READING THE EPIGENETIC STATE OF CHROMATIN ALTERS ITS ACCESSIBILITY

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 D. Gibson, B.S., M.S.
Graduate Program in Physics

The Ohio State University
2016

Dissertation Committee:
Professor Michael G. Poirier, Advisor
Professor Ralf Bundschuh
Professor Fengyuan Yang
Assistant Professor Comert Kural
Professor Michael Barton
Abstract

The eukaryotic genome is organized into a structural polymer called chromatin. Ultimately, all access to genetic information is regulated by chromatin including access required for DNA replication, transcription, and repair. The basic repeating unit of chromatin is the nucleosome which is comprised of ~147 bp of DNA tightly wrapped around a protein histone octamer core. The histone octamer is made up of eight proteins: two each of histones H2A, H2B, H3, and H4. Many mechanisms exist to regulate access to DNA but one of pivotal importance is the creation of unique nucleosomes through i) integration of histone variants and ii) deposition of post translational modifications (PTMs). These modifications help comprise the epigenome of a cell. Classically, the two mechanisms by which they function have been through a direct regulation of nucleosome dynamics, or through third party proteins which are able to recognize the variants or PTMs and facilitate work. The library of potential PTMs therefore forms a sort of histone code which regulates access to DNA. This thesis investigates the intersection of these mechanisms to determine whether the act of recognizing epigenetic information alters DNA accessibility.

The primary method used to determine changes in DNA accessibility is though observing the effective binding affinity of a transcription factor to its target site buried within a recombinantly prepared nucleosome which has been modified to carry a PTM and to report on its wrapping state. We find different regulation depending both on the PTM we investigate and the specific PTM-binding protein.

We first investigate the H3K36me3-binding protein PHF1 and find that while the PTM it recognizes, H3K36me3, does not alter DNA accessibility, the binding of its recognition
domain and N-terminal domain can illicit a change of DNA accessibility of $8 \pm 2$-fold. This means that 8 times less DNA binding protein is required to occupy its target site if the nucleosome is bound by PHF1.

Second, we move to another PTM-binding protein, LEDGF/P75. Though LEDGF/P75 recognizes the same PTM as PHF1, we find that rather than increasing accessibility, LEDGF/P75 decreases accessibility to DNA. Additionally, we find that LEDGF/P75 prefers binding nucleosomes which simultaneously contain the histone variant H3.3 and its target PTM over those containing canonical H3.2 and the PTM.

Finally, we shift from H3K36me3, a mark correlated to active transcription, to H3K9me3, which is correlated with compact, silent chromatin. The PTM-binding protein we investigate is Swi6 which, upon binding, initially increases DNA accessibility to nucleosomes by a factor of $2 \pm 0.3$. We further find that Swi6 maintains a roughly constant ratio of accessibility between modified and unmodified nucleosomes of $3.2 \pm 0.4$.

The findings presented here form a foundation for the study of the direct impact of the binding of PTM-recognition proteins. Future work will establish mechanistic insight into the alteration of DNA access within chromatin by PTM binding.
ACKNOWLEDGMENTS

I would like to begin this thesis by thanking my advisor, Prof. Michael Poirier. Your guidance has allowed me to expand my knowledge and abilities farther than I would have thought possible.

Thank you to my committee, Prof. Ralf Bundschuh, Prof. Fengyuan Yang, Prof. Comert Kural, and Prof. Michael Barton for your guidance throughout my graduate career and especially in critiquing this thesis.

I would like to thank all the people who have taught and helped me with this research. Dr. Marek Šimon was critical in starting Prof. Poirier’s lab. Dr. Paula Vivas gave me my first introduction into both biochemistry and single molecule experiments. Dr. Alex Mooney taught me more than I’d care to admit about experiments, logic, and coding. Anyone in the Poirier lab owes a debt to Dr. Justin North, who developed the basic chemical techniques used to fluorescently label histones used in this thesis as well as a huge number of other protocols used in the lab. Dr. Morgan Bernier shaped the way I see chemistry and biochemistry and taught me to be a careful scientist. Dr. Yi Luo laid the groundwork for single molecule experiments in the lab and the high standards by which we now operate. My lab contemporaries—Matt Brehove, Ben Donovan, Michael Darcy, Khan Cox, Amber Byrum—worked tirelessly to help take care of the lab and never shied away from offering advice or help.

I would especially like to acknowledge the hard work and contributions made by the undergraduates I’ve worked with. Andrew Slater was pivotal in establishing accurate means
of taking fluorometer data. Michelle Scott has worked with me to some extent on each project detailed in this thesis as well as a few which are not included here. I would not have accomplished as much without them.

I would like to thank my collaborators throughout my thesis research. Prof. Catherine Musselman, Dr. Jovylyn Gatchalian, and Prof. Tatiana Kutateladze have been wonderful in providing advice and protein for my work on PHF1. Matt Plumb and Prof. Mamuka Kvaratskhelia provided protein and background for my LEDGF/P75 project. I am thankful to Stefan Isaac and Prof. Geeta Narlikar for help and plasmids in my efforts to express and purify Swi6.

Finally, I would like to thank my friends and family. Mom, Jim, Michelle, Alex, and Jeremy, thank you for your understanding, support, and encouragement over the years. It means more than you can know.
VITA

July 7, 1989 ........................................Born—Phoenix, Arizona

December, 2010 ....................................B.S., Arizona State University, Tempe, Arizona

July, 2013 ............................................M.S., The Ohio State University, Columbus, Ohio

Publications

M.D. Gibson, J. Gatchalian, A. Slater, T.G. Kutateladze, M.G. Poirier. The N-terminal domain of PHF1 is important in controlling nucleosome accessibility. Nucleic Acids Res. (in review).


**Fields of Study**

Major Field: Physics

Studies in histone modification binding proteins: Michael G. Poirier
Table of Contents

Abstract ..................................................... ii
Acknowledgments ................................................ iv
Vita ............................................................. vi
List of Figures .................................................. xiv
List of Tables ................................................... xvii
List of Abbreviations ............................................. xvii

Chapters

1 Introduction .................................................. 1
  1.1 Chromatin and the Nucleosome .............................. 2
    1.1.1 Euchromatin and Heterochromatin ....................... 3
    1.1.2 Nucleosome Structure ................................... 4
    1.1.3 Nucleosome Dynamics ................................... 5
    1.1.4 The Histones ........................................... 5
  1.2 Histone Post Translational Modifications .................... 7
    1.2.1 Modifications Alter Nucleosome Dynamics and Equilibrium 9
    1.2.2 The Histone Code ..................................... 10
  1.3 Summary .................................................. 11

2 Ligand Binding .............................................. 12
  2.1 Introduction .............................................. 13
    2.1.1 Connections to Thermodynamics ......................... 14
  2.2 Binding of a Ligand to a Receptor .......................... 14
    2.2.1 The Hill Equation .................................... 15
    2.2.2 Ligand Binding in the Presence of Co-Factors ........... 16
  2.3 Transcription Factor Binding Within the Nucleosome ........ 19
    2.3.1 3-State Model of Transcription Factor Invasion Into the Nucleosome 20
  2.4 Looping Interactions ..................................... 22

3 Preparation of nucleosomes with Multiple Labels ................ 24
  3.1 Design .................................................. 24
  3.2 Order of Labeling ........................................ 26
  3.3 Verification of Product ................................... 26
4 Fluorescence Methods .................................................. 28
  4.1 Förster Resonant Energy Transfer ............................... 28
    4.1.1 A Brief History ............................................ 29
    4.1.2 A Classical Derivation .................................. 31
    4.1.3 Construct Design ......................................... 36
    4.1.4 Determination of FRET Efficiency ....................... 37
    4.1.5 Special Concerns with Labeled Histone Octamer and DNA .... 39
  4.2 Single Molecule Techniques ..................................... 43
    4.2.1 TIRF Microscopy .......................................... 44
    4.2.2 Zero-mode waveguides ................................... 48
    4.2.3 Single Molecule Sample Preparation ...................... 52
    4.2.4 Data Acquisition ........................................ 54
    4.2.5 Data Analysis ........................................... 55
  4.3 Anisotropy ....................................................... 59

5 PHF1 Directly Modulates Nucleosome Accessibility While Binding to its Target H3K36me3 .................................................. 62
  5.1 Introduction ..................................................... 63
  5.2 Results .......................................................... 64
    5.2.1 Tudor Binds to the H3K_{36}me3 Nucleosome .......... 64
    5.2.2 Tudor Has a High Affinity for H3K_{36}me3-NCP ........ 68
    5.2.3 Tudor Associates With dsDNA ............................ 69
    5.2.4 The Model of the PHF1 Tudor/H3K_{36}me3-NCP Complex .... 70
    5.2.5 Interaction With H3K_{36}me3 Stabilizes a More Open Form of NCP 73
  5.3 Methods .......................................................... 74
    5.3.1 DNA Constructs and Protein Purification ................. 74
    5.3.2 MLA Generation ........................................... 76
    5.3.3 HADDOCK Modeling ....................................... 76
    5.3.4 Nucleosome Reconstitution ................................ 77
    5.3.5 Pulldown Assays .......................................... 77
    5.3.6 NMR Spectroscopy ......................................... 78
    5.3.7 Nucleosome and LexA Preparation for FRET Measurements 79
    5.3.8 FRET Measurements ....................................... 80
  5.4 Discussion ....................................................... 81

6 PHF1 Tudor and N-terminal Domains Synergistically Target Partially Unwrapped Nucleosomes to Increase DNA Accessibility .................. 83
  6.1 Introduction ..................................................... 84
  6.2 Results .......................................................... 86
    6.2.1 The PHF1 NTD is Partially Structured When Linked to Tudor .... 86
    6.2.2 PHF1 Tudor-NTD Does Not Shift Nucleosomes Into a Largely Unwrapped State. 88
    6.2.3 The PHF1 NTD Enhances the Impact of the Tudor Domain on Nucleosome Accessibility .................. 90
    6.2.4 The Increased Influence of PHF1-Tudor with the NTD on Nucleosome Accessibility is Specific to H3K_{36}me3 Nucleosomes .......... 91
6.2.5 PHF1 Tudor-NTD Increases the DNA-Binding Protein Occupancy Within the Nucleosomes by Nearly an Order of Magnitude. 93
6.2.6 The Entire PHF1 NTD is Necessary to Enhance DNA Accessibility Within the Nucleosome. 94
6.2.7 A Four-State Binding Model Predicts that PHF1 Preferentially Interacts with Unwrapped Nucleosomes by Nearly an Order of Magnitude. 96
6.2.8 The PHF1 Tudor Domain Regulates Occupancy of DNA-Binding Complexes Within the Nucleosome by Increasing the Dissociation Rate. 98
6.3 Methods 100
6.3.1 PHF1 Constructs and Purification 100
6.3.2 Nuclear Magnetic Resonance 100
6.3.3 MLA Deposition on Histones 101
6.3.4 Preparation of FRET-Labeled Histone Octamer 101
6.3.5 Preparation of Labeled DNA 102
6.3.6 Preparation of Nucleosomes 102
6.3.7 Preparation of LexA 103
6.3.8 Ensemble FRET Measurements 103
6.3.9 Electromobility Shift Assay 104
6.3.10 Single Molecule FRET Measurements 104
6.4 Discussion 105
7 LEDGF binding protects nucleosomal DNA 107
7.1 Introduction 107
7.2 Results 108
7.2.1 LEDGF Does not Directly Alter Nucleosome Dynamics 108
7.2.2 The Methylation of H3.3K36 Increases DNA Accessibility 109
7.2.3 The Methyl-Recognizing PWWP Domain Does not Impact DNA Accessibility 110
7.2.4 Full Length LEDGF Protects Nucleosomes From Invasion by Transcription Factors 111
7.2.5 LEDGF Binds to Nucleosomes Containing H3.3 with Roughly a 10-fold Higher Affinity 112
7.3 Methods 112
7.3.1 Protein Preparation 112
7.3.2 Preparation of Labeled DNA 113
7.3.3 Nucleosome Preparation 113
7.3.4 FRET Measurements 113
7.4 Discussion 114
7.5 Future Directions 114
8 Swi6 Can Function to Increase Accessibility to Nucleosomal DNA 116
8.1 Introduction 116
8.2 Results 117
8.2.1 Swi6 Binds Oligomerizes on Both Modified and Unmodified Nucleosomes, As Previously Reported 117
8.2.2 Binding of Swi6 does not impact FRET Efficiency 118
8.2.3 H3K9\textsubscript{C}me3 Slightly Increases DNA Accessibility .................. 119
8.2.4 Swi6 Alters DNA Accessibility to Nucleosomes Containing
H3K\textsubscript{C}9me3 Differentially Depending on Concentration ............ 120
8.2.5 Swi6 Largely Does Not Change DNA Accessibility to Unmodified Nu-
cleosomes ................................................................. 122
8.2.6 Swi6 Maintains Relative Accessibility Between Unmodified Nucleo-
somes and those Containing H3K\textsubscript{C}9me3 .................................. 123

8.3 Methods ................................................................. 124
8.3.1 Protein Preparation .................................................. 124
8.3.2 Preparation of Labeled DNA ........................................ 124
8.3.3 Nucleosome Preparation ............................................ 124
8.3.4 Electromobility Shift Assay ........................................ 125
8.3.5 Anisotropy Measurements .......................................... 125
8.3.6 FRET Measurements ................................................. 125

8.4 Discussion ............................................................... 126
8.5 Future Directions ....................................................... 127

9 Concluding remarks and future directions ........................................ 128
9.1 Summary of Results .................................................... 128
9.2 Suggested Future Studies ............................................... 131
9.3 Closing Remarks ....................................................... 132

Bibliography ............................................................... 133

Appendices
A Purification of LexA ................................................... 147
A.1 Buffers ................................................................. 147
A.1.1 Buffer A ............................................................. 147
A.1.2 Buffer B ............................................................. 148
A.1.3 Buffer B + 200mM Sodium Chloride .............................. 148
A.1.4 Buffer B + 800mM Sodium Chloride .............................. 148
A.1.5 Buffer C ............................................................. 148
A.1.6 Buffer D ............................................................. 149
A.1.7 Polyethyleneimine (Polymin P) ................................... 149
A.1.8 2M potassium phosphate dibasic ................................ 150
A.1.9 0.5M potassium phosphate buffer, pH 7.0 .................... 150
A.2 Protein Expression ................................................... 150
A.2.1 Trial Induction ...................................................... 150
A.2.2 Bulk Induction ...................................................... 152
A.2.3 Harvesting Cells .................................................... 154
A.3 Protein Purification .................................................. 155
A.3.1 Cell Lysis .......................................................... 155
A.3.2 Cell Debris Pelleting ............................................. 155
A.3.3 Polyethyleneimine Precipitation ................................. 156
A.3.4 Salting Out Purification ......................................... 157
A.3.5 HiTrap Heparin Affinity Purification .................................................. 159
A.3.6 Hydroxyapatite Ion Exchange Purification ............................................. 160
A.3.7 Determining Fractions and Dialysis ..................................................... 164
A.4 Protein Storage ..................................................................................... 165
A.5 LexA Verification .................................................................................. 165

B Purification of HP1 Homolog Swi6 ................................................................. 167
B.1 Sequences ............................................................................................. 167
B.1.1 Swi6 WT ......................................................................................... 167
B.1.2 Swi6^{AcidX} .................................................................................. 169
B.1.3 Swi6^{LoopX} .................................................................................. 170
B.2 Buffers ................................................................................................ 171
B.2.1 Protease Inhibitors .......................................................................... 171
B.2.2 10×PBS .......................................................................................... 172
B.2.3 Lysis Buffer .................................................................................... 172
B.2.4 Elution Buffer .................................................................................. 172
B.2.5 SE Buffer ....................................................................................... 173
B.2.6 AX Buffer ....................................................................................... 173
B.2.7 Storage Buffer ................................................................................ 173
B.3 Protein Expression .............................................................................. 174
B.3.1 Transformation of Swi6 Expression Vectors ......................................... 174
B.3.2 Trial Induction ............................................................................... 175
B.3.3 Bulk Induction ............................................................................... 177
B.3.4 Harvesting Cells ............................................................................. 179
B.4 Protein Purification .............................................................................. 180
B.4.1 Cell Lysing .................................................................................... 180
B.4.2 Ni-NTA Purification ....................................................................... 181
B.4.3 Size Exclusion Purification ............................................................... 185
B.4.4 TEV Cleavage of the His-Tag .............................................................. 187
B.4.5 Anion Exchange Purification .............................................................. 188
B.5 Protein Storage ................................................................................... 189
B.5.1 Storage Buffer ............................................................................... 190
B.5.2 Stability ......................................................................................... 190

C Site-Specific Deposition of Trimethyllysine Analog ........................................ 192
C.1 Materials ............................................................................................ 192
C.2 Labeling ............................................................................................. 193
C.3 Storage ............................................................................................... 194
C.4 MALDI-TOF Verification ..................................................................... 194

D Purification of Multiply Labeled Histone Octamers for Use in Fluorescence Assays ......................................................... 196
D.1 Procedure ......................................................................................... 196

E Light Path Shutters for Triclops ................................................................. 198
E.1 Parts List ............................................................................................ 200
F Triclops ................................................................. 201
  F.1 Excitation Path .................................................... 201
  F.2 Emission Path .......................................................... 202
F.3 Triclops Parts ........................................................... 204
  F.3.1 Major Components .................................................. 204
  F.3.2 Excitation Parts .................................................... 205
  F.3.3 Emission Parts ..................................................... 206
F.4 Calibration .............................................................. 207

G Copyright Permissions .................................................... 208
  G.1 Figure 1.1 Permission to Reuse ................................. 208
  G.2 Figure 1.3 Permission to Reuse ................................. 209
  G.3 Figure 2.3 Permission to Reuse ................................. 210
# List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1 Cartoon of DNA organization within eukaryotic organisms</td>
<td>2</td>
</tr>
<tr>
<td>1.2 Structure of the nucleosome</td>
<td>4</td>
</tr>
<tr>
<td>1.3 Structure of known histone Post Translational Modifications</td>
<td>8</td>
</tr>
<tr>
<td>1.4 Cartoons showing the two dominant mechanisms of histone post translational modifications impacting DNA accessibility</td>
<td>9</td>
</tr>
<tr>
<td>2.1 Example free energy landscape showing multiple states and the relevant energy differences</td>
<td>14</td>
</tr>
<tr>
<td>2.2 Non-cooperative binding isotherm</td>
<td>16</td>
</tr>
<tr>
<td>2.3 Distribution of transcription factors distances from the nearest nucleosome dyad</td>
<td>19</td>
</tr>
<tr>
<td>2.4 Three state model of transcription factor invasion into the nucleosome</td>
<td>20</td>
</tr>
<tr>
<td>3.1 3-state nucleosome invasion model with fluorophores indication the level of expected FRET efficiency for each state</td>
<td>25</td>
</tr>
<tr>
<td>3.2 FRET labeling placement methodology and design</td>
<td>25</td>
</tr>
<tr>
<td>3.3 MALDI-TOF mass spectrometry of H3.3K\textsubscript{C36me3} labeling reaction</td>
<td>27</td>
</tr>
<tr>
<td>4.1 Dipole-dipole interaction coordinate system</td>
<td>32</td>
</tr>
<tr>
<td>4.2 Distance dependence of FRET efficiency as a function of Förster distance</td>
<td>36</td>
</tr>
<tr>
<td>4.3 Crystal structure of a nucleosome with highlighted fluorophore labeling sites</td>
<td>37</td>
</tr>
<tr>
<td>4.4 Total internal reflection microscopy</td>
<td>44</td>
</tr>
<tr>
<td>4.5 Characteristic penetration depth of an evanescent field</td>
<td>46</td>
</tr>
<tr>
<td>4.6 The theoretical characteristic length of attenuation as a function of aperture size</td>
<td>49</td>
</tr>
<tr>
<td>4.7 Theoretical intensity maps of the zero mode waveguide at both 532 nm and 638 nm</td>
<td>49</td>
</tr>
<tr>
<td>4.8 Initial results from zero mode waveguide fabrication and experiments</td>
<td>51</td>
</tr>
<tr>
<td>4.9 Single Molecule flow cell assembly</td>
<td>53</td>
</tr>
<tr>
<td>4.10 Example FRET data from single molecule experiments</td>
<td>57</td>
</tr>
<tr>
<td>4.11 Lifetime analysis of idealized trace data</td>
<td>58</td>
</tr>
</tbody>
</table>
4.12 Anisotropy response of 147bp DNA molecules fluorescently labeled with fluorescein with increasing concentration of the viscous liquid glycerol and increasing mobility occlusion by LexA .................................................. 60

5.1 MLA-labeled nucleosomes for nuclear magnetic resonance (NMR) ............... 65
5.2 Tudor interacts with the H3K36me3-nucleosome .................................................. 66
5.3 Tudor interaction with the nucleosome is dependent on methylation at H3K36 68
5.4 Tudor interacts non-specifically with double-stranded DNA (dsDNA) ......... 69
5.5 A model of the Tudor/H3KC36me3-NCP complex ................................. 70
5.6 PHF1 Tudor binding to the trimethylated lysine 36 of histone H3 (H3K-36me3)-nucleosome core particle (NCP) stabilizes a more open form of the nucleosome .................................................. 72
5.7 The larger PHF1 construct containing Tudor and PHD1 further enhances LexA binding .................................................. 75

6.1 The PHD finger protein 1 (PHF1)-N-terminal domain (NTD) is partially structured and does not affect the binding of the adjacent Tudor domain to H3K36me3 .................................................. 85
6.2 PHF1-Tudor NMR chemical shift and structure in complex with H3K36me3 87
6.3 Experimental constructs definition and initial characterization for PHF1 accessibility experiments .................................................. 89
6.4 PHF1 titrations with H3KC36me3 and unmodified nucleosomes ................. 89
6.5 Fluorometer FRET raw data change with increasing LexA concentration . 90
6.6 Changes in accessibility of nucleosomal DNA to LexA with respect to PHF1 binding to H3K36me3 .................................................. 92
6.7 PHF1 does not impact LexA binding to naked DNA .................................. 94
6.8 LexA titrations with saturating levels of PHF1(14–87) and PHF1(28–87) to H3K36me3 and unmodified nucleosomes .................................................. 95
6.9 Four state binding model for a PTM binding protein and a transcription factor 96
6.10 A two dimensional titration of LexA and PHF1 to nucleosomes containing H3K36me3 .................................................. 96
6.11 Single molecule FRET apparatus for use with PHF1 and LexA ................. 98
6.12 LexA binding rates to unmodified and H3K36me3 nucleosomes with or without PHF1(28–87) bound .................................................. 99

7.1 LEDGF does not shift the nucleosome to a largely unwrapped state upon binding .................................................. 108
7.2 H3K36me3 does not impact DNA accessibility unless paired with the histone variant H3.3 .................................................. 110
7.3 The PWWP domain of LEDGF/P75 alone does not impact DNA accessibility within chromatin .................................................. 111
7.4 LEDGF decreases DNA accessibility upon binding and prefers binding H3.3 by roughly 10-fold .................................................. 111

8.1 Initial experiments to confirm Swi6 behaves as expected ................. 118
8.2 Swi6 does not alter nucleosome equilibrium to a largely unwrapped state .. 119
8.3 LexA invasion of nucleosomes containing H3K9me3 and unmodified H3 ... 119
# List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.1 Table of Förster distances</td>
<td>35</td>
</tr>
<tr>
<td>E.1 Parts list for shutter controllers</td>
<td>200</td>
</tr>
<tr>
<td>F.1 Triclops: Major Components</td>
<td>204</td>
</tr>
<tr>
<td>F.2 Triclops: Excitation path components</td>
<td>205</td>
</tr>
<tr>
<td>F.3 Triclops: Emission path components</td>
<td>206</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>A</td>
<td>see Ala</td>
</tr>
<tr>
<td>aa</td>
<td>amino acid</td>
</tr>
<tr>
<td>ac</td>
<td>acetylation</td>
</tr>
<tr>
<td>ADD</td>
<td>ATRX, DNMT3, DNMT3L</td>
</tr>
<tr>
<td>Ala</td>
<td>(also A) Alanine</td>
</tr>
<tr>
<td>Arg</td>
<td>(also R) Arginine</td>
</tr>
<tr>
<td>Asp</td>
<td>(also D) Aspartic Acid</td>
</tr>
<tr>
<td>ATP</td>
<td>adenosine triphosphate</td>
</tr>
<tr>
<td>BAH</td>
<td>bromo-adjacent homology</td>
</tr>
<tr>
<td>bioPEG</td>
<td>biotin difunctional PEG</td>
</tr>
<tr>
<td>BME</td>
<td>β-mercaptoethanol</td>
</tr>
<tr>
<td>bp</td>
<td>basepair</td>
</tr>
<tr>
<td>BSA</td>
<td>bovine serum albumin</td>
</tr>
<tr>
<td>BZA</td>
<td>benzamidine</td>
</tr>
<tr>
<td>CCD</td>
<td>charge-coupled device</td>
</tr>
<tr>
<td>cDNA</td>
<td>complementary DNA</td>
</tr>
<tr>
<td>CHD</td>
<td>chromodomain-helicase-DNA binding</td>
</tr>
<tr>
<td>COT</td>
<td>Cyclooctatetraene</td>
</tr>
<tr>
<td>CSP</td>
<td>chemical shift perturbation</td>
</tr>
<tr>
<td>Cys</td>
<td>(also C) Cysteine</td>
</tr>
<tr>
<td>D</td>
<td>see Asp</td>
</tr>
<tr>
<td>DMF</td>
<td>dimethylformamide</td>
</tr>
<tr>
<td>DMSO</td>
<td>dimethyl sulfoxide</td>
</tr>
<tr>
<td>dsDNA</td>
<td>double-stranded DNA</td>
</tr>
<tr>
<td>DTT</td>
<td>Dithiothreitol</td>
</tr>
<tr>
<td>E</td>
<td>see Glu</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Term</td>
</tr>
<tr>
<td>--------------</td>
<td>------</td>
</tr>
<tr>
<td>MLA</td>
<td>methyl-lysine analog</td>
</tr>
<tr>
<td>Mly</td>
<td>Methionine (also M)</td>
</tr>
<tr>
<td>MPEG</td>
<td>monofunctional PEG</td>
</tr>
<tr>
<td>MWCO</td>
<td>molecular weight cut-off</td>
</tr>
<tr>
<td>MWM</td>
<td>molecular weight marker</td>
</tr>
<tr>
<td>NBA</td>
<td>3-Nitrobenzyl alcohol</td>
</tr>
<tr>
<td>NCP</td>
<td>nucleosome core particle</td>
</tr>
<tr>
<td>NHS</td>
<td>N-hydroxy succinimide</td>
</tr>
<tr>
<td>Ni-NTA</td>
<td>nickel-nitrilotiacetic acid</td>
</tr>
<tr>
<td>NMR</td>
<td>nuclear magnetic resonance</td>
</tr>
<tr>
<td>NPS</td>
<td>nucleosome positioning sequence</td>
</tr>
<tr>
<td>NTD</td>
<td>N-terminal domain</td>
</tr>
<tr>
<td>PALM</td>
<td>photoactivated localization microscopy</td>
</tr>
<tr>
<td>PCR</td>
<td>polymerase chain-reaction</td>
</tr>
<tr>
<td>PDB</td>
<td>protein data bank</td>
</tr>
<tr>
<td>PEG</td>
<td>poly-ethylene glycol</td>
</tr>
<tr>
<td>PEV</td>
<td>position effect variegation</td>
</tr>
<tr>
<td>ph</td>
<td>phosphorylation</td>
</tr>
<tr>
<td>PHD</td>
<td>plant homeodomain</td>
</tr>
<tr>
<td>Phe</td>
<td>Phenylalanine (also F)</td>
</tr>
<tr>
<td>PI</td>
<td>protease inhibitors</td>
</tr>
<tr>
<td>PIFE</td>
<td>protein induced fluorescence enhancement</td>
</tr>
<tr>
<td>PIPES</td>
<td>piperazine-N,N'-bis(2-ethanesulfonic acid)</td>
</tr>
<tr>
<td>PMSF</td>
<td>phenylmethyl-sulfonyl fluoride</td>
</tr>
<tr>
<td>Pol II</td>
<td>RNA Polymerase II</td>
</tr>
<tr>
<td>Polymin P</td>
<td>polyethyleneimine</td>
</tr>
<tr>
<td>PRC2</td>
<td>Polycomb Repressive Complex 2</td>
</tr>
<tr>
<td>Pro (also P)</td>
<td>Proline</td>
</tr>
<tr>
<td>PTM</td>
<td>post translational modification</td>
</tr>
<tr>
<td>PWWP</td>
<td>proline tryptophan tryptophan proline</td>
</tr>
<tr>
<td>R</td>
<td>see Arg</td>
</tr>
<tr>
<td>RITS</td>
<td>RNA-induced transcriptional silencing</td>
</tr>
<tr>
<td>RNA</td>
<td>ribonucleic acid</td>
</tr>
<tr>
<td>S</td>
<td>see Phe</td>
</tr>
<tr>
<td>SD</td>
<td>standard deviation</td>
</tr>
<tr>
<td>SEM</td>
<td>scanning electron microscopy</td>
</tr>
<tr>
<td>SIM</td>
<td>structured illumination microscopy</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>smFRET</td>
<td>single molecule FRET</td>
</tr>
<tr>
<td>smTIRF</td>
<td>single molecule TIRF</td>
</tr>
<tr>
<td>SNR</td>
<td>signal-to-noise ratio</td>
</tr>
<tr>
<td>ssDNA</td>
<td>single-stranded DNA</td>
</tr>
<tr>
<td>STED</td>
<td>stimulated emission depletion</td>
</tr>
<tr>
<td>STORM</td>
<td>stochastic optical reconstruction microscopy</td>
</tr>
<tr>
<td>TBE</td>
<td>tris boric acid EDTA</td>
</tr>
<tr>
<td>TCEP</td>
<td>tris(2-carboxyethyl)phosphine</td>
</tr>
<tr>
<td>TEV</td>
<td>tobacco etch virus</td>
</tr>
<tr>
<td>TF</td>
<td>transcription factor</td>
</tr>
<tr>
<td>TIRF</td>
<td>total internal reflection fluorescence</td>
</tr>
<tr>
<td>Trolox</td>
<td>6-hydroxy-2,5,7,8-tetramethylchroman-2-carboxylic acid</td>
</tr>
<tr>
<td>TROSY</td>
<td>transverse relaxation-optimized spectroscopy</td>
</tr>
<tr>
<td>Trp (also W)</td>
<td>Tryptophan</td>
</tr>
<tr>
<td>Tyr (also Y)</td>
<td>Tyrosine</td>
</tr>
<tr>
<td>UV</td>
<td>ultraviolet</td>
</tr>
<tr>
<td>UV-VIS</td>
<td>spectrophotometry</td>
</tr>
<tr>
<td>V</td>
<td>see Val</td>
</tr>
<tr>
<td>Val (also V)</td>
<td>Valine</td>
</tr>
<tr>
<td>W</td>
<td>see Trp</td>
</tr>
<tr>
<td>WD</td>
<td>tryptophan aspartic acid</td>
</tr>
<tr>
<td>wt</td>
<td>wild-type</td>
</tr>
<tr>
<td>Y</td>
<td>see Tyr</td>
</tr>
<tr>
<td>zf-CW</td>
<td>zinc finger CW</td>
</tr>
<tr>
<td>ZMW</td>
<td>zero-mode waveguide</td>
</tr>
</tbody>
</table>
Chapter 1

INTRODUCTION

All cells must carefully regulate levels of expression for each gene throughout the cell cycle. At differing times certain gene expression levels must be increased or activated while others must be repressed or even fully silenced. In multi-cellular organisms the complexity of this process intensifies; it includes not only cell cycle-based timing but also differentiation and gene regulation by cell-to-cell signaling. Many diseases can trace their origin back to improperly expressed genes. A fundamental understanding of gene regulation will lead to cures to these kinds of genetic diseases and inform courses of action for others. The regulatory mechanisms of genes form a system which controls the access to the library of possibilities represented in the genome. These mechanisms make up the epigenome.

DNA contains all the information necessary to encode all proteins and RNA that make up an organism as well as maintain proper proportion of these myriad biomolecules to maintain the health of the organism. However, DNA does not provide critical explanations for easily observed phenomena such as cell differentiation. Epigenetics is required to store the current state of the cell.

Epigenetic data is any information stored in a cell not contained in the sequence of the DNA within the cell. This information takes numerous forms from the presence of a particular protein in a lipid membrane [1] to specific chemical modifications on DNA and protein amino acid (aa) residues [2]. In this work I focus on PTMs to the protein core of the basic subunit of chromatin, the nucleosome. I investigate the impact of sensing the epigenetic state through
Figure 1.1: DNA in eukaryotic organisms is compacted first into nucleosomes which then self associate to form higher order chromatin structure. Chromatin is broadly divided into the largely inaccessible heterochromatin and the active euchromatin. Image is reproduced from Sha and Boyer [3].

protein-PTM binding events on one of the key functions of the nucleosome—regulation of access to DNA. The mechanisms described may play a key role in regulating gene expression.

1.1 Chromatin and the Nucleosome

Eukaryotes have a size problem. In addition to containing more genes in general than prokaryotic organisms, eukaryotic genomes also contain significant amount of non-coding DNA [4]. This results in a large distribution of genome sizes among eukaryota but on the whole translates to significantly larger genomes when compared to prokaryotes [1, 5]. What’s more this DNA must be packaged into a nucleus which, for humans, is roughly 6 µm in diameter [1]. Geometrically, this is like packing all of the hair a person grows in 13
years into a stability exercise ball—probably possible, but definitely cramped.

This problem is solved through ordered compaction of DNA into chromatin. DNA is first wrapped into nucleosomes and then into higher order structures (see Figure 1.1) culminating in either mitotic chromosomes or interphase chromatin. Nucleosomes alone comprise a ~10-fold compaction of genetic material [6, 7] and each cell contains ~20 million. Background of the nucleosome and chromatin will comprise the rest of the chapter.

1.1.1 Euchromatin and Heterochromatin

Chromatin is divided into two broad categories: euchromatin and heterochromatin. Euchromatin makes up ~94% of the human genome and is the portion of the genome targeted by the human genome project [8]. Euchromatin represents the subset of chromatin which is less densely packaged, is relatively gene-rich, and contains all genes which are actively expressed. In contrast, Heterochromatin is gene-poor and comprised of DNA which is not actively expressed. There are two forms of heterochromatin i) constitutive and ii) facultative. Constitutive heterochromatin is essentially a static presence in cells and is made up of architectural structures such as the cetrometeres and telomeres. It is highly repetitive and was thought to be transcriptionally silent. However, recent studies have shown transcriptional activity which is blocked from further processing by RNA-induced transcriptional silencing (RITS) pathways [9–12]. Facultative heterochromatin represents the section of the genome which has been intentionally silenced and does not display the same repetitive nature of constitutive heterochromatin. The interplay between these forms of chromatin allows for control of DNA expression on a global scale. The precise mechanisms are not well understood, but large-scale chromatin regulation is an active field of study.

This work, however, focuses not on the large scale, but on the small—the nucleosome. Heterochromatin and euchromatin are each made up of nucleosomes, discussed in the next section, and are differentiated by marks which will be discussed in sections 1.1.4 and 1.2.
Figure 1.2: The nucleosome is comprised of ~147 basepair (bp) of DNA wrapped ~1.7 times around a protein octamer core. These octamer proteins are some of the most highly conserved proteins in biology [1]. The crystal structure of octamer from *Xenopus laevis* complexed with 601 DNA described in Lowary and Widom [13] (protein data bank (PDB) ID: 1kx5)

1.1.2 Nucleosome Structure

The nucleosome is the basic repeating subunit of chromatin [1]. It is comprised of a histone octamer protein core made up of two repeats of each of the core histones: H2A, H2B, H3, and H4 further discussed in section 1.1.4. This roughly cylindrical core is wrapped about 1.7 times by between 145-149 bp of dsDNA [14]. Nucleosomes are then repeatedly spaced 10-90 bp apart [14, 15] leading to a ‘beads on a string’ motif (see Figure 1.1). Packaging into nucleosomes represents a ~10-fold compaction of the genome. This compaction limits access to DNA which will be further covered in the next section.
1.1.3 Nucleosome Dynamics

The nucleosome is not a static structure. It is capable of undergoing several different thermally or enzymatically driven processes to create a dynamic environment capable of regulating access to genomic DNA. This regulation is important for cellular processes such as transcription, DNA replication, and DNA repair. As mentioned in section 1.1.2, the wrapping of DNA into nucleosomes blocks access to that DNA. This occurs due to steric occlusion of DNA processing machinery, typically large proteins. These large proteins are, on average, physically blocked from their target sites due to the stable nucleosome equilibrium. However, the dynamic nature of nucleosomes and chromatin allows for the regulation of the level of accessibility of any given part of the genome. These dynamics are mainly controlled through adenosine triphosphate (ATP)-dependent chromatin remodelers and histone PTMs [2, 16]. For example, nucleosomes are often depleted from promoter regions of active genes in order to allow access to these promoter sites [17]. Additionally, chromatin remodeling complexes can slide nucleosomes along the genome to i) regulate inter-nucleosome spacing (important for nucleosomes following a promoter) or ii) expose an otherwise blocked protein-DNA interaction site [18, 19]. Finally, even without displacement, nucleosomes can facilitate their own invasion through thermal unwrapping fluctuations which transiently provide access to normally-occluded DNA [20–25]. It is this last mechanism and its regulation that the bulk of this work investigates.

1.1.4 The Histones

The histones comprising the core of the nucleosome, H2A, H2B, H3, and H4, are each highly basic (positively charged) and mostly globular containing a long N-terminal tail (amino-terminal), a ‘histone fold’ 3 α helix domain, and a relatively short C-terminal (carboxyl-terminal) [2]. These histone fold domains enable formation of heterodimers between H2A and H2B as well as a stable H3 and H4 tetramer. In vivo nucleosomes are formed first by deposition of the H3/H4 tetramer to form a tetrasome and through two assembly reactions
each heterodimer is added to form a full nucleosome [26]. Both the well structured octamer core and the flexible histone tail domains are targets for modifications. These PTMs are known to influence structural dynamics [21, 27] as well as behave as signals to chromatin associated proteins [28]. Amino acid sequence analysis shows high conservation from yeast to humans suggesting that most amino acids are important to octamer structure, function, or as a potential PTM target [2]. As we will discuss later in this section and work, small alterations to aa content can lead to altered chromatin and nucleosome equilibria and dynamics.

While each octamer contains two each of H2A, H2B, H3, and H4, there are a number of variants which have known importance for H2A and H3. H2B is markedly deficient in variants while H4 has no known variants [29]. The reason for the lack of H2B and H4 variants is unclear.

H2A variants, however, are the most numerous and include H2A.Z, H2A.X, MacroH2A, and H2A-Bbd. MacroH2A is specifically localized to inactive X chromosomes in vertebrates [30] while H2A-Bbd is located almost exclusively in vertebrate active X chromosomes [31]. The precise function of these modifications is not fully understood, but H2A.Z, conserved across evolution [29], is known to be expressed throughout the cell cycle [29, 32] and is strongly correlated with the first nucleosome after the nucleosome depleted region following a promoter [21, 33–36]. It has been suggested that H2A.Z is associated with destabilizing nucleosomes [37] but is also known to play a critical role in heterochromatin formation [38].

There are five known H3 variants: H3.1, H3.2, H3.3, H3.4, and centromeric H3 (often called CENP-A [39] in mammals but known by other names in other organisms). H3.1 and H3.2 are both considered canonical H3 and differ by only one amino acid with both having similar chemical properties [29]. These canonical variants are expressed only during S-phase [29, 36]. In contrast, H3.3 is expressed throughout the cell cycle and localizes to active regions of the genome [36, 40]. Notably, H3.3 contains only four mutations relative to H3.2 which are highly conserved and result in mechanistic alterations. For example, evidence indicates that H3.3 acts to destabilize nucleosomes, especially when complexed with heterodimer containing H2A.Z [38, 41–43]. However, others have found no such destabilization but
rather a decrease in compaction on a larger nucleosome array scale which counteracts H2A.Z-mediated compaction [38, 44]. CENP-A conserves the critical histone fold domain but has highly divergent N-terminal tails [29]. Interestingly, CENP-A is both necessary and sufficient to cause centromeric activity to any genomic region regardless of DNA sequence [45–47]. Deposition and control of these histone variants represent the first level of chromatin regulation of DNA accessibility. Research is ongoing to completely elucidate the mechanisms of such regulation and to fully understand the impact of variants on nucleosome structure and dynamics.

1.2 Histone Post Translational Modifications

In addition to incorporating histone variants, the cell can regulate DNA accessibility through the mechanism of applying PTMs. There are currently over 500 known post translational modifications located throughout the nucleosome [2, 48]. The number of possible combinations for an individual nucleosome far exceeds the 20 million or so in each cell. In principle, each nucleosome may be unique. Unfortunately the majority of genome-wide mapping efforts are necessarily ensemble measurements and we can say very little about the exact makeup of a specific nucleosome. Efforts are moving toward these kinds of capabilities [49, 50] and combinations of PTMs are beginning to be understood in vitro [51].

Post translational modifications are chemical modification done on an amino acid level after translation of a protein. Figure 1.3 shows the chemical structure alterations to the nine aas which can be modified. There are a large number of known modifications shown in Figure 1.3, but the most common three are methylation, acetylation, and phosphorylation [52]. Each PTM uniquely changes the chemical properties of the aa to which it is applied. For example, acetylated lysine (Kac) removes the positive charge typically associated with lysine at neutral pHs. In this manner, the behavior of individual histones can be tweaked or individual nucleosomes can be used to recruit proteins to very specific locations through recognition of the unique properties of the PTM in conjunction with the neighboring aas.
Figure 1.3: The amino acids which can be modified after translation and the chemical structure changes for all the known modifications performed on histones. Electric charge changes are especially highlighted but all modifications change the bulk of the aa residue in addition to the potential changes to hydrophobicity and charge. The most common PTMs are acetylation (ac), methylation (me) and phosphorylation (ph). Figure reprinted with permission from Huang et al. [2]. Copyright 2015 American Chemical Society.
These two general mechanisms represent the range of effects a given PTM can elicit. They are i) the direct alteration mechanism and ii) the histone code mechanism. These are discussed in the following sections.

1.2.1 Modifications Alter Nucleosome Dynamics and Equilibrium

The most direct method of action for a PTM is to directly impact chromatin or nucleosome dynamics or equilibria. This mechanism is reviewed in Bowman and Poirier [21]. Briefly, the physical and chemical changes caused by a PTM can alter the interaction energy if placed on an aa which is important to binding. For example, Lys 16 on histone H4 acetylation is known to reduce chromatin compaction [21, 27, 53]. On the nucleosome level lysines and tyrosines which directly interact with DNA also increase DNA accessibility by decreasing the affinity of the histone octamer at that location to DNA. Examples of these include H3K56ac and H3Y41ph.
1.2.2 The Histone Code

In contrast to direct alteration of dynamics and equilibria, the histone code mechanism utilizes third parties in order to alter chromatin dynamics. This mechanism relies on a party of outside proteins which deposit, recognize, and remove histone PTMs to dynamically alter the chromatin landscape (called ‘Writers’, ‘Readers’, and ‘Erasers’, respectively). The PTMs themselves behave by either i) increasing the affinity of proteins which recognize them or ii) by inhibiting the binding of a protein to chromatin. These interactions then alter the behavior of the cell [54]. In addition, the interactions need not be first order. There are examples of complex interaction pathways in which recognition of one mark impacts the state of another. For example, the protein HP1 is known to recruit numerous PTM depositing proteins and has even been seen to recruit RNA polymerase complexes [9] to genomic regions containing H3K9me3.

The presence of different kinds of PTMs typically correspond to different behaviors in chromatin. For example, acetylation typically correlates with transcriptional activation while deacetylation is correlated with repression [55–57]. Failure of each individual step—writing, reading, or erasing—has been associated with disease, such as cancer [58–60].

Writers, Readers, and Erasers

The Histone Code depends on the site-specific deposition, recognition, and removal of PTMs from specific nucleosomes. This is possible due to the high specificity shown by protein domains which recognize PTMs. For example, Tudor; plant homeodomain (PHD); malignant brain tumor (MBT); proline tryptophan tryptophan proline (PWWP); tryptophan aspartic acid (WD); ATRX, DNMT3, DNMT3L (ADD); zinc finger CW (zf-CW); bromo-adjacent homology (BAH); and chromodomain-helicase-DNA binding (CHD) are all known to specifically recognize methylated lysine residues [2]. Each individual example of these domains, however, is unique and may be specifically evolved to interact with a particular methylated lysine residue on the nucleosome. This kind of specificity is common across writers, readers,
and erasers. The end result is the ability for the same PTM to correlate, depending on the aa it is deposited on, to regions of transcriptional repression or activation as in the example of lysine methylation. However, location dependence is general across all types of PTMs [2, 28].

1.3 Summary

The regulation of DNA accessibility within the nucleosome and chromatin is a complex field governed by many interweaving factors. These factors include the specific makeup of the chromatin itself, but also the environment of the cell itself with regards to the abundance of chromatin associated proteins or chromatin remodelers. In this work, I investigate the extent to which the two distinct mechanisms of PTM regulation—direct regulation and the histone code mechanism—overlap. I combine techniques to measure the accessibility of DNA with the chemical techniques required to generate nucleosomes which are homogeneously modified in order to determine the impact of PTM readers on DNA accessibility in a purified environment.

I will begin by introducing the necessary theoretical and experimental background in chapters 2, 3 and 4 and move on the descriptions of the experiments themselves in chapters 5, 6, 7 and 8. Finally, I present future directions in chapter 9.
Chapter 2  
Ligand Binding

The work presented here relies heavily on a few key theoretical points which will be discussed in detail in the following chapter. The first will be describing the simplest case of a ligand associating with a receptor. We will then expand this to describe cooperativity and allostery: regulatory mechanisms used by biological organisms to fine-tune response to stimuli. These mechanisms change an otherwise meandering concentration dependence to a binary dose dependence—critical in many systems such as hemoglobin oxygen uptake; hemoglobin must fully bind oxygen in the lungs where concentration is high and fully release it when in regions of the body with only marginally lower oxygen concentrations. Finally, we will develop the understanding of looping and branching interactions which further expand the set of response functions available to the cell. These theories will be of critical importance in chapters 5, 6, 7 and 8.

The final section of this chapter covers previous work related to accessing DNA which is occluded by the nucleosome. Extensive research has been conducted into this protection [24, 25, 61] and has been expanded into the regulation of nucleosome occlusion by PTM deposition [21, 23, 51].
2.1 Introduction

The natural language of biomolecular interactions is one of statistical mechanics and thermodynamics. Of principle interest is an understanding of the relative probability of states as well as the rates of transition between these states. An example free energy landscape as a function of a general reaction coordinate is shown in Figure 2.1. The relative probability of the two states is given by the ratio of Boltzmann weights of each state:

\[
\frac{P(A)}{P(B)} = \frac{e^{\frac{G_A}{k_BT}}}{e^{\frac{G_B}{k_BT}}} = e^{-\frac{\Delta G_{AB}}{k_BT}}. \tag{2.1}
\]

Here, the only variables are the change in free energy between states, \(\Delta G_{AB}\), and the measure of thermal energy in the system, \(k_B T\), where \(T\) is the temperature. For biological systems temperature does not (usually) change much beyond \(\sim 300 \pm 10 K\), meaning that the thermal energy is approximately constant and given by \(k_B T \approx 4.11 \text{ pN nm} \approx 25 \text{ meV}\). It is common to discuss energies involved in interactions in terms of multiples of \(k_B T\) which is a measure of the relative probability of different states. For example, in the reaction shown in Figure 2.1, the energy difference between the states is \(\Delta G_{AB} = -2k_B T\) yielding a relative probability of \(P(A)/P(B) = e^2 \approx 7.4\)—there are roughly 7 times the number of A molecules than B molecules.

If we consider the more complicated reaction

\[ A + B \rightleftharpoons AB \tag{2.2} \]

where two different molecules bind to form a single molecular complex, we must now consider the free energy gained due to pulling a \(B\) molecule out of solution to bind to \(A\). Equation 2.1 now becomes the probability that \(A\) is bound to \(B\):

\[
\frac{P(AB)}{P(A)} = \frac{c}{c_0} e^{-\frac{(G_{\text{bind}} - \mu)}{k_BT}}. \tag{2.3}
\]

\(G_{\text{bind}}\) is the free energy lost by binding the two molecules together, \(\mu\) and the prefactor \(c/c_0\) is a measure of the chemical potential energy gained by \(B\) due to its association with \(A\). \(c\) is
the concentration of $B$ and $c_0$ a characteristic concentration for a dilute solution—typically ~1 M.

### 2.1.1 Connections to Thermodynamics

We typically think in terms of concentrations, but the ratio of molecules is equal to the concentration ratio of species. For a simple state change such as that shown Figure 2.1, the ratio of the concentration of these two molecules (given by Equation 2.1) is called the equilibrium constant, $K_{eq} = [A]/[B]$ (square brackets indicate a concentration). For a chemical reaction of the form of Equation 2.2 the ratio of dissociated molecules over the associated molecules gives a dissociation constant,

$$
K_d \equiv \frac{1}{c_0} e^{-\frac{(G_{bind} - \mu)}{k_B T}} = \frac{[A][B]}{[AB]}.
$$

These thermodynamic quantities ($K_{eq}$ and $K_d$) are bulk measures of the underlying statistical mechanical quantities given by Equations 2.1 and 2.3 which form the necessary foundation of the theoretical background to be covered in the rest of the chapter.

### 2.2 Binding of a Ligand to a Receptor

The binding of ligands is often used as a signal by the cell to alter its state. For example, a ligand binding to a ligand-gated ion channel can shift the channel’s equilibrium to favor an
open state [62]. However, while the affinity of a given binding site (and thus, the concentration at which binding occurs) can be tuned, the form of concentration dependence in the binding probability to a single, independent site cannot (see Equation 2.2.1).

For this work, it will be important to consider the association of PTM binding proteins as well as protein-DNA interactions either directly targeting histones or targeting DNA that is sterically occluded by the nucleosome. These kinds of interactions are well known to be crucial to cellular function including DNA repair, transcription, silencing, and nuclear organization [9, 21, 22, 63–66].

2.2.1 The Hill Equation

The shape of the concentration dependence of binding is determined entirely by the stoichiometry of the reaction. Following Phillips et al. [4], the reaction

$$L + R \rightleftharpoons LR$$

(2.4)

describes the binding of a generalized ligand ($L$) to a general receptor ($R$) forming the complex $LR$. The fraction of receptors bound at any time is trivially given by the number of receptors in the $LR$ state divided by the total number of receptors

$$\Theta = \frac{N_{LR}}{N_R + N_{LR}} = \frac{[LR]}{[R] + [LR]}.$$

Substituting the dissociation constant of this reaction, $K_d = c[R]/[LR]$ (where $c$ is the free ligand concentration to avoid notational overlap in later sections), Equation 2.5 yields

$$\Theta = \frac{1}{1 + K_d/c},$$

(2.5)

a non-cooperative binding isotherm. Notice that as long as $c$ does not change significantly as ligands bind receptors, $K_d$ is the concentration at which half of all receptors are bound. As previously alluded, the fraction of bound receptor is given by Equation 2.5 as long as each ligand-receptor binding is an independent event. However, the transition range of $\Theta$ is roughly two to three orders of magnitude, where the receptors are transitioning from a free
state to a bound state (see Figure 2.2). If a biological organism relied on independent events to determine all behavior, huge concentration swings would be necessary to significantly change receptor occupation—a potential problem if a response is either inherently binary or a process necessitates sensitivity to ligand concentration.

2.2.2 Ligand Binding in the Presence of Co-Factors

Cooperativity and allostery are biology’s answer to the need for a more digital response. Cooperative interactions are ones in which the binding of one ligand to its receptor alters the energies involved in binding another of the same ligand. Allostery likewise is the regulation of energies by different ligands. These changes are achieved by stabilization of alternative conformations which either encourage or discourage binding events.

The simplest way to include cooperativity in the process given by 2.4 is to simply require the receptor to simultaneously bind two ligands:

\[ L + L + R \rightleftharpoons LRL. \]

The dissociation constant is \( K_d^2 = c^2[R]/[LRL] \), allowing us to write down the probability of the doubly bound state as

\[ \Theta_2 = \frac{[LRL]}{[R] + [LRL]} = \frac{1}{1 + (K_d/c)^2}. \]  

(2.6)
2.6 can be safely generalized for \( n \) ligands to

\[
\Theta_n = \frac{1}{1 + \left( \frac{K_d}{c} \right)^n};
\]

in practice, the parameter \( n \), is called the Hill Coefficient [67]. This model, first introduced by AV Hill in 1910 [67], limits binding to an all-or-nothing event which, upon reflection, is far too limiting to remain physical. The model still correctly defines the general shape of binding and so is often used to determine the level of cooperativity in a reaction by fitting the Hill Coefficient. In real use, however, the coefficient is never a whole number, a symptom of the inadequacies of the model.

Unlike Hill cooperativity, intermediate states are inherent to the Monod-Wyman-Changeux (MWC) model [62, 68, 69]. The model was laid down in two papers which helped coin the term allostery [70, 71]. Monod, Wyman, and Changeux [71] predict the binding of a ligand to one receptor in a multi-receptor complex stabilizes an alternate quaternary structure with a different ligand binding affinity which produces the cooperative or allosteric effect. This stabilization is often referred to as a shift in the equilibrium as the equilibrium ratio of one state to the other changes with ligand concentration. Central to the theory is that this shift is not due to changes in the energetics of the quaternary structure or by access to an otherwise disallowed state induced by ligand binding. The shift is truly due to a stabilization of an otherwise energetically unfavorable state due to ligand binding.

The MWC model depends on the existence of two structural states of the protein in question. Originally called the \( T \) (tense) and \( R \) (relaxed) states, these states, when in equilibrium, have a constant ratio of concentrations, \( L = \frac{[T]}{[R]} \), determined by the energy difference between the states. Cooperative behavior occurs when both states are capable of binding a ligand, but do so with different affinities.

For the single site MWC molecule, which is most relevant for the rest of this work, the protein can be in either the \( T \) or \( R \) state, and the binding site can be either occupied or unoccupied. The energy of any given state can be broken down into the energy of the protein’s conformation (defined as \( \epsilon_T \) and \( \epsilon_R \) for tense and relaxed states, respectively) and
the energy required to bind ligand to receptor. This ligand binding itself is a function of both the binding energy, $\epsilon_b^{(T)}$ (tense) or $\epsilon_b^{(R)}$ (relaxed), and of the chemical potential energy required to remove the ligand from solution, $\mu$. For a dilute solution, $\mu = \mu_0 + k_B T \ln \frac{c}{c_0}$, where $\mu_0$ is a characteristic chemical potential and $c_0$ is a characteristic concentration, usually 1M for protein in water.

These energies lead to the grand partition function:

$$Z = e^{-\beta \epsilon_T} + e^{-\beta (\epsilon_R^{(R)} - \mu)} + e^{-\beta (\epsilon_T^{(T)} - \mu)}$$

$$Z = e^{-\beta \epsilon_R} \left( 1 + \frac{c}{c_0} e^{-\beta (\epsilon_b^{(R)} - \mu_0)} \right) + e^{-\beta \epsilon_T} \left( 1 + \frac{c}{c_0} e^{-\beta (\epsilon_b^{(T)} - \mu_0)} \right)$$

$$Z = (1 + c/K_R) + L \left( 1 + c/K_T \right), \quad (2.8)$$

where $\beta^{-1} \equiv k_B T$, as usual, and the definitions of the $K_T$ and $K_R$ (the dissociation constants for the $T$ and $R$ state, respectively) are given by the law of mass action (analogous to $K_d$ in Equation 2.5). We have also introduced the previously defined conformational equilibrium constant $[71] L \equiv e^{-\beta (\epsilon_T - \epsilon_R)}$. Equation 2.8 is the thermodynamic expression of the statistical mechanics partition function in the previous equation.

The probability of the $R$ state is an enumeration of the Boltzmann factors of all microstates in which the protein is in the $R$ state ($R$ state, bound and unbound ligand) divided by the partition function, Equation 2.8.

$$P_R = \frac{e^{-\beta \epsilon_R} \left( 1 + \frac{c}{c_0} e^{-\beta (\epsilon_b^{(R)} - \mu_0)} \right)}{e^{-\beta \epsilon_T} \left( 1 + \frac{c}{c_0} e^{-\beta (\epsilon_b^{(T)} - \mu_0)} \right) + e^{-\beta \epsilon_R} \left( 1 + \frac{c}{c_0} e^{-\beta (\epsilon_b^{(R)} - \mu_0)} \right)}$$

$$P_R = \frac{(1 + c/K_R)}{(1 + c/K_R) + L \left( 1 + c/K_T \right)}, \quad (2.9)$$

This obfuscates the underlying energetics in favor of thermodynamics and constants traditionally used in chemistry. It should be noted that $L$, $K_R$, and $K_T$ all depend on temperature, though the range of temperatures relevant to most biological systems is very small. At equilibrium, $L$ is the ratio $[T]/[R]$ and the dissociation constants, $K_T$ and $K_R$, are the ligand
Figure 2.3: Most transcription factor (TF) binding sites reside within nucleosomes. (a) The distribution of TF binding site distances to the nearest nucleosome dyad (black circles). The black line represents the expected distribution of TF binding sites if they were randomly distributed throughout the genome. The dyad and entry/exit region are defined as 0-36 bases and 37-74 bases, respectively. (b) Percentage of measured vs. randomly distributed TF binding sites based off the regions described in a. Figure reproduced with permission from North et al. [66].

half occupancy concentrations for a protein trapped in the the T and R states, respectively.

We can also determine the macroscopic fraction of receptors bound to ligand, identical to the average number of particles bound to any one receptor:

\[
\Theta_1 = \langle N_{\text{bound}} \rangle = \frac{1}{\beta} \frac{\partial \log Z}{\partial \mu} = c \frac{\partial \log Z}{\partial c} = \frac{c/K_R + Lc/K_T}{(1 + c/K_R) + L(1 + c/K_T)}
\]

(2.10)

where the subscript 1 on \( \Theta_1 \) denotes a one site molecule.

Equation 2.9 already has built into it the indirect regulation of the T and R states. Increasing the concentration of ligand allows the unfavorable R state to compensate for its higher energy by binding ligand and eventually this species dominates.

2.3 Transcription Factor Binding Within the Nucleosome

The majority of the following work quantifies regulation of the affinity of DNA-binding proteins to their recognition sequence which has been buried within a nucleosome. Thermal fluctuations allow for nucleosome ‘breathing,’ the transient exposure of the first bps of the entry-exit region of the nucleosome. The first \~20 bp of DNA are relatively loosely bound
Figure 2.4: Three state model of a DNA binding protein (orange) trapping a thermally unwrapped nucleosome in an open conformation. The DNA target site (orange) is accessible to the protein despite its typical occlusion by the histone octamer (HO) (yellow).

before major H2A contact points 28-30 bp into the nucleosome [72, 73]. These bases are regularly solvent exposed and allow for site-specific binding despite their location within a nucleosome [20, 24, 74–76].

In vivo the regulation of this transient site exposure has been proven to be critical to cellular function [18, 19, 66, 77, 78]. Mapping of TF sites to nucleosome positioning data reveals that ~60% of TF sites are buried within nucleosomes, 31% within the dyad, and 33% within the entry-exit regions (see Figure 2.3)[66]. The TF sites located within the entry-exit region are only transiently accessible and are poised for regulation through nucleosome breathing.

2.3.1 3-State Model of Transcription Factor Invasion Into the Nucleosome

The in vivo motivation presented in Figure 2.3 can be applied to in vitro measurements as well [22, 24, 25, 61, 66] which have largely focused on a branching, three state model to represent the physical process of DNA site exposure and subsequent recognition (see Figure 2.4). Thermal fluctuations allow for transient, short-lived states of accessible DNA which high-affinity DNA-binding proteins bind; this traps the nucleosome in an unwrapped state for an extended period. This model has been confirmed through restriction enzyme digests, TF binding studies, and even direct observations [24, 53, 79].

The arguments for writing the quantitative model is similar to that used in Equation 2.2.1. We begin by writing the equilibrium and dissociation constant of the reactions (using the state notation of Figure 2.4) and using \([L]\) to represent the concentration of DNA binding
protein (typically LexA, in this thesis).

\[ K_{eq} = \frac{[II]}{[I]} \]

\[ K_{D}^{Nuc} = \frac{[L][II]}{[III]} \]

We can now write the fraction of nucleosomes that are in the closed state as a function of the concentrations of each state and finally in terms of the DNA binding protein concentration, \([L]\), \(K_{eq}\), and \(K_{D}^{Nuc}\):

\[ \Theta = \frac{[I]}{[I] + [II] + [III]} \]

\[ \Theta = \frac{[I]}{[I] + K_{eq}[I] + K_{eq}[I][L]^{Nuc}} \]

\[ \Theta = \frac{1}{1 + K_{eq} + \frac{K_{eq}[L]}{K_{D}^{Nuc}}} \]  \hspace{1cm} (2.11)

Solving for the concentration of protein that will bind half of the nucleosomes, \(S_{1/2}\), we find that

\[ S_{1/2} = K_{D}^{Nuc} \left( \frac{1}{K_{eq}} - 1 \right) \quad \text{and} \quad K_{eq} \ll 1 \]

\[ S_{1/2} \approx \frac{K_{D}^{Nuc}}{K_{eq}} \]  \hspace{1cm} (2.12)

The dissociation constant of a DNA binding protein is apparently linearly proportional to the fraction of time that it is available to be bound. Of course, this is assuming the affinity of that protein is unperturbed by proximity to the nucleosome. This assumption was shown to be radically untrue by Luo et al. [22] and we will expand upon this in chapter 6.

Breathing fluctuations exist with unwrapping dwell times on the order of 100 ms and re-wrapping times of \(-1 \text{ ms}\) [53, 75, 80, 81]. These transition rates are too fast to observe easily [81]. However, the development of this three state model lays the foundation to measure DNA access by quantifying the concentrations at which DNA-binding proteins can access their recognition sequences. This represents a powerful tool in determining the impact of epigenetic regulation of DNA accessibility and therefore on the ability of DNA processing
machinery such as RNA polymerases or DNA repair proteins to transcribe or repair DNA. These techniques, therefore, are the primary means I use to determine the effect of PTM binding protein recognition of their target on regulation of DNA access.

2.4 Looping Interactions

The final model that must be discussed is one of the form

\[ A + B + C \rightleftharpoons AB + C \]
\[ \downarrow \quad \downarrow \]
\[ AC + B \rightleftharpoons ABC \]

These kinds of reactions are known as looping reactions and have some interesting properties. If we consider first the components, \( A; B; \) and \( C \), and the final product, \( ABC \), then we will see that there are two possible paths which lead from reactants to products

\[ A + B + C \rightleftharpoons AB + C \]
\[ \downarrow \quad \downarrow \]
\[ AC + B \rightleftharpoons ABC. \]

However, because energy is a function of the state, the total change in energy between these pathways must be identical.

\[ \Delta G_{A,B,C \rightarrow AB,C} + \Delta G_{AB,C \rightarrow ABC} = \Delta G_{A,B,C \rightarrow AC,B} + \Delta G_{AC,B \rightarrow ABC} \]

or equivalently:

\[ \Delta \Delta G^{AC,B,B \rightarrow ABC}_{A,B,C \rightarrow AB,C} = \Delta \Delta G^{AB,C \rightarrow ABC}_{A,B,C \rightarrow AC,B}. \] (2.14)

Information about the change in affinity in \( B \) due to \( C \) binding gives you information about the change in affinity in \( C \) due to \( B \) binding. This is entirely due to the physical necessity for diverging pathways to yield consistent results and is analogous to the famous Kirchhoff loop rules in basic circuit analysis.
Also analogous to circuits, one can utilize the conservation of particles to determine fluxes through each pathway. If the complexes $A$ and $AC$ have a different affinity for $B$, then, with limiting levels of $C$, the two pathways have fundamentally different responses with respect to the concentration of $B$. The top pathway, called the induced fit pathway, is characterized by first binding of the ligand and substrate. This interaction then induces a change in the substrate, in this example the binding of ligand $C$. The new induced state has an increased affinity of $A$ to $B$. The bottom pathway, termed the conformational selection pathway is similar to that discussed in section 2.2.2. This pathway allows for at least a small fraction of $A$ to always exist as $AC$ and the binding of $B$ ‘selects’ the more energetically favorable conformation which in turn converts $A$ molecules to the $AC$ state in order to maintain equilibrium for that part of the reaction. The difference between the two mechanisms is apparent when considering the effective rate ($k_{\text{eff}}$) of generation of the $ABC$ complex [82]:

\[
\text{Induced Fit:} \quad k_{\text{eff}} = k_r^C + \frac{k_f^C}{1 + \frac{K_d}{[B]}}
\]  

\[
\text{Conformational Selection:} \quad k_{\text{eff}} = k_f^C + \frac{k_r^C}{1 + \frac{[B]}{K_d}}
\]

where $k_f^C$ and $k_r^C$ are the forward and reverse rates of the left reaction in Equation 2.13 which converts between $A$ and $AC$. $K_d$ is the dissociation constant for ligand binding to the appropriate complex ($A$ for induced fit and $AC$ for conformational selection). The differences in these models and the behavior of given systems can be measured if one can track all of the rates and number of particles in each state [82, 83].
Chapter 3

PREPARATION OF NUCLEOSOMES WITH MULTIPLE LABELS

The majority of the experiments performed in this work required the preparation of homogenous nucleosome samples containing a total of three chemically ligated labels to the histone proteins or DNA. It is a non-trivial task to ensure high labeling efficiency while ensuring minimal off-target effects. In this chapter, I describe the methodology behind generating the triply-labeled nucleosomes used for PTM-binding, Förster resonance energy transfer (FRET)-based assays. Detailed protocols for each labeling reaction can be found in Appendixes C and D.

3.1 Design

To be sensitive to unwrapping fluctuations and subsequent TF binding, we need a method to measure the distance between the free end of DNA with the TF site nearby and the HO. FRET provides a nanometer scale ‘ruler’ which can perform such a measurement. Specifics on FRET will be covered in chapter 4 but for now it is enough to know that the measurement relies on the distance between two fluorophore labels. Thus, the measurement will require two specific labels: i) on the DNA near a TF site, and ii) somewhere on the HO an appropriate distance away to ensure a large change in signal during an unwrapping event. The model for these types of experiments is shown in Figure 3.1. Reasonable locations for
Figure 3.1: Figure 2.4 with the intended fluorophore placements in each state. The unwrapping fluctuation which exposes the TF site simultaneously increases in the average distance between FRET pairs. This causes a decrease in the efficiency of energy transfer between the fluorophores which can be detected and used to infer TF occupancy.

Figure 3.2: The solved crystal structure of the nucleosome allows for intelligent placement of fluorophore labels to ensure a high initial FRET efficiency that will decrease upon thermal unwrapping fluctuation and subsequent TF binding. The Cy3 Cy5 fluorophore pair has a theoretical Förster radius of 54 Å. The distance between fluorophores as predicted by the crystal structure is 16.6 Å. PDB ID: 1KX5.

these labels are i) at the 5’end of the DNA and ii) at H2AK119. This aa residue is a part of the H2A C-terminal tail and according to the crystal structure (see Figure 3.2 red label) is located near the entry-exit region close enough to the DNA to produce a strong FRET signal (see Figure 3.2 inset). Labeling the H2A has the added benefit of creating a fiduciary marker to ensure complete nucleosomes. As mentioned in chapter 1, the H2A/H2B heterodimer is more loosely bound to the nucleosome complex so if disassembly occurs, it is reasonable to think that heterodimer loss would be a good indication of this.

The final piece of the puzzle for these experiments is deposition of a PTM or PTM-mimic.
Due to the difficulty in obtaining homogeneously modified protein samples, a more robust (and cheaper) solution is the deposition of a mimic. Simon et al. [84] detail an easily installed and accurate mimic of methylated lysines which can be site-specifically installed and robustly mimics the behavior of mono-, di-, or tri-methylated lysine residues termed a methyl-lysine analog (MLA).

Labeling protocols for the FRET and MLA labels are detailed in Appendices D and C, respectively.

3.2 Order of Labeling

Each labeling reaction described in the previous section requires the addition or mutation of a reactive group on the molecule being labeled. DNA can be safely labeled while isolated from the protein molecules because full HO and DNA molecules can be reconstituted to form nucleosomes. However, the labels which must be attached to the histones themselves pose a problem. Each method of labeling (maleimide and nucleophilic addition for fluorophore and MLA, respectively) requires the mutation of a Cys at the desired location of the label. For full octamer whichever labeling reaction was performed first would also label the site of the second. To get around this issue it is necessary to first label the MLA in unfolding conditions then separately refold H3/H4 tetramer and H2A/H2B dimer. The dimer is then fluorophore labeled through a maleimide reaction and the labeled tetramer and dimer are combined and purified yielding pure, doubly-labeled HO. This HO is then reconstituted with fluorophore labeled DNA and purified yielding homogeneous nucleosome samples for subsequent experimentation.

3.3 Verification of Product

Each labeling step requires purification and verification of labeling. In the case of DNA labeling this is again simple. The label itself can be selected for during a reverse-phase high pressure liquid chromatography system (HPLC) purification, ensuring 100% labeling
Figure 3.3: MALDI-TOF mass spectrometry showing the rough efficiency of an MLA deposition on H3 lysine 36 (mutated to a cysteine). The result is semi-quantitative because protein desorption is not uniform across all species of molecule and is sensitive to relatively small changes in molecular composition. For this histone, H3.3K36C, the unlabeled +2 peak (meaning the protein’s total charge is $2e^+$) has a mass/charge ratio of 7570 Da/e$^+$ and the labeled peak, a ratio of 7613.5 Da/e$^+$. 

The labeling efficiency of the other fluorophore label can likewise be evaluated by ultraviolet-visible spectrophotometry (UV-VIS spectrophotometry). The only real challenge in terms of verification is presented by the MLA label. Verification of labeling efficiency is performed via semi-quantitative matrix-assisted laser desorption/ionization time of flight (MALDI-TOF) mass spectrometry as shown in Figure 3.3.
Chapter 4

**Fluorescence Methods**

FRET is the non-radiative, physical transfer of one chromophore’s (the ‘donor’) excited state energy to a neighboring ‘acceptor’ chromophore which has a non-zero probability of sharing a dipole resonance frequency. This idea will be expanded upon at the end of section 4.1.2. The interaction is highly distance dependent and happens to occur, for visible light, over a range important to biology and molecular interactions, on the order of 5 nm. This is critical because FRET has become an extremely important tool in biophysics to prove co-localization between two molecules or to measure the distance between two points on a macromolecule. The sensitivity and stability of chromophores also enables FRET to become an important tool in single molecule studies where individual molecules are tracked in real time to observe the states explored by a system.

### 4.1 Förster Resonant Energy Transfer

FRET is typically quantified in terms of the FRET efficiency. This efficiency is the fraction of energy absorbed by the donor which is transferred through FRET to the acceptor:

\[
E = \frac{k_T}{k_o + k_T}
\]

(4.1)

where \(E\) is the FRET efficiency, \(k_T\) is the rate of energy transfer, and \(k_o\) is the total rate of de-excitation of the donor molecule through pathways other than FRET (e.g. radiation,
triplet-state transition, photo-destruction, etc.). The FRET efficiency is used as a readout of the state of the interaction as well as act as a nanoscale ‘ruler.’ At the single molecule level, advancements in chromophore brightness, photostability, and lifetime have enabled measurement of a large range of interaction timescales [53, 85–89].

Below I give an overview of the development of FRET theory. I will then touch on the more important theoretical points involved in resonant energy transfer.

4.1.1 A Brief History

Though practical applications of FRET began after 1950 [90], the first observations occurred only a few decades after the concepts of EM fields were solidified [90], and initial attempts at a theoretical explanation hinge upon classical EM theory. However, these theories were being developed concurrently with those of Heisenberg, Dirac, and Schrödinger [91–93] and so later attempts utilize the full power of quantum mechanics to determine quantitative models for transition rates and time-averaged energy transfer efficiencies.

Solution of the Photosynthesis Problem

Earlier experiments and theories exist but we will begin with the experiments that led to the complete theory of FRET. The crux of the issue had to do with the efficiency of photosynthesis at various wavelengths. Plants undergoing photosynthesis absorb light through specialized pigments. The absorbed energy is then used to reduce molecules in reaction centers. The most commonly known pigment, chlorophyll, is directly bound to reaction centers and so can use direct contact to transfer energy absorbed from a photon to the reactive center. Some plants such as Chroococcus additionally absorb light through accessory pigments. The pigment, phycocyanin, is water-soluble and unbound to reactive centers. A problem arises when the photosynthetic efficiency of light absorbed by phycocyanin is measured. Emerson and Lewis [94] found that roughly 90% of light energy absorbed by phycocyanin was utilized in photosynthesis with 1–2% fluorescent emission from the molecule. This is
roughly equivalent to the efficiency of chlorophyll, in direct communication with the reactive centers. A freely floating molecule could never produce this conversion efficiency by random collision: the reaction centers are too dilute and spatially restricted and the fluorescence lifetime of the accessory pigment is far too short to account for the transfer [90]. An explanation relaying on radiation and reabsorption by chlorophyll fails as well where the efficiency of such a transfer roughly given by

\[ E = F \sigma n R. \]

Here \( F \) is the fluorescence efficiency of phycocyanin, \( \sigma \) is the cross-section per molecule of chlorophyll, \( n \) is the concentration of chlorophyll in the cell, and \( R \) the dimensions of the cell. Estimation yields \( -10^{-3} \) [95]. We are left with the explanation of a non-radiative energy transfer referred to by Arnold and Oppenheimer [95] as ‘internal conversion.’ The word choice here is suggestive.

Indeed, FRET has a direct analog in the form of internal conversion of gamma rays [90]. An excited atomic nucleus can undergo spontaneous emission of a gamma ray. This emission shifts the nucleus from one energy level to another. However, this multipole transition can also occur by transferring energy directly to one of the orbital electrons through a non-radiative process. The highly energetic electron is then ejected from the atom. The rate of gamma emission to electron emission is defined as the internal conversion coefficient. This insight allowed Arnold and Oppenheimer [95] to solve the problem of photosynthetic efficiency with an energy transfer rate from phycocyanin to chlorophyll of

\[ k_T(r) = \frac{\sigma \tilde{a}^2 C_1^3}{4\pi hv^3 r^6}. \]

It is important to note that Arnold and Oppenheimer’s [95] treatment of FRET is fully compatible with Försters later theories. In fact, they can be easily converted between each other. The difference is that Förster worked to generalize the theory and to make it accessible rather than to solve a single problem, and so took additional steps to ensure generality [90].

While Arnold and Oppenheimer solved the problem of high photosynthetic efficiency from accessory pigments earlier than did Förster, the latter still became interested in energy trans-
fer because of the efficiency of photosynthesis [90]. Förster began with classical derivations of his energy transfer theory [96] and later a quantum mechanical one which allowed for a fully quantitative theory of FRET [97].

4.1.2 A Classical Derivation

The full derivation of FRET necessarily involves a quantum mechanical treatment, but the insight gained from such an endeavor is not necessary for the purposes used here. Therefore, I will present a derivation which determines the rate of energy transfer from one fluorescent molecule to another, but does not yield any insight into special cases such as dependence on selection rules. For my work this is hardly a loss. Where appropriate, I will inject insight derived from Förster’s quantum mechanical derivation. Instead of following Förster [97] we will repeat the later classical derivation by Kuhn [98, 99].

With the ultimate goal in mind, Kuhn first describes the power radiated by the donor molecule [98]:

$$L = \frac{|\mu_D|^2 \omega^4 n}{3c^3}$$

(4.2)

He then considers the power absorbed by a nearby acceptor, $L_A(r)$, when a distance $r$ from the donor and relate these powers to other properties of the fluorophores:

$$\frac{q}{q_A} = \frac{L_A(r)}{L}, \quad E = 1 - \frac{q_A}{q} = \frac{1}{1 + L/L_A(r)}, \quad \frac{q_A}{q} = \frac{k_T}{k_e}$$

(4.3)

Here $q$ is the quantum efficiency of the donor without an acceptor, $q_A$ is the quantum yield with the acceptor nearby, $E$ is the efficiency of energy transfer from the donor to the acceptor, $k_T$ is the rate of energy transfer, and $k_e$ is the rate of fluorescence in the absence of the acceptor.

Naturally the power absorbed by the acceptor molecule, $L_A(r)$, will depend on the intensity of the electric field produced by the donor. Further, because fluorescence is due to an induced dipole moment during excitation, we begin by investigating the field of an oscillating dipole,
Figure 4.1: The coordinate system definitions for dipole-dipole interactions. \( \theta_D \) is the angle between the donor’s dipole moment, \( \vec{\mu}_D \), and the vector, \( \vec{r} \), connecting the donor moment to the acceptor dipole moment. The coordinates \( \vec{r}, \hat{x}_A, \) and \( \hat{y}_A \) are centered at the location of the acceptor dipole. \( \hat{r} \) is a unit vector in the direction of \( \vec{r} \), \( \hat{y}_A \) lies in the plane of \( \vec{\mu}_D \) and is normal to \( \hat{r} \). Finally, \( \hat{x}_A \) is normal to both \( \hat{r} \) and \( \hat{y}_A \). The angle \( \phi_A \) is the angle between \( \hat{y}_A \) and the projection of \( \vec{\mu}_A \) on the \( \hat{y}_A/\hat{r} \) plane and \( \theta_A \) is the angle between \( \hat{x}_A \) and \( \vec{\mu}_A \).

here, conceptually, the donor molecule [100, 101]

\[
E_\theta = \frac{\mu_0 \sin \theta}{n^2} \left( \frac{1}{r^3} - \frac{i k}{r^2} - \frac{k^2}{r} \right) e^{i \omega (t - \frac{nr}{c})} \]
\[
E_r = \frac{2 \mu_0 \cos \theta}{n^2} \left( \frac{1}{r^3} - \frac{i k}{r^2} \right) e^{i \omega (t - \frac{nr}{c})} \]

(4.4)

where the wave number, \( k = \frac{2 \pi}{\lambda} \), \( \lambda \) is wavelength, and \( \theta \) is the angle between the dipole axis and the vector connecting the donor to the acceptor fluorophores (see Figure 4.1). Famously, in the far field the radiative \( \frac{1}{r} \) term dominates. FRET, however, is a near field interaction so we are concerned with the near field term only. Selecting only the \( \frac{1}{r^3} \) terms in Equation 4.4 yields the time-dependent electric field near a donor molecule as:

\[
\tilde{E}_{n.f.D}(t) \approx \frac{\mu_0}{n^2 r^3} \left( 2 \cos \theta_D \hat{r} + \sin \theta_D \hat{\theta} \right) e^{i \omega (t - \frac{nr}{c})}. \]

(4.5)

Consider a second molecular dipole (the acceptor fluorophore) with some orientation and distance relative to the first. The field strength seen by the acceptor due to the donor is
simply $|E_{0,D}|_A = \mu_A \cdot \vec{E}_{nf,D}$. To follow convention this is rewritten as

$$|E_{0,D}|_A = \frac{\mu_0}{\pi \varepsilon_0} \kappa \quad \quad \quad (4.6)$$

$$\kappa \equiv 2 \cos \theta_D (\hat{r} \cdot \hat{\mu}_A) + \sin \theta_D (\hat{\theta}_D \cdot \hat{\mu}_A) \quad \quad \quad (4.7)$$

The power absorbed by the acceptor molecule, $L_A(r)$, for the electric field produced by the donor is assumed by Kuhn [99, 102] to be the same as if the electric field originated as a light wave.

$$L_A(r) = \alpha |E_{0,D}(r)|_A^2 \quad \quad \quad (4.8)$$

Kuhn [99] argues that the coefficient $\alpha$ is given by

$$\alpha = \frac{3cn \ln(10)}{8\pi N_A} \varepsilon_A. \quad \quad \quad (4.9)$$

The argument is this:

The extinction coefficient, $\varepsilon_A$ is defined by

$$-dI = \varepsilon_A IC \ln(10) dl \quad \quad \quad (4.10)$$

where $I$ is the intensity of light passing through an absorber of concentration $C$ of a thickness $dl$. If the absorber is randomly oriented vs isotropically oriented in the direction of the $\vec{E}$ field (the average vs. the maximum absorption) the absorption increases by a factor of 3 in the latter case (average $\theta_A$ and $\phi_A$ over all angles versus aligned with $\vec{E}_{nf,D}$). Finally, inserting the relationship of $I$ to $|\vec{E}|$ Equation 4.10 now becomes

$$-dI = \frac{3cn \ln(10)}{8\pi} |\vec{E}|^2 C \varepsilon_A dl. \quad \quad \quad (4.11)$$

Because $-dI$ is the rate of absorption by the molecules in the unit layer $dl$, $L_A$ can be had by averaging over the number of molecules in that layer, $C N_A dl$.

$$L_A = \frac{-dI/dl}{CN_A}. \quad \quad \quad (4.12)$$

By plugging Equation 4.11 into Equation 4.12 and subsequently into Equation 4.8, Equa-
tion 4.9 emerges and the rate of energy absorption becomes:

\[ L_A = \frac{3cn\ln(10)}{8\pi N_A} \varepsilon_A |E_{0,D}|^2_A \]  (4.13)

and finally plugging into Equation 4.6, we arrive at the power transfer from a donor fluorophore to an acceptor.

\[ L_A = \frac{3c\varepsilon_A |\tilde{\mu}_A|^2 \kappa^2 \ln(10)}{8\pi n^3 N_A \nu^6}. \]

By substituting \( \omega = 2\pi \nu \) we can now write, using the above equation, Equation 4.3, and Equation 4.2 the distance at which the power absorbed by the acceptor equals the power emitted by the donor, \( L_a = L \)

\[ R_0 = \left( \frac{9 \ln(10) \kappa^2 c^4 \varepsilon_A}{128\pi^5 N_A n^4 \nu^4} \right)^{\frac{1}{6}}. \]  (4.14)

This expression is the same as that derived by Förster [96, 97, 102] in the ‘narrow band limit’ [98]. The final form involves replacing \( \varepsilon_A/\nu^4 \) by the spectral overlap integral:

\[ J(\nu) \equiv \int_0^\infty \frac{\varepsilon_A(\nu)}{\nu^4} f_D(\nu) d\nu \]  (4.15)

Where \( f_D(\nu) \) is the normalized fluorescence spectrum of the donor [98, 102]. One must also multiply Equation 4.14 by \( q \), the quantum yield of the donor in absence of the acceptor, which takes care of other non-radiative de-excitation pathways. The final form of \( R_0 \) is:

\[ R_0 = \left( \frac{9 \ln(10) \kappa^2 c^4 \phi_D J(\nu)}{128\pi^5 N_A n^4} \right)^{\frac{1}{6}}. \]  (4.16)

This equation has been implemented in a spreadsheet which will calculate the expected \( R_0 \) given the emission and absorption spectra of the donor and acceptor, respectively (for \( J(\nu) \), \( \phi_D \), \( \kappa^2 \), and \( n \) (note \( n = 1.4 \) in the file as suggested by Clegg [103] for fluorophores in a nucleic acid environment). The spreadsheet can be found on the Poirier Lab share drive at /path/to/drive/ALUMNI/mgibson/thesis/supplemental/FRETCalcTheory.gnumeric. Several relevant FRET pairs have been calculated there already. We can also see from
Table 4.1: A table of Förster distances for various donor/acceptor pairs. Values are from Lee, Lee, and Hohng [105]. Many more pairs can be found in a table compiled by Wu and Brand [106]. Additionally, a gnumeric spreadsheet is available on the Poirier lab share drive at: /path/to/drive/ALUMNI/mgibson/thesis/supplemental/FRETCalcTheory.gnumeric. This spreadsheet will calculate the spectral overlap integral and $R_0$ given a few experimentally determined constants.

<table>
<thead>
<tr>
<th>Donor Fluorophore</th>
<th>Acceptor Fluorophore</th>
<th>Förster distance ($R_0$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cy3</td>
<td>Cy5</td>
<td>54 Å</td>
</tr>
<tr>
<td>Cy3</td>
<td>Cy7</td>
<td>38 Å</td>
</tr>
<tr>
<td>Cy5</td>
<td>Cy7</td>
<td>62 Å</td>
</tr>
</tbody>
</table>

Equation 4.16 $L_A/L = (R_0/r)^{1/6}$. Yielding the classic equations

$$E = \frac{1}{1 + \left(\frac{r}{R_0}\right)^6}$$

$$k_T = k_e \left(\frac{R_0}{r}\right)^6$$

This classical derivation is satisfying in its simplicity. The energy transfer merely originates from one dipole being in the near field of another. What is lost in this classical treatment when compared to the quantum mechanical is in these last steps of inserting an overlap integral. This integral represents both the broadening of energy levels due to solvent interactions as well as the inherent requirement at the quantum mechanical level that during the time of energy transfer, the resonance frequencies of each dipole must match. The spectral overlap integral is a means of determining that probability and Förster emphasized the importance of this resonance matching in his original papers [90, 96, 97, 102].

The distance dependence of FRET ultimately arises from its dipole-dipole interaction nature, yielding a highly switch-like phenomena (see Figure 4.2). The advantages (and disadvantages) for experimentation here will be discussed more thoroughly in sections 4.1.3, 4.1.4 and 4.1.5.
Figure 4.2: The FRET efficiency (fraction of energy absorbed by a donor molecule radiated by an acceptor molecule) as a function of distance between the two molecules. FRET plotting in terms of $R_0$ highlights the switch-like behavior of FRET over molecular distances. Some of the first experiments to observe FRET were de-polarization experiments in viscous media [97, 104] where de-polarization occurred when concentration of fluorophores yielded a mean distance less than the Förster distance, $(R_0$, for reference, $1\text{nm} \approx 1\text{molecules/\text{nm}^3}$).

### 4.1.3 Construct Design

The principle concerns when integrating FRET readouts into biological samples is (i) to attempt to minimize the impact of the fluorophore and any mutations necessary on the behavior of the molecule, (ii) to ensure that the placement of the fluorophores is within the Förster distance of the fluorophores (see Figure 4.2 and Table 4.1), and (iii) that the molecules have either a fixed orientation with respect to one another or are free to rotate, allowing for either direct calculation or motional averaging of $\kappa^2$ (see Equation 4.7).

An example construct used in many of the experiments described in the following chapters is shown in Figure 4.3 and uses these design elements. The distance between the green Cy3 and red Cy5 fluorophores is predicted to be only 16.6 Ångström (see chapter 3).

The nucleosome FRET construct depicted in Figure 4.3 is designed to probe the state of DNA wrapped around a nucleosome. The rates discussed in section 2.3.1 of unwrapping and re-wrapping lead to an equilibrium $k_{\text{eq}} = k_{\text{re-wrap}}/k_{\text{unwrap}}$ of nucleosomes being wrapped ~99% of the time. The FRET labeled nucleosome will therefore yield a high FRET efficiency on average. However, when the nucleosome does fluctuate open to its transiently accessible
state, the FRET efficiency will decrease, and a TF has the opportunity to bind to its recognition site. By measuring the characteristics of this TF binding, information about the accessibility and the energies involved in the interaction are elucidated.

4.1.4 Determination of FRET Efficiency

There are several experimentally viable ways to determine a FRET efficiency (see Equations 4.3 and 4.17), but most require two separate samples: one with an acceptor and one without. Indeed, the most accurate means is suggested by Equation 4.17. If one measures the fluorescence lifetime of a donor, \( k_e \), with and without an acceptor, a very accurate measure of the FRET efficiency can be calculated through the reduced fluorescence lifetime due to a non-zero energy transfer de-excitation pathway.

For our purposes, however, an accurate measure is not necessary. We will be concerned only with changes in FRET efficiency and a constant, repeatable offset is of no concern and the experimental weight of doubling all sample preparation is both daunting and unnecessary. Clegg [103] describes multiple methods of determining FRET efficiency through fluorescence spectra of both the donor and acceptor. One method, \((\text{Ratio})_A\), is particularly suited to the experimental design used for the FRET assays described in this work.

\((\text{Ratio})_A\) essentially compares the emission of the acceptor due to direct excitation to the emission due to energy transfer through an excited donor molecule to the acceptor. Cor-
rection due to donor bleed-through into the acceptor emission wavelength and differences in energy absorption yield the final (Ratio) equation:

$$\varepsilon = \frac{N\varepsilon^A_{Aex} F^A_{Dex} - \varepsilon^A_{Dex}}{\varepsilon^D_{Dex} d^+}.$$  \hspace{1cm} (4.18)

$\varepsilon$ is the extinction coefficient of either the acceptor (superscript $A$) or the donor (superscript $D$) at either the acceptor excitation wavelength (subscript $A_{ex}$) or the donor excitation wavelength (subscript $D_{ex}$). These factors account for differences in energy absorption of the two molecules. $F$ is the fluorescence intensity of the acceptor (superscript $A$) when directly excited (subscript $A_{ex}$) or excited through FRET (subscript $D_{ex}$). $F^{A^*}_{Dex} = F^A_{Dex} - F^D_{Dex}$ ($F^D_{Dex}$ is fluorescence of the donor through donor excitation) which corrects for donor bleed-through into the acceptor emission wavelengths. $d^+$ is the donor labeling efficiency and $N$ is a correction having to do with the number fraction of acceptors which can undergo FRET. Historically, for nucleosomes with their HO acceptor labeled $N = 2$ while for systems with single acceptors and single donors $N = 1$. This factor will be discussed further in section 4.1.5.

To perform a (Ratio) measurement three fluorescence spectra must be obtained.

- Donor emission spectrum. This spectrum is fit to the donor portion of the following spectrum to allow for subtraction of the donor bleed-through to the acceptor signal.
  - Note: The multiplicative factor used to fit this spectrum to the following can be used as a readout for protein induced fluorescence enhancement (PIFE) (not discussed).

- Donor and acceptor emission spectrum during Donor excitation. This is the only spectrum which will actually allow for FRET.

- Acceptor emission spectrum taken immediately following the donor and acceptor spectrum. This spectrum measures (somewhat indirectly) the fluorescence intensity of the same sample should 100% energy transfer occur.

The FRET efficiency obtained from the (Ratio) method can be used in an attempt to calculate the average spatial separation of the donor/acceptor pair using Equation 4.17.
However, it should be noted that inherent to the calculation of $R_0$ potential assumptions are made. Namely, it is unlikely that one knows to a high degree of accuracy either the dipole orientation factor, $\kappa^2$, or the index of refraction, $n$, in the vicinity of the fluorophore once it is closely attached to a biomolecule. One further concern dealing with the prefactor $N$ will be covered in the next section.

4.1.5 Special Concerns with Labeled Histone Octamer and DNA

In the case of nucleosomes (and, indeed, any complex labeled with multiple acceptors or donors) it is necessary to consider the effect of the multiple, nearby FRET pairs. For example, the labeling scheme depicted in Figure 4.3 features two acceptor molecules for every donor molecule. We will search for a general expression for the $N$ term in Equation 4.18 for the situation of two acceptors per donor. The methods discussed here can be expanded to $m$ donors near $n$ acceptors. We begin by looking at the (Ratio)$_A$ method and altering it for our purposes. The methodology involved is in taking the ratio of the acceptor fluorescence excited through FRET and acceptor fluorescence through direct excitation. Clegg [103] gives the necessary expressions as

$$ \text{(Ratio)}_A = \frac{F_{Aex}^A(\nu, \nu')}{F_{Aex}^A(\nu, \nu'')} $$  \quad (4.19)

$$ F_{Dex}^A(\nu, \nu') = [S] \left( \varepsilon^D(\nu') \phi^A(\nu)a^+a + \varepsilon^A(\nu') \phi^A(\nu)a^+ \right) $$  \quad (4.20)

$$ F_{Aex}^A(\nu, \nu'') = [S] \left( \varepsilon^A(\nu'') \phi^A(\nu)a^+ \right) $$  \quad (4.21)

where $\nu$ is the acceptor emission frequency, $\nu'$ the donor excitation frequency, and $\nu''$ the acceptor excitation frequency. $\varepsilon(\nu)$ is the extinction coefficient of either the acceptor or donor at the frequency $\nu$, $\phi^A(\nu)$ is the quantum yield of the acceptor at the frequency $\nu$, and $[S]$ is the concentration of the fluorescently labeled substrate, nucleosomes here. Equation 4.18 is obtained by plugging Equations 4.20 and 4.21 into Equation 4.19. It is clear that both $F_{Dex}^A$ and $F_{Aex}^A$ must be altered in the case of multiple acceptors.

First we consider $F_{Aex}^A$. Here the only problem is the concentration of $[S]$ does not represent
the concentration of fluorophores being excited. Equation 4.21 therefore becomes

\[ F^A_{Acx}(\nu, \nu') = n[S] \left( \varepsilon^A(\nu') \phi^A(\nu) a^+ \right) \] (4.22)

where \( n \) is the number of acceptors per substrate molecule (2 for nucleosomes labels as in Figure 4.3). This is the origin of the correction factor of 2 typically used in the literature. The argument is that (i) the far fluorophore is too far to FRET and (ii) any given fluorophore has a 50% chance of being able to undergo FRET and so the efficiency calculated by \((\text{Ratio})_A\) will be small by a factor of 2 (there is also a term in \( F^A_{Dex} \) which contributes to this mathematically, but the rational for the other factor of \( n \) is identical). If the far fluorophore truly does not FRET this would be the end, but we have evidence at the single molecule level of two-step photobleaching events of acceptors undergoing FRET and the distance between the donor and far fluorophores are on the order of 5 nm according to the crystal structure. Clearly indicating at least a small efficiency of transfer to each fluorophore. We must continue our edits to \((\text{Ratio})_A\) by considering the FRET efficiency to both fluorophores.

In principle each acceptor can FRET with the single donor. The rates involved with this energy transfer are

\[ E_N = \frac{k^N_T}{k_o + k^N_T} \] (4.23)
\[ E_F = \frac{k^F_T}{k_o + k^F_T} \] (4.24)

where \( N \) and \( F \) denote the near and far fluorophore, respectively, \( k_T \) and \( k_o \) are defined as in Equation 4.1. These equations are accurate for singly labeled nucleosomes. However, when considering molecules which can undergo both transitions the denominators in Equations 4.23 and 4.24 need to be altered:

\[ E_{N,2} = \frac{k^N_T}{k_o + k^N_T + k^F_T} \]
\[ E_{F,2} = \frac{k^N_T}{k_o + k^N_T + k^F_T}. \]
We can write the above expressions in terms of Equations 4.23 and 4.24:

\[
E_{N,2} = \frac{E_N (1 - E_F)}{1 - E_N E_F}
\]
\[
E_{F,2} = \frac{E_F (1 - E_N)}{1 - E_N E_F}.
\] (4.25)

Which will be useful later to write the final equations in terms of the singly labeled nucleosome FRET efficiencies.

We must also consider the possibility of homofret between the two acceptors once they are excited

\[
E_H = \frac{k^H}{k_{o-r} + k_r + k^H} \phi_A(H(\nu)) = \frac{k_r}{k_{o-r} + k_r + k^H} \phi_A(A(\nu))(1 - E_H).
\] (4.26)

\(\phi^A\) is the acceptor radiative quantum efficiency and the subscript \(o-r\) explicitly removes the radiative pathway from all other de-excitation pathways for consideration here.

Given Equations 4.23, 4.24, 4.25 and 4.26 we are now ready to write the expression for acceptor fluorescence emission when excited through FRET.

\[
F_{Dex}^A(\nu, \nu') = [S] \left( \varepsilon^D(\nu') \phi^A(\nu) d^+ \left( a^+ a^- (E_N + E_F) + (a^+)^2 (E_{N,2} + E_{F,2}) \left( \frac{1}{1 - E_H} \left( \frac{\phi^A_H(\nu)}{\phi^A(\nu)} \right) \right) + 2 a^+ \varepsilon^A(\nu') \phi^A(\nu) \right) \right) \] (4.27)

The terms here replacing \(E\) in Equation 4.20 represent the different stoichiometries and placements of acceptors possible for a 1 donor, 2 acceptor system. The first term, \(a^+ a^- (E_N + E_F)\) is the term due to singly labeled HO with \(a^- = 1 - a^+\). The \((a^+)^2\) term is the doubly labeled HO case. The \(E_H\) factor allows for infinite regression of homofret between the two acceptors and the \(\phi\) correction factor accounts for the difference in fluorescence quantum yield if a fluorophore is undergoing FRET. Finally, the term \(2 a^+ \varepsilon^A(\nu') \phi^A(\nu)\) accounts for the possible direct excitation of two acceptors due to the incoming light. Plugging Equations 4.25 and 4.26 into Equations 4.22 and 4.27 and finally into Equation 4.19 yields the

41
final expression

\[
\frac{1}{2} \left( (E_N + E_F) + a^+ \left( \left( \frac{E_N + E_F - 2E_N E_F}{1 - E_N E_F} \right) - (E_N + E_F) \right) \right) = \\
\left( \frac{\varepsilon^A(\nu'')(\text{Ratio})_\lambda - \varepsilon^A(\nu')}{d'\varepsilon^D(\nu')} \right)
\]

(4.28)

where we have substituted \( n = 2 \) for this case in Equations 4.22 and 4.27. The right hand side of this equation corresponds to the right hand side of Equation 4.18 but considers corrections due to multiple acceptor labels. Looking at the limits of this equation, it matches the expected behavior both at negligible far FRET

\[
\lim_{E_F \to 0} \left( \frac{1}{2} \left( (E_N + E_F) + a^+ \left( \left( \frac{E_N + E_F - 2E_N E_F}{1 - E_N E_F} \right) - (E_N + E_F) \right) \right) \right) = \frac{E_N}{2}
\]

with \( n = 2 \) as in Equation 4.18. In the simpler case considered by Fábián et al. [107]

\[
\lim_{E_F \to E_N} \left( \frac{1}{2} \left( (E_N + E_F) + a^+ \left( \left( \frac{E_N + E_F - 2E_N E_F}{1 - E_N E_F} \right) - (E_N + E_F) \right) \right) \right) = \frac{2E_N}{2(1 + E_N)},
\]

which is the case for one donor undergoing FRET to two identical acceptors. In addition to these limits, Equation 4.28 also provides theoretical backing for the somewhat paradoxical experimental observation that the right hand side of Equation 4.18, traditionally seen as simply the FRET efficiency, decreases with increasing labeling efficiency. This, however, is precisely the behavior of Equation 4.28. In fact, by measuring the FRET efficiency at low labeling efficiencies and the dependence on labeling efficiency (it is predicted to have a linear response), \( E_N \) and \( E_F \) could be measured. When the FRET response as a function of labeling efficiency is fit to a line, \( y = mx + b \), the near and far FRET efficiencies will be given by:

\[
E_N = b - \frac{\sqrt{(b - 1)(b + m - 1)(b^2 + bm + m)}}{b + m - 1} \\
E_F = b + \frac{\sqrt{(b - 1)(b + m - 1)(b^2 + bm + m)}}{b + m - 1}
\]

(4.29)

The complications leading to Equation 4.28, in addition to concerns discussed at the end of section 4.1.4, make the use of \( \Delta \text{FRET} \), the total change in FRET over a given experiment, critical to interpretation of FRET experiments from nucleosomes.
4.2 Single Molecule Techniques

Fluorescence microscopy, as opposed to bright field microscopy, forms an image from photons emitted by fluorophores rather than reflection, scattering, or absorption of an external light source. Vibrational relaxation of fluorophores produces a stokes shift which allows for filtering of excitation light vs fluorescent light and increased contrast compared to bright field microscopy. For this reason, fluorescence microscopy is often the tool of choice for the imaging of small molecules. The technique requires specific labeling of the structure or molecule of interest and excitation of the fluorophore through a narrow range of wavelengths. The fluorescent emission can then be filtered from excitation light and collected to form a high resolution image while minimizing noise from the unlabeled background environment. Biological samples are ideal targets for fluorescence microscopy as they pose the unique challenges of being generally transparent, complex, and highly concentrated. Fluorescence microscopy can help make sense out of the disorder observed in whole-cell imaging by highlighting individual locations of interest as well as expanding microscopy techniques to distance scales far below what is possible through traditional bright field light microscopy.

The simplest form of fluorescence microscopy is through epi-illumination [108]. Epi-illumination lights the sample through the objective, which has the benefit over trans-illuminated bright field microscopy in that it does not excite fluorophores which are not being observed. However, a significant amount of energy is deposited to regions of the sample which lie outside the focal plane of the objective. Fluorophores which are out of plane still absorb light and still fluoresce, which leads to an increase in background signal. Several methods have been developed to improve on this most basic design. Some focus on increasing contrast by reducing the light collected from out of plane fluorophores, such as confocal microscopy [109, 110]. Other methods reduce the number of fluorophores excited entirely through limiting the excitation volume such as multi-photon interactions or stimulated emission (stimulated emission depletion (STED), structured illumination microscopy (SIM)). Yet more methods use computational or chemical methods to further reduce the number of simultaneously
Figure 4.4: (A) Diagram of Snell’s law. An ray of light is incident on an interface between two materials with indexes of refraction $n_1$ and $n_2$ at an angle $\theta_1$ relative to the surface normal is reflected at an angle $\theta_r$ and transmitted at a refracted angle, $\theta_2$, according to Snell’s Law. (B) Cartoon depicting the incoming and outgoing excitation light during objective-based total internal reflection fluorescence (TIRF). Light is input on the far edge of the objective back focal plane which create a high angle of attack at the quartz/water interface shown above the objective (black). (C) Excitation path in prism-based TIRF. A prism is placed on top of a quartz slide with an aqueous sample beneath in a flow cell. Light is then bent via a prism to yield an angle of incidence greater than the critical angle for total internal reflection. An objective is then focused through the flow cell onto the quartz-water interface. In (B) and (C) an evanescent field is produced which can excite fluorophores near the surface due to total internal reflection from the quartz-water interface and omnidirectional fluorescence is observed in real time with very low background.

Excited fluorophores (stochastic optical reconstruction microscopy (STORM), photoactivated localization microscopy (PALM)). Each of these techniques improve the spatial resolution of fluorescence microscopy, some of which improve resolution beyond the diffraction limit and are termed super-resolution. However, each of these techniques severely limits the number of photons detected per molecule in order to achieve their high resolution. These techniques are most used in imaging of live cells where separation of fluorophores cannot be achieved by any means other than increased resolution. If the concentration of fluorophores can be controlled other means, such as with in vivo measurements, single molecule TIRF (smTIRF) microscopy can be used.

4.2.1 TIRF Microscopy

TIRF microscopy utilizes the evanescent wave created during total internal reflection (discussed below). The shallow penetration of the evanescent wave (∼100 nm) lends itself readily
to fluorescence microscopy as only fluorophores very near the interface are excited yielding drastically reduced background noise when compared to other means of excitation. The trade-off is that the excitation field is so shallow surface tethering is necessary to observe fluorophores for any extended period. Tethering always introduces concerns over surface effects but control for these effects is possible and the smTIRF technique opens doors to single molecule imaging of fast protein kinetics at the 10s of ms timescale.

There are two distinct methods for generating the total internal reflection which characteristic of TIRF which are expanded upon after a theoretical overview of the evanescent wave relied upon by TIRF microscopy.

**Evanescent Wave**

We consider the classic Snell’s law problem of a light beam incident on an interface between two materials as in Figure 4.4A. In simplified Ray optics, the light is split at the interface between a reflected wave and a refracted wave whose directions are determined by the law of reflection and Snell’s law, respectively (using the notation of Figure 4.4).

\[
\theta_1 = \theta_r \quad \text{(Law of Reflection)}
\]
\[
n_1 \sin \theta_1 = n_2 \sin \theta_2. \quad \text{(Snell’s Law)}
\]

If we think about these rays as plane waves, the electric field is given as a sum of exponentials

\[
\tilde{E} = \tilde{E}_0 e^{i(\tilde{k}\cdot r - \omega t)}
\]
\[
\tilde{E}_0 e^{i(k_x \sin \theta_2 x + k_z \cos \theta_2 z - \omega t)}
\]

Where \( E_0 \) is the amplitude of the electric field, \( k_x \) is the amplitude of the transmitted \( \tilde{k} \) in the x direction (right in Figure 4.4A, \( k_z \) is the amplitude in the z direction (down in Figure 4.4A and \( \theta_2 \) is as defined in Figure 4.4A. It becomes clear that if \( \tilde{k} \) becomes imaginary, \( \tilde{E} \) gains an exponential decay in that direction. The value of \text{arccos} is purely imaginary for angles greater than 90° and so the \( \tilde{E} \) field decays exponentially in the z direction. This is the origin of the evanescent wave taken advantage of in smTIRF.
Figure 4.5: The characteristic penetration depth (the depth into the second medium which yields a 1/e decrease in the field intensity) as a function of the angle of incidence on the interface. The media here are taken as Quartz ($n = 1.47$) and water ($n = 1.33$). The depth is also dependent on wavelength and values are shown for wavelengths of 450 nm to 700 nm.

In order to produce an evanescent wave all that is required is an angle greater than or equal to the critical angle, $\sin \theta_c = n_2/n_1$. Because $\theta_c$ is limited to between 0 and 90°, $n_1$ must be greater than $n_2$. For biological samples in aqueous buffer solutions, $n_2 = 1.33$. Because my experiments excite fluorophores through a quartz slide ($n_1 = 1.47$), $\theta_c = 64.79^\circ$. However, the depth of the evanescent wave depends on the wavelength, the indices of refraction and the actual angle of diffraction as well. See Figure 4.5 for the penetration depth as a function of angle of incidence and wavelength for the interface discussed here.

**Objective TIRF**

The first of the two means of producing a TIRF field passes the excitation light through the same objective used to collect fluorescent light. A laser is focused off-center on the back focal plane of the objective which produces a large angle of incidence at the coverslip-water interface (see Figure 4.4). The farther off center, the larger the angle of incidence. The largest angle of light which is focused by an objective (which is also the largest angle de-focused) is given by

$$\text{NA} = n_0 \sin \theta_m$$  \hspace{1cm} (4.30)

where $n_0$ is the index of refraction of the immersion solution the objective is designed for and $\theta_m$ is the maximum angle which is focused by the objective. Clearly a higher NA is desirable as it not only increases the light gathered from fluorescence, but it also increases
the angle of incidence of the TIRF illumination which, as shown by Figure 4.5, reduces the
background light as well by reducing the depth of the evanescent wave.

A major negative of objective-based TIRF is the requirement of a dichroic to filter excitation
light from emission light. This is a necessity of exciting from the same side as you collect—at
some point these paths must diverge and this necessitates a loss of photons. Typically
this is accomplished through a dichroic mirror which bounces excitation light and passes
fluorescent (a long-pass filter). However, a technique dubbed micro-mirror TIRF has made
this configuration obsolete by replacing the wavelength-dependent mirror with very small
full silver mirrors. These micro-mirrors are placed to only reflect at the extremes of the
objective, minimally reducing fluorescence yield while completely blocking both input and
output of excitation light (see Figure 4.4B, mirrors are placed to intercept the cyan laser
light)[85]. This advancement has removed many of the disadvantages of objective-based
TIRF and improves signal-to-noise ratio (SNR) to that of prism-based systems.

Prism-Based TIRF

Prism-based TIRF, in contrast to objective TIRF utilizes a new optical element placed above
the sample and light is passed top-down as in trans-illumination. Figure 4.4C displays a
schematic of the TIRF production in a prism-based system. The prism is placed on top of
the microscope slide and light is incident upon it to create total internal reflection at the
quartz-water interface as before. The prism is necessary to alter the beam-path and allow for
total internal reflection. The angle of the incident light to the horizontal, $\theta_3$, is dependent
on the prism angle $\alpha$ and the angle light is incident on the prism relative to the horizontal,
$\phi$. After performing the necessary geometry and applying section Snell’s Law, we end up
with:

$$\theta_3 = \alpha - \arcsin \left( \frac{n_1}{n_2} \sin \left( \phi - 90^\circ + \alpha \right) \right)$$

where $n_1$ is the index of refraction of air and $n_2$ is the index of refraction for the prism, 1.457. For the geometries typically used in our smTIRF microscopes the final incidence
angle, $\theta_3$, is 80° and relatively insensitive to the incoming angle, $\phi$. The black dotted line
on Figure 4.5 highlights the characteristic penetration depth for the systems used in these studies which ranges from \( \sim 75 \text{–} 100 \text{ nm} \).

There are no restrictions on the NA of the objective, however a larger NA still corresponds to more fluorophore signal so is preferable. This is the orientation chosen for the home-build systems used in my work because we have found it produces a higher SNR with more robust sample preparation.

### 4.2.2 Zero-mode waveguides

Zero-mode waveguides (ZMWs) are nanoapertures formed in a conductor [87]. If the aperture is smaller than the wavelength of light used to illuminate it, it behaves as a waveguide attenuator. This attenuation can be tuned such that an even smaller excitation volume is achieved than is possible through TIRF-based illumination; tens of zeptoliters of excited volume are attainable through the use of ZMWs. By calculating the typical volume containing a single fluorophore at various concentrations, this excitation volume is likely to contain less than one molecule for any concentration up to tens or hundreds of micromolar. ZMW technology expands single molecule fluorescence co-localization to the regime which includes the \( K_d \)s of PTM binding proteins.

**Theoretical Behavior**

Waveguides are well studied electromagnetic phenomena. If we consider the ZMW to be a macroscopic cylindrical waveguide, the behavior of light traveling through the waveguide is calculable. For TE modes the \( z \)-dependence of the electric field is given by:

\[
E_\theta(r, \theta, z) \propto e^{\pm i \beta z}
\]

\[
\beta = k^2 - \left( \frac{p_{\nu n}'}{a} \right)^2,
\]

(4.31)

Here, \( \beta \) is the longitudinal component of the wavevector, \( k \). The factor \( p_{\nu n}' \) is the \( n \text{th} \) zero of the derivative of the \( \nu \text{th} \)-order Bessel function of the first kind, \( J'_{\nu}(x) \), and \( a \) is the radius
Figure 4.6: The theoretical characteristic length of attenuation as a function of aperture size and wavelength from 450 nm (blue) to 700 nm (red) in 25 nm steps. The approximate average size of the ZMW produced here is shown as a dotted line and corresponds to ~40 nm characteristic length.

Figure 4.7: (A and B) Intensity of the electric field perpendicular to the polarization of light and direction of travel. Light is incident perpendicular to the waveguide and shown for 532 nm (A) and 638 nm (B) wavelengths. (C and D) A zoom in and rescale of light intensity from the bottom of the gold surface of the waveguide to 200 nm above it for 532 nm (C) and 638 nm (D). Attenuation is \( \geq 10000 \).
of the cylinder. The dominant mode of the wave is the lowest energy corresponding to a non-trivial solution. This is the $\nu = 1, n = 1$ mode where $p'_{11} \approx 1.8142$.

The equation for $\beta$ in Equation 4.31 can be used to determine a cutoff wavelength. Longer wavelengths will propagate and shorter wavelengths will be attenuated. For the wavelengths we are interested in, 473 nm, 532 nm, and 638 nm, a diameter of 139 nm, 156 nm, and 187 nm, respectively, are the maximum acceptable to produce an attenuative behavior. The characteristic depth of penetration as a function of ZMW diameter is shown in Figure 4.6.

In addition to this analytical treatment, numerical electrodynamics can be performed to determine the expected behavior of a given ZMW. The application MEEP [111] was used to simulate a cylindrical, metallic nanoaperture on the surface of a quartz slide submerged in water. The metal used is a theoretical perfect conductor. Plane wave light was incident on the quartz-water interface and the simulated intensity of light is shown in Figure 4.7. For this geometry, MEEP calculated the relative intensity before and after the waveguide to be reduced by at least a factor of 10,000 for the position and time slices shown in Figure 4.7.

The theoretical treatment considered thus far neglects i) the microscopic nature of the physics involved and ii) the presence of a fluorophore to excite. The microscopic scale allows for plasmon excitation in the metallic coating. Numerous studies have investigated the impact of gold near a fluorophore [87, 112–114]. These studies find that the rate of fluorescence is dramatically increased in a geometry-dependent manner. This improves the fluorescence lifetime and the FRET response of nearby fluorophores [112, 114]. For the geometry proposed here, up to 1000-fold enhancement can be expected [112]. The enhanced lifetime and signal allows for faster acquisition rates over longer lengths of time. This plasmonic enhancement coupled with the extremely small excitation volumes attainable make ZMWs extremely appealing.
Initial Production Results

The significant drawback of ZMWs is the involved process required to produce them. The small patterns necessitate e-beam lithography; however, lithography is complicated by the insulating nature of the quartz slides used in microscopy. Finally, liftoff of the photoresist from a small aperture is non-trivial. However, these problems can be overcome to realize efficient production of ZMWs.

The production process developed in collaboration with Nanotech West here at OSU has begun to solve these problems. Initial results from ZMWs produces are shown in Figure 4.8. Thus far, nanoapertures have been successfully fabricated and confirmed by SEM (Figure 4.8A). Additionally, cleaning protocols have been developed which leave the ZMWs intact and reduce background fluorescence to an acceptable level (Figure 4.8B). A surface passification protocol is detailed in Kinz-Thompson et al. [87] which must be repeated before experiments determining the efficacy of the ZMWs can be undertaken.
4.2.3 Single Molecule Sample Preparation

Flow Cell Design and Preparation

Experiments are performed in microchannel flow cells with 8 channels per microscope slide. Each flow cell surface is coated with polyethylene glycol (PEG) to pacify the slide surface against nonspecific interactions. A fraction of the PEG is labeled with a biotin molecule so that a biotin-streptavidin-biotin linkage tethers biotin labeled DNA to the surface. The methods used for preparation of these flow cells and experimental methodologies are based on previously described methods [115] and outlined below.

Flow cells are formed from a three-layer stack as shown in Figure 4.9. Eight channels are cut in Parafilm (Bemis PM966) which is then sandwiched between a passivized glass cover-slip (Fisher 12-544-E) and a quartz slide (G. Finkenbeiner) with drilled inlet and outlet holes. These holes may be drilled with a CO$_2$ laser engraver or Starlight 115005 diamond drill bit to a diameter of 0.029 inches. Carefully clean and dry beakers for use in each step.

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.

Begin cleaning the slides with sonication in toluene followed by ethanol for 20 min each. Thoroughly rinse in water after each sonication to remove the solvents. Slides and coverslips are then cleaned in piranha solution for 30 min until the reaction completes and sonicated in 1M sodium hydroxide for 20 min to append hydroxyl groups to the silica surface. These hydroxyls serve as ligation sites for silane in the next step. Rinse the slides and coverslips thoroughly in water and dry overnight at 75°C.

Silane is the first layer of coverage of the sticky quartz/glass surface necessary to minimize non-specific binding of fluorophores to the surface and therefore minimize surface effects during single molecule experiments. The linkage is done through a dehydration reaction that forms an ether linkage between the silane and quartz/glass surface. The reaction
is performed in Acetone which limits the self-quenching of silane. 8 mL of 2% (3-aminopropyl)triethoxysilane (MP Biomedicals 215476680) are added to 500 mL of acetone while swirling to efficiently coat the slides and coverslips with silane. Finally the reaction is quenched in a solution of 50% acetone 50% water and the slides are rinsed thoroughly with water and incubated at 75°C for 6 h.

PEG is covalently linked to the previously deposited silane via an NHS-ester reaction. This is accomplished through a mixture of monofunctional PEG (MPEG) (5 kDa, Laysan Bio, mPEG-SVA-5000) and biotin difunctional PEG (bioPEG) (5 kDa, Laysan Bio, Biotin-PEG-SVA-5000) which has been further functionalized with a biotin to provide a specific attachment point for fluorescently labeled samples.

A ratio of 1:100 bioPEG:MPEG has been found to be ideal when pacifying surfaces. This provides more than enough linkage sites and allows fine control of fluorophore density to be adjusted by concentration of sample injected during the experiment. The process of PEGylating a surface is complex and has been thoroughly explained in other resources [79, 116].

Cleaned slides and coverslips are finally assembled into fully formed flow cells by sandwiching the two around a strip of Parafilm with flow channels cut into it and adhered with heat. Prepare Parafilm by cutting channels in line with previously drilled inlet/outlet holes on the quartz slides. Place one PEGylated quartz slide on a clean, dry glass plate with the PEG side facing up. Place a single sheet of prepared Parafilm without paper backing 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 4.9. Press the coverslip slightly to temporarily adhere it to the slide. Permanently adhere the flow cell together by placing it on a glass plate heated to 90°C on a hot plate. Allow the Parafilm to turn partially clear while gently pressing the flowcell together. Do not allow the Parafilm to completely melt.

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. Flow cells are dismantled by soaking in acetone for 15 minutes and gently prying the coverslip and slide apart.

**Imaging Buffer**

A glucose-oxidase-based oxygen scavenging system is used in the imaging buffer to increase fluorescence lifetime before photobleaching. In addition to this system, triplet state quenchers are used to improve fluorophore stability and inhibit fluorophore blinking. These quenchers are Cyclooctatetraene (COT) and 3-Nitrobenzyl alcohol (NBA) and Trolox.

Final concentrations of each are

- Glucose: 1.6% (w/v)
- Glucose Oxidase: 450 ng/µL
- Catalase: 22 ng/µL
- COT: 0.0115% (v/v)
- NBA: 0.0132% (v/v)
- Trolox: 2 mM pH 8.0.

The remainder of the imaging buffer is matched to the salt, glycerol, crowding agents, surfactants, etc. required by the experiment and details are provided where appropriate.

**4.2.4 Data Acquisition**

The order of operations to prepare a flow cell for imaging is designed to minimize surface interactions and background fluorescence. Channels are first flushed with a wash buffer
appropriate to the experiment. Next a blocking buffer containing 1 mg/mL bovine serum albumin (BSA) is injected into the channel in an attempt to non-specifically bind to any exposed charged or hydrophobic sections of the surface. Next, the biotinylated PEG molecules are charged by flowing in a buffer containing 20 µg/mL streptavidin. Because each streptavidin contains binding sites for four biotin molecules, these are used as a strong linkage between the fluorescent sample and the surface of the slide. Finally the fluorescent sample is injected into the flow channel and incubated for several minutes to ensure binding before being washed out to remove background fluorescence. Imaging buffer is then injected into the flow cell and it is imaged on the TIRF microscope.

The fluorescent nucleosome samples used in this work for single molecule experiments differ from those used in bulk assays in that a 75 bp linker is added to the nucleosome. This linker extends the nucleosome ~25 nm from the surface of the slide. The idea here is to minimize interactions between the surface and the nucleosome and better measure solution state equilibrium dynamics of the nucleosome. Of course, surface interactions are always a worry and so any single molecule data is always compared to bulk data (see section 4.2.5).

Videos are acquired using a pre-calibrated image splitter (see section F.4). The EM-gain is set to ~600 and exposure adjusted to a timescale appropriate for the events being observed. For LexA binding to its target site at bases 8-27 within the nucleosome, this is 50 ms. The acceptor fluorophore is directly excited for ~50 frames before switching to donor excitation until the field of view is significantly photobleached. Laser powers are adjusted to yield strong signal to noise ratios while maximizing video length. The initial acceptor excitation will be used to identify molecule locations for analysis of FRET traces.

4.2.5 Data Analysis

Fluorophore detection

Molecules are first identified by the presence of an acceptor. The image is split into fluorophore channels and the acceptor channel is averaged over the acceptor excited frames.
This yields a very low noise image which can be analyzed for molecule locations. A flood filling algorithm identified local maxima and minima. The height of these maxima relative to the local minima is sorted and the vertex of the sorted data is determined which gives a threshold value for classifying a maxima as a fluorescent molecule.

Once molecules have been identified the fluorescence intensity of a 3×3 pixel square surrounding the maxima is averaged and measured in time to yield a fluorescence trace. In order to identify molecules which possess all fluorophores, the median value of the trace is calculated for each channel. Traces with a median above a set limit are marked as evincing signal for that fluorophore. In addition to the presence of all fluorophores, changes in fluorophore signals are also of interest. In the case of FRET experiments as used in this work, the Cy3 and Cy5 fluorescence intensity fluctuations must be anti-correlated to indicate a change in state. Correlation coefficients are calculated for each trace.

**Molecule Trace Selection**

Traces likely exhibiting fluctuations can be semi-automatically selected for further processing, but it is 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 requires human input to individually select for sections of traces before moving on to the computationally intensive determination of idealized FRET states (see Figure 4.10A).

**Idealized Traces and Dwell Time Distributions**

vbFRET [89], a MATLAB package provided by Dr. Ruben Gonzalez, is used to generate idealized FRET traces from selected, truncated traces. In preparation for import into vbFRET, the background of each fluorescent signal is subtracted and a pseudo FRET trace is calculated for each molecule using $E = \frac{IA}{IA + ID}$, i.e. the intensity of the acceptor di-
Figure 4.10:  (A) Raw data trace from a single molecule FRET experiment. Cy3 emission is shown in green and Cy5 in red. The truncation points are indicated which cut out the acceptor excitation portion of the experiment as well as the end photobleaching event. Between the truncation points clear anticorrelated fluctuations exists indicating transitions from a long lived high FRET state to a short lived low FRET state. (B) FRET traces (purple) and Viterbi paths (black) which show the idealized path as calculated by vbFRET. The plots show the evolution of a low FRET state with increasing TF concentration of TF (no TF top, 5 µM Lexa middle, 50 µM bottom).
Figure 4.11: 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.
vided 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 4.10B. States 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 4.11A–B. In order to use all the information in our dataset and avoid using an arbitrary histogram bin size, cumulative sum distributions are produced and fit to find characteristic dwell times as shown in Figure 4.11C. The characteristic dwell times correspond to the binding and unbinding rates for each concentration. It is important to show that the high state dwell times are inversely proportional to protein concentration to confirm the fluctuations are in fact a result of protein binding and that low state dwell times do not depend on protein concentration (Figure 4.11D).

Site Occupancy and Comparison to Bulk Data

It is possible to calculate from idealized traces an excellent analogue to bulk fluorescence measurements. If the concentration-dependent FRET response is normalized to span from zero to one, the plot measures the fraction of nucleosomes unbound by TFs. The fraction of time spent in the unbound state (for FRET this is the high FRET state) averaged over all molecules directly measures the fraction of nucleosomes unbound at a given concentration of TF. The two should match; If not, there may be an issue selecting molecules or at some other point in the single molecule pipeline.

4.3 Anisotropy

The dipole nature of chromophores yields polarized light if considering a single molecule which is individually exited. This point leads to an interesting effect when these fluorophores are immobilized. Polarized excitation light incident of stationary fluorophores will yield polarized fluorescence. Therefore, the level of isotropy of fluorescent light is a direct measure
Figure 4.12: (A) Fluorescein-labeled 147 bp DNA is rotationally confined by increasing viscosity due to increasing glycerol concentration. (B) steric occlusion by LexA binding to its target site on the DNA which has been complexed in to a nucleosome. As the concentrations increases, the rotational diffusion constant of the fluorophore decreases and the isotropy of the fluorophore’s emission also decreases resulting in an increase in the polarization signal. (C) Crystal structure of the nucleosome highlighting the position of the fluorescein molecule used as a reporter in B). This labeling position allows for a reduction in rotational freedom upon binding of LexA and, as will be shown in chapters 7 and 8, PTM readers

of the mobility (specifically rotational freedom) of fluorophores. Some of the earliest experiments of this polarization was through fluorophores immobilized in highly viscous media [104] (see Figure 4.12A). However, the technique has now been refined to probing biomolecular interactions [117–119]. The key feature is a loosely attached fluorophore with a relatively long excitation lifetime. This loose attachment and long lifetime allows for exploration of many rotational conformations prior to fluorescence. Taken in aggregate, without rotational limiting, highly anisotropic fluorescence is expected. The measure of binding comes into play with addition of some ligand. The ligand must bind near enough the fluorescent attachment point to limit the number of states accessible to the fluorophore. Mobility restriction due to ligand binding has the effect of increasing the level of polarization measured through fluorescence. This technique gives simple, fluorescence based (and therefore sensitive) access to biomolecular interactions over a wide range of concentrations (nM to tens of µM).

In the case of the nucleosome, the tumbling time of the large, ~200 kDa nucleosome is far too slow to exhibit motional averaging of fluorescence polarization. Therefore, a loosely attached fluorophore that can be limited in its access upon ligand binding is perfectly suited for anisotropy measurements as in Figure 4.12B. Measurement of PTM binding has been
successfully performed with fluorescein labeled DNA incorporated into the nucleosome [120, 121].

As a practical note. Measurements of the anisotropy of fluorescence require two polarizers. One of these polarizers acts on the incoming excitation light and the other on the fluorescence light. Fluorescence intensity is measured at four positions, two with the polarizers parallel (vertical and horizontal) and two with them anti-parallel (vertical/horizontal and horizontal/vertical). These four intensities form the definition of anisotropy:

\[
\langle r \rangle = \frac{I_{VV} - GI_{VH}}{I_{VV} + 2GI_{VH}} \quad (4.32)
\]

\[
G \equiv \frac{I_{HV}}{I_{HH}}
\]

The factor \(G\) is called the ‘G factor’ and corrects for potential differences in the detection efficiency of the different light polarizations. In principle, this must only be measured once, but measurements are fast and typically automated. Conversion from anisotropy to polarization is done by

\[
P = \frac{3\langle r \rangle}{2 + \langle r \rangle}. \quad (4.33)
\]

Polarization numbers are typically quite small so it is common to define a unitless quantity

\[mP \equiv \frac{P}{1000}.
\]
Chapter 5

PHF1 Directly Modulates Nucleosome Accessibility While Binding to its Target H3K36me3

Reproduced with modifications from:

The Tudor domain of human PHF1 recognizes trimethylated lysine 36 of histone H3 (H3K36me3). This interaction modulates the methyltransferase activity of the Polycomb Repressive Complex 2 (PRC2) complex and has a role in retention of PHF1 at DNA damage sites. We have previously determined the structural basis for the association of Tudor with a methylated histone peptide. Here we detail the molecular mechanism of binding of the Tudor domain to the H3K_{36}me3-NCP. Using a combination of transverse relaxation-optimized spectroscopy (TROSY) NMR and FRET, we show that Tudor concomitantly interacts with H3K36me3 and DNA. Binding of the PHF1 Tudor domain to the H3K_{36}me3-NCP stabilizes the nucleosome in a conformation in which the nucleosomal DNA is more accessible to DNA-binding regulatory proteins. Our data provide a mechanistic explanation for the consequence of reading the active mark H3K36me3 by the PHF1 Tudor domain.
5.1 Introduction

Human PHF1 is found in major nuclear regulatory complexes. PHF1 is an accessory component of the PRC2 that methylates lysine 27 of histone H3 and is required for gene silencing (reviewed in [122]). Upon genotoxic stress, PHF1 localizes to the sites of DNA damage through the association with DNA damage repair enzymes Ku70/Ku80 ([123]). PHF1 contains an N-terminal Tudor domain that binds to H3K36me3 [124–126]. This interaction inhibits the methyltransferase activity of PRC2 and has a role in the retention of PHF1 at the sites of dsDNA breaks [124]. We have previously determined the structural basis of Tudor association with a peptide corresponding to the H3K36me3 tail (residues 31–40 of H3) [124]. However, the question remains as to how the Tudor domain binds to H3K36me3 in the context of the full nucleosome and even further, the chromatin fiber. This is an especially pressing question given that the position of this histone mark would place Tudor just adjacent to the nucleosome core, which may greatly alter the interaction.

Although recent pioneering studies have begun shedding light on the mechanistic aspects for the association of effector domains with their target histone PTMs on the nucleosome [120, 127–130], there is still a dearth of such reports. This is in large part due to experimental limitations. Most of the detailed structural information regarding the nucleosome complexes has been obtained by X-ray crystallography. However, because the N-terminal histone tails protrude from the nucleosome core and are solvent accessible and flexible, it has proven difficult to study interactions of effectors with the modified tails using crystallographic methods. Currently only four nucleosome/effector complexes have been successfully crystallized, and all reveal interactions with the particle core (and unmodified H4 tail in the case of the Sir3 BAH domain) [131–137]. As a result, various interactions with PTMs have been studied primarily using histone peptides, leaving a large gap in our knowledge of how these proteins bind to histone PTMs within the framework of the full NCP.

NMR spectroscopy is an ideal tool to address this question as it provides solution-state atomic resolution information and is very powerful in the detection and characterization of
intermolecular interactions, even in very flexible constructs [138, 139]. However, it remains challenging to produce modified nucleosomes in amounts sufficient for NMR analysis. Moreover, owing to the large size of the nucleosome/effector complexes, sophisticated labeling and advanced experimental methods are necessary to obtain quality data. Yet, progress is being made in this regard and exciting studies have recently demonstrated the influence of histone PTMs on the nucleosome and on the association with co-factors [127, 129]. The issues with the NCP size and amount can be overcome by studying the nucleosome/effector interactions using FRET, which requires only nanomolar concentrations of the NCPs. FRET has previously been used to demonstrate that the entry/exit regions of nucleosomal DNA transiently unwrap from and rapidly rewrap onto the histone core surface [20, 75, 140]. This phenomenon, which is often referred to as the DNA end ‘breathing’ motion, has a fundamental role in nucleosome remodelling, DNA replication and gene transcription.

Here we describe the mechanism for binding of PHF1 Tudor to the H3K$_C$36me3-NCP and characterize the effect of this interaction on targeting a specific DNA site within the NCP by a protein effector. We show that in addition to recognizing H3K36me3, the Tudor domain associates with DNA, and such multiple contacts may account for a higher affinity of Tudor for the H3K$_C$36me3-NCP. Interaction of the PHF1 Tudor domain with the H3K$_C$36me3-NCP facilitates in vitro binding of the DNA-specific protein LexA to the DNA site that is inaccessible in the fully wrapped nucleosome, thus suggesting a shift in equilibrium towards a more open form of the nucleosome. Together, our findings provide a possible mechanistic explanation for the consequence of recognition of the H3K36me3 mark by the PHF1 Tudor domain.

5.2 Results

5.2.1 Tudor Binds to the H3K$_C$36me3 Nucleosome

To determine the mechanism for recognition of the nucleosome by the Tudor domain of PHF1, we reconstituted the NCP carrying a MLA at position 36 of histone H3 (H3K$_C$36me3)
(Figure 5.1). We generated the H3(C110A,K36C) mutant and subjected it to alkylation by (2-bromoethyl)trimethylammonium bromide to yield the N-methylated aminoethylcysteine derivative. The MLA histone was refolded with unmodified recombinant H2A, H2B and H4, and the resulting histone octamer was assembled onto a Widom 601 DNA sequence through slow desalting. Because the size of the H3K$_{36}^{\text{me3}}$-NCP/Tudor complex (~205 kDa) is beyond the limit of conventional standard deviation (SD), we generated the double $^2$D- and $^{15}$N-labeled Tudor domain and characterized its binding to H3K$_{36}^{\text{me3}}$-NCP using $^1$H, $^{15}$N-TROSY-heteronuclear single-quantum coherence (HSQC) experiments carried out at the elevated temperature of 310 K, on a 900 MHz spectrometer equipped with a cryogenic probe.

The $^1$H, $^{15}$N-TROSY-HSQC spectra were collected on the free protein and in the presence of increasing concentrations of H3K$_{36}^{\text{me3}}$-NCP and overlayed (Figure 5.2a). Titration of the H3K$_{36}^{\text{me3}}$-NCP up to 0.1mM induced significant changes in the Tudor amide resonances, which could be grouped into the three sets. One set of resonances showed a substantial decrease in intensity and disappeared completely upon addition of 0.05 mmol/dm$^3$ of the nucleosome. The second set of cross-peaks gradually moved, revealing the fast-to-intermediate exchange regime on the NMR timescale. The third set was perturbed to a lesser degree. At a 1:1 molar ratio there was, on average, an ~80% decrease in resonance intensity, indicating
Figure 5.2: (a) $^1$H, $^{15}$N-TROSY-HSQC spectral overlays of wild-type Tudor in the presence of increasing concentrations of H3K$_C^{36}$me3-NCP (molar ratio is shown to the left). (b) A plot of the decrease in resonance intensity ($1-I/I_0$) induced by the H3K$_C^{36}$me3-NCP as a function of Tudor residue. Decreases of greater than 85, 90 and 96% are shown in tan, salmon and mauve, respectively. An asterisk indicates a missing or unassigned peak. The double asterisk indicates value below shown threshold. (c) Residues demonstrating significant decreases in resonance intensity are mapped onto a surface representation of Tudor with the H3K36me3 peptide shown as sticks in yellow. Residues of the H3K$_C^{36}$me3 peptide and the Tudor domain are labeled using a three-letter code and a one-letter code, respectively. (d) $^1$H, $^{15}$N HSQC spectral overlays of wild-type Tudor in the presence of increasing concentrations of H3K$_C^{36}$me3 peptide. (e) A plot of the extent of normalized chemical shift change induced by the H3K$_C^{36}$me3 peptide as a function of Tudor residue, with changes greater than the average, average plus $1/2$ or average plus 1 standard deviation shown in tan, salmon and mauve, respectively. An asterisk indicates a missing or unassigned peak. (f) $^1$H, $^{15}$N-TROSY-HSQC spectral overlay of Y47A Tudor mutant in the presence of increasing concentrations of the H3K$_C^{36}$me3-NCP (molar ratio shown to the left).
a robust interaction between Tudor and the H3K_{C36me3}-NCP.

Plotting the intensity decrease \((1 - I/I_0)\) for each Tudor residue allowed us to identify the binding interface (Figure 5.2b). Mapping the residues with rapidly disappearing amide signals onto the structure of the Tudor domain revealed an extensive patch at one of the open ends of the five-stranded \(\beta\)-barrel (Figure 5.2c, magenta). Many of these residues are directly involved in the interaction with H3K_{36me3} [12]. Particularly, the aromatic cage, where trimethylated Lys36 is bound (PHF1 residues W41, Y47, F65 and F71), the adjacent hydrophobic patch underneath Pro38 of the peptide (PHF1 residues L45 and L46) and the acidic groove (PHF1 residues E66, D67 and D68) were significantly perturbed upon addition of the H3K_{C36me3}-NCP. This binding interface matched well to the binding surface mapped based on chemical shift changes observed in the \(^{15}\)N-labeled Tudor domain as the H3K_{36me3} peptide was titrated in under similar conditions, that is, 310 K at 900 MHz (Figure 5.2d, e).

These results suggest that PHF1 Tudor binds to the methylated H3 tail of H3K_{C36me3}-NCP in a manner largely similar to how it associates with the histone peptide. However, several residues of PHF1 were perturbed by the nucleosome to a greater extent, specifically those that descend from the hydrophobic patch along the \(\beta2\) strand of the barrel, as well as from the acidic patch along the \(\beta4\) strand. In addition, R58 and E59 at the opposite end of the \(\beta\)-barrel were affected only upon binding to the H3K_{C36me3}-nucleosome, thus suggesting additional contacts with the intact NCP.

Recognition of the trimethylated Lys36 residue remains the major driving force for binding of the PHF1 Tudor domain to H3K_{C36me3}-NCP. We tested the Y47A mutant of Tudor previously found to have impaired histone binding activity (Figure 5.2f). Addition of H3K_{C36me3}-NCP to the \(^{15}\)N-labeled Y47A mutant led to only a small and uniform decrease in resonance intensity, which is indicative of a very weak and non-specific association. These data reveal that a robust interaction with the NCP occurs only when Tudor is capable of binding to the H3K36me3 mark.
5.2.2 Tudor Has a High Affinity for H3K\textsubscript{36}me\textsubscript{3}-NCP

To assess the strength of the PHF1–NCP interaction, we carried out pulldown assays with GST-fusion Tudor domain (Figure 5.3). GST-Tudor was incubated first with increasing concentrations of H3K\textsubscript{36}me\textsubscript{3}-NCP and then with glutathione-agarose beads. After centrifugation, the nucleosome fraction bound to the protein was detected by western blot using an anti-H3 antibody (Figure 5.2a). Densitometry analysis of the bound H3K\textsubscript{36}me\textsubscript{3}-NCP yielded an apparent $K_d$ of $\sim1.3\,\mu M$ (Figure 5.2b), and revealed that binding of the PHF1 GST-fusion Tudor domain to H3K\textsubscript{36}me\textsubscript{3}-NCP is stronger than its binding to the H3K\textsubscript{36}me\textsubscript{3} peptide alone ($K_d = 36\,\mu M$, [124]). Pulldowns using Y47A mutant of GST-Tudor and the H3K\textsubscript{36}me\textsubscript{3}-NCP or wt GST-Tudor and unmodified NCP showed much weaker associations. In agreement with NMR data, these results indicate that recognition of H3K\textsubscript{36}me\textsubscript{3}s is the key requirement for PHF1 Tudor to bind the nucleosome; however, additional contacts with the NCP beyond those with the histone tail contribute to the complex’s formation.
5.2.3 Tudor Associates With dsDNA

To explore whether the PHF1 Tudor domain makes contacts other than to the histone tail sequence in the nucleosome, we tested its ability to bind DNA. Titration of a 10 bp dsDNA fragment into $^{15}$N-labeled Tudor induced substantial chemical shift changes in the protein (Figure 5.4a, top). In contrast, titration of a single strand of the same DNA fragment did not lead to resonance perturbations, indicating that this interaction is specific for dsDNA and likely depends on the presence of a major/minor groove (Figure 5.4a, middle). Analysis of the chemical shift perturbations (CSPs) afforded a $K_d$ of 201 $\mu$M for binding of the Tudor domain to the 10 bp dsDNA. Furthermore, the 601 DNA sequence, which we used to reconstitute
the H3K$_{C36me3}$-NCP, caused large CSPs in the Tudor domain, which were generally similar in direction to those seen upon addition of the 10 bp dsDNA; however, the presence of ~14 major/minor grooves in the 601 DNA construct, and thus non-stoichiometric association, precluded a straightforward quantitative analysis of this interaction (Figure 5.4a, bottom). Nevertheless, the pattern of CSPs inferred that the Tudor domain binds to dsDNA in a non-specific manner. A plot of the resonance perturbations as a function of Tudor residue revealed that the residues involved in the interaction with DNA lay along the $\beta$-barrel sides and around the H3K36me3s binding pocket (Figure 5.4b, c). Subsequent titration of the H3K36me3s peptide into the 10 bp dsDNA-bound Tudor domain caused additional CSPs in the protein, and the intermediate exchange regime indicated a stronger interaction with H3K36me3s (Figure 5.4d). Together these data suggest that within the intact NCP the Tudor domain can concomitantly interact specifically with the methylated histone H3 tail as well as non-specifically with the nucleosomal DNA.

5.2.4 The Model of the PHF1 Tudor/H3K$_{C36me3}$-NCP Complex

We generated a model of the PHF1 Tudor domain bound to the H3K$_{C36me3}$-NCP using the docking program HADDOCK. A three-body docking protocol utilizing Tudor/H3K36me3s
(PDB ID: 4HCZ), the histone octamer with H3 cut at His39, and DNA (PDB ID: 3LZ0) was applied. A total of ∼64,000 restraints between the octamer and DNA were used, and Pro38 of the H3K36me3s histone peptide was constrained to a bonding distance with His39 of the octamer either near the entry point of DNA or the exit point of DNA. Docking calculations of Tudor at either entry/exit point yielded very similar results, and comparison of the two models suggests that two Tudor domains can simultaneously associate with the NCP symmetrically trimethylated at Lys36 on both H3 tails Figure 5.5.

The model suggests that Tudor is stabilized at the H3K<sub>C</sub>36me3-NCP through concomitant interactions with the trimethylated histone tail and the minor groove and backbone of the nucleosomal DNA Figure 5.5. The Tudor domain is positioned between the two gyres of the DNA superhelix with the hydrophobic patch and β1 and β2 strands aligned with the minor groove of the DNA helical segment nearest the dyad. In addition, there are a few contacts between the residues in and around the Lys36me3 binding pocket and the phosphate backbone of the entry/exit DNA. The model is in good agreement with CSPs observed upon titration of dsDNA and the H3K<sub>C</sub>36me3-NCP, and although the full histone tail was not included in the docking algorithm, the orientation of the residues N-terminal to Lys36me3 suggests that some of the additional CSPs seen upon titration of the H3K<sub>C</sub>36me3-NCP may be due to interactions between these histone residues and the Tudor domain in the context of the entire NCP.

We also note that during the docking calculation the DNA was extensively restrained on the octamer core. Our numerous attempts to examine the effect of the Tudor association on the NCP breathing dynamics by applying limited or no restraints at the entry/exit ends of DNA have failed. Nevertheless, modeling using the fully restrained NCP shows that the Tudor domain does not lock the DNA into a wrapped position at either orientation, implying that the entry/exit DNA can be readily lifted off the core.
Figure 5.6: FRET measurements of the influence of PHF1 Tudor on TF binding. (a) Diagram of the 601L DNA molecule. (b) A two-state model of TF binding to its target site within the nucleosome. (c) Crystal structure of the NCP (PDB ID: 1KX5) that indicates the positions of H3 (orange); MLA at H3K36C (dark red); Cy5 at H2AK119C (purple); Cy3 at the 5’ end of 601L (green); and the LexA target site (cyan). (d) Fluorescence spectra with Cy3–Cy5-labeled H3K36me3-NCPs with 3 μM of LexA and either 0 μM (black), 10 μM (red) and 600 μM (purple) of PHF1 Tudor. The Cy3 emission increases as the Cy5 emission decreases indicating a decrease in FRET efficiency. (e) FRET efficiencies of LexA titrations with H3K36me3 (black) and unmodified wt nucleosomes (red). FRET efficiencies were done in duplicate and fit to a non-cooperative binding curve with S_{1/2} of 3,000 ± 600 nM and 2,400 ± 900 nM for H3K36me3 and unmodified C110A nucleosomes, respectively. (f,g) FRET efficiencies of PHF1 Tudor titrations with (f) and without (g) 3 μM LexA and either H3K36me3 (black) or unmodified NCPs (red). Error bars represent a standard deviation based on three experiments (PHF1 titrations) or two experiments (LexA titrations).
5.2.5 Interaction With H3K<sub>C</sub>36me3 Stabilizes a More Open Form of NCP

What is the consequence of binding of the PHF1 Tudor domain to H3K36me3s in the NCP? The location of H3K36me3s in the nucleosome readily positions this mark itself, as well as the PHF1 Tudor docking to this PTM, to influence nucleosomal DNA unwrapping/rewrapping or DNA accessibility and binding of regulatory proteins to DNA sites wrapped onto the nucleosome. We therefore investigated the impact of recognition of H3K36me3s by the PHF1 Tudor domain on nucleosome unwrapping and binding of the DNA-specific repressor protein LexA using FRET Figure 5.6. H3K<sub>C</sub>36me3-NCPs were prepared with the 147 bp 601 nucleosome positioning sequence, this time with the LexA binding sequence replacing bases 8–27 (601L, Figure 5.6a–c). The Cy3 donor fluorophore was attached to the 5′ end of the 601L DNA molecule adjacent to the LexA site, and the Cy5 acceptor fluorophore was attached to H2A(K119C) (Figure 5.6c). This positions Cy3 in the proximity of one of the Cy5 fluorophores so that there is significant energy transfer from fully wrapped nucleosomes (Figure 5.6d), whereas a reduction in FRET efficiency implies a decrease in nucleosome wrapping [20, 65].

We first examined the effect of H3K<sub>C</sub>36me3 on nucleosome wrapping using LexA and Cy3–Cy5-labeled NCPs. The concentration of LexA required to bind 50% of the nucleosomes (S<sub>1/2</sub>) is inversely proportional to the probability the LexA target site being unwrapped from the nucleosome for LexA binding [65] (Figure 5.6b). Thus a reduction of S<sub>1/2</sub> induced by H3K<sub>C</sub>36me3 would imply an increase in partial nucleosome unwrapping. Upon titration of LexA into Cy3–Cy5-labeled unmodified NCP or H3K<sub>C</sub>36me3-NCP we observed a decrease in FRET, due to LexA binding and stabilization of the unwrapped state. We found the LexA S<sub>1/2</sub> to be 3,000 nM and 2,400 nM for H3K<sub>C</sub>36me3 and unmodified nucleosomes, respectively (Figure 5.6e). The similar S<sub>1/2</sub> values indicate that H3K<sub>C</sub>36me3 itself does not directly impact nucleosome unwrapping.

We next titrated Tudor into 50 nM of unmodified NCP or H3K<sub>C</sub>36me3-NCP in the presence of 3 µM (S<sub>1/2</sub>) LexA. We found that above 30 µM of PHF1 Tudor, the FRET efficiency is
substantially reduced for nucleosomes containing H3K36me3, whereas the FRET efficiency for unmodified nucleosomes remains unchanged (Figure 5.6f). These data indicate that Tudor binding to H3K36me3 disturbs the nucleosome, facilitating LexA binding to its target sequence buried within the nucleosome. This ability to facilitate LexA binding was further enhanced when a larger PHF1 construct, containing the Tudor domain and adjacent PHD1 finger, was examined (Figure 5.7a). Because the PHD1 finger has no effect on binding of the Tudor domain to H3K36me3s (Figure 5.7b) and shows no histone-binding activity [126], these results suggest that the increase in size of the NCP-interacting macromolecule produces the greater effect. Addition of PHF1 Tudor to the H3K36me3-nucleosomes in the absence of LexA resulted in no significant change in the FRET efficiency, implying that the presence of the DNA-binding target protein is essential for the enhancement of the DNA accessibility (Figure 5.6g). Together, our findings support the previously proposed idea that TF binding to a target site may expedite TF binding to a second target site further into the nucleosome [25, 61], and provide the first example of such a heteromolecular augmentation (Figure 5.7c).

5.3 Methods

5.3.1 DNA Constructs and Protein Purification

The wt PHF1 Tudor domain (residues 14–87 or residues 28–87) and Tudor-PHD1 (residues 14–140) were cloned from full-length hPHF1 (obtained from Open Biosystems). Point mutant Y47A was generated by site-directed mutagenesis using the Stratagene QuickChange XL kit. Wt and mutant proteins were expressed in *Escherichia coli* BL21(DE3) pLysS cells grown in lysogeny broth (LB) or 15NH4Cl-supplemented M9-minimal media (grown in D2O for purposes of NMR nucleosome binding studies) and induced with isopropyl-β-D-1-thiogalactopyranoside (IPTG). Bacteria were harvested by centrifugation and lysed by sonication. The unlabeled, 15N-labeled and 2D/15N-labeled GST-fusion proteins were purified using glutathione Agarose 4B beads (Fisher). The GST tag was either cleaved with
Figure 5.7: (a) FRET efficiencies of the PHF1 Tudor (black) and Tudor-PHD1 (orange) titrations with H3K36me3 nucleosomes in the presence of 3 μM LexA. Error bars represent a standard deviation based on three experiments. (b) Superimposed $^1$H, $^{15}$N HSQC spectra of PHF1 Tudor-PHD1 recorded as H3K36me3 peptide was titrated in. An almost identical pattern of chemical shift changes was observed in the Tudor domain linked to the PHD1 finger, as compared with the changes seen in this domain alone (see [124]). Arrows indicate the most notable changes. Resonances corresponding to the PHD1 finger were unperturbed, implying that PHD1 is not involved in this interaction. (c) Four-state model of PHF1 Tudor and LexA binding to nucleosomes.
Prescission protease, or left for the purposes of pulldowns, in which case the GST-fusion protein was eluted off the glutathione Agarose beads using 0.05 M reduced L-glutathione (Sigma Aldrich).

*Xenopus* histone proteins H2A, H2B, H3(C110A), H3(C110A, K36C) and H4 were expressed in *E. coli* BL21(DE3) pLysS cells grown in 2×YT media and induced with IPTG. Bacteria were harvested by centrifugation and lysed by sonication. Proteins were extracted from inclusion bodies, purified over ion exchange resin and lyophilized.

Thirty-two repeats of the 601 Widom sequence were cloned into the pJ201 plasmid. The plasmid was purified in high yield primarily following the protocol outlined in [141]. The individual sequences were released from the plasmid using EcoRV and purified away from the parental plasmid using polyethylene glycol precipitation [142].

### 5.3.2 MLA Generation

Histone H3(C110A,K36C) point mutant was generated by site-directed mutagenesis using the Stratagene QuickChange XL kit and purified as described above. The histone H3K36me3 was generated by alkylation of H3K36C with (2-bromoethyl) trimethylammonium bromide following the protocol outlined in [84] and Appendix C. After desalting, the protein was dialysed into water and re-lyophilized.

### 5.3.3 HADDOCK Modeling

HADDOCK modeling was performed via the webserver interface [143, 144]. The crystal structure Tudor in complex with the H3K36me3 peptide (PDB ID: 4HCZ) including peptide residues 31–38 was docked onto the 601-NCP crystal structure (PDB ID: 3LZ0) with the H3 tail cut at His39. A three-body docking protocol was followed, in which the HO was taken as a single chain, the nucleosomal DNA as the second chain and the Tudor in complex with H3K36me3 peptide as the third chain. Extensive unambiguous restraints were generated between the HO and DNA and the peptide residue Pro38 and the HO residue His39 were
restrained to a bonding distance. Semi-flexibility options for the DNA and HO were set to ‘none’, and the Tudor/peptide was set to ‘automatic’. The final model was chosen which had the best HADDOCK score and best fit for all restraints. In the final structure, a covalent bonding distance was enforced between Pro38 and His39. Calculations were run for Tudor at both of the histone H3 tails in the NCP.

5.3.4 Nucleosome Reconstitution

Equimolar ratios of H2A, H2B, H4 and either H3(C110A) or H3(C110A, K36me3) were mixed and refolded into 2 mM NaCl, 10 mM Tris pH 7.5 and 5 mM β-mercaptoethanol to form the HO [141]. The HO was purified by size exclusion chromatography on a sephacryl S-300 column. Purified HO was mixed at a 1.2:1 (DNA:HO) ratio with purified 601 DNA and desalted according to protocol [141] into a final buffer of 20 mM Tris pH 7.5, 150 mM NaCl and 5 mM β-mercaptoethanol. The nucleosome was further purified from free DNA by sucrose gradient.

5.3.5 Pulldown Assays

GST-fusion PHF1 Tudor was incubated with reconstituted H3K36me3- or unmodified nucleosomes in binding buffer containing 20 mM Tris (pH 7.5), 150 mM NaCl and 0.2% Triton x-100 for 2 h at 4°C ([128]). Glutathione-agarose beads were added and further incubated for 1 h at 4°C. Free nucleosome was washed away by washing 4 times with 20 mM Tris (pH 7.5), 275 mM NaCl and 0.2% Triton x-100. Bound nucleosome was detected by western blot using an anti-H3 antibody (1:3,000 dilution) (Abcam, AB1791). A control of nucleosome with beads alone (at the equivalent of a 0:10 molar ratio) was used to assess for non-specific binding. Blots were analyzed using ImageJ and normalized to account for non-specific binding. The apparent $K_d$ was calculated using the equation:

$$\text{Pixels} = \frac{\text{Pixels}_{\text{max}}}{2[P]} \left( ([L] + [P] + [K_d]) - \sqrt{([L] + [P] + [K_d])^2 - 4[P][L]} \right)$$
where \([L]\) is concentration of the nucleosome, \([P]\) is the concentration of Tudor, Pixels is the normalized pixel count calculated in ImageJ and Pixels\(_{\text{max}}\) is the count at saturation.

For control, GST pulldown with H3K\(_C\)36me3-NCP was analyzed using western blot, which was first probed with an anti-H3 antibody (Abcam ab1791) and then stripped with strip buffer (1.5% (w/v) glycine, 0.1% (w/v) SDS, 1% Tween-20, pH 2.2) for 1 h, re-blocked, and probed with anti-GST-horseradish peroxidase antibody (GE Life Sciences RPN1236V) overnight.

### 5.3.6 NMR Spectroscopy

NMR experiments were collected on Varian 900 MHz and 600 MHz spectrometers equipped with a cryogenic probes at the University of Colorado School of Medicine NMR core facility. \(^1\)H, \(^{15}\)N HSQC or TROSY-HSQC experiments were carried out at 298 K (peptide and 10 bp DNA) or 310 K (nucleosome and 601 DNA) on \(^{15}\)N- or \(^2\)D/\(^{15}\)N-labeled Tudor (wt or mutant) or Tudor-PHD1 in 20 mM Tris pH 6.8 and 150 mM NaCl. Spectra were recorded in the presence of increasing concentrations of H3K36me3 peptide (synthesized by the University of Colorado Peptide Core Facility), 10 bp ssDNA, 10 bp dsDNA, 601 DNA or reconstituted nucleosomes. \(K_d\) value for the Tudor-dsDNA interaction was calculated by a nonlinear least-squares analysis in Kaleidagraph using the equation:

\[
\Delta\delta = \frac{\Delta\delta_{\text{max}}}{2[P]} \left( ([L] + [P] + [K_d]) - \sqrt{([L] + [P] + [K_d])^2 - 4[P][L]} \right)
\]

where \([L]\) is the concentration of dsDNA, \([P]\) is the concentration of the protein, \(\Delta\delta\) is the observed normalized chemical shift change and \(\Delta\delta_{\text{max}}\) is the normalized chemical shift change at saturation, calculated as

\[
\Delta\delta = \sqrt{\Delta\delta_1^2 + \Delta\delta_{15\text{N}}}^2
\]

where \(\delta\) is the chemical shift in p.p.m. in \(^1\)H and \(^{15}\)N dimensions.
5.3.7 Nucleosome and LexA Preparation for FRET Measurements

Nucleosomal DNA, 601L was prepared by polymerase chain-reaction (PCR) from a plasmid containing the LexA binding site at bases 8–27 with the Cy3-labeled oligonucleotide, Cy3-CTGGAGATACTGTATGAGCATACAGTACAATTGGTC and the unlabeled oligonucleotide, ACAGGATGTATATATCTGACACGTGCCTGGAGACTA. The Cy3-labeled oligonucleotide was labeled using a Cy3-N-hydrosyssuccinimide (NHS) ester (GE Healthcare) at a 5′ amino group and purified by reverse phase HPLC on a 218TPTM C18 (Grace/Vydac) column. *Xenopus* histones were expressed and purified as previously described [141]. All H3 histones contained the H3C110A mutation. LexA was expressed and purified by known methods [145] described in Appendix A.

H3K_c36me3 was prepared as reported [23, 84] and described in Appendix C. Briefly, recombinant H2AK119C and H2B were refolded into heterodimer and then labeled with Cy5-maleimide (GE Healthcare). Heterodimer was first reduced in 10 mM tris(2-carboxyethyl)phosphine (TCEP) pH 7.1 for 0.5 h, and then dialysed against 5 mM sodium PIPES pH 6.1, 2 mM NaCl. After reclaiming, heterodimer was purged under argon at 4°C for 30 min, and 2 mM HEPES pH 7.1 was bubbled with argon at 4°C for 5 min. The Cy5-maleimide was dissolved in anhydrous dimethylformamide to a concentration of 22 mM and added to heterodimer in a 7.5-fold molar excess. Sample was gently rotated for 1 h at room temperature and overnight at 4°C. The heterodimer was then purified by gel filtration on a Superdex 200 column. The labeling efficiency for the reaction was 88% as measured by ultraviolet–visible absorption. H3K_c36me3 and unmodified H3 were separately refolded with H4 into tetramer. H3–H4 tetramer in 0.5 mM potassium phosphate (pH 7.5) with 2 mM NaCl was then combined with 10% excess Cy5-labeled H2AK119C-H2B heterodimer in 5 mM pipes (pH 6.1) with 2 mM NaCl and incubated at 4°C overnight to form histone octamer. The resulting HO was purified by gel filtration on a Superdex 200 column. Nucleosomes were reconstituted by double dialysis with 10% excess 601L DNA into 0.5 mM potassium phosphate (pH 7.5) with 1 mM benzamidine (BZA). Reconstituted nucleosomes were purified by 5–30% sucrose gradient.
5.3.8 FRET Measurements

All FRET efficiency measurements were determined from spectra acquired by a Horiba Scientific Fluoromax 4. Samples were excited at 510 and 610 nm and the photoluminescence spectra were measured from 530 to 750 nm and 630 to 750 nm for donor and acceptor excitations, respectively. Each wavelength was integrated for 1 s, and the excitation and emission slit width were set to 5 nm with 2 nm emission wavelength steps. FRET measurements were computed through the (ratio)\(A\) method [103].

LexA titrations were carried out in 75 mM NaCl, 0.1 mM potassium phosphate pH 7.5, 11.5 mM Tris-HCl pH 7.5, 0.00625% IGEPAL and 0.00625% TWEEN20 with 10 nM nucleosomes. PHF1 Tudor and Tudor-PHD1 titrations were carried out with or without 3 µM LexA in 75 mM NaCl, 62.5 µM potassium phosphate pH 7.5, 15.25 mM Tris pH 7.5, 0.00625% IGEPAL, 0.00625% TWEEN20 with 50 nM nucleosomes. The nucleosome concentration was increased to 50 nM to compensate for increased background caused by the addition of PHF1. The change in nucleosome concentration does not impact the FRET measurements because the (ratio)\(A\) method is insensitive to fluorescence intensity [103], and does not impact LexA and PHF1 Tudor binding as both concentrations are significantly below the microMolar concentrations required for LexA and PHF1 Tudor binding. FRET values in each titration were normalized to the FRET efficiency in the absence of the titrant. Titrations were fit to

\[
E = \frac{E_f - E_0}{1 + \frac{S_{1/2}}{C}} + E_0,
\]

where \(E\) is the FRET efficiency at concentration \(C\) of the titrant, \(E_0\) the efficiency in the absence of the titrant, \(E_f\) the efficiency at high titrant concentration and \(S_{1/2}\) is the inflection point. Errors in Figure 5.6 and Figure 5.7 represent a standard deviation based on three experiments (PHF1 titrations) or two experiments (LexA titrations). Errors for fits represent 68% confidence bounds.
5.4 Discussion

In the past few years, our knowledge of how reader domains bind PTMs on histone tails has grown significantly; however, very few studies have investigated the mechanism of these associations in the context of the full nucleosome [120, 127–130, 146]. We have previously determined the structural basis for the interaction of the PHF1 Tudor domain with a H3K36me3s peptide [124]. In this work, we characterize binding of PHF1 Tudor to the intact H3KC36me3-NCP. We find that in addition to its histone-binding capabilities, PHF1 Tudor interacts non-specifically with a dsDNA. The multiple interactions may account for the higher affinity of Tudor for the H3KC36me3-NCP as compared with its affinity for the histone H3K36me3s peptide.

Dual association with H3K36me3s and nucleosomal DNA has recently been reported for the PWWP domain of lens epithelium-derived growth factor (LEDGF/P75) ([128, 129]). Whereas PWWP binds to a DNA fragment or the H3K36me3s peptide with low affinities (K$d$ of 150 µM and 11–17 mM, respectively), bipartite binding to the H3KC36me3-nucleosome results in a strong interaction (K$d$ of 1.5 µM) [128, 129]. Particularly, interaction with nucleosomal DNA was found to be responsible for a 104-fold enhancement in binding affinity and drives the recognition of the H3KC36me3-NCP [128, 129]. Much like PWWP, the PHF1 Tudor domain binds to the H3KC36me3-NCP with a 1.3 µM affinity; however, its tight association with the H3K36me3s peptide alone (K$d$ of 36 µM) suggests that the recognition of methylated Lys36 is the driving force for the Tudor–H3KC36me3-NCP interaction.

Nucleosomes have been shown to undergo spontaneous conformational fluctuations, in which a stretch of the entry/exit point DNA lifts off the histone core, providing transient access to the occluded regions of DNA for DNA-binding regulatory proteins and protein complexes [20, 75, 76, 140]. The unwrapping and following rewrapping processes are fast, with spontaneous opening occurring within ~250 ms, and are widespread, with the DNA ends being unwrapped in 1–5% of NCPs at any point [20, 75, 140]. On average, the NCP stays unwrapped only ~10–50 ms before rewrapping [75]. Within this short interval, some regulatory proteins are
able to bind their respective target DNA sequences, but this time is not sufficient for RNA Polymerase II (Pol II) to advance by even 1 bp [75]. It takes on average ∼6–7 s for Pol II to elongate through the nucleosome, or ∼2 s to process a ∼40–50 bp linker DNA [75]. Consequently, the NCP undergoes many cycles of rapid unwrapping and rewrapping before Pol II fully moves onto the nucleosome.

Our data demonstrate that interaction of the PHF1 Tudor domain with the H3K_{36}me3-NCP facilitates binding of LexA to its target DNA sequence buried within the nucleosome, which is inaccessible in the fully wrapped NCP. This indicates a shift in equilibrium towards a more open form of the nucleosome and hindering the rewrapping event in the presence of the DNA-binding protein. HADDOCK analysis of the Tudor–H3K_{36}me3-NCP interaction suggests a mechanism for this shift. Because the Tudor domain binds H3K36me3s between the two gyres of the DNA superhelix and does not lock the DNA entry point, once the DNA lifts off, the bound PHF1 Tudor may sterically preclude rapid DNA rewrapping, thus increasing the DNA accessibility to its ligand protein. As the H3K36me3s mark is found in actively transcribed gene bodies, it will be interesting to explore the effect of the PHF1–H3K36me3s interaction on the elongation activity of Pol II and whether this effect can be altered by chromatin fiber density and nucleosome positioning. Furthermore, it will be important to establish the significance of recognition of one H3K36me3s mark or two H3K36me3s marks on a single NCP. Last, future studies are also necessary to determine whether the nucleosomal changes imparted by Tudor association with H3K36me3s may have a role in PHF1 functions, particularly in modulating the methyltransferase activity of PRC2 and mediating PHF1 accumulation at DNA damage sites.
Chapter 6

PHF1 Tudor and N-terminal Domains Synergistically Target Partially Unwrapped Nucleosomes to Increase DNA Accessibility

Reproduced with modifications from:
Matthew D. Gibson, Jovylyn Gatchalian, Andrew Slater, Tatiana G. Kutateladze, Michael G. Poirier. (2016) PHF1 Tudor and N-terminal domains synergistically target partially unwrapped nucleosomes to increase DNA accessibility. NAR (In Review)

The Tudor domain of human PHF1 recognizes trimethylated lysine 36 on histone H3 (H3K36me3). PHF1 relies on this interaction to regulate PRC2 methyltransferase activity, localize to DNA double strand breaks, and mediate nucleosome accessibility. Here we investigate the impact of the PHF1 N-terminal domain (NTD) on the Tudor domain interaction with the nucleosome. We show that the NTD is partially ordered when it is natively attached to the Tudor domain. Through a combination of FRET and single molecule studies, we find that the increase of DNA accessibility within the H3K36me3-containing nucleosome, instigated by the Tudor binding to H3K36me3, is dramatically enhanced by the NTD. We demonstrate that this nearly order of magnitude increase is due to preferential binding of PHF1 to partially unwrapped nucleosomes, and that PHF1 alters DNA-protein binding within
the nucleosome by increasing dissociation rates. These results highlight the potency of a PTM-binding protein to regulate DNA accessibility and underscores the role of the novel mechanism by which nucleosomes control DNA-protein binding through increasing protein dissociation rates.

6.1 Introduction

The organization of eukaryotic DNA into nucleosomes sterically occludes DNA-binding complexes that regulate DNA processing including transcription [34], replication [147] and repair [148]. PTMs of the histone proteins that wrap DNA into nucleosomes regulate numerous aspects of nucleosome function [148]. There are over 500 different histone PTMs [2, 48] located throughout the nucleosome, which can function individually or in combination [49]. Histone PTMs function by two general mechanisms: (i) the histone code mechanism [149] where a histone PTM provides a specific binding site for recruiting chromatin modifying complexes including chromatin remodelers and histone modifying enzymes, and (ii) the nucleosome dynamics mechanism [21, 27] where a histone PTM directly alters chromatin properties including chromatin compaction, nucleosome unwrapping and nucleosome stability. Recently, we reported that these mechanisms can function in combination where the binding of the Tudor domain of PHF1 to H3K36me3 increases nucleosomal DNA accessibility [64]).

PHF1 is a 456 aa protein (Figure 6.1A) that is involved in transcriptional regulation [150, 151] and DNA repair [123]. It contains an N-terminal domain (aa 2–28) of unknown function, a Tudor domain (aa 28–87) that specifically recognizes H3K36me3 [124], and two PHD domains (aa 87–140 and aa 189–240) that may facilitate interaction with Enhancer of zeste homolog 2 (EZH2) [152], the methyltransferase subunit of Polycomb Repressive Complex 2 (PRC2) [153, 154]. Interaction with H3K36me3 stabilizes PHF1 at DNA double strand breaks together with the DNA repair complexes PARP1 and Ku70-Ku80 [124]. Separately, binding of PHF1 Tudor to H3K36me3 results in a reduction of H3K27 trimethylation by the gene silencing complex PRC2 [124, 155, 156]. Our finding that binding of PHF1 Tudor to H3K36me3 increases DNA accessibility [64] suggested that PHF1 may function to facilitate
Figure 6.1: (A) Diagram showing the known domains in PHF1. (B) $^1$H, $^{15}$N HSQC spectral overlays for PHF1 2–28 (teal), PHF1 14–87 (orange), and PHF1 2–87 (black). (C) $^1$H, $^{15}$N HSQC spectral overlays of wild type PHF1 NTD-Tudor upon titration with H3K36me3 peptide. The spectra are color-coded according to the protein:peptide molar ratio.
transcription or prevent spreading of repressive chromatin by more than one mechanism. However, the mechanistic basis by which PHF1 Tudor increases accessibility and how additional PHF1 domains influence this function remains unknown.

Here, we report on the cooperativity between the NTD and the adjacent Tudor domain of PHF1 in the regulation of nucleosome accessibility. We used NMR to find that the NTD is partially structured when linked to the neighboring Tudor and that the NTD does not enhance Tudor binding to H3K36me3. We used ensemble FRET measurements to determine that the combination of NTD with the Tudor domain, PHF1(2–87) (or NTD-Tudor), dramatically increases nucleosome accessibility by nearly an order of magnitude. We demonstrated this increase is due to an 8-fold preference of PHF1 NTD-Tudor to bind partially unwrapped nucleosomes. Using single molecule FRET (smFRET) measurements, we found that the increase of DNA accessibility is due to a reduction of DNA-binding protein dissociation rate instead of an increase in DNA unwrapping that would increase the DNA-protein binding rate. This finding highlights our recently reported novel mechanism for the regulation of nucleosome accessibility by altering DNA-binding protein dissociation. Importantly, these studies reveal that PTM binding domains can function not only to recruit their host proteins and protein complexes to chromatin, but can also directly regulate nucleosome dynamics.

6.2 Results

6.2.1 The PHF1 NTD is Partially Structured When Linked to Tudor.

PHF1 contains a 28-residue tail N-terminal to the Tudor domain, for which the structure and biological role have not been investigated. To examine the structural conformation of the PHF1 NTD, we generated an isolated NTD construct, produced uniformly 15N-labeled protein, and recorded its $^1$H, $^{15}$N HSQC spectrum. A low dispersion of amide resonances in the spectrum, particularly in the $^1$H dimension, indicated that an isolated N-terminal tail of PHF1 is largely disordered (Figure 6.1B, green). However, when NTD was physically
Figure 6.2:  (A) Superimposed $^1$H, $^{15}$N HSQC spectra of the PHF1 NTD-Tudor (2–87) region collected upon titration with H3K36me3 peptide. Spectra are color coded according to the protein:peptide molar ratio (inset). The pattern of chemical shift changes is essentially identical to the pattern of changes observed upon titration of H3K36me3 peptide to the PHF1 Tudor(14–87) domain [124], indicating that the NTD does not influence this activity of Tudor. (B) The structure of the PHF1 Tudor (pink) in complex with H3K36me3 peptide (yellow). The N-terminus of the Tudor domain is indicated.

linked to the Tudor domain in the construct containing residues 2–87 of PHF1, a significant dispersion of all amide resonances was observed, indicating that the entire NTD-Tudor region (Figure 6.1B, black spectrum) becomes structured. A substantial overlap of resonances in $^1$H, $^{15}$N HSQC spectra of PHF1(2–87) and PHF1(14–87) suggested that Tudor can promote a more rigid conformation even in the partially truncated NTD (Figure 6.1B compare orange and black spectra).

To determine the effect of the NTD on the interaction of the Tudor domain with H3K36me3, we carried out NMR titration experiments. Gradual addition of the H3K36me3 peptide (residues 31–40 of H3) to NTD-Tudor(2–87) resulted in large changes in chemical shifts of the protein (Figure 6.1C). As expected, resonances of the Tudor residues were perturbed utmost, particularly those involved in the formation of the H3K36me3-binding pocket, such as Y47 and E66. However, resonances of several NTD residues, including D17 and W18, which were previously assigned in PHF1(14–87) [124], were also perturbed. The pattern of chemical shift changes was essentially identical to that of seen upon titration of H3K36me3 peptide into the PHF1(14–87) sample, indicating that the binding affinity of Tudor for H3K36me3 is not affected by the presence of the entire NTD (Figure 6.2A). Because the N-
terminal region of Tudor along with the adjacent NTD are positioned on the side opposite to the H3K36me3-binding site of Tudor (Figure 6.2B), the observed perturbations likely point to a conformational change that may accompany the interaction rather than a direct involvement of the NTD in binding of the Tudor domain to H3K36me3.

6.2.2 PHF1 Tudor-NTD Does Not Shift Nucleosomes Into a Largely Unwrapped State.

PHF1-Tudor makes direct contacts with H3 residues 32–40 [124]. Since H3 residues 36–40 are located between the 2 DNA gyres in the DNA entry-exit region of the nucleosome, the nucleosome may need to partially unwrap for PHF1-Tudor to bind the H3K36me3 tail. Using FRET [103], we previously found that the Tudor domain alone, PHF1(28–87), does not induce detectable changes to nucleosome unwrapping [64]. Still, we considered the possibility that the PHF1-NTD could influence the binding of PHF1-Tudor on nucleosome unwrapping.

We prepared nucleosomes that contains the Widom 601 nucleosome positioning sequence [13] and a Cy3 label at one of the 5′ends (Figure 6.3A). The nucleosome is also labeled with the Cy5 fluorophore at H2A(K119C) (Figure 6.3B), so fully wrapped nucleosomes undergo efficient Cy3-Cy5 FRET. We prepared and purified Cy3-Cy5 labeled nucleosomes with and without a trimethyl-lysine mimic at lysine 36 of H3 (H3K36me3). This modification has been shown to accurately mimic trimethyl-lysines [84]. We then titrated PHF1(2–87), PHF1(14–87) and PHF1(28–87) with Cy3-Cy5 labeled nucleosomes. We find that each of these proteins do not have a significant effect on the FRET efficiency (Figure 6.4). This result indicates that binding by these PHF1 constructs does not shift the nucleosome into a predominately unwrapped state.
Figure 6.3:  (A) Diagram of DNA molecules used to reconstitute nucleosomes. DNA used for ensemble (top) and single molecule (bottom) measurements. (B) Crystal structure of the nucleosome core particle (PDB ID: 1KX5) that indicates the positions of H3 (orange); MLA at H3K36me3 (dark red); Cy5 at H2AK119C (purple); LexA binding site (teal). (C) FRET efficiencies of LexA titrations with H3K36me3 (red) and unmodified (black) nucleosomes. Titrations were performed in triplicate and fit to a noncooperative binding isotherm. (D,E) FRET efficiencies of PHF1(2–87) (black), PHF1(14–87) (purple), and PHF1(28–87) (blue) titrations with constant [LexA]–S_{1/2} to H3K_{36}me3 (D) or unmodified (E) nucleosomes. Error bars represent a standard deviation based on three experiments.

Figure 6.4: Titrations of PHF1(2–87) (left), PHF1(14–87) (middle), and PHF1(28–87) (right) with H3K_{36}me3 (black) and unmodified (red) nucleosomes. Direct changes in FRET are not observed so any unwrapping caused by PHF1 binding is too small to be detected directly with this technique.
Figure 6.5: Emission spectra for Cy3/Cy5 FRET with 0 (black), 300 nM (blue), or 3 µM (red) LexA. The peak at 570 nm corresponds to Cy3 emission which increases with increasing [LexA] and the 670 nm peak corresponds to Cy5 emission which decreases with increasing [LexA]. This leads to the decrease in FRET efficiency calculated in Figure 6.3C.

6.2.3 The PHF1 NTD Enhances the Impact of the Tudor Domain on Nucleosome Accessibility.

We previously determined that the PHF1-Tudor domain increased DNA accessibility within the nucleosome even though it does not shift the nucleosome into a predominately unwrapped state [64]. We therefore investigated if the PHF1-NTD influences the impact of PHF1 on DNA accessibility. To do this, we relied on the Cy3-Cy5 labeled nucleosomes where the LexA target site is inserted into the Widom 601 sequence from the 8th-27th base pairs, which positions the site near the Cy3 label. LexA can bind to its site when the nucleosome is partially unwrapped [20]. So, as LexA is titrated with Cy3-Cy5 labeled nucleosomes, the FRET efficiency reduces as LexA binds to its site trapping the nucleosome in a partially unwrapped state (Figure 6.3C and Figure 6.5). The normalized change in FRET efficiency (∆FRET) as a function of LexA concentration fits to a non-cooperative binding isotherm, \[ E = \frac{1}{1 + \frac{S_{1/2}}{[LexA]}} \]. [LexA] is the concentration of LexA and \( S_{1/2} \) is the concentration of LexA that changes the FRET efficiency by 50% of the total change. A reduction in the \( S_{1/2} \) indicates that the accessibility within the nucleosome has increased.

To investigate the influence of PHF1(2–87), PHF1(14–87) and PHF1(28–87), we prepared Cy3-Cy5 labeled nucleosomes that contained H3K_{36}me3 with LexA at a concentration equal to the measured \( S_{1/2} \) for binding the nucleosome. Therefore, without a PHF1 truncation the normalized ∆FRET due to LexA binding is 0.5. We then titrated each of the PHF1
truncations and measured the normalized ΔFRET (Figure 6.3D). If the PHF1 truncation does not influence LexA binding then the normalized ΔFRET remains at 0.5. We find that the normalized ΔFRET decreases as the concentration of PHF1 construct increases, which indicates that PHF1 increases nucleosome accessibility as we previously reported [64]. By comparing titrations with PHF1(2–87), PHF1(14–87) and PHF1(28–87), we find that the addition of the PHF1-NTD significantly increases the influence of PHF1 on DNA accessibility. In fact, the normalized ΔFRET saturates in the PHF1(2–87) titration fits to binding isotherm with an $S_{1/2} = 16 \pm 5\mu M$. These results confirm that PHF1-Tudor increases nucleosome accessibility and that the PHF1-NTD dramatically enhances this increase in accessibility.

6.2.4 The Increased Influence of PHF1-Tudor with the NTD on Nucleosome Accessibility is Specific to H3K$_C$36me3 Nucleosomes.

The influence of the PHF1-Tudor domain on nucleosome accessibility is highly specific to H3K$_C$36me3 [64]. Since the NTD significantly enhances the increase of nucleosome accessibility induced by PHF1-Tudor, we investigated if the combination of the PHF1 Tudor and N-terminal domains retained H3K$_C$36me3 specificity or if this enhancement was independent of this modification. To do this, we carried out separate PHF1(2–87), PHF1(14–87) and PHF1(28–87) titrations with unmodified Cy3-Cy5 labeled nucleosomes and a constant LexA concentration that is equal to the $S_{1/2}$ (Figure 6.3E). These titrations have a negligible shift in the normalized ΔFRET from 0.5, while H3K$_C$36me3 nucleosomes have a significant reduction of the normalized ΔFRET, as discussed above (Figure 6.3D). This indicates that the increase in nucleosome accessibility induced by the combination of PHF1 Tudor and NTD remains specific to H3K$_C$36me3 nucleosomes.
Figure 6.6: (A, B) FRET efficiencies of LexA titrations with (orange and blue) or without (red or black) PHF1 2–87 held at a constant 100 μM with H3K36me3 (A) or unmodified (B) nucleosomes. (C) LexA $S_{1/2}$ binding to H3K36me3 (orange) or unmodified (blue) with 100 μM PHF1 2–87 relative to no PHF1 2–87. (D) LexA $S_{1/2}$ shifts as a function of PHF1 2–87 concentration with H3K36me3 nucleosomes. LexA $S_{1/2}$ is normalized relative to the $S_{1/2}$ in the absence of PHF1 2–87.
6.2.5 PHF1 Tudor-NTD Increases the DNA-Binding Protein Occupancy Within the Nucleosomes by Nearly an Order of Magnitude.

After determining that the PHF1 Tudor with the NTD significantly increased the accessibility of H3K\textsubscript{36}me3 containing nucleosomes, we quantified the change in nucleosome accessibility. To do this, we carried out the LexA titrations with Cy3-Cy5 labeled nucleosomes that contain H3K\textsubscript{36}me3 and a constant PHF1(2–87) concentration of 100 \(\mu\text{M}\) (Figure 6.6A). We chose this concentration because the impact of PHF1(2–87) on nucleosome accessibility is nearly saturated (Figure 6.3D). This LexA titration determines the \(S_{1/2}\) of the LexA binding within H3K\textsubscript{36}me3 nucleosomes in the presence of PHF1(2–87). By comparing the \(S_{1/2}\) of LexA binding to nucleosomes with (\(S_{1/2}^{\text{K36me3}, \text{PHF1}(2–87)} = 0.2 \pm 0.05 \mu\text{M}\)) and without (\(S_{1/2}^{\text{K36me3}, \text{no PHF1}(2–87)} = 1.5 \pm 0.3 \mu\text{M}\)) PHF1(2–87), we determined the relative binding probability to be \(S_{1/2}^{\text{K36me3}, \text{no PHF1}(2–87)}/S_{1/2}^{\text{K36me3}, \text{PHF1}(2–87)} = 8 \pm 2\). This can be converted to a change in binding free energy by:

\[
\Delta \Delta G_{\text{PHF1}(2–87)} = -k_B T \ln \left( \frac{S_{1/2}^{\text{K36me3}, \text{PHF1}(2–87)}}{S_{1/2}^{\text{K36me3}, \text{no PHF1}(2–87)}} \right)
\]

\[
\Delta \Delta G_{\text{PHF1}(2–87)} = -2.0 \pm 0.2k_B T.
\]

\(k_B\) is the Boltzmann constant and \(T\) is room temperature. We controlled for LexA binding being directly impacted by PHF1(2–87) by detecting LexA binding to naked DNA with EMSA(Figure 6.7), and find that PHF1(2–87) does not impact LexA binding. Interestingly, PHF1(2–87) impact on nucleosome accessibility is larger than the changes induced by other histone PTMs in the nucleosome entry-exit region such as H3K56ac [23] and H3Y41ph [51], and DNA sequence changes in the first 7 base pairs of the nucleosome [66]. However, combinations of these factors such as both H3K56ac and H3Y41ph increase accessibility similarly to that of PHF1(2–87) [51].

Our results shown in Figure 6.3 indicate that H3K\textsubscript{36}me3 is required for the PHF1 dependent increase in nucleosome accessibility of 8 fold. To investigate this PTM depen-
Figure 6.7: DNA electromobility shift assay (EMSA) upon LexA binding with the PHF1 truncations used in this work. The concentration of DNA in each lane is 1 nM and concentrations of PHF1 are 0 (top left), 100 µM PHF1(2–87) (top right), 100 µM PHF1(14–87) (bottom left), 100 µM PHF1(28–87) (bottom right). LexA binding to naked DNA is unaffected by the presence of any of the three test PHF1 truncations.

dence further, we carried out LexA titrations with unmodified nucleosomes and either with (S1/2, PHF1(2–87) = 1.5 ± 0.7µM) or without (S1/2, no PHF1(2–87) = 2.0 ± 0.3µM) 100 µM PHF1(2–87) (Figure 6.6B). We find that with unmodified nucleosomes there is no significant change in the LexA S1/2, which confirms that the PHF1 induced changes in nucleosome accessibility requires H3K36me3.

6.2.6 The Entire PHF1 NTD is Necessary to Enhance DNA Accessibility Within the Nucleosome.

Our observation that PHF1(2–87) has a larger impact than PHF1(14–87) and PHF1(28–87) on nucleosome FRET efficiency with constant LexA (Figure 6.3D) suggested to us that the PHF1 NTD increases the impact of PHF1 on nucleosome accessibility. To investigate this we carried out LexA titrations with 100 µM PHF1(14–87) and 100 µM PHF1(28–87)
Figure 6.8: FRET efficiencies of LexA titrations with (orange and blue) or without (red or black) PHF1(14–87) (A, C) or PHF1(28–87) (B, D) held at a constant 100 µM with H3K_C36me3 (A, B) or unmodified (C, D) nucleosomes.
Figure 6.9: A four-state model of TF binding to its target site within the nucleosome and PHF1 binding to H3K36me3.

Figure 6.10: A two dimensional titration of LexA and PHF1 to nucleosomes containing H3K36me3 shows a clear decrease in the $S_{1/2}$ upon addition of PHF1 to titrations of LexA. The results displayed here are represented by each LexA titration’s $S_{1/2}$ in Figure 6.6D. Each bar represents the average of three experiments, standard deviations are not shown but are used to calculate the error bars of the $S_{1/2}$ represented in Figure 6.6D.

separately. We find that PHF1(14–87) and PHF1(28–87) decreases the LexA $S_{1/2}$ by $2.7 \pm 0.8\Delta G_{K36me3}^{PHF1(14-87)} = 0.99 \pm 0.30k_BT$, and $S_{1/2}$ by $2.8 \pm 0.9\Delta G_{K36me3}^{PHF1(28-87)} = 1.03 \pm 0.34k_BT$, respectively (Figure 6.6C and Figure 6.8). This implies that both PHF1 truncations have a similar 3-fold impact on DNA accessibility within the nucleosome. Given the similar impact of PHF1(14–87) and PHF1(28–87) strongly suggests that the entire NTD is required for it to impact nucleosome accessibility.

**6.2.7 A Four-State Binding Model Predicts that PHF1 Preferentially Interacts with Unwrapped Nucleosomes by Nearly an Order of Magnitude.**

To further understand how PHF1 impacts nucleosome accessibility, we carried out separate LexA titrations with a range of PHF1(2–87) concentrations (See Figures 6.10 and 6.6D).
We then modeled the combined interactions of PHF1(2–87) and LexA with nucleosomes using a 4 state model (Figure 6.9). We chose to use a four state model where we combine the unwrapped LexA-free nucleosome states with the unwrapped LexA-bound nucleosome states for three reasons. (i) We cannot differentiate between a unwrapped nucleosome with and without either LexA or PHF1(2–87) bound. (ii) The equilibrium between the fully wrapped nucleosome and the partially unwrapped nucleosome state is much less than 1 [20, 53, 75, 76]. So the unwrapped LexA-free states are low probability states and should not contribute significantly to the population of states. (iii) This is the minimal model for three interacting molecules.

This model predicts that the LexA $S_{1/2}$ changes as

$$S_{1/2} = \frac{K^L_D K^{PL/P}_D (K^P_D + [P])}{K^P_D K^{PL/P}_D + K^L_D [P]}$$

where $K^L_D$ is the LexA dissociation constant without PHF1 and $K^{PL/P}_D$ is the LexA dissociation constant with PHF1 bound to the nucleosomes. We fit the change in LexA $S_{1/2}$ for increasing PHF1(2–87) concentrations to this functional form and find that $K^{PL/P}_D / K^L_D = 8 \pm 2$ (similar to the ratio previously found) and that $K^P_D$, the PHF1(2–87) dissociation constant without LexA, is $20 \pm 10 \mu M$, which is similar to our previous measurement with PHF1(28–87). An important aspect of this 4 state model (Figure 6.9) is that the free energy difference between the nucleosome bound with LexA and PHF1(2–87) (State LP) and the unbound nucleosome state (State 0) must be the same whether LexA (State L) or PHF1(2–87) (State P) binds first. Therefore, $\Delta G_{PL-L} + \Delta G_{L-0} = \Delta G_{PL-P} + \Delta G_{P-0}$ and since $\Delta G = k_B T \ln(K_D)$ then

$$\frac{K^{PL/P}_D}{K^L_D} = \frac{K^{PL/L}_D}{K^P_D}.$$ 

This implies that the relative binding affinity of LexA to nucleosomes with and without PHF1(2–87) is equal to the relative binding affinities of PHF1(2–87) with and without LexA. We separately determined that PHF1(2–87) does not influence LexA binding to its target site within DNA (Figure 6.7), which controls for a direct interaction between PHF1(2–87) and LexA. Therefore, we conclude that PHF1(2–87) binds partially unwrapped nucleosomes.
Figure 6.11: (A) Schematic of a single molecule flow cell. (B) Representative single molecule FRET traces with 0 (top), 100 nM (middle), and 1 µM (bottom) LexA. The evolution of a low FRET state is due to LexA trapping the nucleosome in an unwrapped state. (C) Rate constants of LexA binding to H3K_C36me3 (on: orange, off: blue) and unmodified (on: red, off: black). (D,E) Comparison of on (D) and off (E) rates of LexA with no PHF1, with 30 µM PHF1(14–87), and with 30 µM PHF1(28–87). Rates are shown for H3K_C36me3 nucleosomes relative to unmodified nucleosomes. Off rates are fit to a flat line. On rates are fit to lines with 0 y-intercept. Error bars represent error in fits to cumulative sum distributions (see section 4.2.5).

8-fold higher than fully wrapped nucleosomes, which in turn significantly enhances DNA accessibility within the entry-exit region of the nucleosome.

6.2.8 The PHF1 Tudor Domain Regulates Occupancy of DNA-Binding Complexes Within the Nucleosome by Increasing the Dissociation Rate.

PHF1 could regulate DNA accessibility by two nonexclusive mechanisms. (i) PHF1 could increase the probability that a DNA site is exposed for binding, which would increase the
binding rate and occupancy. (ii) PHF1 could decrease protein dissociation from its site once it is bound. We recently showed that nucleosomes can dramatically increase transcription factor dissociation from their target sites within nucleosomes [22] increasing protein occupancy. To determine which of these mechanisms PHF1 uses to increase LexA occupancy within the nucleosome, we carried out smFRET [115]. We used the same nucleosome constructs that were used in the ensemble measurements except that we included a 75 base pair extension on the opposite side to where the LexA binding site is located (Figure 6.3A). Here we detect single LexA binding and dissociation events to and from nucleosomes that are tethered to a quartz surface (Figure 6.11A) by observing changes in smFRET (Figure 6.11B) [79]. By quantifying the dwell times in the high and low FRET states and then fitting the cumulative sum of the dwell time (section 4.2.5), we can determine the binding and dissociation rates of LexA. In the absence of PHF1, we find that with unmodified and H3K\textsubscript{C}36me3 nucleosomes the binding rate increases linearly as a function of LexA concentration and the dissociation rate is constant as we previously reported for unmodified nucleosomes [22]. In addition, the magnitude of the binding and dissociation rates without PHF1 are nearly identical (Figure 6.12). This implies that H3K\textsubscript{C}36me3 alone does not increase accessibility within the nucleosome, which is consistent with our ensemble FRET measurements (Figure 6.3C, [64]).

Figure 6.12: Rates of LexA binding and dissociation (on and off, respectively) to unmodified (red, black) or H3K\textsubscript{C}36me3 (orange, blue) nucleosomes without (left) or with 30 \(\mu\)M PHF1(28–87) (right). Off rates (black and blue) are fit to lines with zero slope and on rates to lines with zero y-intercept. Errors represent error in cumulative sum fits (see section 4.2.5).
We then carried out smFRET experiments with PHF1(2–87) to determine its impact on both LexA binding and dissociation rates. Unfortunately, we found that in the presence of PHF1(2–87) tethered nucleosomes were not stable (data not shown). This prevented us from using smFRET measurements to determine the impact of PHF1(2–87) on LexA binding dynamics. However, we did find that 30 µM PHF1(14–87) and PHF1(28–87) did not destabilize tethered nucleosomes. Therefore, we carried out smFRET measurements of LexA binding with 30 µM of PHF1(14–87) and PHF1(28–87), separately. With either PHF1(14–87) or PHF1(28–87), we find that the LexA binding rates are linear with increasing LexA and that the dissociation rates are constant (Figure 6.11C). We then compared the binding and dissociation rates of LexA with PHF1(14–87) or PHF1(28–87) to no PHF1. We find that the LexA binding rates are not impacted by either 30 µM PHF1(14–87) or PHF1(28–87) (Figure 6.11D), while the LexA dissociation rates were reduced by a factor of 2.0±0.2 and 1.7±0.2 with PHF1(14–87) and PHF1(28–87), respectively (Figure 6.11E). These results indicate that the PHF1 Tudor domain increases LexA occupancy by decreasing the dissociation rate. This highlights how controlling nucleosome induced protein dissociation can be used to regulate DNA accessibility within the nucleosome.

6.3 Methods

6.3.1 PHF1 Constructs and Purification

PHF1 14-87 and 28-87 were cloned as previously described [124]. PHF1 2-28 and 2-87 were generated by amplification of construct of interest from PHF1 complementary DNA (cDNA) (Open Biosystems) with BamHI and XhoI restriction sites and ligation into pGex6P1 expression vector. Proteins were expressed and purified, essentially as described [124].

6.3.2 Nuclear Magnetic Resonance

NMR experiments were recorded at 298 K on a Varian INOVA 600 MHz spectrometer equipped with a cryogenic probe. Chemical shift perturbation experiments were carried
out using 0.1 mM uniformly $^{15}$N-labeled protein in 50 mM Tris pH 6.9, 150 mM NaCl, 2 mM Dithiothreitol (DTT) with 5% D$_2$O. The $^1$H, $^{15}$N HSQC spectra were recorded in the presence of increasing concentrations of histone H3K36me3 (31–40) peptide (synthesized by the University of Colorado Denver Biophysics Core Facility.)

6.3.3 MLA Deposition on Histones

Histone H3.2(C110A, K36C) was generated by site-directed mutagenesis using a Stratagene QuickChange lightning kit. The histone was then expressed in *E. coli* and purified as previously described [157]. The MLA was deposited on the sole cysteine present in the mutant H3.2(C110A, K36C) following previously reported protocols [84] and described in Appendix C. Briefly, 5 mg of Histones are unfolded in 980 µL of Alkylation Buffer (1 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) pH 7.8, 4 mM Guanidine-HCl, 10 mM D/L-Methionine) for 1 h. After unfolding, histones were reduced in 6.66 mM DTT for 1 h at 37 ºC. 100 mg of (2-bromoethyl) trimethylammonium bromide was added to the histones and the reaction is allowed to proceed for 5 h at 50°C while covered and stirring. The reaction was quenched with 50 µL of 14.3 µM β-mercaptoethanol (BME). Histones were then dialysed against 3 mM BME, lyophilized, resuspended in water and dried via vacuum concentration. Labeling efficiency was confirmed by MALDI-TOF mass spectrometry.

6.3.4 Preparation of FRET-Labeled Histone Octamer

Human histones H2A(K119C), H2B, H3.2(C110A), H3.2(C110A, K36C), and H4 were expressed and purified as previously reported [157]. Octamer as then refolded as described in Appendix D. Briefly, histone heterodimer, H2A(K119C) and H2B, were then refolded separately from tetramer, H3.2(C110A) and H4 or H3.2(C110A, K36C) and H4, by dialysis from unfolding buffer (7 mM Guanidine-HCl, 10 mM Tris-HCl pH 7.5, 10 mM DTT) to refolding buffer (5 mM piperazine-N,N′-bis(2-ethanesulfonic acid) (PIPES), 2 mM NaCl).

Heterodimer was then labeled with Cy5-maleimide (GE Healthcare) and purified as previ-
ously reported [22, 23, 51, 64]. Heterodimer and tetramer were then combined to a molar ratio of 1:2.2 tetramer:heterodimer. The proteins were allowed to complex overnight at 4°C while gently rotating. The resulting octamer was then purified by size exclusion on a Superdex 200 column. Octamer was concentrated in an 30 kDa molecular weight cut-off (MWCO) Amicon Ultra (Millipore) and stored on ice.

6.3.5 Preparation of Labeled DNA

Nucleosomal DNA was prepared by PCR from a plasmid containing the 601 sequence with a LexA binding site located at bases 8–27. PCR primers for ensemble experiments were the Cy3-labeled oligonucleotide, Cy3-CTGGAGATACCTGTATGAGCATACAGTACAATTGGTC and the unlabeled reverse primer ACAGGATGTATATCTGACACGTGCCTGGAGACTA. Single molecule experiments used the biotinylated reverse primer Biotin-CGCA TGCTGCAGACGCGTTACGTATCG which extends the 147 bp 601 sequence with a 75 bp linker and provides the biotin attachment point for use in single molecule experiments. Cy3-labeled oligos were labeled with Cy3 NHS ester (GE healthcare) at an amino group attached to the 5’end of the DNA oligo. Primers are then HPLC purified on a 218TP C18 column (Grace/Vydac). After PCR, dsDNA molecules were purified by HPLC on a MonoQ (GE Healthcare) ion exchange column.

6.3.6 Preparation of Nucleosomes

Cy3-labeled DNA and Cy5-labeled histone octamer were combined and reconstituted by double salt dialysis and purified by 5–30% sucrose gradient as previously described [23]. Fractions containing nucleosomes were then collected and concentrated with a 30 kDa MWCO Amicon Ultra (Millipore).
6.3.7 Preparation of LexA

LexA was expressed and purified by known methods [145] and described in detail in Appendix A. Briefly, LexA is expressed in *E. coli*. LexA is separated from genomic DNA and the proteome by polyetheleneimine (Sigma) precipitation followed by salting out LexA with ammonium sulfate. LexA was resuspended in buffer B (20 mM Potassium Phosphate pH 7, 0.1 mM Ethylenediaminetetraacetic acid (EDTA), 10% glycerol, 1 mM DTT) + 200 mM NaCl and purified by a linear gradient to B + 800 mM NaCl over either a cellulose phosphate or HiTrap Heparin HP Column (Ge Healthcare, cellulose phosphate discontinued). Final LexA purification was performed over a Hydroxyapatite column and dialysed into storage buffer (10 mM PIPES pH 7.0, 0.1 mM EDTA, 10% \(v/v\) glycerol, 200 mM NaCl).

6.3.8 Ensemble FRET Measurements

Ensemble FRET efficiencies were determined from spectra acquired by a Horiba Scientific Fluoromax 4. Samples were excited at 510 and 610 nm and the fluorescence spectra were measured from 530 to 750 and 630 to 750 nm for donor and acceptor excitations, respectively. Each wavelength was integrated for 1 s, and the excitation and emission slit width were set to 5 nm bandwidth with 2 nm emission wavelength steps. FRET measurements were computed through the (Ratio)\(_A\) method (Equation 4.18, [103, 145])

\[
E = \frac{E_f - E_0}{1 + \frac{S_{1/2}}{C}} + E_0
\]

where \(E\) is the FRET efficiency at concentration \(C\) of LexA, \(E_f\) is the FRET efficiency at high concentration of titrant and \(E_0\) is the efficiency in the absence of the titration and \(S_{1/2}\) is the inflection point. Errors represent a standard deviation based on three experiments.
Fit errors represent 68% confidence bounds. Figure 6.3D is fit to the equation

$$S_{PHF1}^{1/2} = \frac{K_D^{PL/P}(K_D^P + [P])}{2K_D^P K_D^{PL/P} + (K_D^L + K_D^{PL/P})[P]},$$

and Figure 6.6D is fit to the equation

$$S_{LexA}^{1/2} = \frac{K_D^{L} K_D^{PL/P}(K_D^P + [P])}{K_D^P K_D^{PL/P} + K_D^{L}[P]},$$

where $S_{LexA}^{1/2}$ and $S_{PHF1}^{1/2}$ are the inflection points of LexA and PHF1 titrations respectively as predicted by the four state model presented. $K_D^P$, $K_D^L$, $K_D^{PL/P}$ are the dissociation constants of PHF1(2–87) to nucleosomes, LexA to nucleosomes, and LexA to a PHF1(2–87)-nucleosome complex, respectively.

### 6.3.9 Electromobility Shift Assay

DNA containing the LexA recognition site was incubated at 1 nM with 0–1000 nM LexA and no PHF1, 100 µM PHF1(2–87), 100 µM PHF1(14–87), or 100 µM PHF1(28–87) in 0.5×TE for 10 min at room temperature and resolved by EMSA with a native 5% polyacrylamide gel in 0.3×tris boric acid EDTA (TBE).

### 6.3.10 Single Molecule FRET Measurements

Single molecule measurements were performed as previously described in sections 4.2.3 and 4.2.5. Briefly, biotinylated sample nucleosomes were allowed to incubate in flow cells at room temperature for 5 min and washed out with imaging buffer containing the desired concentration of LexA and PHF1 truncation. Imaging buffer was 50 mM Tris pH 8.0, 75 mM NaCl, 10% v/v glycerol, 0.005% v/v TWEEN20, 0.1 µg/mL BSA, 2 mM 6-hydroxy-2,5,7,8-tetramethylchroman-2-carboxylic acid (Trolox), 0.0115% v/v Cyclooctatetraene, 0.012% v/v NBA, 1.6% w/v glucose, 450 µg/mL glucose oxidase, 22 µg/mL catalase.

Time series were selected and generated by custom software and fit to a two-state step function by the vbFRET [89] Matlab program. Idealized time series were further analyzed
by custom software to determine dwell-time distributions.

6.4 Discussion

In this study, we quantitatively investigated the combined influence of the PHF1 N-terminal and Tudor domains on nucleosome accessibility, and find that together, they increase accessibility by nearly an order of magnitude. We then made two key observations about how PHF1 increases accessibility. First, we find that PHF1(2–87) increases nucleosome accessibility due to preferential binding to partially unwrapped nucleosomes that contain H3K36me3. This appears to be the first report of a histone PTM-recognizing protein that shows selectivity for a structurally altered nucleosome state. The location of H3K36 is important for the observed effect because (i) this residue protrudes between two sections of the nucleosomal DNA molecule near the entry-exit region (Figure 6.3B) and (ii) the PHF1 Tudor domain interacts with amino acids 32–39 of the H3 tail. While modeling indicates that the PHF1 Tudor domain can bind to fully wrapped nucleosomes containing H3K36me3 [64], our results indicate that PHF1 preferentially binds partially unwrapped nucleosomes. Also, the orientation of PHF1 Tudor bound to a fully wrapped nucleosome positions the PHF1 NTD away from the nucleosome so it cannot interact directly with the nucleosome (Figure 6.2). This combined with our observation that the PHF1 NTD is important for increasing nucleosome accessibility suggests that the NTD likely interacts with the nucleosome only when it is unwrapped, which in turn enhances PHF1 binding.

This finding suggests that PHF1 binding could function synergistically with other chromatin factors that impact DNA accessibility. We and others have reported that histone PTMs in the DNA entry-exit region [23, 158–161] and DNA sequence [66, 72, 162, 163] can individually and in combination [51, 66] increase nucleosome accessibility. Perhaps, these PTMs and DNA sequence can function to regulate PHF1 binding to nucleosomes and/or function with PHF1 to regulate DNA accessibility. In addition, since some ATP dependent chromatin remodeling complexes partially unwrap nucleosomes [164], our results suggest that these remodeling complexes could also facilitate PHF1 binding to H3K36me3 contain-
ing nucleosomes. Finally, there are numerous histone PTM binding domains that recognize H3K36me3 [165]. Our findings suggest that other H3K36me3 binding proteins may also impact nucleosome accessibility and that other factors that regulate DNA unwrapping within the entry-exit region of the nucleosome could regulate H3K36me3 recognition within the nucleosome.

Our second key observation here is that PHF1 Tudor enhances nucleosome accessibility by decreasing the dissociation rate of a DNA-binding protein. We recently reported that nucleosomes regulate DNA-protein interactions by increasing the rate of protein dissociation from DNA by at least three orders of magnitude [22, 165]. This work appears to be the first reported chromatin regulator controlling nucleosome accelerated protein dissociation, demonstrating the potential for regulating this nucleosome function. It will be interesting to determine if other H3K36me3 binding proteins also regulate nucleosome accessibility by decreasing DNA-protein dissociation.

Finally, PHF1 is linked to transcription regulation and DNA repair, which both require enhanced DNA accessibility within chromatin. PHF1 can protect chromatin from transcriptional repression by PRC2 through its interaction with H3K36me3 [124, 155, 156, 166, 167], implying that PHF1 can facilitate transcription. PHF1 is also localized to chromatin sites containing a DNA double strand break together with the DNA repair regulators PARP1 and Ku70-Ku80 [124]. Future studies are needed to determine how the two separate PHF1 properties: (i) interacting with regulatory complexes such as PRC2 and PARP1 and (ii) enhancing DNA accessibility, function together and if these two properties are functionally linked.
Chapter 7

LEDGF binding protects nucleosomal DNA

The work discussed in this chapter was done in collaboration with Matt Plumb and Mamuka Kvaratskhelia, who provided all LEDGF/P75 molecules.

7.1 Introduction

LEDGF/P75, a PTM reader protein which specifically recognizes H3K$_{36}$me3, is essential for the integration of human immunodeficiency virus 1 (HIV-1) cDNA into the human genome [168–171]. Due to its targeting of H3K$_{36}$me3, LEDGF/P75 preferentially steers HIV-1 integration complexes to active regions of chromatin [169]. The C-terminal portion of LEDGF/P75 is specifically bound by lentiviral integrases and is termed the integrase binding domain (IBD). The biological role within humans for this domain is not well understood, however, the structural and mechanistic role of the LEDGF/P75-integrase interaction has been well studied [172, 173]. At the other end of LEDGF/P75 is a single PWWP domain which has been isolated and studied thoroughly in its own right [128, 174]. It is known to specifically bind H3K36me3 with an extremely high affinity in the 100s of nanoMolar [128]. Additionally, numerous studies investigating deletion and chimeric mutants have determined that LEDGF/P75’s PWWP domain is instrumental in determining the location of HIV-1 integration [175–177].
Despite the large amount of work which has gone into understanding HIV-1 integration, little work has attempted to understand the impact of LEDGF/P75 itself on chromatin and DNA accessibility. The high binding affinity of the LEDGF/P75 PWWP domain combined with the presence of AT-hooks to nonspecifically bind DNA is suggestive of the possibility to alter chromatin dynamics or conformation. Here, we investigate the impact of LEDGF/P75 binding on single nucleosomes which have been labeled with H3K36me3. We further demonstrate that LEDGF/P75 has a significantly higher affinity to H3.3K36me3 over H3.2K36me3. This distinction may further illuminate the pattern of HIV-1 integration seen on the genome.

7.2 Results

7.2.1 LEDGF Does not Directly Alter Nucleosome Dynamics

LEDGF/P75 is a 60 kDa protein which contains not only the PWWP domain which localizes it to active regions of the genome, but also two AT-hook domains which non-specifically bind DNA [178]. These two features could conceivably cooperate to trap the nucleosome in an altered state upon binding. We therefore prepared nucleosomes as described in chapter 3 which contained i) a fluorescence donor Cy3 molecule located on a 5′ end of the Widom 601 positioning sequence [13] ii) a fluorescence acceptor Cy5 ligated to the histone octamer at H2AK119C iii) a MLA label located at H3K36C producing a trimethyllysine analog to

![Figure 7.1: LEDGF does not shift the nucleosome to a significantly unwrapped state upon binding.](image)
serve as a binding target for LEDGF/P75’s PWWP domain and iv) a LexA recognition site located at bases 8–27 near the Cy3 donor fluorophore. In addition to these nucleosomes, we prepared nucleosomes which lacked the H3K36C to investigate behavior of LEDGF/P75 on unmodified nucleosomes.

To determine if the binding of LEDGF/P75 alters the nucleosome conformation enough to expose a significant amount of DNA, we performed a titration of LEDGF/P75 to i) H3K36me3 modified nucleosomes and ii) unmodified nucleosomes and observed FRET efficiency. Any change in energy transfer would indicate a shift in the equilibrium distance between fluorescence donor and acceptor and therefore a change in the level of wrapping of the nucleosome. As can be seen in Figure 7.1, no such change is visible over the concentrations which would be expected for LEDGF/P75 binding considering PWWP binding affinity has been measured at ~100 nM [128]. This, however, does not necessitate the conclusion that nucleosomes wrapping is not altered by LEDGF/P75 binding. As is shown in Figure 3.2, the predicted distance between fluorophore centers is much less than the Förster radius for Cy3-Cy5 FRET pairs (see Table 4.1). It is possible that a small shift in the end-to-end distance in either direction would produce no measurable change in FRET.

7.2.2 The Methylation of H3.3K36 Increases DNA Accessibility

Previously in chapter 5 we have seen that, despite the added bulk of trimethylation, H3K36me3 does not alter DNA accessibility within the nucleosome. Repeating the experiments here for use with LEDGF/P75 yields the same results (Figure 7.2A). However, when these experiments were repeated with H3.3K36me3, we detect a repeatable difference in accessibility of 1.7 ± 0.2-fold, corresponding to a free energy difference between bound, unmodified H3.3-containing nucleosomes and bound, H3.3K36me3-containing nucleosomes of 0.5 ± 0.1 k_BT.

As discussed in section 1.1.4, H3.3 is distinguished by only four aa mutations relative to H3.2. These mutations are i) A32S; ii) S88A; iii) V90I; and iv) M91G. Of these mutations, ii–iv are located in the histone octamer core and are located at the tetramer-heterodimer
interface. Only A32S is located at the tail and is likely to interact cooperatively with H3K36me3 to increase accessibility. Interestingly, it appears that neither of these alterations to H3.2 are sufficient to change accessibility as unmodified H3.3 produces a similar $S_{1/2}$ to unmodified H3.2: $1.1 \pm 0.1 \mu M$ and $1.2 \pm 0.2 \mu M$, respectively. Further experiments are necessary to acquire mechanistic detail in this apparent cooperative activity between these two epigenetic markers.

### 7.2.3 The Methyl-Recognizing PWWP Domain Does not Impact DNA Accessibility

In the case of PHF1, even the small methyl-recognition domain was sufficient to alter nucleosome dynamics. To test whether this was the case for LEDGF/P75, we purified the PWWP domain of LEDGF/P75 alone (aa 1–93). We then titrated increasing concentration of LEDGF/P75 PWWP against a constant concentration of nucleosomes and LexA and measured the efficiency of energy transfer between fluorophores on the nucleosomes. A change in the FRET efficiency indicates a conformational change due causing a change in LexA occupancy of its target site. LexA concentrations were kept at the $S_{1/2}$’s determined in section 7.2.2. The results seen in Figure 7.3 clearly show no detectable change in LexA occupancy due to PWWP binding, previously been determined to occur at $\sim 100 \text{ nM}$ [128].
Figure 7.3: PWWP does not significantly alter LexA occupancy to either nucleosomes containing H3.2 (A) or H3.3 (B) regardless of whether or not the K\textsubscript{C}36me3 modification is present.

Figure 7.4: Invasion of Nucleosomes containing either H3.2K\textsubscript{C}36me3 and H3.2K36 (A) or H3.3K\textsubscript{C}36me3 and H3.3K36 (B) by LexA is altered by LEDGF. While K\textsubscript{C}36me3 is associated with active transcription, LEDGF interaction decreases the ability of outside factors to simultaneously access nucleosomal DNA.

It appears that either i) LEDGF/P75 does not induce alterations in DNA accessibility or ii) the PWWP domain alone is insufficient to produce an effect.

7.2.4 Full Length LEDGF Protects Nucleosomes From Invasion by Transcription Factors

Unlike PHF1, it is possible to purify full length LEDGF/P75 [179–181]. We therefore repeated experiments from the previous section using full LEDGF/P75 instead of only the
methylation-recognition domain, PWWP. As shown in Figure 7.4A both modified and unmodified nucleosome respond to LEDGF/P75 binding with an increased FRET efficiency indicating a decrease in DNA accessibility to DNA. It is interesting that the effect is modification-independent indicating the binding orientation and affinity of LEDGF/P75 is insensitive to PWWP binding in the buffer conditions of the experiment (75 mM sodium chloride). Further experiments are necessary to determine whether this insensitivity remains in higher, more physiological monovalent salt concentrations.

7.2.5 LEDGF Binds to Nucleosomes Containing H3.3 with Roughly a 10-fold Higher Affinity

Recent findings have determined that another PWWP domain found in the ZMYND11 protein specifically targets nucleosomes containing both H3.3 and K36me3 [182]. We were interested in whether or not the LEDGF/P75 and its PWWP domain prefer binding to nucleosomes containing H3.3K\text{\textsubscript{36}}me3 over H3.2K\text{\textsubscript{36}}me3. As an initial experiment, LEDGF/P75 titrations were carried out with constant LexA and nucleosome background such that \([\text{LexA}] = S_{1/2}\). Figure 7.4 clearly displays a decreased LEDGF/P75 concentration necessary to produce similar decreases in accessibility while maintaining insensitivity to methylation state. This method of measuring LEDGF/P75 affinity is indirect and potentially inaccurate, however, Figure 7.4B displays an \(S_{1/2}\)-10-fold lower than Figure 7.4A. Further experiments utilizing a more direct measure of LEDGF/P75 binding are necessary to determine the difference in affinity between nucleosomes comprised of H3.2 and H3.3.

7.3 Methods

7.3.1 Protein Preparation

LexA used in this study was purified as detailed in Appendix A. Histones were prepared as previously described [157].
7.3.2 Preparation of Labeled DNA

DNA used in this study was prepared by PCR from a plasmid containing the 601 sequence with a LexA binding site located at bases 8–27. PCR primers for ensemble experiments were the Cy3-labeled oligonucleotide Cy3-CTGGAGATACTGTATGAGCATACAGTAATTGGTCA and the unlabeled reverse primer ACAGGATGTATATATCTGACACGTGCCTGGAGACTA. Cy3-labeled oligos were labeled with Cy3 NHS ester (GE healthcare) at an amino group attached to the 5′end of the DNA oligo. After PCR, dsDNA molecules were purified by HPLC on a MonoQ (GE Healthcare) ion exchange column.

7.3.3 Nucleosome Preparation

Cy3-labeled DNA and Cy5-labeled histone octamer were combined and reconstituted by double salt dialysis and purified by 5–30% sucrose gradient as previously described [23]. Fractions containing nucleosomes were then collected and concentrated with a 30 kDa MWCO Amicon Ultra (Millipore).

7.3.4 FRET Measurements

Ensemble FRET efficiencies were determined from spectra acquired by a Horiba Scientific Fluoromax 4. Samples were excited at 510 and 610 nm and the fluorescence spectra were measured from 530 to 750 and 630 to 750 nm for donor and acceptor excitations, respectively. Each wavelength was integrated for 1 s, and the excitation and emission slit width were set to 5 nm bandwidth with 2 nm emission wavelength steps. FRET measurements were computed though the (Ratio)A method (Equation 4.18, [103, 145])

\[ E = \frac{E_f - E_0}{1 + \frac{S_{1/2}}{c}} + E_0 \]

Titrations were carried out in 75 mM NaCl, 0.1 mM Potassium phosphate pH 7.5, 11.5 mM Tris-HCl pH 7.5, 0.00625% v/v Tween20. LexA titrations were fit to
where $E$ is the FRET efficiency at concentration $C$ of LexA, $E_f$ is the FRET efficiency at high concentration of titrant and $E_0$ is the efficiency in the absence of the titration and $S_{1/2}$ is the inflection point. Errors represent a standard deviation based on three experiments. Fit errors represent 68% confidence bounds.

7.4 Discussion

Thus far, in this study we have determined the direction of impact of LEDGF/P75 binding on nucleosomal DNA accessibility. We have yet to perform experiments which will allow for a quantitative measure of the change in DNA accessibility, which are planned in the near future. In contrast to the Tudor domain of PHF1 (chapters 5 and 6), the PTM-recognition domain of LEDGF/P75, PWWP, does not alter nucleosome dynamics. Both PHF1-Tudor and LEDGF/P75-PWWP bind both non-specifically to DNA [64, 128] and specifically to H3K36me3. However, the different results indicate a yet-unknown mechanistic difference between the two domains in the conformational requirements of binding H3K36me3 and the nucleosomal DNA. Structural work is required to elucidate these causes.

The increase in affinity to H3.3-containing nucleosomes may provide further insight into LEDGF/P75’s role in HIV-1 integration. It is thought that LEDGF/P75 targets integration to active regions of the genome through its interaction with H3K36me3. However, because it binds significantly tighter to nucleosomes containing H3.3K36me3, it may be reasonable to strengthen the targeting statement to regions marked as active by both H3.3 and K36me3. Mechanistic studies into the increased affinity of LEDGF/P75 to H3.3-containing nucleosomes will further develop this finding.

7.5 Future Directions

To solidify the initial findings described here further experiments are necessary. First, quantification and confirmation of the findings shown in Figure 7.4 is necessary. Additionally, mutations to W21 and A51 are known to significantly inhibit PWWP recognition of K36me3.
Titrations utilizing these mutations must be performed to investigate the impact on the PWWP domain on DNA accessibility and LEDGF/P75 targeting to H3.3.

The H3.3 variant has two distinct regions which are mutated relative to H3.2. These are a single mutation in the histone tail and three point mutations located in the heterodimer/tetramer interface. Chimeras are being developed and purified to identify which of these regions is responsible for the increased affinity of LEDGF/P75 to H3.3. Additionally, anisotropy measurements are being conducted to directly measure the binding of LEDGF/P75. These experiments will allow for the deconvolution of binding and accessibility—something we have not yet been able to do with any other PTM-binding protein. These experiments will prove to be an important crucible to any model describing LEDGF/P75 binding and DNA accessibility.
Chapter 8

Swi6 Can Function to Increase Accessibility to Nucleosomal DNA

8.1 Introduction

The assembly of DNA into nucleosomes is essential for compaction and regulation of the eukaryotic genome, but neither compaction nor regulation stops there. A complex system of mechanisms exists to further compact nucleosomes into higher order structures. The largest scale of these structures comprises the two basic forms of chromatin. Euchromatin is gene-rich, generally accessible, and heavily transcribed, whereas heterochromatin is largely gene-poor, condensed, and transcriptionally silenced. Heterochromatin assembly at repetitive elements such as centromeres and telomeres is common across all eukaryotes and has been shown to be critical to many important cellular processes such as chromosome segregation, recombination, nuclear organization, and long-range chromatin interaction [183–185]. Additionally, facultative heterochromatin located at specific genes is thought to be essential for heritable gene silencing during development [186–188].

A core feature of heterochromatin is methylation on Lysine 9 on Histone H3 (H3K9me3) and the specific targeting of this PTM by Heterochromatin Protein 1 (HP1) [189, 190]. In fact, often di- and tri-methylation of H3K9 is used to define heterochromatic regions [191]. The HP1-H3K9me3 interaction is thought to mediate many functions of heterochromatin
through the recruitment of a diverse set of regulators [189, 190, 192], facilitate heterochromatic silencing, and cause heterochromatin spread [120, 193, 194]. Although HP1 proteins were originally identified in *Drosophila melanogaster* through implication in position effect variegation (PEV) [195, 196], the erroneous spreading of heterochromatin to nearby euchromatin, it has become clear in recent years that HP1 proteins have many additional nuclear functions including transcriptional activation and elongation, sister chromatid cohesion, chromosome segregation, telomere maintenance, DNA repair, and RNA splicing [183, 197–210].

In fission yeast, *Schizosaccharomyces pombe*, two HP1 paralogs have been identified, Swi6 and Chp2. Swi6 is the more abundant of the two HP1 homologs [190, 211] and has been associated with many factors including histone deacetylases (HDACs) and, paradoxically, anti-silencing factors [9]. Here we investigate the impact of Swi6 binding on accessibility of DNA protected by its incorporation into a nucleosome in an attempt to elucidate the apparently bimodal behavior of Swi6.

### 8.2 Results

#### 8.2.1 Swi6 Binds Oligomerizes on Both Modified and Unmodified Nucleosomes, As Previously Reported

Multiple previous studies describe some of the behavior of Swi6 with respect to chromatin and mono-nucleosomes [9, 121, 183, 212]. Canzio et al. [120] in particular show i) the specificity of swi6 binding to modified and unmodified mononucleosomes and ii) the presence of multiple binding events in a purified nucleosome/Swi6 environment.

To verify the Swi6 made for further experiments, we repeated several experiments found in Canzio et al. [120]. To complete the experiments we prepared both unmodified nucleosomes and nucleosomes labeled at H3K9 with a trimethyllysine analog. These nucleosomes contain the Widom 601 positioning sequence [13] and a fluorescein label at one of the 5’ends of the DNA as shown in Figure 4.12C. These results are summarized in Figure 8.1. The activity
Figure 8.1: (A) EMSAs showing Swi6 binding to mononucleosomes both without (top) and with (bottom) H3K$_C$9me3. Fluorescence anisotropy results showing increased polarization of a fluorescein molecule ligated to one of the 5’ends of the DNA in a nucleosome as in Figure 4.12. Both of these experiments display a slight preferential binding to nucleosomes containing H3K$_C$9me3 over unmodified H3.2 and closely match the results from Canzio et al. [120].

of Swi6 was first verified by EMSA in Figure 8.1A. These data show a ~3-fold preference to H3K$_C$9me3 nucleosomes over those without the modification. Further, the hypershift seen at high concentrations of Swi6 indicate multiple binding events. These multiple binding events are confirmed by fluorescence anisotropy.

8.2.2 Binding of Swi6 does not impact FRET Efficiency

Swi6 is known to alter rotational freedom of a fluorophore attached to mononucleosomes as well as directly interact with both the H3 tail and core [121, 192] (See Figure 8.1). The tail extends between the two DNA gyres and binding of Swi6 could conceivably require the nucleosome to partially unwrap. However, we have not previously seen this kind of equilibrium shift due to any PTM recognition protein.

We prepared both H3K$_C$9me3 and unmodified nucleosomes which have been FRET labeled
Figure 8.2: Swi6 does not shift the nucleosome equilibrium to a largely unwrapped state.

Figure 8.3: Trimethylated lysine nine on histone H3.2 increases DNA accessibility over the apo state.

as described in chapter 3 so that fully wrapped nucleosomes undergo efficient FRET which decreases upon unwrapping. As shown in Figure 8.2 we find that Swi6 does not produce a FRET decrease when bound to nucleosomes at similar concentrations which produce a polarization change and gel hypershift as seen in Figure 8.1. This indicates that the binding of Swi6 does not shift the equilibrium state of the nucleosome to a largely unwrapped state. However, as was the case previously, these results preclude neither a small, undetectable by FRET shift in the unwrapping equilibrium nor a change in the DNA accessibility.

8.2.3 H3K9\textsubscript{me3} Slightly Increases DNA Accessibility

Histone tails are known to interact with both intra- and inter-nucleosomal DNA [213]. These interactions have been shown to be important in chromatin compaction and in regulation
of the amount of DNA sequestered into a given nucleosome. Additionally, PTMs have been shown to alter both chromatin compaction and nucleosome wrapping [213, 214]. Acetylation and phosphorylation are the typical PTMs which impact tail dynamics, but little research has looked into tail methylation impact.

To investigate this impact and to determine a baseline for accessibility of H3Kc9me3 nucleosomes, we carried out titrations of LexA binding to its target site buried within the nucleosome as in chapters 5 and 6. Previously these titrations for H3.2 have yielded indistinguishable response curves as in Figures 5.6 and 6.3. However, as can be seen in Figure 8.3, the presence of H3Kc9me3 increases the accessibility of DNA by a small, but repeatable 1.6 ± 0.4 fold, or a 0.5 ± 0.3 $k_B T$ free energy difference between the LexA bound states for each nucleosome. In addition to this finding, these data provide concentrations for each nucleosome which results in half of the nucleosomes being bound by LexA. FRET is sensitive to changes in this percentage and so changes in DNA accessibility can be measured by first priming the nucleosomes with an amount of LexA equal to the given nucleosome’s $S_{1/2}$.

### 8.2.4 Swi6 Alters DNA Accessibility to Nucleosomes Containing H3Kc9me3 Differentially Depending on Concentration

The biological function of Swi6, particularly its association with heterochromatin, led us to believe that investigations into nucleosome accessibility changes upon Swi6 binding would show protection of nucleosomal DNA. We studied these changes by observing LexA occupancy changes as a function of Swi6 concentration as shown in Figure 8.4. LexA concentration was held constant at each nucleosome’s respective $S_{1/2}$ any deviation from 50% occupancy is revealed as a shift in the FRET efficiency. Because Swi6 does not alter the FRET of the nucleosome upon binding at any concentration, any change in FRET represents a change in DNA accessibility.

Accessibility measurements shown in Figure 8.4 reflect the multiple binding modes detected in section 8.2.1. Initially and surprisingly, FRET efficiency decreases, indicating an increase of LexA occupancy and therefore an increase in DNA accessibility. This increased accessi-
Figure 8.4: (A) Swi6 initially increases DNA accessibility with an $S_{1/2} = \pm$ before decreasing it at concentrations greater than $\sim 3$ µM. (B and C) Titration of LexA with constant Swi6 quantitatively reveal the change in accessibility due to Swi6 binding to nucleosomes either with (B) or without (C) the methyllysine mark at H3K9. (D and E) Bar charts summarizing the impact of Swi6 at various concentrations on DNA accessibility. Modified nucleosomes (D) exhibit a saturated de-protection over the 10-100 nM range before reversing behavior with a return to initial levels of accessibility at 10 µM Swi6. Unmodified nucleosomes (E) do not show a significant change over any Swi6 concentration tested.
ibility occurs with an $S_{1/2}$ of $8.0 \pm 4nM$ (Figure 8.4A) indicating an extremely high binding affinity of Swi6 to nucleosomes for a PTM reader and corresponds to the affinity of the first binding event observed in Figure 8.1.

These experiments were quantified by titrations of LexA at constant Swi6 concentration. This technique enables measurement in the $S_{1/2}$ of LexA binding to its target site within the nucleosome. The value of the $S_{1/2}$ reflects the effective accessibility of the DNA occluded by the nucleosome (Figure 8.4B). Figure 8.4D summarizes these findings. Enhancement is saturated at 30 nM at a factor of $1.9 \pm .4$ increase in accessibility (compare to $2 \pm .3$ at 300 nM).

The secondary binding events in Figure 8.1 are determined by Canzio et al. [120] as oligomerization of Swi6 on a nucleosome substrate. At concentrations corresponding to this transition, DNA accessibility also undergoes a change in behavior from increasing DNA accessibility to decreasing it. We were unable to investigate concentrations greater than 10 µM, but at this concentration DNA accessibility has returned to free nucleosome levels for nucleosomes containing H3K$_{C9}$me3. It is likely that, were higher concentrations possible, the protection of DNA would further increase.

### 8.2.5 Swi6 Largely Does Not Change DNA Accessibility to Unmodified Nucleosomes

In section 8.2.1 we determined, as was originally reported by Canzio et al. [120], that Swi6 prefers binding nucleosomes containing H3K$_{C9}$me3 over unmodified nucleosomes with a specificity of only $\sim 3$-fold. Therefore, we were very interested in determining if the increases in DNA accessibility reported in the previous section were mirrored in unmodified nucleosomes.

We performed Swi6 and LexA titrations as previously with unmodified nucleosomes. As shown in Figure 8.4, for concentrations of Swi6 which produce clear binding in Figure 8.1 we see no detectable increase in DNA accessibility. This indicates that interaction with
the methylated H3 tail is necessary to alter nucleosome dynamics. At high concentrations, Figure 8.4A shows a protection of DNA which may be confirmed in Figure 8.4E where the LexA $S_{1/2}$ ratio is $S_{1/2, \text{No Swi6}} / S_{1/2, 10 \mu M \text{ Swi6}} = 0.66 \pm 0.17$. The effect is slight, but given higher Swi6 concentrations, more definitive conclusions may be drawn. Given current results, it is likely that the increase in accessibility requires H3K9me3 while the decreased accessibility at high Swi6 concentrations does not depend on the methyl state of the nucleosome.

8.2.6 Swi6 Maintains Relative Accessibility Between Unmodified Nucleosomes and those Containing H3K$_C$9me3

Swi6 in vivo is known to be involved in the spread of heterochromatin. This behavior is not well understood but it does indicate that it is likely that Swi6 is often in regions of the genome which are not fully methylated. Therefore, we decided to investigate the behavior of Swi6 on nucleosomes containing H3K$_C$9me3 relative to those which are unmodified. The
results of these comparisons are shown in Figure 8.5. It is clear the relative accessibility of each nucleosome is held roughly constant over each Swi6 concentration measured. The total change in free energy between modified and unmodified are $1.4 \pm 0.2$, $1 \pm 0.1$, and $1.1 \pm 0.2 k_B T$ for 30 nM, 300 nM, and 10 µM Swi6, respectively. Taking the $0.5 \pm 0.3 k_B T$ inherent in H3Kc9me3 vs unmodified nucleosomes into account this results in a free energy change in LexA binding of $0.9 \pm 0.3$, $0.5 \pm 0.3$, and $0.6 \pm 0.3 k_B T$ for each concentration, respectively.

8.3 Methods

8.3.1 Protein Preparation

LexA used in this study was purified as detailed in Appendix A. Swi6 used was purified as described in Appendix B. Histones were prepared as previously described [157].

8.3.2 Preparation of Labeled DNA

DNA used in this study was prepared by PCR from a plasmid containing the 601 sequence with a LexA binding site located at bases 8–27. PCR primers for ensemble experiments were the Cy3-labeled oligonucleotide Cy3-CTGGAGATACTGTATGAGCATACAGTACAATTGGTC and the unlabeled reverse primer ACAGGATGTATATCTGACACGTCCTGGAGACTA. Cy3-labeled oligos were labeled with Cy3 NHS ester (GE healthcare) at an amino group attached to the 5’end of the DNA oligo. After PCR, dsDNA molecules were purified by HPLC on a MonoQ (GE Healthcare) ion exchange column.

8.3.3 Nucleosome Preparation

Cy3-labeled DNA and Cy5-labeled histone octamer were combined and reconstituted by double salt dialysis and purified by 5–30% sucrose gradient as previously described [23]. Fractions containing nucleosomes were then collected and concentrated with a 30 kDa MWCO Amicon Ultra (Millipore).
8.3.4 Electromobility Shift Assay

Nucleosomes containing either H3K\textsubscript{C}9me3 or unmodified H3 were incubated at 1 nM with 0–10 µM Swi6 in 0.5×TE, 75 mM NaCl for 5 min before separation on a native 0.3×TBE 5% polyacrylamide gel.

8.3.5 Anisotropy Measurements

Anisotropy experiments were determined by fluorescein emission acquired by a Tecan M1000 plate reader. Z-height was automatically determined and parallel and perpendicular fluorescence was used to calculate polarization by the equation:

\[
\langle r \rangle = \frac{I_{\text{par}} - GI_{\text{perp}}}{I_{\text{par}}2GI_{\text{perp}}} \quad \text{(Anisotropy)}
\]

\[
P = \frac{3\langle r \rangle}{2 + \langle r \rangle} \quad \text{(Polarization)}
\]

where \(I_{\text{par}}\) is the fluorescence intensity parallel to the excitation polarization, \(I_{\text{perp}}\) is the fluorescence intensity perpendicular to the excitation polarization, and \(G\) is the G-factor for the instrument determined by assuming 0 polarization for free DNA in 0.5×TE.

8.3.6 FRET Measurements

Ensemble FRET efficiencies were determined from spectra acquired by a Horiba Scientific Fluoromax 4. Samples were excited at 510 and 610 nm and the fluorescence spectra were measured from 530 to 750 and 630 to 750 nm for donor and acceptor excitations, respectively. Each wavelength was integrated for 1 s, and the excitation and emission slit width were set to 5 nm bandwidth with 2 nm emission wavelength steps. FRET measurements were computed though the (Ratio)\textsubscript{A} method (Equation 4.18, [103, 145])

Titrations were carried out in 75 mM NaCl, 0.1 mM Potassium phosphate pH 7.5, 11.5 mM
Tris-HCl pH 7.5, 0.00625% v/v Tween20. LexA titrations were fit to

\[ E = \frac{E_f - E_0}{1 + \frac{S_{1/2}}{C}} + E_0 \]

where \( E \) is the FRET efficiency at concentration \( C \) of LexA, \( E_f \) is the FRET efficiency at high concentration of titrant and \( E_0 \) is the efficiency in the absence of the titration and \( S_{1/2} \) is the inflection point. Errors represent a standard deviation based on three experiments. Fit errors represent 68% confidence bounds.

8.4 Discussion

In this study, we have investigated the effect of Swi6 binding on DNA accessibility. We have been able to decouple Swi6 binding from Swi6’s influence on DNA accessibility and find that DNA accessibility is dependent both on Swi6 concentration and methylation state. For nucleosomes possessing the H3K9me3 PTM, DNA accessibility increases in the nanoMolar regime and decreases at microMolar concentrations. Further, the relative accessibility between modified and unmodified nucleosomes remains constant over all concentrations measured.

Recent studies have implicated Swi6 in both small interfering RNA transcribed from heterochromatin and in activation of particular gene loci [9]. The initial findings reported here suggest a mechanism for Swi6 transcriptional enhancement. In addition to selective targeting of nucleosomes containing H3K9me3, we find that off-target binding to unmodified nucleosomes alters accessibility with a constant difference relative to modified nucleosomes. The findings described here suggest that the cell can utilize Swi6 differentially to both increase and decrease accessibility to regions of the genome containing H3K9me3 by controlling Swi6 expression levels. This may help to answer the outstanding questions related to the apparent dichotomy of Swi6 function. The oligomerization of Swi6 described by Canzio et al. [120] helps to understand the decrease in accessibility observable at the same concentrations. Oligomerization may bring multiple nucleosomes into close contact to reduce DNA
accessibility. More experiments are required to investigate this potential function of Swi6.

8.5 Future Directions

Questions remain with regard to DNA accessibility regulation by Swi6. Investigations into the specific domains of Swi6 which induce accessibility changes should be undertaken. This would include experiments with Swi6\textsuperscript{LoopX}, which removes aas which mimic an ARK loop present in H3; Swi6\textsuperscript{CageX}, which produces a PTM-binding deficient chromodomain; and Swi6\textsuperscript{AcidX} which mimics human HP1\textalpha before post translational modification. Each of these mutants also alter the dimerization and isodesmic $K_d$ associated with Swi6 and will yield further mechanistic information on Swi6’s role in altering nucleosome dynamics.

In addition to these experiments, it is necessary to determine the impact of Swi6 on a di-nucleosome construct capable of measuring accessibility in both the linker region and buried within one of the two nucleosomes. Experiments involving these constructs can answer the questions posed at the end of section 8.4. Additionally, it is possible to prepare nucleosome arrays which house i) fully modified nucleosomes, ii) fully unmodified nucleosomes, and iii) one unmodified and one H3\textsuperscript{\textalpha}me3 nucleosome. These arrays mimic i) nucleosomes buried in heterochromatin ii) nucleosomes buried in euchromatin, and iii) nucleosomes at the border between heterochromatin and euchromatin allowing for determination of Swi6 behavior which enables heterochromatin spread.
Chapter 9

CONCLUDING REMARKS AND FUTURE DIRECTIONS

There are hundreds of combinations of histone PTMs and the PTM-recognizing proteins which bind them. Many of these proteins utilize one or multiple of a set of PTM-recognizing domains to specifically bind their target. This work has focused on only three proteins utilizing three distinct PTM-recognition domains each targeting trimethylated lysine. Despite the similarity in PTM location and chemistry, the impact of binding for each protein is unique. Here, I summarize the results of each set of studies and state the implications of each individually and as a whole. I then go on to suggest future studies which expand upon the results stated here.

9.1 Summary of Results

PHF1

We began by establishing the binding properties of the PHF1 Tudor domain to nucleosomes. We determined the binding affinity of Tudor to nucleosomes through a GST pulldown experiment and establish that the Tudor domain of PHF1 increases accessibility to nucleosomal DNA. We then expanded these initial studies to more of the PHF1 protein to include the NTD. With this addition and an expanded experiment set, we were able to determine that i) the Tudor domain induces a partial structuring of the NTD; ii) NTD-Tudor enhances
accessibility by nearly an order of magnitude; iii) the first 12 aa of the NTD are necessary for the full increase in DNA accessibility; iv) PHF1 prefers binding an open nucleosome conformation; v) PHF1 regulates DNA accessibility by altering the off rate of TFs bound within the nucleosome.

The implications of these findings are significant. This is the first reported case of a PTM reader with a higher affinity to an altered nucleosome structure. This opens the door for both induced fit and conformational selection models to begin to explain behaviors seen in other PTM readers. For example, it is common for PTM-recognition domains to show μM affinity to their target PTM. It is assumed that cooperative binding to DNA or other parts of the histone drive occupancy in vivo. With the availability of conformational selection as a model for binding, these μM affinities may prove to be the affinity to a loose-binding conformation rather than the preferred conformation.

Next, Luo et al. [22] report a novel regulatory mechanism by which the nucleosome limits access to DNA not by steric occlusion, but through competition with TFs. Nucleosomes increase dissociation rates of DNA-binding proteins bound to DNA contained in the nucleosome by a factor of 1000. Here, we have found that the binding of PHF1 to the nucleosome, reduces that factor of 1000 by at least a factor of 2. This is the first reported example of the potential regulation of the newly-discovered mechanism of nucleosome occupancy. Thus, nucleosomes can not only block DNA, but also drive proteins off DNA and we now have evidence that both of these mechanisms are regulated by PTMs and PTM-binding proteins.

LEDGF

The LEDGF/P75 studies are very preliminary in that significant future experiments are required to gain quantitative and mechanistic understanding of impacts of LEDGF/P75 recognition of H3K36me3. However, significant insight is still gained. First, the increased protection observed during LEDGF/P75 binding proves recognition by different proteins produce different results. It is logical that a PTM-binding event to H3K36me3 would increase accessibility because the mark is associated with active chromatin which must often be
sequentially cleared of histones for transcription. However, in the case of LEDGF/P75 we observe behavior which logically opposes the what would be expected at a PTM associated with active regions. Clearly the precise combination of PTM and PTM reader or at least PTM recognition domain alters the response of chromatin to binding. It will be interesting to attempt to decrypt this code in the future.

The second result is that LEDGF/P75 targets not only H3K36me3 but specifically H3.3K36 me3. This finding has major implications for understanding both chromatin recognition and HIV-1 integration. This appears to be only the second example of a PTM reader specifically recognizing a modified variant H3. For HIV-1 a better understanding of the mechanisms involved in targeting integration is invaluable. Everything from genomic assays to structural studies can benefit from knowledge of an increased affinity to H3.3.

**Swi6**

While still preliminary, our investigations into Swi6 have also produced interesting results. First is the concentration-dependent regulation of DNA accessibility. Further experiments as discussed in section 8.5 are necessary to fully grasp the implications of this finding but it seems reasonable to expect that the low concentration limit increased accessibility will persist for nucleosome arrays which are not fully modified. If this is the case, it sheds light on the mechanism Swi6 utilizes to act both as a transcriptional activator and mediator of heterochromatin spread—in regions of low methylation, the behavior is fundamentally different.

We also describe a constant accessibility change between methylated and unmodified nucleosomes regardless of Swi6 concentration. The biological role of this is not immediately obvious but it is possible that this behavior may be a clue into Swi6’s function at heterochromatin boundaries. It is known to be pivotal in heterochromatin spread but the mechanism by which this is achieved is not known. This finding points to a possible explanation. If Swi6 decreases accessibility of DNA wrapped in unmodified nucleosomes more than those carrying the H3K9me3 modification, Swi6 can be initially silencing unmodified regions found
at the boundaries of heterochromatin as well as recruit external factors for more permanent silencing.

### 9.2 Suggested Future Studies

Several of the implications discussed here are possibilities which require further experiments to confirm. In the cases of LEDGF/P75 and Swi6 further experiments utilizing these proteins are detailed in their respective chapters. However, in addition to these experiments already detailed, expansion to other PTM recognition domains, other PTMs beyond methylation, and other regions beyond the entry-exit region should be undertaken.

Different domains and PTMs go hand-in-hand. The work involved in assessing the impact on a new recognition protein is significant. However, each additional protein will build an understanding and it is likely that eventually different domains or combinations of domains will emerge with particular effects on chromatin. An understanding of this level will allow prediction of the effect of new proteins. Beyond understanding of methylation-recognition domains future studies should investigate the impact of recognition of other PTMs. In this regard we are collaborating with the Cole lab at Johns Hopkins to characterize mimics of acyl modifications which will enable site-specific chemical labeling very similar in final product to that acquired by the methods defined by Simon et al. [84]. These mimics, and others like them, will hopefully open up the plethora of domain which recognize the many other chemical modifications on histones, of particular interest are proteins containing acetyllysine-recognizing bromodomains for their implication in cancer.

In addition to expanding studies to other proteins and modification types, studies should be expanded to different regions of chromatin. Thus far all experiments have been conducted on mononucleosomes with the target binding site at one particular location in the entry-exit region of the nucleosome. The intricacies of position-dependence of DNA accessibility is not fully understood in terms of location within a given nucleosome or in terms of positioning in a linker region. It can be seen from Figure 2.3 that linker regions are selected for TF
binding sites relative to a random distribution. What’s more is that small changes in the angle of the DNA at the entry-exit region of the nucleosome can produce large changes in compaction of chromatin and thus large changes in accessibility of linker DNA. These studies necessitate nucleosome arrays which further expand the space of possible experiments. It is naïve to think that homogeneously modified nucleosomes or nucleosome arrays yield all the information necessary to understanding accessibility changes due to PTM recognition. For example, in the case of Swi6 it is precisely the boundaries of modifications which are the most interesting. Arrays allow the possibility for studying these boundaries by ligating unmodified and modified array halves together. Studies on the accessibility of both modified and unmodified nucleosomes could then be carried out in a similar manner to that discussed in this work.

Finally, we have yet to directly observe binding of a PTM reader to its recognition site. Utilization of new means of generating extremely small excitation volumes such a zero-mode waveguides will enable breaking a 1 µM solution fluorophore barrier. PTM readers could then be fluorescently labeled and co-localization or FRET-based techniques could then be used to measure the on and off rates of histone readers directly while in equilibrium. The insights gained by these future experiments will greatly increase our understanding of the epigenetic regulation of DNA accessibility.

9.3 Closing Remarks

Chromatin regulation of DNA accessibility is an extremely import aspect of the biology of eukaryotes. The work presented here provides information on a new regulatory mechanism in that PTM readers alter the state of the chromatin to which they bind. In the case of PHF1 we have determined the mechanism which produces the change in accessibility is the regulation of a novel mechanism which was described only two years earlier. Overall, these studies provide the foundation for the future understanding of the direct impact of PTM readers on chromatin accessibility.


T. Förster. “Energiewanderung und Fluoreszenz”. In: Naturwissenschaften 33.6 (1946), pp. 166–175. ISSN: 00281042. DOI: 10.1007/BF00585226 (Referenced on pages 31, 34, 35).


Appendix A

PURIFICATION OF LEXA

This protocol is based on Little et al. [145]. It was adapted for use in the Poirier lab by Justin North in 2007 and further adapted to overcome discontinued chromatography media by me in 2015.

A.1 Buffers

Filter and degas all buffers.

A.1.1 Buffer A

- 50 mM Tris-HCl pH 8.0 (Sigma T1503)
- 0.5 mM EDTA (Sigma E1644)
- 1 mM DTT (Sigma D9779)
- 200 mM sodium chloride (Sigma S3014)

Buffer A + 10% Sucrose

Add to Buffer A (before adding dH₂O to final volume) sucrose (Sigma S0389) to 10% w/v (100 g per L final volume) and fill to the final volume with dH₂O.
**Buffer A + 10% Sucrose + 1mM PMSF**

Add to Buffer A + 10% sucrose (before adding dH$_2$O to final volume) 100 mM phenylmethanesulfonyl fluoride (PMSF) (Sigma P7626) in dimethylformamide (DMF) (Sigma 227056) while rapidly mixing. Some PMSF may aggregate out. When thawed PMSF is active for ~2 h. This buffer does not need to be filtered.

**A.1.2 Buffer B**

- 20 mM potassium phosphate pH 7.0
- 0.1 mM EDTA (Sigma E1644)
- 1 mM DTT (Sigma D9779)
- 10% glycerol (Sigma G5516)

**Buffer B + 500mM Sodium Chloride**

Add to Buffer B (before adding dH$_2$O to final volume) sodium chloride (Sigma S3014) to 500 mM final concentration and fill to the final volume with dH$_2$O.

**A.1.3 Buffer B + 200mM Sodium Chloride**

Add to Buffer B (before adding dH$_2$O to final volume) sodium chloride (Sigma S3014) to 500 mM final concentration and fill to the final volume with dH$_2$O.

**A.1.4 Buffer B + 800mM Sodium Chloride**

Add to Buffer B (before adding dH$_2$O to final volume) sodium chloride (Sigma S3014) to 800 mM final concentration and fill to the final volume with dH$_2$O.

**A.1.5 Buffer C**

- 10% glycerol (Sigma G5516)
• 1 mM DTT (Sigma D9779)
• 0.5 mM calcium chloride dihydrate (Sigma C3306)

**Buffer C + 50mM Potassium Phosphate**

Add to Buffer C 0.5 mM potassium phosphate (section A.1.9) to 50 mM potassium phosphate final concentration.

**Buffer C + 400mM Potassium Phosphate**

Add to Buffer C 0.5 mM potassium phosphate (section A.1.9) to 400 mM potassium phosphate final concentration.

**A.1.6 Buffer D**

• 10 mM PIPES (Sigma )
• 0.1 mM EDTA (Sigma E1644)
• 10% glycerol (Sigma G5516)
• 200 mM sodium chloride (Sigma S3014)
• pH to 7.0 with sodium hydroxide and hydrochloric acid

**A.1.7 Polyethyleneimine (Polymin P)**

For 10 mL of 10% polyethyleneimine (Polymin P):

• In a 50 mL conical bottom tube, weigh out 2.12 g Polymin P (Sigma P3143).
• Add dH₂O to 8mL
• pH with sodium hydroxide and hydrochloric acid
• Add dH₂O to 10 mL total volume.

Note: Polymin P can be stored for 24 h at 4°C.
A.1.8 2M potassium phosphate dibasic

- Potassium phosphate dibasic (Sigma P9666)

A.1.9 0.5M potassium phosphate buffer, pH 7.0

- Potassium phosphate monobasic (Sigma P9791)
- Potassium phosphate dibasic (Sigma P9666)
- pH to 7.0 with potassium phosphate monobasic (down) and potassium phosphate dibasic (up).

For 1 L add 21.8 g monobasic and 77.6 g dibasic potassium phosphate and add dH2O to 1 L.

A.2 Protein Expression

A.2.1 Trial Induction

In this section we will determine which colonies made during a transformation produce the highest yield of LexA.

Materials

- Autoclaved 5 mL 2×YT growth media
- 50 mg/mL (1000×) ampicillin
- 34 mg/mL (1000×) chloramphenicol
- 100 mM IPTG (isopropyl-β-D-1-thiogalactopyranoside Sigma I6758)
- 80% autoclaved glycerol in dH2O
- 12% sodium dodecyl sulfate (SDS)-poly-acrylamide gel electrophoresis (PAGE) gel
- 1×Tris-Glycine-SDS buffer
- 6×SDS loading dye
- Gel stain
- Gel destain
- 0.5×TE

**Induction**

1. Innoculate 3-5 single transformed BL21(DE3)pLysS colonies previously transformed into separate tubes of 5 mL of 2×YT + 50 µg/mL ampicillin + 34 µg/mL chloramphenicol and shake at 280 rpm, 37°C until the OD₆₀₀nm,₁cm = 0.2 for a given clone.
2. Mix 750 µL parts culture for each clone with 250 µL part 80% glycerol to make a 20% glycerol stock.
3. Immediately flash freeze in liquid nitrogen and store at -80°C.
4. Allow the remaining culture for each clone to continue to shake at 280 rpm, 37°C until OD₆₀₀nm,₁cm = 0.5
5. Transfer 100-200 µL to a 1.5 mL microcentrifuge tube on ice for a pre-induction time point.
6. Induce the remaining culture with 100 mM IPTG to 0.2 mM final concentration.
7. Allow induced culture to continue to grow at 280 rpm, 37°C.
8. Each hour up to 2 transfer 100-200 µL of culture to a microcentrifuge tube on ice.

**SDS-PAGE analysis**

1. Discard the remaining culture after a given clone has been induced for the full time series.
2. Centrifuge each time point at 2,500 g for 1 min. Remove supernatant and suspend each pellet in 10 µL of 0.5×TE; Add 10 µL 6×SDS loading dye; boil at 95°C for 5 min; centrifuge at 2,500 g for 10 sec and store on ice
3. Prepare at 12% SDS-PAGE gel. Load 10 µL of each time point onto gel reserving a lane for a molecular weight ladder. Run in 1×Tris-Glycine-SDS per standard protocol.
4. Stain and destain the gel
5. Determine the best colony based on expression levels relative to the background proteome. (see Figure A.1)
6. Discard glycerol stocks of colonies NOT chosen.

A.2.2 Bulk Induction

Materials

- Autoclaved 5 mL 2×YT growth media
- Autoclaved 250 mL Erlenmeyer flask containing 60 mL 2×YT growth media
- Autoclaved 1 L Erlenmeyer flask containing 600 mL 2×YT growth media (as many as desired)
- 50 mg/mL (1000×) ampicillin
- 34 mg/mL (1000×) chloramphenicol
- 1 mM IPTG (isopropyl-β-D-1-thiogalactopyranoside Sigma I6758)
- 12% SDS-PAGE gel
- 1×Tris-Glycine-SDS buffer
- 6×SDS loading dye
- Gel stain
- Gel destain
- 0.5×TE

Expression

1. Inoculate 5 mL 2×YT growth tube with 50 µL ampicillin and 34 µL chloramphenicol with a scraping from the desired glycerol stock made in section A.2.1. Mix by swirling to homogenize the sample.
2. Perform a serial dilution resulting in 1/100, 1/1000, and 1/10000 dilutions of the original 1/1 growth tube into 5 mL 2×YT + 50 µg/mL ampicillin + 34 µg/mL chloramphenicol.
3. Shake each tube overnight at 280 rpm, 37°C.

4. Determine which tube has grown to optimal conditions, 0.2 < OD$_{600\text{nm,1cm}<0.5}$. This may require waiting. If so, prewarm a 250 mL Erlenmeyer flask containing 60 mL 2×YT + 50 μg/mL ampicillin + 34 μg/mL chloramphenicol by shaking at 280 rpm, 37°C.

5. Inoculate 250 mL Erlenmeyer flask with the appropriate culture and allow to grow until 0.2 < OD$_{600\text{nm,1cm}<0.5}$ by shaking at 280 rpm, 37°C.

6. Pre-warm desired number of 1 L Erlenmeyer flasks containing 600 mL 2×YT.

7. When the 250 mL flasks have reached an appropriate OD, 0.2 < OD$_{600\text{nm,1cm}<0.5}$, add antibiotics to the 1 L Erlenmeyer flasks to a final concentration of 50 μg/mL ampicillin + 34 μg/mL chloramphenicol and transfer equal volumes of the culture to each 1 L flask.

8. Continue to grow at 280 rpm, 37°C until OD$_{600\text{nm,1cm}=0.5}$.

9. Prepare a 0 h time point as before in section A.2.1.

10. Induce each 1 L Erlenmeyer flask with 0.5 mM final concentration of 1 mM IPTG.

11. Continue to grow for 2 h Taking time points each hour.

12. Immediately move on to section A.2.3.

13. While going through the protocol as described in section A.2.3, prepare and analyze an SDS gel as described in section A.2.1 and verify the bulk expression (see Figure A.1).
A.2.3 Harvesting Cells

Materials

- 500 mL centrifuge tubes and rotor
- Buffer A + 10% sucrose + 10% PMSF (section A.1.1)
- Lysozyme (Sigma L6876)

Procedure

1. Once cultures have finished growth, rapidly chill each by immersing the flasks in an ice-water slurry.
2. Fill and balance 500 mL bottles and centrifuge at 4,000 g, 4 °C for 10 min.
3. If the supernatant is cloudy, repeat the centrifugation.
4. If the supernatant is clear, decant off the supernatant and refill with culture and repeat centrifugation until a cell pellet is isolated.
5. Invert bottles on a paper towel and drain for 5 min to remove residual media. ~8 g cell paste per liter of culture can be expected.
6. Resuspend each pellet in 20 mL Buffer A + 10% sucrose + 1 mM PMSF. Keep the 500 mL centrifuge bottle on ice during resuspendtion.
7. Transfer resuspended cells in to a 50 mL tube on ice.
8. Prepare Buffer A + 10% sucrose + 1 mM PMSF + 10 mg/mL lysozyme by dissolving solid lysozyme in Buffer A + 10% sucrose + 1 mM PMSF. Centrifuge at 20,000 g, 4°C for 2 minutes to remove aggregates.
9. Add 20 µL of Buffer A + 10% sucrose + 1 mM PMSF + 10 mg/mL lysozyme per 1 mL of resuspended cells to a final lysozyme concentration of 0.2 mg/mL.
10. Incubate the cells on ice for 1 h.
11. Flash Freeze the cells in liquid nitrogen and store at -80°C. Cell pellets are stable for up to 2 months before purification.
A.3 Protein Purification

A.3.1 Cell Lysis

Materials

- 37°C water bath
- 1 M magnesium chloride (Sigma)

Procedure

1. Thaw cell pellets at room temperature with occasional mixing by inversion.
2. Add 1 M magnesium chloride to 10 mM final concentration and mix thoroughly by inversion.
3. Heat capped cell suspensions in 37°C water bath. Continuously immerse and mix cells gently by hand. Equilibrate to 37°C for 5 min.
4. Incubate cell pellets on ice for 30 min. During this time, the cells will fully lyse.
5. Immediately proceed to section A.3.2

A.3.2 Cell Debris Pelleting

Materials

- Buffer A + 10% sucrose section A.1.1
- 30 mL Syringe
- 50 mL Round bottom tube (min volume = 20 mL, max = 34 mL)

Procedure

1. Dilute lysate 2-fold in Buffer A + 10% sucrose to decrease viscosity.
2. pipette lysate up and down with a 30 mL syringe (without a needle) to initially shear genomic DNA and lower the viscosity of the lysate.

3. Centrifuge at 30,000 g, 4°C for 30 min in 50 mL round bottom tubes well balanced to minimize shaking.

4. In a beaker on ice, pool together supernatant (including viscous, but clear, liquid near the brown debris pellet) from each culture showing good expression as determined from section A.2.2 to produce Fraction I.

5. Immediately proceed to section A.3.3

A.3.3 Polyethyleneimine Precipitation

This stage precipitates genomic DNA through high concentrations of Polymin P.

Materials

- 10% Polymin P (section A.1.7)
- Stir plate at 4°C
- 50 mL Round bottom tube (min volume = 20 mL, max = 34 mL)
- Fraction I

Procedure

1. Stir Fraction I at 4°C.

2. Over the course of 15 min slowly add 10% Polymin P to a final concentration of 0.362%.

3. Continue to stir at 4°C for 10 min. DNA will precipitate out.

4. Centrifuge at 15,000 g for 10 min 50 mL round bottom tubes.

5. Decant supernatant into a new 250 mL beaker on ice producing Fraction II. Precipitate is genomic DNA.

6. Proceed immediately to section A.3.4.
A.3.4 Salting Out Purification

Materials

- Ammonium sulfate (Sigma A4418)
- Stir plate at 4°C
- 50 mL Round bottom tube (min volume = 20 mL, max = 34 mL)
- Buffer A + 10% sucrose (section A.1.1)
- Buffer B + 500 mM sodium chloride (section A.1.2)

Procedure

1. Stir Fraction II at 4°C.
2. Over the course of 15 min slowly add 0.4 g ammonium sulfate per mL to Fraction II. Rapid addition of ammonium sulfate will cause increased local concentration and premature precipitation of protein.
3. Continue to stir at 4°C for 20 min.
4. Centrifuge at 20,000 g, 4°C for 20 min in 50 mL round bottom tubes.
5. Decant off and discard supernatant.
6. Dissolve pellet in 1× the original volume before precipitation in buffer A + 10% sucrose.
7. Repeat precipitation in steps 1-5.
8. Dissolve pellet in a small volume of Buffer B + 500 mM sodium chloride, approximately 5-10 mL per gram of cell paste from the original harvest.
9. Dialyze against 3 changes of 1 L Buffer B + 500 mM sodium chloride for at least 8 h per change with the last proceeding overnight.
10. Centrifuge dialyzed protein in 15 mL tube at 1,000 g for 10 min.
11. Collect supernatant to form Fraction III.
Figure A.2: FPLC chromatogram of LexA elution from HiTrap Heparin HP column. Fractions between tags 6 and 8 were pooled to produce Fraction IV.
A.3.5 HiTrap Heparin Affinity Purification

Materials

- Buffer B
- Buffer B+200 mM sodium chloride
- Buffer B+800 mM sodium chloride
- fast protein liquid chromatography system (FPLC)
- 50 mL superloop
- HiTrap Heparin HP 5 mL column (GE 17-0407-01)
- 12% SDS-PAGE gel
- 1×Tris-Glycine-SDS buffer
- 6×SDS loading dye
- Gel stain
- Gel destain

Purification

1. Dilute Fraction III 2.5-fold with Buffer B to yield a final concentration of 200 mM sodium chloride.
2. Load diluted Fraction III onto the superloop
3. Inject the sample onto the column at 0.75 mL/min while collecting flow through in case LexA does not bind the column.
4. Wash the column with ∼1.5 column volumes with Buffer B + 200 mM sodium chloride until the ultraviolet (UV) spectrum returns to baseline. Collect flow through.

5. Elute the protein with flow rate of 1 mL/min over a linear gradient in Buffer B from 200 mM sodium chloride to 800 mM sodium chloride over 35 mL followed by 10 mL of Buffer B+800 mM sodium chloride to wash the column. Collect 0.56 mL fractions during the gradient and wash. Expect elution at ∼500 mM sodium chloride.

6. Identify fractions containing LexA through SDS-PAGE analysis as in section A.2.1

7. Pool fractions containing LexA to form Fraction IV

A.3.6 Hydroxyapatite Ion Exchange Purification

Materials

- CHT ceramin hydroxyapatite Type II, 20 µm particle size (Biorad 158-2200)
- 200 mM potassium phosphate dibasic (section A.1.8)
- 1 M sodium hydroxide, filtered and degassed
- 500 mM potassium phosphate buffer, pH 7.0 (section A.1.9)
- Buffer C + 50 mM potassium phosphate, pH 7.0 (section A.1.5)
- Buffer C + 400 mM potassium phosphate, pH 7.0 (section A.1.5)
- C10 column and adapters
- Amersham P1 pump
- 50 mL Superloop (18-1113-82)
- 0.1 M sodium hydroxide + 10 mM potassium phosphate buffer

Pre-Cycling and Packing

1. Weigh out an appropriate amount of CHT Ceramic Hydroxyapatite. Use 9 mL Bed volume of hydrated hydroxyapatite per liter of starting LexA culture. 1 g of CHT Ceramic Hydroxyapatite yields ∼2.3 mL bed volume once hydrated.
2. Gently stir in 200 mM potassium phosphate dibasic to make a 10% w/v slurry. Use a plastic or glass pipette, metal will produce fines.

3. Allow the slurry to settle for 10 min

4. Fill the C10 column (without the top) with 5 mL of 200 mM potassium phosphate dibasic. Allow the buffer to flow out of the bottom until it is 10% full, then plug the outlet.

5. Note the volume of the settled hydroxyapatite and pipette off excess until the total volume is 120% of the settled resin volume.

6. Swirl the slurry and pipette the mixture into the C10 column

7. Allow the slurry to settle for 5 min.

8. Open the column outlet and allow fluid to flow until supernatant level is just above the resin level and plug the outlet. Do not allow the resin to dry.

9. Repeat steps 6-8 until the slurry is used or the resin is within 1 cm of the top of the column.

10. Insert the top C10 flow adapter at an angle to prevent air bubbles.

11. Lower the top adapter until the support screen is about 1 mm from the resin

    Note: The initial resistance of the top flow adapter is due to the press-fit o-ring, make sure the flow from the top is open to avoid over-pressurizing the bed and apply force liberally until the adapter o-ring slips into the column. Then carefully press the adapter to the correct position without crushing the resin.

12. Plug the column inlet and secure the top adapter position.

13. Prime the P1 peristaltic pump with 200 mM potassium phosphate dibasic.

14. Pump 5 column volumes of 200 mM potassium phosphate dibasic through the column at 2 mL/min.

15. Plug the column outlet.

16. Remove the inlet from the P1 pump and lower the top adapter to 1 mm above the final resin height.

17. Plug the column inlet.
Column Equilibration and Waking a Stored Column

A stored column can be re-used, but should first go through the pH cycling described here. Additionally, a column just packed needs to go through the same pH cycling.

1. Wash the column using a P1 peristaltic pump with 1 column volume of dH$_2$O. Flow at 1 mL/min.
2. Flow 5 column volumes of 1 M sodium hydroxide at 1 mL/min.
3. Wash the column with 1 column volume of dH$_2$O at 1 mL/min.
4. Pre-elute the column with 3 column volumes of 500 mM potassium phosphate buffer, pH 7.0 at 1 mL/min.
5. Stop flow and plug the inlet and outlet of the column.

Sample Loading and Elution

1. Equilibrate a FPLC with Buffer C + 50 mM potassium phosphate and Buffer C + 400 mM potassium phosphate.
2. Equilibrate a 50 mL superloop with Buffer C + 50 mM potassium phosphate.
3. Equilibrate hydroxyapatite column with at least 3 column volumes of Buffer C + 50 mM potassium phosphate until UV absorbance and conductivity is constant.
4. Dilute Fraction IV 2.5-fold with Buffer C + 50 mM potassium phosphate to bring sodium chloride concentration down to < 200mM.
5. Load the sample onto the superloop.
6. Inject the sample onto the column at 1.5 mL/min while collecting flow through in case LexA does not bind.
7. Wash the column with 1.5 column volumes of Buffer C + 50 mM potassium phosphate at 1 mL/min or until the UV absorbance returns to baseline. Again, Collect this flowthrough.
Figure A.4: FPLC chromatogram of LexA elution from hydroxyapatite column. Fractions between tags 1 and 2 were pooled and dialyzed against Buffer D.
8. Elute the sample over 4 column volumes linear gradient from Buffer C + 50 mM potassium phosphate to 400 mM potassium phosphate at 1 mL/min flow rate.

9. Collect 1 mL fractions and store on ice.


11. Before running another protein on the hydroxyapatite column: repeat section A.3.6 and Equilibrate with 3 column volumes of Buffer C + 50 mM potassium phosphate.

12. To store the hydroxyapatite column: repeat section A.3.6 but instead of step 4, store the column with 3 column volumes of 0.1 M sodium hydroxide + 10 mM potassium phosphate.

13. Store the column at 4°C away from light.

A.3.7 Determining Fractions and Dialysis

Materials

- 12% SDS-PAGE gel
- 1×Tris-Glycine-SDS buffer
- 6×SDS loading dye
- Gel stain
- Gel destain
- Buffer D
Procedure

1. Determine Fractions from section A.3.6 which contain LexA as in section A.2.1.
2. Pool fractions containing LexA and dialyze against 3 changes of 1 L each of Buffer D at 4°C for at least 8 h per change with the last change going overnight.

A.4 Protein Storage

LexA has proven to be sensitive to storage conditions with respect to buffers. It is known to be stable in Buffer D at -80°C for years so this is the recommended long term storage buffer.

Materials

- 0 kDa, 15 mL MWCO amicon ultra
- liquid nitrogen

Procedure

1. After dialysis into protein storage buffer, concentrate the protein in an amicon ultra to no more than 0 μM. When concentrating, be sure to mix the very concentrated protein at the bottom of the concentrator every 0 min to avoid aggregation. Check the concentration often to ensure the protein is not over concentrated. Aggregation begins at 0 μM LexA.
2. Quantify the LexA concentration by UV-VIS spectrophotometry. LexA is used as a dimer with a total molecular weight of 2 Da and $\varepsilon_{275\text{nm}} = 13940 \text{cm}^{-1} \text{M}^{-1}$.

A.5 LexA Verification

LexA activity needs to be verified before use. This is done through EMSA as shown in Figure A.6. Prepare a titration from 0.3 nM LexA to 1 μM LexA with a constant 1 nM
Figure A.6: EMSA showing LexA binding to free DNA confirming LexA activity. DNA concentration in the gel is 1 nM and LexA binds stoichiometrically to it, indicating the $K_d$ is less than 1 nM. Lanes are of increasing concentration of LexA: 0 nM, 0.3 nM, 1 nM, 3 nM, 10 nM, 30 nM, 100 nM, 300 nM, 1 µM LexA.

Concentration of fluorescently labeled DNA containing the LexA recognition sequence. Allow each binding reaction to proceed for 5 min and run on 0.3×TBE, 5% native polyacrylamide gel. Binding should occur stoichiometrically as the actual $K_d$ of LexA is less than the 1 nM of DNA present in the experiment. This indicates the activity/quantification of LexA is good.
Appendix B

Purification of HP1 Homolog Swi6

Swi6 is one of two HP1 homolog proteins in *Schyzosaccharomyces pombe*, the other being CHP2. For an introduction to Swi6, see chapter 8. The interest primarily lies in its shared homology with other HP1 proteins and the fact that it does not require modification to function as opposed to HP1α. The protocol described here is used to purify recombinantly grown Swi6 with an N-Terminal His-tag.

This protocol is based on the protocols of Daniele Canzio from the Narlikar lab.

B.1 Sequences

B.1.1 Swi6 WT

Vector: pET30a, KanR
MW: 43633
Extinction Coefficient: ε_{280nm} = 41370cm^{-1}M^{-1} \times (0.9481cm^{-1}mg/mL^{-1}):

1   atgcaccatc atcatcatca tttttcttgt ctggtgccac gcggttctgg tatgaaagaa
61  accgctgctg ctaatccga acgccagcac atggacagcc cagatctggg taccgacgc
121 gacgacaagg ccatggctga tatcggatcc gagaacctgt acttccaaag aggaatgaag
181  aagggggtg ttgatctta tgggcccttc tctactttca aaacgatcagt tatgacgac
The highlighted portion is translated to:

```
  1      HHHSHSSG   LVPRGSMKE   TAAKFERQH   MDSPDLTDD   DDKAMIGS   ENLYPGCMK
  61     KGGVRSYRRS   STSKRSVIDD   DSEPELSMT   KEAIASHKAD   SGSSDNEVES   DHESTSSSKK
 121    LKENAKEEEO   GEDEEEYEV   VEKVLKRMA   RKGGGYEYLL   KWEGYDDPSD   NTWSEADCS
 181     GCKQLIEAYW   NEHGRPEPS   KRKRATRPPK   PEAKEPSPKS   RKTEDKHDK   DSNEKIEDVN
 241    EKTIKFADS   QEEFENNGPP   SGQPQNHIES   DNEKSPSQQK   ESNESEDIQI   AETPSNVTPK
 301    KKPSPEVPKL   PDNRELTVQK   VENYDSWDEL   VSSIDTIERK   DDGTLEIYLT   WKNAGISHHP
 361     STITNKKCPQ   KMLQFYESHL   TFRENE-
```

The Swi6 sequence is in light gray, the His-tag is shown in orange, and the TEV protease cleavage site is shown in teal. TEV cleaves between the Q/G.
B.1.2 Swi6\textsuperscript{AcidX}

Swi6\textsuperscript{AcidX} is Swi6(E74-80A)

Vector: pET30a, Kan\textsuperscript{R}

MW: 43241

Extinction Coefficient: $\varepsilon_{280nm} = 41370\text{cm}^{-1}\text{M}^{-1}$ (0.9567/cm/mg/mL):

1 atgcaccatc atcatcatca ttcttctggt ctgggtccac gcggttctgg tatgaaagaa
61 accgctgctg cttaatgctga acgccagcac atggacagcc cagatctggg taccgagcac
121 gacgacaagg ccatggtctga tatcggtacctc gagaacctgt acttcccaag aggaatgaag
181 aaagggagtt ttgatcttta tcggtcctcc tctacttca aaagacagt tattgacgac
241 gattcggaac cagagttacc aagcatgaca aaggagctta ttctcttcctc taagaacagac
301 tctgggtcgt cgacaaatgc agtggtaaagc gccatgaaatga gtaaatcttc gtcgaagaa
361 ttgaaagaaa atgccaaaga agaagaagga ggacgaccag caccagacac agcatatgtt
421 gtgaaaaaagg tttaaaaaaca ccgtatgccc agaaaaggtg gaggctatga atacctttttg
481 aaatgaagaag gtatgacgta tccagctgtat aatcatgga gttcgaagac cgattgtagc
541 gttgcaaac agtggatagac agcatattgg aatgaacttg gaggagaagc agaaccctct
601 aaaaagaaga ggactgcaaa cccttaaaa acagtggaaac cccgaagcag aggacctcct accaaatctt
661 cgtaaaactg atgaagataaa acagcacaac gattccatag aaaaatgtaga atgtgtgaat
721 gagaaaaacta ttaaatctgc cgataatctg caagaggaat ttaatggaac cggccccactt
781 tccggtcagc cgaatgttca cattgaactt gacatgagt ctaaatctcc ttcacaaaaa
841 gattcgaaggg aagttgagaa tttcataata gcagagaccc ctgcaatgtt tactccgaa
901 aagagccta gtccggagaat tcccaactt ccggacaaaca gaggctgac gtttaaacaa
961 gtgaaaaaact cattctcggg ggaagacttg gtagctagca tcgataaaatt tgaacagaaaa
1021 gatgatggaaa ccttggaat tattctgact tggaagaacg gtgcataact ccatacatcct
1081 agtactatca ccaataaaaa atgtctcctg aagatgcttc aattttatga aagccactta
1141 acgtwccgtg aaaaatgaata a
The Swi6 sequence is in light gray, mutations to the Swi6 gene is shown in magenta, the His-tag is shown in orange, and the TEV protease cleavage site is shown in teal. TEV cleaves between the Q/G.

### B.1.3 Swi6\textsuperscript{LoopX}

Swi6\textsuperscript{LoopX} is Swi6(R93A,K94A)

**Vector:** pET30a, Kan\textsuperscript{R}

**MW:** 43491

**Extinction Coefficient:** $\varepsilon_{280nm} = 41370 \text{cm}^{-1} \text{M}^{-1}$ (09512/cm/\text{mg/mL})

```
1  MHRRHHSSG LVPRGSMKE TAAKFERQH MDSPDLGTD DDKAMADIGS ENLYEQCGMK
61  KGGVRSYRSS STSKRSVIDD DSEPILPSMKE KEAIASHKAD SG SSDNEVES DHESSSSKK
121 LKENAKEEG GAAAAAYV VEKVLKVRMA RKGGGYEYLL KWEGYDDPSD NTWSSEADCS
181 GCKQLIEAYW NEHGRPERS KRKTARPKK PEAKEPSPKS RKTDEDKHDK DSEEKIEDVN
241 EKTIKFADKS QEEFNEGPP SQGPNHIES DNEKSPSQK ESNEEDIQIQ AETPSNVTPK
301 KKPSPEVPKL PDNRELTVKQ VENYDSWEDL VSSIDTIERK DDGTIYVL TFKNGAISHHP
361 STITNKKCPQ KMLQEYESHL TFSNE-
```

```
1  atgcaccatc atcatcwtca ttctcttggt ctgggtccac gcgggtctgg tatgaaagaa
61  accgctgctg ctaataattcga agcccacagc agtgacagcc cagatctttg gagagtacagc
121 gacgacaagg ccattggctga tatccgatcc gagaacctgt actcccaagg aggaatgaag
181 aaaggaggtg ttgctatctta tggcgttcc ttctacttca aacgatcgctg tattgacgac
241 gattcgcaag caggtttacc aagcatgaca aagaggtcta ttgctttctca taagcagac
301 tttgggtctt ggtcaatagc agttgaaagc gatcatgaa gtaatccttc gtcgaagaaa
361 ttgaaagaaa atgcccaaga agaagagaaga ggagaagaga aagaagagga tgaatgttt
421 atagaaaaaa ctttaaaaca cctgttaggct ggacgaggtg gaggcaatag ataccttttg
481 aatggggaag gttaatatgga cccgatgtgat aatacatgga ttgcaagac cgattgtagc
541 ggttgcaaac agttgataga acatatttgg aatgaatcag gaggaagacc agaacctttttc
601 aaaaagaaaga ggactgcaag acctaaagag cccgaagcaaaa aggagccttc accaaagtct
```
The Swi6 sequence is in light gray, mutations to the Swi6 gene is shown in magenta, the His-tag is shown in orange, and the TEV protease cleavage site is shown in teal. TEV cleaves between the Q/G.

B.2 Buffers

B.2.1 Protease Inhibitors

1. Leupeptin (1 mg/mL in dimethyl sulfoxide (DMSO) aliquot, Sigma L2884)
2. Pepstatin A (1 mg/mL in DMSO aliquot, Sigma P5318)
3. PMSF (100 mM in DMF aliquot, Sigma P7626)
**B.2.2 10×PBS**

1. 137 mM sodium chloride
2. 2.7 mg potassium chloride (Sigma P9541)
3. 4.3 mM disodium phosphate (Sigma S3264)
4. 1.3 mM potassium phosphate monobasic (Sigma P9791)

**B.2.3 Lysis Buffer**

1. 1×PBS (section B.2.2)
2. 300 mM potassium chloride
3. 10% glycerol (Sigma G5516)
4. 7.5 mM imidazole (Sigma I5513)
5. 1 mM PMSF
6. 20 µg/mL Pepstatin A
7. 20 µg/mL Leupeptin
8. pH to 7.5 with hydrochloric acid/sodium hydroxide
   - Note: once protease inhibitor(s) (PIs) are added, keep frozen unless the buffer is being used (has protein in it)
   - Note: for 10 mM ATP and 10 mM magnesium chloride add to adenosine 5′-triphosphate disodium salt and magnesium chloride to 10 mM.

**B.2.4 Elution Buffer**

1. 25 mM HEPES (Sigma H3034)
2. 100 mM potassium chloride
3. 10% glycerol
4. 500 mM imidazole
5. pH to 7.5 with hydrochloric acid/sodium hydroxide
B.2.5 SE Buffer

1. 25 mM HEPES
2. 300 mM potassium chloride
3. 10% glycerol
4. 2 mM DTT (Sigma 43816)
5. pH to 7.5 with hydrochloric acid/sodium hydroxide

B.2.6 AX Buffer

1. 50 mM HEPES
2. 10% glycerol
3. 2 mM DTT
4. pH to 8.0 with hydrochloric acid/sodium hydroxide
   • Note: for AX + 40 mM potassium chloride add potassium chloride to 40 mM
   • Note: for AX + 1.5 M potassium chloride add potassium chloride to 1.5 M

B.2.7 Storage Buffer

1. 50 mM HEPES
2. 300 mM sodium chloride
3. 10% glycerol
4. 2 mM DTT
5. pH to 8.0 with hydrochloric acid/sodium hydroxide
B.3 Protein Expression

B.3.1 Transformation of Swi6 Expression Vectors

Materials

- LB+Agarose culture plate
  - LB+Agarose culture plates impregnated with 50 µL 1000×kanamycin (for long term storage)
  - LB+Agarose culture plates impregnated with 50 µL 1000×kanamycin and 50 µL 1000×chloramphenicol (for expression)
    
    Note: Add antibiotics ~1 h before plating cells and allow to dry at room temperature under a flame.

- * LB Broth (Sigma L3152)
- * Agar (Sigma Fisher BP1423)
- * kanamycin (Sigma 60615)
- * chloramphenicol (Sigma C0378)

- Autoclaved 5 mL 2×YT growth media
  
  For 1 L 2×YT:
  - 16 g Tryptone (Sigma T9410)
  - 10 g Yeast Extract (Sigma Y1625)
  - 5 g sodium chloride (Sigma S3014)

- Host cells
  - DH5α for storage OR
  - Rosetta(DE3)pLysS for expression (Millipore 71401-3)

- 80% autoclaved glycerol in dH₂O

Long Term Storage (DH5α glycerol stock)

1. Transform by standard procedures.
2. Plate onto LB+Agarose culture plate and incubate at 37°C for 16 hours before selecting single colonies
3. Inoculate single colonies into 5 mL LB + 50 µL 1000×kanamycin + 50 µL 1000×chloramphenicol.
4. Shake at 280 rpm, 37°C until the optical density measures between 0.2-0.5 at 600 nm through a 1 cm path length (OD$_{600nm,1cm}$)
5. Mix 750 µL culture for each clone with 250 µL 80% glycerol to make a 20% glycerol stock.
6. Immediately flash freeze in liquid nitrogen and store at -80°C.
7. Harvest remaining cells by mini-prep

**Expression Host (Rosetta(DE3)pLysS)**

1. Trasform by standard procedures.
2. Plate onto LB+Agarose culture plate and incubate at 37°C for 16 hours before selecting single colonies
3. Proceed immediately to section B.3.2

**B.3.2 Trial Induction**

In this section we will determine which colonies picked during section B.3.1 produce the highest yield of Swi6.

**Materials**

- Autoclaved 5 mL 2×YT growth media
- 50 mg/mL (1000×) kanamycin
- 34 mg/mL (1000×) chloramphenicol
- 100 mM IPTG (isopropyl-β-D-1-thiogalactopyranoside Sigma I6758)
- 80% autoclaved glycerol in dH$_2$O
- 12% SDS-PAGE gel
- 1×Tris-Glycine-SDS buffer
- 6×SDS loading dye
- Gel stain
- Gel destain
- 0.5×TE

**Induction**

1. Innoculate 3-5 single transformed Rosetta(DE3)pLysS colonies from section B.3.1 into separate tubes of 5 mL of 2×YT + 50 µg/mL kanamycin + 34 µg/mL chloramphenicol and shake at 280 rpm, 37°C until the OD_{600nm,1cm} = 0.2 for a given clone.
2. Mix 750 µL parts culture for each clone with 250 µL part 80% glycerol to make a 20% glycerol stock.
3. Immediately flash freeze in liquid nitrogen and store at -80°C.
4. Allow the remaining culture for each clone to continue to shake at 280 rpm, 37°C until OD_{600nm,1cm} = 0.5
5. Transfer 100-200 µL t a 1.5 mL microcentrifuge tube on ice for a pre-induction time point.
6. Induce the remaining culture with 100 mM IPTG to 0.5 mM final concentration.
7. Allow induced culture to continue to grow at 280 rpm, 37°C.
8. Each hour up to 4 transfer 100-200 µL of culture to a microcentrifuge tube on ice.

**SDS-PAGE analysis**

1. Discard the remaining culture after a given clone has been induced for the full time series
2. Centrifuge each time point at 2,500 g for 1 min. Remove supernatant and suspend each pellet in 10 µL of 0.5×TE; Add 10 µL 6×SDS loading dye; boil at 95°C for 5 min; centrifuge at 2,500 g for 10 sec and store on ice
Figure B.1: 12% SDS-PAGE gel of a Swi6\textsuperscript{WT} trial induction. The lanes are: molecular weight marker (MWM); culture 1: 0 h, 1 h, 2 h, 3 h, 4 h; culture 2: 1 h, 2 h, 3 h, 4 h; culture 3: 1 h, 2 h, 3 h, 4 h. In this case each culture grew similarly so culture 1 was selected.
Growth appears to stagnate after 2 hours so the bulk expression is chosen to proceed for 2 hours after induction.

3. Prepare at 12% SDS-PAGE gel. Load 10 µL of each time point onto gel reserving a lane for a molecular weight ladder. Run in 1×Tris-Glycine-SDS per standard protocol.

4. Stain and destain the gel

5. Determine the best colony based on expression levels relative to the background proteome. (see Figure B.1)

6. Discard glycerol stocks of colonies NOT chosen.

**B.3.3 Bulk Induction**

**Materials**

- Autoclaved 5 mL 2×YT growth media
- Autoclaved 250 mL Erlenmeyer flask containing 60 mL 2×YT growth media
- Autoclaved 1 L Erlenmeyer flask containing 600 mL 2×YT growth media (as many as
Figure B.2: 12% SDS-PAGE gel of a Swi6\textsuperscript{WT} trial induction. The lanes are: MWM; flask 1: 0 h, 1 h, 2 h; blank lane; flask 2: 1 h, 2 h. Each flask shows IPTG dependent and strong expression of Swi6\textsuperscript{WT}

- 50 mg/mL (1000×) kanamycin
- 34 mg/mL (1000×) chloramphenicol
- 1 m IPTG (isopropyl-β-D-1-thiogalactopyranoside Sigma I6758)
- 12% SDS-PAGE gel
- 1×Tris-Glycine-SDS buffer
- 6×SDS loading dye
- Gel stain
- Gel destain
- 0.5×TE

Expression

1. Inoculate 5 mL 2×YT growth tube with 50 µL kanamycin and 34 µL chloramphenicol with a scraping from the desired glycerol stock made in section B.3.1. Mix by swirling to homogenize the sample.
2. Perform a serial dilution resulting in 1/100, 1/1000, and 1/10000 dilutions of the original 1/1 growth tube into 5 mL 2×YT + 50 µg/mL kanamycin + 34 µg/mL chloramphenicol.

3. Shake each tube overnight at 280 rpm, 37°C.

4. Determine which tube has grown to optimal conditions, 0.2 < OD₆₀₀nm,₁cm<0.5. This may require waiting. If so, prewarm a 250 mL Erlenmeyer flask containing 60 mL 2×YT + 50 µg/mL kanamycin + 34 µg/mL chloramphenicol by shaking at 280 rpm, 37°C.

5. Inoculate 250 mL Erlenmeyer flask with the appropriate culture and allow to grow until 0.2 < OD₆₀₀nm,₁cm<0.5 by shaking at 280 rpm, 37°C.

6. Pre-warm desired number of 1 L Erlenmeyer flasks containing 600 mL 2×YT.

7. When the 250 mL flasks have reached an appropriate OD, 0.2 < OD₆₀₀nm,₁cm < 0.5, add antibiotics to the 1 L Erlenmeyer flasks to a final concentration of 50 µg/mL kanamycin + 34 µg/mL chloramphenicol and transfer equal volumes of the culture to each 1 L flask.

8. Continue to grow at 280 rpm, 37°C until OD₆₀₀nm,₁cm = 0.5

9. Prepare a 0 h time point as before

10. Induce each 1 L Erlenmeyer flask with 0.5 mM final concentration of 1 mM IPTG.

11. Continue to grow for 2 h Taking time points every hour.

12. Immediately move on to section B.3.4.

13. While going through the protocol as described in section B.3.4, prepare and analyze an SDS gel as described in section B.3.2 and verify the bulk expression (see Figure B.2).

### B.3.4 Harvesting Cells

**Materials**

- Lysis buffer

  **Note:** the Protease Inhibitors in this buffer are only good for ~1–2 hours thawed and are very expensive. Make only what is needed immediately before use.
Procedure

1. Harvest cells by spinning at 4000 g for 10 min at 4°C.
2. Pour off the supernatant if clear, if not, repeat the centrifugation.
   • Note: 1 L of cells yields ~8 g cell paste
3. Resuspend the cell paste in 10 mL Lysis buffer.
4. Transfer resuspended cells to a 50 mL conical bottom tube and watch the bottle with another 10 mL of Lysis buffer.
5. Flash freeze in liquid nitrogen and store at -80°C until purification up to 3 months later.

B.4 Protein Purification

B.4.1 Cell Lysing

Materials

• Qsonica Sonicator with microtip
• Cell pellets from section B.3.4 suspended in 20 mL Lysis buffer

Procedure

1. Remove frozen cell pellets from -80°C freezer
2. Thaw at room temperature with occasional swirling/inversion in an attempt to homogenize thawing
3. Once thawed, place on ice
4. Using a microtip in a Qsonica sonicator lyse cells by the following program while on ice:
   • Amplitude 6
   • Program duration 1 min
5. Repeat the above program a total of 6 times (6 min) total sonication time) for each cell pellet. Sonicate alternating pellets after each program to minimize local heating due to sonication.

- The cells will become slightly less viscous and brown as they lyse.

6. Immediately following lysing, pellet the lysed cell debris for 20 min at 25,000 g, 4°C.

- Decant supernatant into a new centrifuge tube.

7. Spin supernatant for an additional 25 min at 40,000 g, 4°C.

8. Decant supernatant into a 50 mL conical bottom tube and store on ice. Pool to form Fraction I.

B.4.2 Ni-NTA Purification

Materials

- nickel-nitrilotiacetic acid (Ni-NTA) agarose beads
- Lysis buffer
- ATP disodium salt
- magnesium chloride
- Lysis buffer + 10 mM ATP + 10 mM magnesium chloride (section B.2.3)
- Elution buffer (section B.2.4)
- C10 column with appropriate caps and tubes
- Fraction I
- 12% SDS-PAGE gel
- 1×Tris-Glycine-SDS buffer
- 6×SDS loading dye
- Gel stain
- Gel destain
• 10 kDa, 15 mL MWCO amicon ultra

Prewash Ni-NTA Agarose

1. Pipette 1/3 mL Ni-NTA per liter of original cell growth volume into a 50 mL conical bottom tube.
2. Add 2× volume of Lysis buffer to the Ni-NTA beads.
3. Invert to mix
4. Centrifuge at 5,000 g, 4°C for 5 min
5. Turn the tube 180° in the centrifuge and repeat the centrifugation
6. Decant supernatant and discard
7. Repeat the wash steps a total of three times finishing with decanted, Ni-NTA agarose beads equilibrated with lysis buffer

Bind to Ni-NTA Agarose

1. Add Fraction I to decanted Ni-NTA beads and resuspend.
2. Rotate on a lab rotisserie for 30 min at 4°C.

Remove Chaperone Proteins

1. Pellet the Ni-NTA resin by centrifugation at 5,000 g, 4°C for 5 min.
2. Decant supernatant and save for later analysis
3. Resuspend the beads in 1× the original volume of lysis buffer + 10 mM ATP + 10 mM magnesium chloride.
4. Rotate on a lab rotisserie for 30 min at 4°C.

Prepare C10 Column and FPLC

Perform the following while the lysate is going through its last incubation with lysis buffer + 10 mM ATP + 10 mM magnesium chloride.
Figure B.3: FPLC chromatogram of Swi6 elution from Ni-NTA agarose column. Fractions between tags 1 and 2 were pooled to produce Fraction II.
Figure B.4: 12% SDS-PAGE gel of fractions eluted from Ni-NTA agarose gel. The His-tag which binds the nickel resin will later be cleaved off. All fractions were pooled and further purified by Size Exclusion.

1. Clean a C10 column, 1 bottom adapter, and 1 adjustable plunger adapter suitable for use on an FPLC
2. Set FPLC up by equilibrating line A into lysis buffer and line B into elution buffer
3. Once incubation with Lysis buffer + 10 mM ATP + 10 mM magnesium chloride is finished centrifuge the Ni-NTA at 5,000 g, 4°C for 5 min.
4. Ensure the outlet of the C10 column is plugged and fill with ∼1 mL of lysis buffer.
5. Pack the C10 column with bound Ni-NTA beads by pipetting. Use flow to help pack the column but do not allow the resin to dry.
6. Assemble the column onto an FPLC at 4°C for purification.
7. Program:
   - 100% Lysis buffer at 1.5 mL/min for 2mL
   - Immediately change to 100% elution buffer
   - flow for 2 column volumes while collecting fractions
8. Verify Ni-NTA purification by 12% SDS-PAGE (see Figures B.3 and B.4)
9. pool fractions containing Swi6 for form Fraction II.
10. if Fraction II is larger than the appropriate volume for the Superdex Hilo& 600 column (∼32mL), concentrate in a 10 kDa MWCO, 15 mL amicon ultra. Centrifuge at 2,500 g in 20 min increments. Mix to prevent precipitation of the higher concentrated sample near the bottom of the concentrator. Repeat centrifugation until the total volume is less than 10% of the bead volume.
B.4.3 Size Exclusion Purification

Materials

- SE buffer (size exclusion buffer, section B.2.5)
- HiLoad 26/600 Superdex 200
- 50 mL Superloop
- 12% SDS-PAGE gel
- 1×Tris-Glycine-SDS buffer
- 6×SDS loading dye
- Gel stain
- Gel destain
- 10 kDa, 15 mL MWCO amicon ultra
- Fraction II

Purification

1. Equilibrate the column and superloop into SE buffer. This will take ~a day.

2. Load Fraction II onto the superloop at ~1 mL/min

3. Flow for 250 mL at 1 mL/min

4. Collect 2 mL fractions from ~30 mL to ~200 mL (see Figure B.5).

SDS-PAGE Verification

1. Verify size exclusion purification by 12% SDS-PAGE (see Figure B.6)

2. Pool fractions containing Swi6 for form Fraction III.

3. If necessary, concentrate Fraction III to 10 mL in a 10 kDa, 15 mL MWCO amicon ultra at 2,500 g, 4°C mixing every 20 minutes to prevent aggregation.
Figure B.5: FPLC chromatogram of Swi6 elution from HiLoad Superdex column. Fractions between tags 1 and 2 were pooled to produce Fraction III.
Figure B.6: 12% SDS-PAGE gel of fractions eluted from size exclusion column. All fractions corresponding to Swi6 were pooled and further purified (see section B.4.3).

B.4.4 TEV Cleavage of the His-Tag

Materials

- tobacco etch virus (TEV) protease
- AX buffer + 40 mM potassium chloride (section B.2.6)
- 6-8 kDaMWCO dialysis tubing
- 12% SDS-PAGE gel
- 1×Tris-Glycine-SDS buffer
- 6×SDS loading dye
- Gel stain
- Gel destain
- Fraction III

Cleavage

1. Measure the total protein mass by UV absorption.
2. Add TEV protease to a final concentration ratio of 1:100 (w/w) TEV protease:Swi6.
3. mix by inversion
4. Dialyze the protein against AX buffer + 40 mM potassium chloride at 4°C overnight.
5. The following morning reclaim to form Fraction IV
SDS-PAGE verification

1. Verify TEV protease cleavage by 12% SDS gel (see Figure B.7)
2. If cleavage is not complete, add more TEV protease and fresh DTT, and allow the reaction to proceed for 1 h at room temperature. Repeat gel to verify the His-tag has been cleaved off.
3. Immediately proceed Figure B.4.5.

B.4.5 Anion Exchange Purification

Anion exchange chromatography is used to purify cleaved Swi6 from TEV protease and cleaved His-tag (see Figure B.8).

Materials

- AX buffer (anion exchange buffer, section B.2.6)
- AX buffer + 40 mM potassium chloride (section B.2.6)
- AX buffer + 1.5 M potassium chloride (section B.2.6)
- HiTrap Q HP 5mL
- 50 mL Superloop
- 12% SDS-PAGE gel
- 1×Tris-Glycine-SDS buffer
- 6×SDS loading dye
• Gel stain
• Gel destain

Anion Exchange

1. Equilibrate the column and superloop in AX buffer
2. Dilute Fraction IV from 40 mM potassium chloride by 5-fold to ~8 mM potassium chloride.
3. Load Fraction IV on to superloop at 1 mL/min.
4. Elution Program:
   • Flow continuously at 5 mL/min
   • Inject onto the HiTrap Q HP column and collect flow through
   • Elute with a linear gradient from AX buffer to 80% AX buffer + 1.5 M potassium chloride.
   • Hold at 80% AX buffer + 1.5 M potassium chloride for 2 column volumes.
   • Collect 2 mL fractions
   • Swi6 will elute at ~40% AX buffer + 1.5 M potassium chloride.

SDS-PAGE verification

1. Verify purity of Swi6 once more by SDS-PAGE.

B.5 Protein Storage

Swi6 can be stored immediately following anion exchange purification. Alternatively, the protein may be concentrated using a 10 kDa, 15 mL MWCO amicon ultra, centrifuge at 2,5000 g, 4°C and buffer exchanged by repeated dilution into storage buffer. Aliquot, and flash freeze in liquid nitrogen for storage at -80°C.
B.5.1 Storage Buffer

- 50 mM HEPES
- 300 mM sodium chloride
- 10% glycerol
- 2 mM DTT
- pH to 8.0 with sodium hydroxide and hydrochloric acid

B.5.2 Stability

Swi6 has proven to be quite a stable protein. It remains active on ice for months and is insensitive to salt conditions and pH. Tests have been done from 40 mM potassium chloride to 600 mM potassium chloride and pHs from 6.0 to 8.0.
Figure B.8: FPLC chromatogram of Swi6 elution from HiTrap Q HP column. Fractions between markers 1 and 2 were pooled and represent fully purified Swi6.
Appendix C

SITE-SPECIFIC DEPOSITION OF TRIMETHYLYSINE ANALOG

This protocol is based on Simon et al. [84]. It was adapted for use in the Poirier lab by Justin North in 2010 and further adapted to improve labeling efficiency by me in 2012.

It is assumed in this protocol to have previously purified 5-10 mg aliquots of the histone to label mutated such that the sole Cys residue is located at the desired position of the MLA.

C.1 Materials

- (2-bromoethyl)-trimethylammonium bromid (Sigma 117196)
- Alkylation Buffer
  - 1 M HEPES
    - Note: 310 mM HEPES (Sigma H3034) and 690 mM glshepes sodium salt (Sigma H7006) should yield the correct pH of 7.8. Adjust with HEPES (down) and HEPES sodium salt (up).
  - 10 mM D/L-Methionine (Sigma M9500)
  - 4 M Guanidine-HCl (MP Biomedicals 820539)
  - Flash freeze and store at -80°C in 1.5 mL aliquots
- 333 mM DTT (Invitrogen 15508-013)
Note: Store DTT as a dry powder in a dessicator. Otherwise the activity will decrease overtime and the correct concentrations to add to the reaction will need to be determined empirically.

C.2 Labeling

1. Resuspend 5-10 mg histone to be labeled in 980 µL of alkylation buffer and allow to unfold for 1 h.
2. While unfolding the histones, set up a heated stir plate with a 1.5 mL tube block. Place a 1.5 mL tube with a stir bar, thermocouple, and parafilm cap to secure the thermocouple on the block. Cover from light and air currents. Heat the thermocouple to 50°C while stirring at maximum speed (typically temperature set to ~55-60°C results in 50°C). Check the temperature often over the next few hours to ensure a stable 50°C.
3. Mix the unfolded histone and transfer to a clean 1.5 mL tube.
4. Add 20 µL 333 mM DTT and allow to reduce at 37°C for 1 h.
5. Transfer the histones to a new 1.5 mL tube containing 100 mg (2-bromoethyl)trimethylammonium bromide and a flea stir bar. Place in the heated block and replace the light/air cover.
6. Allow the reaction to proceed for 2.5 h.
7. Add another 20 µL 333 mM DTT to the reaction.
8. Allow the reaction to proceed an additional 2.5 h.
9. Quench the reaction with 50 µL 14.3 mM BME.
10. Dialyse against at least 4 1.8 L changes of 3 mM BME.
11. Lyophilize the protein
12. Resuspend the protein in dH$_2$O and measure the concentration via UV-VIS spectrophotometry absorbance.
13. Aliquot into 1–10× (0.1–1 mg for H3, equal molar for other histones) aliquots
14. Vacuum concentrate the samples to dryness and store at -80°C.
C.3 Storage

Histones harboring MLA labels appear to be stable for years at -80°C. However, once resuspended they appear to lose their methyl groups on the order of weeks. It has been determined that storage of the MLA-labeled sample in 0.5 mM potassium phosphate pH 7.5 mitigates this lose of label and extends the lifetime of the sample to typical levels. It appears the free amines in Tris attack the methyl groups and remove them so any buffer used to store MLA-labeled histones should be free of amines.

C.4 MALDI-TOF Verification

Labeling efficiency must be verified by MALDI-TOF as discussed in section 3.3. Histones should be diluted to no more than 0.1 mM sodium chloride before mass spec. A ratio of 2:1 matrix:histone should be used to deposit a sample on a mass spectrometry plate.

The verification of labeling efficiency is at best semi-quantitative but a small relative labeled vs unlabeled peak indicates a successful labeling reaction (see Figure C.1).
Figure C.1: MALDI-TOF mass spectrometry showing the rough efficiency of an MLA deposition on H3 lysine 36 (mutated to a cysteine). The result is semi-quantitative because protein desorption is not uniform across all species of molecule and is sensitive to relatively small changes in molecular composition. For this histone, H3.3K36C, the unlabeled +2 peak (meaning the protein’s total charge is 2e+) has a mass/charge ratio of 7570 Da/e+ and the labeled peak, a ratio of 7613.5 Da/e+.
Appendix D

PURIFICATION OF MULTIPLY LABELED HISTONE OCTAMPERS FOR USE IN FLUORESCENCE ASSAYS

Octamer may be purified with multiple Cysteines as long as no one protein contains two Cysteines. First, individual histones may be maleimide labeled in unfolding condition. However, in many cases, notably when the label is applied to this histone tails, labeling efficiency is increased if labeling is done after folding the histone into a complex. For this reason it is recommended to design nucleosomes such that one Cysteine is located on the tetramer and one on the dimer. In the examples used in this thesis, MLAs were deposited on the H3 during unfolding conditions and the fluorescent Cy5 molecule on H2AK119C after refolding into H2A/H2B dimer.

D.1 Procedure

1. Refold H2A/H2B and H3/H4 separately. The MLA should already be deposited, but the fluorophore should be deposited while folded, if possible.
2. Label the desired complex with the fluorescent label by the typical protocol.
3. Purify the fluorescently labeled complex from free dye by size exclusion as if it were octamer.
4. Combine tetramer and dimer to a final molar ratio of 1:2.2 tetramer:dimer.
5. Allow to rotate at 4°C on a lab rotisserie overnight.

6. Purify the octamer by size exclusion replacing refolding buffer with 5 mM PIPES pH 6.1 2 M sodium chloride.

7. Store the MLA labeled octamer in a 5 mM PIPES pH 6.1 2 M sodium chloride.
Appendix E

Light Path Shutters for Triclops

Each laser microscope in the lab has been outfitted with shutters produced by Electro-Optical Products Corp. The circuit which controls the shutter must be built before the shutter can be used. The design is identical to that supplied by Electro-Optical Products Corp but the layout shown in Figure E.1 is my own. Two or Three shutter controllers can be placed on a single prototyping board using these plans.

The circuit uses TTL logic to control whether the shutter is open or closed. $V_{\text{in}}$ requires a 12 V power supply with a minimum of 150 mA current capacity. $V_{\text{sig}}$ is the TTL logic input with TTL high being closed and low being open.
Figure E.1: Circuit layout for laser safety shutters. A schematic for either two (left) or three (right) controllers is provided with the wiring for each displayed below the component placement for each design. For each controller, the 4.7 kΩ resistor is the signal input and the 39 Ω resistor is the power supply input. The actual circuit diagram is available at Electro-Optical Products Corp’s website, http://www.eopc.com/sh10_cad.html.
# E.1 Parts List

<table>
<thead>
<tr>
<th>Product Code</th>
<th>Vendor</th>
<th>Part Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>PN2907A-APCT-ND</td>
<td>Digikey</td>
<td>PN2907A transistor</td>
</tr>
<tr>
<td>CF14JT1K00CT-ND</td>
<td>Digikey</td>
<td>1 kΩ resistor, $\frac{1}{4}$W. Need two per controller circuit.</td>
</tr>
<tr>
<td>CF14JT10K0CT-ND</td>
<td>Digikey</td>
<td>10 kΩ, $\frac{1}{4}$W resistor</td>
</tr>
<tr>
<td>CF14JT4K70CT-ND</td>
<td>Digikey</td>
<td>4.7 kΩ, $\frac{1}{4}$W resistor</td>
</tr>
<tr>
<td>1N5231BFSCT-ND</td>
<td>Digikey</td>
<td>Zener Diode</td>
</tr>
<tr>
<td>PPC39W-2CT-ND</td>
<td>Digikey</td>
<td>39 Ω, 2 W resistor</td>
</tr>
<tr>
<td>1N4933-E3/54GICT-ND</td>
<td>Digikey</td>
<td>Diode</td>
</tr>
<tr>
<td>P5547-ND</td>
<td>Digikey</td>
<td>4,700 µF aluminum capacitor</td>
</tr>
<tr>
<td>TIP100-BP-ND</td>
<td>Digikey</td>
<td>TIP100 MOSFET</td>
</tr>
<tr>
<td>377-2076-ND</td>
<td>Digikey</td>
<td>Prototyping board (large enough for two controllers)</td>
</tr>
</tbody>
</table>

Table E.1: Parts list for shutter controllers
Appendix F

Triclops

A prism-based smFRET three-color microscope is built for this work to expand the capabilities currently available with a simpler, two-color smTIRF microscope.

The major design is described below and sectioned into two major areas. The excitation path (section F.1) is responsible for control of laser light and beam-steering toward ultimate TIRF-field production. The emission path (section F.2) consists of the microscope body itself as well as a custom-made image splitter and electron multiplying CCD (EMCCD) camera.

F.1 Excitation Path

A ray diagram of the excitation path is shown in Figure F.1. There are no major advancements to the excitation path over previous microscopes. Briefly, light is generated by CW, diode-pumped lasers, filtered through narrow bandpass dichroic beam filters, and controlled via electronic shutter. The warm up time for diode-pump lasers is on the order of minutes and prior to reaching steady state higher order modes are present in the emission. By controlling the lasers with shutters we can ensure uniform and predictable excitation of fluorophores while switching excitation wavelengths potentially on the order of 10’s of ms. Each laser is then individually steered via dichroic beamsplitters onto the same path which is then brought up and focused on a Pellin-Broca prism to ultimately create the TIRF field
Figure F.1: (A) Ray diagram of the Excitation portion of Triclops. The inset is 90° relative to the rest and shows the pellin-broca prism responsible for creating the total internal reflection field. (B) Ray diagram of the image splitter/emission portion of Triclops. D3, D4, D3.2, D4.2, F3, F4, and F5 are all interchangeable to allow for different fluorophores to be used on the system. Angles of deflection are exaggerated.

as discribed in section 4.2.1

**F.2 Emission Path**

The excitation side establishes a uniform and consistent TIRF field. Preparation of samples was discussed in section 4.2.3. If we assume that fluorophores are then excited by the TIRF field, the emitted light is then collected by a silicone oil-based objective which combines the high NA of oil objectives with the extreme working distances of water objectives. The microscope body then directs the fluorescent light out of a side port where it enters the image splitter.

Light emerging from the microscope body is focused onto an adjustable slit which defines the physical size of each fluorescence channel down the line. The slit is adjustable to ensure the EMCCD chip is fully covered by the fluorescent image. Light is then re-collimated Figure F.1B L2 and passed onto a system of dichroic beam splitters which shift each color in space to allow for simultaneous measurement of each channel on the EMCCD. Each beamsplitter and filter is placed on a removable mount to allow for different fluorophore schemes in experiments with minimal alignment required.
The first two beamsplitters, D3 and D4 in Figure F.1 are long pass filters which necessitates the rightmost path as the shortest wavelength path. We are currently set up for 5 fluorophores giving a large number of potential fluorophore combinations:

- Alexa Fluor 488, Cy3, Cy5
- Alexa Fluor 488, Cy3, Cy5.5
- Alexa Fluor 488, Cy3, Cy7
- Alexa Fluor 488, Cy5, Cy5.5
- Alexa Fluor 488, Cy5, Cy7
- Alexa Fluor 488, Cy5, Cy5.5
- Alexa Fluor 488, Cy5.5, Cy7
- Alexa Fluor 488, Cy7
- Cy3, Cy5, Cy5.5
- Cy3, Cy5, Cy7
- Cy3, Cy5.5, Cy7
- Alexa Fluor 488
- Cy3
- Cy5
- Cy5.5
- Cy7

Each of these combinations runs from shortest wavelength to longest and can be achieved by simply swapping out beamsplitters and bandpass filters. A significant spectral overlap exists between

- Alexa Fluor 488 and Cy3
- Cy3 and Cy5
- Cy3 and Cy5.5
- Cy3 and Cy7
- Cy5 and Cy5.5
- Cy5 and Cy7
- Cy5.5 and Cy7

and so any of these combinations may be used to probe small scale conformational changes while also tracking a co-localization event or a second conformational change.
After the set of beamsplitters each individual color is imaged onto a unique portion of the camera which, when calibrated, allows for simultaneous tracking of the emission from each fluorophore at a single molecule level.

### F.3 Triclops Parts

#### F.3.1 Major Components

<table>
<thead>
<tr>
<th>Product Code</th>
<th>Vendor</th>
<th>Part Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>IX73</td>
<td>Olympus</td>
<td>μscope: Microscope body</td>
</tr>
<tr>
<td>U2B712</td>
<td>Olympus</td>
<td>Objective: silicone immersion 60x long working distance objective</td>
</tr>
<tr>
<td>iXon ultra 897</td>
<td>Andor</td>
<td>EMCCD: Camera detector</td>
</tr>
<tr>
<td>MBL-III-473-50</td>
<td>UltraLaser</td>
<td>473: 50 mW 473 nm low noise laser</td>
</tr>
<tr>
<td>CL532-50-L</td>
<td>CrystaLaser</td>
<td>532: 50 mW 532 nm low noise laser</td>
</tr>
<tr>
<td>DL638-50</td>
<td>CrystaLaser</td>
<td>638: 50 mW 638 nm low noise laser</td>
</tr>
</tbody>
</table>

Table F.1: Major components of the triclops microscope. The IX73 is a modular microscope body and minimal additions include objective mount, filter cube assembly, and epi-illumination. Excitation lasers were chosen to excite Alexa Fluor 488, Cy3, Cy5, and Cy5.5. A fourth laser of a longer wavelength is helpful for directly exciting Cy7.
### F.3.2 Excitation Parts

<table>
<thead>
<tr>
<th>Product Code</th>
<th>Vendor</th>
<th>Part Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ZET473/10x</td>
<td>Semrock</td>
<td><strong>F1</strong>: 473 nm laser cleanup filter</td>
</tr>
<tr>
<td>FF01-640/14-25</td>
<td>Semrock</td>
<td><strong>F2</strong>: 638 nm laser cleanup filter</td>
</tr>
<tr>
<td>FF02-531/22-25</td>
<td>Semrock</td>
<td><strong>F3</strong>: 532 nm laser cleanup filter</td>
</tr>
<tr>
<td>SH-10-12-MP</td>
<td>Electro-optical Products</td>
<td><strong>S1–3</strong>: Optical shutter. Control circuit described in Appendix E</td>
</tr>
<tr>
<td>PF10-03-P01</td>
<td>ThorLabs</td>
<td><strong>M1–4</strong>: Coated silver mirror</td>
</tr>
<tr>
<td>KS1</td>
<td>ThorLabs</td>
<td><strong>M1 mount</strong>: Tip-tilt mirror mount</td>
</tr>
<tr>
<td>KS2D</td>
<td>ThorLabs</td>
<td><strong>M2 mount</strong>: Tip-tilt precision mirror mount</td>
</tr>
<tr>
<td>H45</td>
<td>ThorLabs</td>
<td><strong>M3 mount</strong>: 45° vertical-bounce mirror mount</td>
</tr>
<tr>
<td>KS6XS</td>
<td>ThorLabs</td>
<td><strong>M4 mount</strong>: Six axis kinetic mount (overkill KS1 recommended)</td>
</tr>
<tr>
<td>FF505-SDi01-25x36</td>
<td>Semrock</td>
<td><strong>D1</strong>: 638 nm short-pass beamsplitter</td>
</tr>
<tr>
<td>NFD01-532-25x36</td>
<td>Semrock</td>
<td><strong>D2</strong>: 532 nm notch beamsplitter</td>
</tr>
<tr>
<td>KM100S</td>
<td>ThorLabs</td>
<td><strong>D1–2 mount</strong>: Tip-tilt square element mount</td>
</tr>
<tr>
<td>LA1509</td>
<td>ThorLabs</td>
<td><strong>L1</strong>: 100 mm focal length plano-convex lens</td>
</tr>
<tr>
<td>PLBC-5.0-79.5-SS</td>
<td>Melles-Griot</td>
<td><strong>Prism</strong>: Pellin-Broca fused silica prism</td>
</tr>
</tbody>
</table>

Table F.2: Components comprising the Excitation path.
## F.3.3 Emission Parts

<table>
<thead>
<tr>
<th>Product Code</th>
<th>Vendor</th>
<th>Part Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>FF01-496/LP-25</td>
<td>Semrock</td>
<td><strong>F4</strong>: 473 nm long pass laser cutout filter</td>
</tr>
<tr>
<td>VA100</td>
<td>ThorLabs</td>
<td><strong>Slit</strong>: Adjustable 1 cm slit</td>
</tr>
<tr>
<td>RSP05</td>
<td>ThorLabs</td>
<td><strong>Slit mount</strong>: Locking rotation mount</td>
</tr>
<tr>
<td>AC254-100-A</td>
<td>ThorLabs</td>
<td><strong>L2</strong>: 100 mm convex-convex de-focusing lens</td>
</tr>
<tr>
<td>PF10-03-01</td>
<td>ThorLabs</td>
<td><strong>M5–8</strong>: Coated silver mirror</td>
</tr>
<tr>
<td>KS1</td>
<td>ThorLabs</td>
<td><strong>M5 mount</strong>: Tip-tilt mirror mount</td>
</tr>
<tr>
<td>LMR1</td>
<td>ThorLabs</td>
<td><strong>M6–7 mount</strong>: Fixed mirror mount</td>
</tr>
<tr>
<td>KS2D</td>
<td>ThorLabs</td>
<td><strong>M8 mount</strong>: Tip-tilt precision mirror mount</td>
</tr>
<tr>
<td>AC254-100-A</td>
<td>ThorLabs</td>
<td><strong>L3</strong>: mm convex-convex re-focusing lens</td>
</tr>
<tr>
<td>NF03-532E-25</td>
<td>Semrock</td>
<td><strong>F5</strong>: 532 nm notch cutout filter</td>
</tr>
<tr>
<td>T570lxpr-UF2 25.5x36</td>
<td>Chroma Tech</td>
<td><strong>Alexa Fluor 488 beamsplitter</strong>: 2 mm thickness long pass filter</td>
</tr>
<tr>
<td>FF01-520/35</td>
<td>Semrock</td>
<td><strong>Alexa Fluor 488 bandpass</strong>: 520 nm centered Alexa 488 emission bandpass</td>
</tr>
<tr>
<td>T635lpxr-UF2</td>
<td>Chroma Tech</td>
<td><strong>Cy3 beamsplitter</strong>: 2 mm thickness long pass filter</td>
</tr>
<tr>
<td>FF01-600/37</td>
<td>Semrock</td>
<td><strong>Cy3 bandpass</strong>: 600 nm centered Cy3 emission bandpass</td>
</tr>
<tr>
<td>T685lpxf-UF2</td>
<td>Chroma Tech</td>
<td><strong>Cy5 beamsplitter</strong>: 2 mm thickness long pass filter</td>
</tr>
<tr>
<td>FF01-661/20</td>
<td>Semrock</td>
<td><strong>Cy5 bandpass</strong>: 661 nm centered Cy5 emission bandpass</td>
</tr>
<tr>
<td>FF01-716/40</td>
<td>Semrock</td>
<td><strong>Cy5.5 bandpass</strong>: 716 nm centered Cy5.5 emission bandpass</td>
</tr>
<tr>
<td>FF01-775/46-25</td>
<td>Semrock</td>
<td><strong>Cy7 bandpass</strong>: nm centered Cy7 emission bandpass</td>
</tr>
<tr>
<td>KM100S</td>
<td>ThorLabs</td>
<td><strong>beamsplitter mount</strong>: Square tip-tilt mount</td>
</tr>
<tr>
<td>LMR1</td>
<td>ThorLabs</td>
<td><strong>Bandpass mount</strong>: Fixed round optic mount</td>
</tr>
<tr>
<td>KB1X1</td>
<td>ThorLabs</td>
<td><strong>Kinetic mount</strong>: High-repeatability detachable magnetic plates</td>
</tr>
</tbody>
</table>

Table F.3: Components comprising the modular emission path.
F.4 Calibration

There is a basic assumption in our data analysis that the image splitter has been calibrated such that each separate color, when overlain, spatial maps to the others. In order to facilitate this, a labVIEW program was written which programmatically overlays different sections of the image. In this way, a calibration slide containing some micro-patterned field which can be aligned with itself.
Permission has been obtained for any figure or extended section of text for which I do not hold the copyright. These permissions are included here for completeness and reference the section or figure for which permission was obtained.

G.1 Figure 1.1 Permission to Reuse

This figure is reused from:


Sha and Boyer [3] is published under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction provided proper citation.
G.2 Figure 1.3 Permission to Reuse

Title: Quantitative Proteomic Analysis of Histone Modifications
Author: He Huang, Shu Lin, Benjamin A. Garcia, et al
Publication: Chemical Reviews
Publisher: American Chemical Society
Date: Mar 1, 2015
Copyright © 2015, American Chemical Society

PERMISSION/LICENSE IS GRANTED FOR YOUR ORDER AT NO CHARGE

This type of permission/license, instead of the standard Terms & Conditions, is sent to you because no fee is being charged for your order. Please note the following:

- Permission is granted for your request in both print and electronic formats, and translations.
- If figures and/or tables were requested, they may be adapted or used in part.
- Please print this page for your records and send a copy of it to your publisher/graduate school.
- Appropriate credit for the requested material should be given as follows: "Reprinted (adapted) with permission from (COMPLETE REFERENCE CITATION). Copyright (YEAR) American Chemical Society." Insert appropriate information in place of the capitalized words.
- One-time permission is granted only for the use specified in your request. No additional uses are granted (such as derivative works or other editions). For any other uses, please submit a new request.

If credit is given to another source for the material you requested, permission must be obtained from that source.
### G.3 Figure 2.3 Permission to Reuse

This Agreement between Matthew D Gibson ("You") and Oxford University Press ("Oxford University Press") consists of your license details and the terms and conditions provided by Oxford University Press and Copyright Clearance Center.

<table>
<thead>
<tr>
<th>License Number</th>
<th>3902091246142</th>
</tr>
</thead>
<tbody>
<tr>
<td>License date</td>
<td>Jul 04, 2016</td>
</tr>
<tr>
<td>Licensed content publisher</td>
<td>Oxford University Press</td>
</tr>
<tr>
<td>Licensed content publication</td>
<td>Nucleic Acids Research</td>
</tr>
<tr>
<td>Licensed content title</td>
<td>Regulation of the nucleosome unwrapping rate controls DNA accessibility:</td>
</tr>
<tr>
<td>Licensed content author</td>
<td>Justin A. North, John C. Shimko, Sarah Javaid, Alex M. Mooney, Matthew A. Shoffner, Sean D. Rose, Ralf Bundschuh, Richard Fishel, Jennifer J. Ottesen, Michael G. Poirier</td>
</tr>
<tr>
<td>Licensed content date</td>
<td>11/01/2012</td>
</tr>
<tr>
<td>Type of Use</td>
<td>Thesis/Dissertation</td>
</tr>
<tr>
<td>Institution name</td>
<td>n/a</td>
</tr>
<tr>
<td>Title of your work</td>
<td>Reading the Epigenetic State of Chromatin Alters its Accessibility</td>
</tr>
<tr>
<td>Publisher of your work</td>
<td>n/a</td>
</tr>
<tr>
<td>Expected publication date</td>
<td>Aug 2016</td>
</tr>
<tr>
<td>Permissions cost</td>
<td>0.00 USD</td>
</tr>
<tr>
<td>Value added tax</td>
<td>0.00 USD</td>
</tr>
<tr>
<td>Total</td>
<td>0.00 USD</td>
</tr>
</tbody>
</table>

**Terms and Conditions**

1. Use of the material is restricted to the type of use specified in your order details.
2. This permission covers the use of the material in the English language in the following territory: world. If you have requested additional permission to translate this material, the
terms and conditions of this reuse will be set out in clause 12.
3. This permission is limited to the particular use authorized in (1) above and does not allow you to sanction its use elsewhere in any other format other than specified above, nor does it apply to quotations, images, artistic works etc that have been reproduced from other sources which may be part of the material to be used.
4. No alteration, omission or addition is made to the material without our written consent. Permission must be re-cleared with Oxford University Press if/when you decide to reprint.
5. The following credit line appears wherever the material is used: author, title, journal, year, volume, issue number, pagination, by permission of Oxford University Press or the sponsoring society if the journal is a society journal. Where a journal is being published on behalf of a learned society, the details of that society must be included in the credit line.
6. For the reproduction of a full article from an Oxford University Press journal for whatever purpose, the corresponding author of the material concerned should be informed of the proposed use. Contact details for the corresponding authors of all Oxford University Press journal contact can be found alongside either the abstract or full text of the article concerned, accessible from www.oxfordjournals.org Should there be a problem clearing these rights, please contact journals.permissions@oup.com
7. If the credit line or acknowledgement in our publication indicates that any of the figures, images or photos was reproduced, drawn or modified from an earlier source it will be necessary for you to clear this permission with the original publisher as well. If this permission has not been obtained, please note that this material cannot be included in your publication/photocopies.
8. While you may exercise the rights licensed immediately upon issuance of the license at the end of the licensing process for the transaction, provided that you have disclosed complete and accurate details of your proposed use, no license is finally effective unless and until full payment is received from you (either by Oxford University Press or by Copyright Clearance Center (CCC)) as provided in CCC’s Billing and Payment terms and conditions. If full payment is not received on a timely basis, then any license preliminarily granted shall be deemed automatically revoked and shall be void as if never granted. Further, in the event that you breach any of these terms and conditions or any of CCC’s Billing and Payment terms and conditions, the license is automatically revoked and shall be void as if never granted. Use of materials as described in a revoked license, as well as any use of the materials beyond the scope of an unrevoked license, may constitute copyright infringement and Oxford University Press reserves the right to take any and all action to protect its copyright in the materials.
9. This license is personal to you and may not be sublicensed, assigned or transferred by you to any other person without Oxford University Press’s written permission.
10. Oxford University Press reserves all rights not specifically granted in the combination of (i) the license details provided by you and accepted in the course of this licensing transaction, (ii) these terms and conditions and (iii) CCC’s Billing and Payment terms and conditions.
11. You hereby indemnify and agree to hold harmless Oxford University Press and CCC, and their respective officers, directors, employs and agents, from and against any and all claims arising out of your use of the licensed material other than as specifically authorized pursuant to this license.
12. Other Terms and Conditions:
v1.4
Questions? customercare@copyright.com or +1-855-239-3415 (toll free in the US) or +1-978-646-2777.