCTAGACAGCTCTAGCAATTGGTCTAGACAGCTCTAGCA
CTGGCGACCCGCTTAAACGCACGTACGCCTCGTCGCCCTCCCCCCGTCTTTTAACCGCCAAGGGGAT
TACTCCCTAGTCTCCAGGCAGTCTCGAGATATATACATCCTGT

126
**SW-Exo**
Cy3-ATATAAGGAGGACACTGGGACATGCATCGGCTGGAGACCCGGAGGGGCTGCCCCTCCGGTCATTTGCTGTAAGACAGCTCTAGCACCCTTAACGCACGTACGCCTGTCCCCCGCTTTTAAACCCGCTTATTACTCCCTAGTCTCCAGGGCAGCGTGCAGATATACATCCTGTATAATGCATAGGGCAGTGAGTGATTTGACGCTAC-Cy5

**WS-16-Exo**
Cy3-ATATAAGGAGGACACTGGGACATGCATCGGCCTGGAGACCCGGAGGGGCTGCCCCTCCGGCTGGAGAAGCTAGGGAAGTACTCTAGCACCCTTAACGCACGTACGCCTGTCCCCCGCTTTTAAACCCGCTTATTGTCTAGCAGCCAATTGAGCAGTGCAGATATACATCCTGTATAATGCATAGGGCAGTGAGTGATTTGACGCTAC-Cy5
Appendix E: Primer Design Considerations

In my work I have had to design many nucleosomal DNA constructs. These involved designing and ordering custom primers used in polymerase chain reaction (PCR). See Figure 10 for a schematic explanation of PCR. There are many considerations to keep in mind when designing these primers.

E.1 General PCR Primer Design

1. Check that primers anneal to the correct ends of DNA. The 5’ end of the forward primer will determine one end of the final product and the 3’ end of the reverse primer will determine the other end of the product.

2. Check that primers don’t end inside a palindromic DNA sequence. Transcription factor binding sequences are often palindromic and primers that end inside these sequences can anneal in two locations causing side-products.

3. Check that there is a G or C base at the 3’ end. The G/C base pairing interaction is stronger than the A/T interaction and it is thus makes for a better polymerase starting point.

4. Check that both internal and end labels have 6-carbon linker attachments. This helps the label rotate so that the effects of orientation on FRET efficiency are averaged out.
5. If you are using an internal label, make sure to use a thiophosphate linker before the last base on the 3’ end of the primer. This assures that error-correcting polymerases such as *pfu* will not backtrack and remove the label. If a internally labeled primer does not have a thiophosphate linker it can still be used with a polymerase that doesn’t have exonuclease activity such as *taq*.

6. If a PCR reaction produces unwanted side-products, raising the annealing temperature or slowing down the rate of temperature change before the annealing step can help alleviate this problem.

7. Forward and reverse primers should have similar annealing temperatures.

8. Primers longer than 60 bases can cause difficulties in synthesis and should be avoided if possible.

### E.2 Design Considerations for Ligated Constructs

Some DNA constructs cannot be made with a single PCR reaction. It can be necessary to make multiple constructs and ligate them together. In this design, both constructs would contain a restriction enzyme cut site. The site would be digested to produce an overhang which would anneal to the other construct. The two annealed constructs would then be attached by DNA ligase. The following are considerations when designing these constructs.

1. Check that the PCR primers anneal correctly and follow the above considerations for general PCR.

2. Check that the segments ligate correctly. There should be a 5’ overhang on one and a complimentary 3’ overhang on the other.

3. Check that the restriction enzyme used doesn’t have cut sites in other parts of the sequence.
4. Check that the restriction enzyme cut site is a few bases away from any labels or modifications which could interfere with digestion.

5. Check that the restriction enzyme cut will decrease the length of the primer by roughly 10% or more so that the digestion can be confirmed by gel shift.

6. During the ligation step there are two strategies for assuring that the digested ends are not re-ligated back on instead of the desired segment. The first is purification either by gel purification or anion exchange chromatography to remove the digested ends. The second is using an excess of the desired extension that out competes the digested ends. This second strategy is useful when the desired extension is composed of DNA oligomers that do not need to be digested to produce a single stranded overhang.

7. Restriction enzyme cut sites are often palindromic. This can cause problems during ligation because it allows digested products to anneal to themselves. This problem can be solved by using a large excess of the desired extension as in (6). More often, this is solved by using a non-palindromic restriction enzyme such as TspRI. This cut site has degeneracy which allows for the design of a non-palindromic overhang.

8. When selecting a restriction enzyme, use the unit definition to calculate the cost to cut 100pmol for each enzyme. This cost can vary widely between enzymes and it is necessary to consider cost in the design.