DIVERSE REGULATION OF ISWI FAMILY ATP-DEPENDENT CHROMATIN REMODELING ENZYMES BY NUCLEOSOME MODIFICATIONS

Geoffrey Paul Dann

A DISSERTATION PRESENTED TO THE FACULTY OF PRINCETON UNIVERSITY IN CANDIDACY FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

RECOMMENDED FOR ACCEPTANCE BY THE DEPARTMENT OF MOLECULAR BIOLOGY

Adviser: Tom W. Muir

September, 2017
Abstract

ATP-dependent chromatin remodelers regulate access to genetic information by controlling nucleosome positions in vivo. However, the mechanism by which remodelers discriminate between different nucleosome substrates is poorly understood. Considering that many chromatin remodeling proteins possess conserved protein domains that interact with nucleosomal features, this study employed a quantitative high-throughput approach, based on use of a DNA-barcoded mononucleosome library, to profile the biochemical activity of human ISWI remodelers in response to a diverse set of nucleosome modifications. The results of this work show that accessory (non-ATPase) subunits of ISWI remodelers can distinguish between differentially modified nucleosomes, directing remodeling activity toward specific nucleosome substrates based on their modification state. Unexpectedly, this investigation led to the discovery that the nucleosome acidic patch is necessary for maximum activity of all ISWI remodelers evaluated, a dependence that also extends to CHD and SWI/SNF family remodelers, suggesting the acidic patch may be generally required for chromatin remodeling. Critically, remodeling activity can be regulated by modifications neighboring the acidic patch, signifying it may act as a tunable interaction hotspot for ATP-dependent chromatin remodelers and, by extension, many other chromatin effectors that engage this region of the nucleosome surface.
Acknowledgements

First and foremost, I would like to thank my adviser, Dr. Tom Muir for providing an outstanding training environment in addition to excellent mentorship throughout my time at Princeton, as well as the other members of my thesis advisory committee, Dr. Fred Hughson and Dr. Ileana Cristea. I would also like to thank current and former Muir lab members. Specifically, I want to thank Dr. Manuel Müller for, in every sense, teaching me how to do science, as well as for his friendship in and outside of the lab. I similarly need to thank Dr. Glen Liszczak for these same reasons, and for his willingness to share his expertise and insights with me during the latter parts of my graduate studies. Drs. Neel Shah, Rob Oslund, Uyen Nguyen, and Lenka Bittova deserve thanks for their initial training and enthusiasm, which played a huge role in me ending up where I am. In addition, I am forever grateful to Dick Smith and Elaine Chapman for providing me the chance to start my journey into scientific research and the continued opportunity to cultivate it.

As perhaps many more memories were made outside of the lab as opposed to inside of it, I would like to thank Joel Tamayo, Dr. Rommel Mathias, Dr. Kenric Hoegler, and Dr. Evan Abbaszadeh for their friendship. Joel, thank you for the Butler days, random and intended nights on Nassau, Reunions, rescuing me from Deer Path and Rollingmead, hanging out with Agnes, granting me V. I. P. status in the Snakehole Lounge, helping maintain our sponsorship by Bombay Sapphire gin, hitting up the ‘tan, concerts in NYC, the trip to Miami, and stealing food. Rommel, thank you for letting me be probably the worst rotation student you ever had and still hanging out with me afterward. Aside from being a great friend, you were and continue to be a great scientific role model. I’ll visit you soon, you mad ****.

It is necessary that I dedicate this thesis to my parents, Paul and Mary Dann, for, without their unconditional support and encouragement, I would not have managed to come close to where I am today. Thank you for providing me with every opportunity to explore my interests such that I can confidently say I am pursuing something that I love. I would also like to thank my brother, Greg, and my sister, Claire, for being the best siblings that anyone could ever ask for. Lastly, I would like to thank Chelsea Martin for balance and for her immediate and continued interest in me and what I do. You can have the ‘h’.
# TABLE OF CONTENTS

Abstract........................................................................................................................................... iv

Acknowledgements.................................................................................................................................. vi

List of Figures and Tables..................................................................................................................... xii

Chapter 1: Introduction.......................................................................................................................... 1

Section 1.1: Genetic Information Storage in Eukaryotes................................................................. 2

Section 1.2: Chromatin Regulation..................................................................................................... 6

Section 1.2.1: Introduction.................................................................................................................. 6

Section 1.2.2: DNA Methylation......................................................................................................... 8

Section 1.2.3: The Nucleosome, Histones, and Histone Variants..................................................... 11

Section 1.2.4: Histone Post-Translational Modifications................................................................. 16

Section 1.3: ATP-dependent Chromatin Remodeling Enzymes....................................................... 23

Section 1.3.1: Introduction.................................................................................................................. 23

Section 1.3.2: SWI/SNF Family........................................................................................................... 29

Section 1.3.3: CHD Family................................................................................................................ 30

Section 1.3.4: INO80 Family.............................................................................................................. 31

Section 1.3.5: ISWI Family................................................................................................................ 32

Section 1.4: High-throughput Chromatin Biochemistry—Catching up with Epigenomics........ 34

Section 1.4.1: Regulation of ATP-dependent Chromatin Remodeling Enzymes by Nucleosome
Modifications........................................................................................................................................ 36

Chapter 2: High-throughput Chromatin Remodeling Assay Development and Nucleosome
Library Construction.......................................................................................................................... 40

Section 2.1: Introduction...................................................................................................................... 41

Section 2.2: Purification and Characterization of ISWI Family Chromatin Remodelers.............. 44

Section 2.3: Design and Development of a High-throughput Nucleosome Remodeling
Assay...................................................................................................................................................... 46

Section 2.4: DNA-barcoded Mononucleosome Library Construction............................................. 49
<table>
<thead>
<tr>
<th>Chapter 2:</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Page</td>
<td>Section</td>
</tr>
<tr>
<td>49</td>
<td>2.4.1: Preparation of Modified Histones Using Protein Chemistry</td>
</tr>
<tr>
<td>64</td>
<td>2.4.2: Preparation of Barcoded 601 DNA</td>
</tr>
<tr>
<td>67</td>
<td>2.4.3: Octamer and Nucleosome Preparation and Library Characterization</td>
</tr>
<tr>
<td>73</td>
<td>2.5: Summary</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Chapter 3:</th>
<th>Diverse Regulation of ISWI Remodelers by Nucleosome Modifications</th>
</tr>
</thead>
<tbody>
<tr>
<td>Page</td>
<td>Section</td>
</tr>
<tr>
<td>74</td>
<td>3.1: Data Summary</td>
</tr>
<tr>
<td>80</td>
<td>3.2: Accessory Subunits Direct ISWI Family Remodeling Activity Toward Unique Substrates</td>
</tr>
<tr>
<td>86</td>
<td>3.3: Histone Binding Properties Do Not Necessarily Correlate with Remodeling Activity</td>
</tr>
<tr>
<td>89</td>
<td>3.4: Histone-DNA Contacts Modulate Remodeling Activity</td>
</tr>
<tr>
<td>91</td>
<td>3.5: Summary</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Chapter 4:</th>
<th>The Acidic Patch as a Dynamically Regulatable Nucleosome Interaction Hotspot</th>
</tr>
</thead>
<tbody>
<tr>
<td>Page</td>
<td>Section</td>
</tr>
<tr>
<td>92</td>
<td>4.1: The Nucleosome Acidic Patch is Critical for Nucleosome Remodeling and is Tunable by Histone PTMs</td>
</tr>
<tr>
<td>100</td>
<td>4.2: Summary</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Chapter 5:</th>
<th>A DNA-barcoded Mononucleosome Library as a General Platform for High-Throughput Chromatin Biochemistry—Ongoing Projects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Page</td>
<td>Section</td>
</tr>
<tr>
<td>101</td>
<td>5.1: Introduction</td>
</tr>
<tr>
<td>105</td>
<td>5.2: b-JIM, a Novel PWWP-domain Containing Protein Functionally Regulates the Chromosomal Kinase JIL-1 in Drosophila melanogaster—with Catherine Regnard, Peter Becker lab, LMU Munich</td>
</tr>
<tr>
<td>113</td>
<td>5.3: Functional Crosstalk Between Histone H2B Ubiquitylation and H2A Modifications and Variants—with Felix Wojcik, Tom Muir lab, Princeton University</td>
</tr>
<tr>
<td>117</td>
<td>5.4: Histone Acetylation Regulates ADP-ribosylation in DNA Damage—with Glen Liszczak and Katharine Diehl, Tom Muir lab, Princeton University</td>
</tr>
<tr>
<td>122</td>
<td>5.5: Summary</td>
</tr>
</tbody>
</table>
Chapter 6: Perspectives........................................................................................................124
Section 6.1: Introduction....................................................................................................125
Section 6.2: ISWI Remodeling Dataset—Other Considerations.................................125
Section 6.3: Considerations Regarding Streamlined Library Preparation................130
Section 6.4: DNA-barcode Mononucleosome Libraries—Untapped Applications........134
  Section 6.4.1: Profiling Histone Eraser Activity..........................................................134
  Section 6.4.2: Profiling H2A.Z Incorporation by the hSRCAP Complex......................135
Section 6.5: Expanding DNA-barcode Chromatin Library Substrate Complexity........136
Section 6.6: Summary.....................................................................................................138

Chapter 7: Methods...........................................................................................................140
Section 7.1: Analytical Methods......................................................................................141
Section 7.2: Production and Purification of RCC1 and Sir3 Constructs........................141
Section 7.3: Production and Purification of Recombinant Histones............................142
Section 7.4: Production of Modified Histones by Protein Semisynthesis....................143
  Section 7.4.1: Preparation of H2AK5ac, H2AK9ac, H2AK13ac, H2AK15ac, and H2AKpolyAc Histones........................................................................................................143
  Section 7.4.2: Preparation of the H2AK118ac Histone..................................................145
  Section 7.4.3: Preparation of H2BK5ac, H2BK11ac, H2BK12ac, H2BK15ac, H2BK16ac, and H2BK20ac Histones........................................................................................................145
  Section 7.4.4: Preparation of H2BK108ac, H2BK116ac, H2BK120ac, H2BK125ac Histones........................................................................................................146
  Section 7.4.5: Preparation of Phosphorylated H3 Histones.........................................146
  Section 7.4.6: Preparation of H3K4me1 and H3K4me2 Histones.........................147
  Section 7.4.7: Preparation of H3R2 Methylated Histones........................................147
  Section 7.4.8: Preparation of the H3K36me3 Histone................................................148
  Section 7.4.9: Preparation of H4R3 Methylated Histones.........................................148
  Section 7.4.10: Preparation of H4K77ac and H4K79ac Histones..............................150
  Section 7.4.11: Preparation of H2BS112GlcNAc......................................................150
Section 7.5: Production and Purification of Modified Histones by Amber Suppression

Section 7.5.1: Preparation of H3K4ac, H3K36ac, H3K37ac, H3K56ac, H3K64ac, H3K79ac, H3K115ac, H3K122ac, and H4K91ac Histones

Section 7.6: Synthesis of LANA Peptides

Section 7.7: Production of Chromatin Remodelers in Sf9 Cells

Section 7.8: DNA Preparation

Section 7.8.1: Barcoded 601 (BC-601) DNA Preparation

Section 7.8.2: Biotinylated MMTV ‘Buffer’ DNA

Section 7.8.3: 192 b.p. DNA Fragment Used in Follow-up Studies on Individual Nucleosomes

Section 7.9: Octamer Formation for the Nucleosome Library

Section 7.10: Nucleosome Reconstitution for the Nucleosome Library

Section 7.11: Analysis of Nucleosome Library Integrity by Antibody Pull-down

Section 7.12: Nucleosome Thermal Mobility Shift Experiments

Section 7.13: Experiments Employing the Nucleosome Library

Section 7.13.1: Remodeling Assays Using the Nucleosome Library

Section 7.13.2: Pull-down Experiments with the RCC1-GST Construct using the Nucleosome Library

Section 7.13.3: Pull-down Experiments with the Sir3-BAH-FLAG-6xHis Construct Using the Nucleosome Library

Section 7.13.4: Experimental Procedures Corresponding to Section 5.2 ‘β-JIM, a Novel PWWP-domain Containing Protein Functionally Regulates the Chromosomal Kinase JIL-1 in Drosophila melanogaster’

Section 7.13.5: Experimental Procedures Corresponding to Section 5.3 ‘Functional Crosstalk Between Histone H2B Ubiquitylation and H2A Modifications and Variants’—Performed by Felix Wojcik

Section 7.13.6: Experimental Procedures Corresponding to 5.4 ‘Histone Acetylation Regulates ADP-ribosylation in DNA Damage’—Performed by Glen Liszczak and Katharine Diehl
Section 7.13.7: Addition of Multiplex Barcodes and Illumina Forward and Reverse Adapter Sequences via PCR........................................................................................................................................169

Section 7.13.8: Illumina Sequencing.......................................................................................................................................................171

Section 7.14: Nucleosome Remodeling Experiments in the Presence of LANA Peptides........171

Section 7.15: Nucleosome Remodeling Rate Analysis for Validation Experiments on Individual Nucleosomes Using a Restriction Enzyme Accessibility Assay.................................................................172

Section 7.16: Nucleosome Remodeling Assays and Rate Analysis for Validation of Acidic Patch Mutant Nucleosomes (H2A E61A, D90A, E92) Using an Electrophoretic Mobility Shift Nucleosome Repositioning Assay.........................................................................................................................172

Section 7.17: Processing of Sequencing Data for Chromatin Remodeling Experiments Using the Nucleosome Library..................................................................................................................................................174

Section 7.18: Processing of Sequencing Data for Binding Experiments Using the Nucleosome Library.........................................................................................................................................................175

Section 7.19: Nucleosome Remodeling Rate Analysis for Library Experiments.......................175

References...............................................................................................................................................................................................................177
LIST OF FIGURES AND TABLES

Chapter 1

Figures

Figure 1.1: Chromatin Organization in Eukaryotic Organisms........................................6
Figure 1.2: Chromatin Regulation..................................................................................8
Figure 1.3: 5-methylcytosine......................................................................................9
Figure 1.4: Nucleosome Structure..............................................................................12
Figure 1.5: Histone Post-translational Modifications................................................18
Figure 1.6: Functions of ATP-dependent Chromatin Remodelers............................25
Figure 1.7: Model for DNA Translocation in Nucleosome Remodeling by ISWI Enzymes...29
Figure 1.8: ISWI Chromatin Remodelers Profiled....................................................38
Figure 1.9: Regulation of ISWI Remodeling by Nucleosome Modifications..............39

Tables

Table 1: ATP-dependent Chromatin Remodeling Families in Humans.......................27

Chapter 2

Figures

Figure 2.1: Restriction Enzyme Accessibility and Gel Mobility Shift Assays for Chromatin Remodeler Analysis........................................................................................................42
Figure 2.2: Nuclear Extraction of Chromatin Remodelers.........................................45
Figure 2.3: Characterization of Recombinant ISWI Chromatin Remodelers..............46
Figure 2.4: A High-throughput Nucleosome Remodeling Assay for ISWI Family Chromatin Remodelers........................................................................................................................................................................49
Figure 2.5: Amber Suppression and Native Chemical Ligation Technologies..............................51
Figure 2.6: Two-piece Ligation Schematic for the Preparation of an H3K4me3 Modified Histone..................................................................................................................................................................................53
Figure 2.7: Characterization of Histones by RP-HPLC/ESI-MS......................................................62
Figure 2.8: A Diverse Library of Modified Nucleosomes.................................................................64
Figure 2.9: Characterization of Barcoded 601 (BC-601) DNA......................................................66
Figure 2.10: Barcoded 601 (BC-601) Nucleotide Sequence..........................................................67
Figure 2.11: Nucleosome Assembly Using a Crude Octamer Preparation..................................69
Figure 2.12: Analysis of the Quality and Integrity of the Nucleosome Library.........................72

Tables

Table 2.1: Histones Used to Construct the Nucleosome Library.................................................55
Table 2.2: Histone Composition of Nucleosome Library Members...........................................70

Chapter 3

Figures

Figure 3.1: Diverse Regulation of ISWI Remodelers by Nucleosome Modifications...............75
Figure 3.2: Nucleosome Remodeling Activity is Negligible in the Absence of ATP...............77
Figure 3.3: Assessment of Modeling Nucleosome Library Remodeling as a 1- vs. 2-phase Exponential Decay......................................................................................................................................................79
Figure 3.4: PCA Analysis of Library Remodeling Data.............................................................81
Figure 3.5: Nucleosome Thermal Mobility Shift Experiments...................................................82
Figure 3.6: Specialization of ISWI Remodelers Toward Diverse Nucleosome Modifications....83
Figure 3.7: Subsets of the Nucleosome Library Drive Differences in Remodeler Activity........84
Figure 3.8: Remodeling of Nucleosomes Containing Modifications Preferred by Histone Recognition Domains.................................................................88
Figure 3.9: Alteration of Histone-DNA Contacts Affects Remodeling Activity..................90

Tables

Table 3.1: Nucleosome Library Members Contributing to Variation in Remodeler Behavior.....85

Chapter 4

Figures

Figure 4.1: The Nucleosome Acidic Patch is Crucial for Remodeling and Regulatable by Histone PTMs..............................................................................................94
Figure 4.2: A Gel Mobility Shift Assay and Inhibition by the LANA Peptide Confirm the Importance of the Acidic Patch.................................................................95
Figure 4.3: Characterization of Recombinant BRG1 and CHD4 Chromatin Remodelers........96
Figure 4.4: Diverse Biochemical Regulation of Acidic Patch Interactors by Nucleosome Modifications.................................................................................................99

Chapter 5

Figures

Figure 5.1: Demonstrated Applications of DNA-Barcoded Nucleosome Libraries to Study Chromatin Effector Proteins........................................................................104
Figure 5.2: JIL-1 and b-JIM Physically Interact to Form a Protein Complex....................107
Figure 5.3: b-JIM and JIL-1 Co-localize In Vivo

Figure 5.4: Pull-down Experiments Using the Nucleosome Library

Figure 5.5: A High-throughput In Vitro Kinase Assay Using the Nucleosome Library

Figure 5.6: H3K36me3 are Preferred Substrates at H3S10 for the b-JIM-JIL1 Kinase Complex

Figure 5.7: A High-throughput In Vitro Ubiquitylation Assay Using the Nucleosome Library

Figure 5.8: Analysis of UBE2A:RNF20/40 Enzymatic Activity Against the Nucleosome Library

Figure 5.9: ChIP-seq Analysis Proximal to Transcription Start Sites Supports Negative Biochemical Cross-talk between H2A.Z and H2BK120ub

Figure 5.10: A High-throughput In Vitro ADP-ribosylation Assay Using the Nucleosome Library

Figure 5.11: Analysis of PARP1 Binding and Enzymatic Activity Against the Nucleosome Library

Figure 5.12: S10 and S28 on Histone H3 are the Major Sites of ADP-ribosylation by PARP1

Chapter 6

Figures

Figure 6.1: Unmodified H3K4 is Important for WICH Remodeling

Figure 6.2: A High-throughput In Vitro Deacetylase Assay Using the Nucleosome Library

Figure 6.3: A High-throughput In Vitro H2A.Z Incorporation Assay Using the Nucleosome Library
Chapter 7

Figures

Figure 7.1: Characterization of LANA Peptides..............................................................152
CHAPTER 1: INTRODUCTION
1.1 Genetic Information Storage in Eukaryotes

Packaging and storage of genetic information by eukaryotes represents an incredible evolutionary accomplishment due to the relatively large amount of material that must be accounted for. In humans, this requires fitting an approximately 6 billion base pair diploid genome into a cell nucleus with dimensions on the order of only ~10 μm. Considering the 3.4 Å distance between base pairs in B-form DNA, which most commonly occurs under conditions of physiological pH and ionic strength (Van Holde, 1989), the total length of the human genome is just over 2 m (6 ft). To fit into the eukaryotic nucleus would require an end-to-end compaction ratio of ~200,000, an achievement made further impressive bearing in mind the fact that DNA naturally repels itself due to the inherent negative charge carried by its phosphate-containing backbone.

Early work by Friedrich Miescher marks what appeared to be the first isolation of nuclei (Miescher, 1871). When these nuclei, originating from white cells from pus on bandages of hospital patients, were extracted with alkaline solutions followed by acid precipitation, the resulting material was observed to have a relatively high phosphorus content, but not close to that of pure DNA. Looking back with current knowledge, this would indicate that what was isolated likely contained DNA, but that it was not the only constituent of the isolate. Subsequent experiments using salmon sperm revealed that this material contained a highly basic, acid soluble component that, when extracted, left behind what today would be considered characteristic of a true DNA preparation (Miescher, 1874). This was thus the first evidence representing what we know today as the physiological form of the eukaryotic genome—chromatin. At the most fundamental level, chromatin roughly consists of a nucleic acid component, DNA, that is complemented by a ubiquitous and highly basic protein component, the histones. Immediately, it is apparent that a basic, and thus highly positively charged substance, would aid in solving the problem of DNA charge repulsion previously mentioned. How histones might participate in resolving the issue of compacting genetic material inside of the nucleus is less clear. Structural studies have since revealed that the four canonical core histones, H2A, H2B, H3, and H4 are
arranged into an octameric protein complex containing two copies of each histone (Luger et al., 1997). 147 base pairs of DNA are then wrapped 1.7 times around the octameric histone core in a left-handed superhelix to form what is known as the nucleosome core particle. Through nucleosome formation, the 147 base pairs of DNA, at a length of ~500 Å, are effectively shortened to the ~100 Å radius of the nucleosome disc. Though this only achieves an approximate 5-fold compaction, much less than that required to compact all genetic material in the human nucleus, it becomes evident how the histones might form the basis of a scaffold to do so.

In the 1930s, experiments that involved staining of DNA and visualization of nuclei by light microscopy revealed that chromatin existed in both loosely packed (light staining) and tightly packed (dense staining) forms (Alberts, 2015). These two forms of chromatin are generally referred to as euchromatin and heterochromatin, respectively. Though details of chromatin composition were relatively scarce at the time, this provided early evidence for the existence of different types of chromatin structures. With the nucleosome serving as the fundamental repeating unit of the chromatin fiber, loosely packed, euchromatic regions of the genome, as measured by accessibility of DNA to nuclease cleavage or transposition (Bell et al., 2011; Buenrostro et al., 2013), are generally assumed to adopt a ‘beads on a string’ type structure that was observed early on under the electron microscope (Olins and Olins, 1974; Oudet et al., 1975; Woodcock et al., 1976). In this type of environment, nucleosomes are separated by “linker” DNA. This helps to define the nucleosome repeat length as the number of base pairs beginning from the start of one nucleosome, around the histone octamer, and through the linker DNA, and to the start of the subsequent nucleosome. Nucleosome arrays that form in human cells generally display a nucleosome repeat length of ~200 base pairs, but this value has been observed to vary depending on cell type as well as genomic location within the same cell (Gaffney et al., 2012; Valouev et al., 2011). A fifth histone, linker histone H1, can occupy the linker DNA between nucleosomes. H1 binds in the vicinity of the nucleosome dyad axis and requires 10-20 base pairs of extranucleosomal DNA to form the ‘chromatosome’ (Harshman et al., 2013), an event that facilitates higher levels of chromatin compaction. Linear arrays of nucleosomes can compact
further to form higher order chromatin structures, a process that depends on the presence of H1 and ionic strength (particularly divalent cations such as Mg$^{2+}$). This higher level of compaction is often referred to as the '30 nm fiber', due to its rod-like shape and diameter of 30 nm as originally observed by electron microscopy (Finch and Klug, 1976). Though how pervasive this structure is in vivo is currently a highly debated topic, in vitro analyses, typically employing X-ray crystallography or cryo-electron microscopy, have been able to paint a high resolution picture of particular ways nucleosomes might compact into higher-order chromatin fibers (McGinty and Tan, 2015). The most recent structural evidence reveals a left-handed two-start helix of repeating tetranucleosomal subunits (Song et al., 2014). Indeed, evidence of a two-start helical structure in compact heterochromatic regions of the genome has also been observed in interphase nuclei of human skin fibroblasts (Risca et al., 2017).

Beyond internucleosomal interactions that result in locally folded chromatin fibers, chromosome conformation studies reveal that the human genome (as well as the genomes of many other metazoans) is, at the three-dimensional level, organized into topologically associated domains (TADs) (Dixon et al., 2012; Nora et al., 2012; Sexton et al., 2012). TADs are independently folded kilobase-to-megabase chromosomal regions that are spatially and functionally isolated from one another (Sexton and Cavalli, 2015). Their assumed regulatory roles are realized through their marked correlation with many aspects of chromatin function such as chemical modification of histones, coordinated gene expression, lamina association, and DNA-replication timing (Sexton and Cavalli, 2015). How and if TADs hierarchically co-associate remains unclear. Current experimental evidence suggests that populations of cells typically exhibit a variety of higher-order chromosomal arrangements (Nagano et al., 2013). Interestingly, TAD organization is broadly conserved in interphase nuclei (Naumova et al., 2013), but such domains “disappear” during mitosis and must be reset by an unknown mechanism upon re-entry into G$_1$.

During mitosis, eukaryotic chromatin adopts its most condensed form within metaphase, just prior to division of replicated genetic material and cytokinesis. Modern research on the ultrastructure of mitotic chromosomes began in the 1970s upon the development of mild methods
for their isolation (Wray and Stubblefield, 1970). At high ionic strength, chromosomes prepared in this way have a length on the order of tens of µm (Adolph et al., 1986). However, the length of highly compact chromosomes viewed in a typical metaphase spread (e.g. for karyotyping) can be a maximum of approximately 10 µM for the longest chromosomes (Estandarte et al., 2016). Other studies similar to those performed by Adolph et al. saw 50 nm fibers that extended radially from a central chromosomal core (Marsden and Laemmli, 1979). These structures had the appearance of coiled 30 nm fibers, and, indeed, at lower ionic strengths unfolded to reveal loops of fibers with a 30 nm diameter. When divalent metal ions were removed by chelation, extended loops of ‘10 nm fibers’, thought to represent linear arrays of nucleosomes were found. Such experiments are also a good example demonstrating that the conditions used to isolate chromosomes have profound effects on what they look like under the microscope. Though the abundance and precise nature of the organization of each of these structures in vivo is still argued, such early experiments revealed that chromatin was hierarchically folded. New techniques developed over the years to study chromatin structure in vivo continue to provide a clearer picture of the eukaryotic genome in its native environment (see Figure 1.1 for a depiction of the different levels of chromatin organization in eukaryotes). Nevertheless, the dense packaging of so much material paradoxically raises a second problem—how is the genetic material physically accessed and interpreted?
1.2 Chromatin Regulation

1.2.1 Introduction

All cells in the human body contain all the required information to produce every protein encoded within the genome, yet they do not. Thus, there must be a mechanism to determine which information is available for interpretation and utilized and which remains obscured. In other words, multi-cellular eukaryotes must repurpose an invariable genome to serve as a genetic blueprint to establish and maintain a myriad of functionally diverse cell types. This must be accomplished by somehow introducing heterogeneity along the DNA double-helix, without
altering the DNA sequence, to make sure that only specific subset of the underlying instructions is available to each cell. Furthermore, differentiated cells with established transcriptional programs must remain able to respond to specific environmental stimuli. How are genes in different cell types selected for expression from an identical genetic library? This field of study is known as 'epigenetics'.

A consequence of organizing eukaryotic genomes into chromatin is the creation of an inherent barrier to vital cellular processes by sequestration of DNA binding sites. Correspondingly, chromosomal transactions such as gene transcription require alterations in chromatin structure so that necessary trans-acting factors can interact with DNA to properly execute their roles (Figure 1.2). Thus, nucleosomes are natural obstacles that compete with DNA binding proteins for genomic access. It has been discovered that chromatin structure, and consequently availability of functional local nucleic acid elements, is controlled by introducing 'signals' into the chromatin fiber that communicate to the underlying DNA sequence, for example, whether a gene is to be transcribed. Such signals include the presence or absence of linker histones, histone sequence variants, an assortment of covalent histone and DNA modifications, and non-histone proteins such as transcription factors. The diversity of this regulatory tool kit allows for virtually unlimited variability in chromatin composition and configuration, permitting the precise multilevel control of gene expression observed in higher organisms.
**Figure 1.2 Chromatin Regulation.** (Left) DNA is packaged inside the eukaryotic nucleus at different densities as indicated by dark and light staining. Cartoons depicting chromatin organization in compact (heterochromatin) and extended (euchromatin) chromatin loci are indicated. (Right) DNA provides ‘data’ encoding instructions for all cellular processes. Chromatin stores and regulates efficient access and retrieval of the ‘data’ contained within DNA.

### 1.2.2 DNA Methylation

DNA methylation is one of the most well-studied epigenetic modifications in vertebrates. This covalent modification occurs at on the cytosine base in the context of cytosine-guanine (5’CpG3’; p represents the phosphate linkage) dinucleotides (Li and Zhang, 2014). The addition of a methyl group at the 5 position of the cytosine pyrimidine ring by DNA methyltransferases (DNMTs) results in formation of 5-methylcytosine (5mC) (Figure 1.3). This methyl mark resides in the major groove of DNA where many regulatory proteins engage the double helix. Thus, CpG methylation may exert its effects through enhancing or diminishing interactions between these proteins (e.g. transcription factors and epigenetic regulators) and genomic loci (Li and Zhang, 2014; Tate and Bird, 1993). In this way, it participates in processes
such as embryonic development and dosage compensation, primarily through its widespread role in transcriptional regulation.

Figure 1.3 5-methylcytosine. The structure of 5-methylcytosine (left) paired with guanine (right) is shown. The methyl group on the number 5 position of the cytosine pyrimidine ring is shown in blue and does not interfere with GC base pairing. Hydrogen atoms and corresponding hydrogen bonds are shown in grey.

Primarily centered around next-generation sequencing technology, methods to detect 5mC genome-wide (Flusberg et al., 2010; Frommer et al., 1992; Gu et al., 2011; Irizarry et al., 2009; Tost and Gut, 2007; Weber et al., 2005) led to the discovery that 70-80% of all CpG sites in mammalian somatic tissues are methylated (Li and Zhang, 2014). Bisulfite sequencing, to date, has been most reliable method for large-scale 5mC identification (Cokus et al., 2008; Gu et al., 2011; Lister et al., 2008). It employs bisulfite modification of cytosine residues, leading to deamination and conversion to uracil (Frommer et al., 1992). A subsequent PCR step further converts them to thymidine. 5mC nucleotides are protected from this process and this property serves as the basis for their detection by DNA sequencing when compared to an untreated reference sample. Single Molecule, Real Time (SMRT®) sequencing is an alternative promising technology developed by Pacific Biosciences® has the potential to revolutionize detection of nucleotide base modifications not limited to 5mC (Flusberg et al., 2010). This method relies on monitoring real-time DNA polymerase elongation on a genomic template. Uniquely labeled fluorescent nucleotides allow identification of each nucleotide being incorporated. Detection of
nucleotide modifications is based on measuring distinct kinetic incorporation signatures. SMRT sequencing also has the advantage of being able to generate extremely long read lengths (>20 kb), which aids significantly in the process of genome assembly in post-sequencing data workup.

The self-complementarity of the CpG dinucleotide made it apparent very early on that DNA methylation might be inherited through cell division and serve as a form of cell memory (Holliday and Pugh, 1975; Riggs, 1975). Following semiconservative replication, DNA would retain methylation on the parent strand with the new daughter strand lacking modification. An enzyme with the ability recognize hemimethylated CpG dinucleotides could then methylate the corresponding cytosine base on the daughter strand. Indeed, this is the case and it is carried out by the ‘maintenance methyltransferase’ Dnmt1 (Bestor and Ingram, 1983). New DNA methylation patterns are created by de novo DNA methyltransferases. After fertilization, global DNA methylation takes place, followed by establishment of de novo DNA methylation patterns that aid in regulating gene expression patterns through formation of CpG islands (Bird et al., 1985) and the well-understood process of ‘genomic imprinting’ (Barlow and Bartolomei, 2014).

DNA demethylation occurs through both active and passive routes, and is initiated by the conversion of 5mC to 5-hydroxymethylcytosine (5hmC) by the Tet family of dioxygenases (Ito et al., 2010; Tahiliani et al., 2009), with iterative oxidation also leading to the formation of 5-formylcytosine and 5-carboxylcytosine (5fC and 5caC, respectively) (He et al., 2011; Ito et al., 2011). Passive DNA demethylation occurs through continual cell division in the absence of Dnmt1 maintenance methyltransferase activity and dilution of 5mC over multiple rounds of DNA replication (Li and Zhang, 2014). This can occur due to the inability of Dnmt1 recognize CpG dinucleotides bearing oxidized forms of 5mC (Valinluck and Sowers, 2007). Active DNA demethylation is an enzymatic process that occurs independent of cell division and relies on thymine-DNA glycosylase to remove 5hmC, 5fC, and 5caC as part of the process of base excision repair (He et al., 2011; Maiti and Drohat, 2011). Alternatively, it is possible that 5caC may be converted back to cytosine by a putative decarboxylase (Li and Zhang, 2014), which has yet to be established.
1.2.3 The Nucleosome, Histones, and Histone Variants

There are 5 classes of histone proteins found in the chromosomes of eukaryotic cells: H2A, H2B, H3, H4, and H1. Two copies of each of the core histones H2A, H2B, H3, and H4 comprise the histone octamer, while H1, the 'linker histone', occupies intervening linker DNA between nucleosomes. Core histones are characterized by being highly abundant, and their synthesis being 'replication-dependent'. Each time a cell divides it must produce roughly $10^8$ molecules of each of the core histones during S-phase of the cell cycle to package the newly replicated genome. The human genome contains 10-20 copies of each core histone gene that occur in two major clusters, a feature likely related to the need to pair their rapid and massive production with DNA duplication (Marzluff et al., 2002). These genes lack introns, and their corresponding mRNAs are not polyadenylated. Instead, the 3' end of histone mRNAs contains a specific stem-loop structure that is critical to coupling their synthesis with S-phase of the cell cycle (Marzluff and Duronio, 2002). Core histones are relatively small proteins, with H4 being the shortest (103 amino acids) and H3 being the longest (135 amino acids). They are also highly basic, all possessing calculated isoelectric points above 10. This property is due to the presence of a very large number of arginine and lysine residues and thus makes the histones very well suited for interaction with negatively charged DNA. One H3-H4 tetramer is viewed to serve as the basis for nucleosome formation as it is the last portion of the histone octamer to dissociate from DNA as nucleosomes are incubated with progressively higher (non-physiological) salt concentrations (Burton et al., 1978). Supporting this notion, H3-H4 tetramers are able protect DNA fragments nearly the size of the those protected by the entire nucleosome core particle from micrococcal nuclease digestion (Burton et al., 1978). Combination with two less stably bound H2A-H2B (Burton et al., 1978) dimers finalizes the histone octamer and the wrapping of 147 base pairs of DNA serves to complete the nucleosome core particle (Luger et al., 1997). Altogether, the nucleosome takes on the shape of a disc, where approximately 147 base pairs of DNA are wrapped around the histone octamer 1.7 times in a left-handed super helix (Figure 1.4). Once folded within the octamer, the core histones all adopt a similar alpha-helical ‘histone fold’
structure flanked by the remaining N- and C-termini. In total, 10 unstructured histone termini, commonly referred to as 'tails' protrude from the nucleosome—1 N-terminus from each of the core histones and 2 additional tails from the C-terminus of H2A.

![Figure 1.4 Nucleosome Structure](image)

**Figure 1.4 Nucleosome Structure.** Atomic-level resolution structure of the nucleosome (PDB: 1KX5). Core histones are color-coded with nucleosomal DNA depicted in silver.

Under conditions of physiological pH and ionic strength nucleosomes are quite stable, showing almost no dissociation (~10%) at 300 mM NaCl (Van Holde, 1989) after 50 hours! H2A-H2B dimers are observed to dissociate over an NaCl concentration range of 0.5-1.5 M, and only once the NaCl concentration approaches 2 M are interactions between the H3-H4 and DNA abolished (Burton et al., 1978). As evidenced by their disruption at high ionic strength and subsequently by high resolution structural studies (Burton et al., 1978; Luger et al., 1997), association between histone octamers and DNA are primarily electrostatic in nature. Dominating such interactions are a suite of arginine residues, inserted into the minor groove all 14 times it faces the histone octamer. The affinity of nucleosomal DNA is sequence specific only in the sense that nucleosome formation favors sequences that can be bent to contour the histone octamer (McGinty and Tan, 2015). Nucleosomal DNA sequences that have AT, TA, AA, and TT
base steps where the minor groove faces the histone octamer are favored (Satchwell et al., 1986). Natural and synthetic DNA oligomers such as the 5S RNA coding sequence, the human α-satellite repeat, and the Widom 601 sequence (Lowary and Widom, 1998), that favor nucleosome formation over random genomic DNA fragments (Thastrom et al., 2004) prominently display this feature. For example, the 601 sequence contains the TA base step at 5 of the 8 possible positions where the minor groove of the DNA faces the histone octamer. The reason for this is that DNA curvature causes notable stress at these ‘pressure points’ and TA nucleotide pairs at these positions are best able to handle such structural distortions (Davey et al., 2002; McGinty and Tan, 2015). Sequence ‘rules’ such as these and others have been used to predict nucleosome positions in vivo (Jiang and Pugh, 2009; Kaplan et al., 2009; Lowary and Widom, 1998; Segal et al., 2006). In summary, nucleosome stability is highly dependent on genetic sequence and must contribute variability to chromatin dynamics across the genome.

In addition to the 4 core, replication-dependent, histones, there are many replication-independent histone sequence variants. These histones can take the place one of the core, replication-dependent, histones in the nucleosome and confer unique chromatin regulatory properties due to their differing amino acid composition (Henikoff and Smith, 2015; Maze et al., 2014). In humans, variants of H2A, H2B, and H3 have all been described. Unlike the core histones, they are present in only 1-3 gene copies, and their mRNAs are all polyadenylated (Marzluff et al., 2002; Maze et al., 2014). Currently, there is not much data regarding the biological function of known core H2B variants (Maze et al., 2014), and interestingly, there exist no variants of core histone H4, indicating its sequence is under strong evolutionary selection (Marzluff et al., 2002).

The two most well characterized variants of core histone H3 are H3.3 and CENP-A (its counterparts generally referred to cenH3 across species). H3.3 is enriched in transcriptionally active chromatin, and, thus, is commonly viewed as a marker for these regions of the genome (Filipescu et al., 2013). Even though the amino sequence of H3.3 differs from canonical H3 by only 5 residues, it exclusively participates in replication-independent nucleosome assembly (Henikoff and Smith, 2015). Chromatin is highly dynamic at transcribed genes, with nucleosome
disassembly and assembly constantly taking place ahead and behind elongating RNA polymerase II, respectively (Teves et al., 2014). H3.3 is also found in telomeric and pericentric chromatin. These findings obviously challenge the notion that H3.3 is a specific marker of active transcription. However, H3.3 has been observed to be a modulator of unexpected transcription, which occurs at both telomeres and centromeres, an event that implies these regions are therefore sites of nucleosome turnover (Drane et al., 2010; Goldberg et al., 2010; Law et al., 2010). This process is carried out by the H3.3-specific Daxx histone chaperone complex and the ATP-dependent nucleosome remodeling protein ATRX (Drane et al., 2010; Goldberg et al., 2010; Law et al., 2010). However, assuming H3.3 serves as the general substrate for nucleosome replacement independent of DNA replication, the fact that it is enriched in active chromatin simply suggests that nucleosome turnover rates are highest in these regions.

CENP-A is a centromeric H3 variant and was discovered to localize specifically to centromeres, replacing H3 in nucleosomes (Palmer et al., 1991). Its role in establishing centromere identity is underscored by the fact that it is found in the genomes of all eukaryotes (Talbert and Henikoff, 2010), and that it serves as much better marker for centromere formation than DNA sequence. As such, fluorescent staining of CENP-A can identify genetic defects like ectopic formation of ‘neocentromeres’ that do not occur on canonical centromeric α-satellite repeat DNA (Amor et al., 2004a). Genetic experiments show essentiality of centromeric H3 variants in kinetochore formation and chromosome segregation (Amor et al., 2004b). Like H3.3, CENP-A has its own histone chaperone to carry out assembly of centromeric chromatin. In mammals, this process is carried out by the protein HJURP (Dunleavy et al., 2009; Foltz et al., 2009).

The histone class with largest number of variants is that of H2A. In addition to core histone H2A, there exist H2A.Z, H2A.X, H2A.Bbd, and macroH2A. Of these H2A variants, the function of H2A.X in DNA double-strand break repair is probably the most well-studied. H2A.X is characterized by a C-terminal amino acid sequence motif (SQEY in humans) that contains a serine, which is phosphorylated at regions of the genome that have experienced a double-strand break (Rogakou et al., 1998). This modified version of the protein is commonly referred to as
Even though detection $\gamma$H2A.X is a hallmark of double-strand breaks, cells in which the gene has been deleted or the target serine mutated can still undergo repair, albeit at a much lower efficiency. Such cells are therefore hypersensitive to genomic insults that result in DNA damage (Celeste et al., 2002; Redon et al., 2003). It is thought that $\gamma$H2A.X plays a role in the recruitment or retention of proteins required for restoring the break site (Morrison and Shen, 2005) or that it may stabilize the severed ends of chromosomes through recruitment of cohesion (Lowndes and Toh, 2005).

H2A.Z appears to play a more varied role than H2A.X in the nucleus. Incorporation of H2A.Z into nucleosomes introduces several structural differences. Localized changes lead to subtle alterations in the interaction the H2A.Z-H2B dimer as well as interactions between H2A.Z and the H3-H4 tetramer (Suto et al., 2000). As the overall nucleosome structure, including the path of the DNA gyres, show little difference when compared to nucleosomes containing canonical H2A, most distinctions are a result of changes in molecular surface features. The most prominent of which is the introduction of an extended nucleosome ‘acidic patch’ (Luger et al., 1997; Suto et al., 2000). This alteration has been shown to affect interactions between a variety of chromatin associated proteins and H2A.Z-containing nucleosomes (Goldman et al., 2010).

Aside from possession of unique biochemical properties, histone H2A.Z has been shown to play a diverse role in chromatin regulation. Genome-wide localization studies show that H2A.Z occupancy is highest at gene promoters and enhancers (Obri et al., 2014), its incorporation and removal carried about by the INO80 family of ATP-dependent chromatin remodeling enzymes (Clapier et al., 2017). H2A.Z function has been linked to both transcriptional activation and repression, RNA polymerase II elongation, heterochromatin antisilencing, cell-cycle control, DNA replication, DNA damage repair, chromosome segregation, and genome integrity (Henikoff and Smith, 2015).

The roles of other H2A variants, macroH2A and H2A.Bbd, are much less clear, although they at least appear to participate in the human dosage compensation mechanism. In addition to a canonical histone-fold domain and N- and C-terminal tails macroH2A contains a C-terminal macrodomain over 200 amino acids in length and is found to be enriched on facultative, inactive
regions of the X chromosome in females (Chadwick and Willard, 2004). It is therefore thought to function in gene repression. On the other hand, H2A.Bbd, with ‘Bbd’ standing for ‘Barr body deficient’, is undetectable on the inactive X chromosome, hence its name. It forms inherently less stable nucleosomes (Arimura et al., 2013; Bao et al., 2004). Consistent with a function distinct from that of macroH2A, H2A.Bbd is found at transcription start sites (Soboleva et al., 2011).

Histone variants represent the most fundamental mode of variation in chromatin structure—alteration of nucleosome composition by histone replacement. As previously described, such alterations have definable effects on nucleosome structure, affecting their intrinsic biochemistry as well as that of nucleosome interacting proteins. *In vitro* studies are essential to shed light on possible mechanisms by which histone variants carry out their ascribed biological roles. With involvement in such a wide variety of essential nuclear processes, it comes as no surprise that their dysfunction is implicated in a variety developmental and disease phenotypes (Maze et al., 2014). A number of mutations have recently been found in H3.3 by sequencing tumors from several brain and bone cancers (Schwartzentruber et al., 2012) (Behjati et al., 2013; Wu et al., 2012). A subset of these mutations involve replacement of lysine with methionine. Such substitutions have been shown contribute to oncogenesis by affecting the function of histone methyltransferases that act on these residues, therefore perturbing the gene regulatory networks they are involved in (Weinberg et al., 2017).

### 1.2.4 Histone Post-Translational Modifications

Chromatin is not only a generic packaging mechanism for DNA, but is also a scaffold to regulate virtually all chromosomal transactions. Eukaryotes must be able package their genome, reducing its dimensions many orders of magnitude to fit inside the nucleus, while retaining the ability to precisely unpackage parts of it at a moment’s notice and elicit specific responses to diverse external stimuli. Additionally, multicellular eukaryotes must manage a universal set genetic material in such a manner that it be interpreted in as many ways as to instruct and maintain as many necessary cellular phenotypes. This demands a mechanism to introduce a
degree of heterogeneity into the chromatin fiber far beyond the capability of histone variants (discussed in Section 1.2.3). Post-translational modification of histones is a well-characterized mode of biological regulation with the potential to handle the indexing of a suitably large number chromatin states (Bannister and Kouzarides, 2011). Already by 1984 had acetylation, methylation, phosphorylation, ADP-ribosylation, ubiquitylation, and glycosylation been reported to occur on histones (Van Holde, 1989). The sheer number and combinatorial complexity of histone post-translational modifications (PTMs, also referred to as 'marks') allows for a practically limitless number of unique chromatin modification landscapes. Application of cutting edge mass spectrometry techniques to chromatin biology have since led to the total reporting of nearly 30 unique chemical moieties existing on almost every modifiable histone residue (Huang et al., 2014; Janssen et al., 2017). Such modifications include a host of lysine acylations (acetylation, formylation, propionylation, butyrylation, 2-hydroxyisobutyrylation, crotonylation, malonilation, succinilation), lysine ubiquitylation, lysine mono di and trimethylation, serine and threonine acetylation, arginine mono and dimethylation (both symmetric and asymmetric), arginine citrullination, serine, threonine, tyrosine, and histidine phosphorylation, tyrosine hydroxylation, serine and threonine O-GlcNAcylation, and lysine and glutamate ADP-ribosylation (Figure 1.5).
Figure 1.5 Histone Post-translational Modifications. Chemical structures of all reported human, mouse, and rat histone post-translational modifications as in Huang et al. 2014. Positive (blue) and negative (red) electrostatic properties at physiological pH are indicated.

Among known histone PTMs, acetylation and methylation of lysine residues are the most well studied. Acetylation of histones was first described in 1963 (Phillips, 1963) and has been found to exist on all 4 core histones as well as linker histone H1. Two major types of histone acetylation occur in vivo. It is most recognized as it exists on the ε-amino group of lysine side chains, however, it can also occupy the α-amino group of N-terminal serine residues in H1, H2A, and H4 (Allfrey et al., 1964; Phillips, 1963, 1968). Although there have also been reports of O-acetylation on serine, threonine, and tyrosine amino acid side chains (Britton et al., 2013), only acetylation of amino groups present within histones will be discussed here. Not long after the discovery of histone acetylation it was proposed to be linked to transcriptional activation.

Pioneering work by Allfrey and colleagues led them to state, “It may be suggested that specificity in DNA-histone binding, alterable by acetylation of the histone, can influence the rate of RNA synthesis. This would allow a means of switching-on or –off RNA synthesis at different times, and at different loci on the chromosomes” (Allfrey et al., 1964). The discovery of a histone acetyltransferase enzyme in *Tetrahymena* that was homologous to yeast Gcn5 (a previously
identified and putative transcriptional co-activator) was the first evidence of a direct link between histone modification and transcriptional regulation (Brownell et al., 1996). Considering this, it is immediately evident that acetylation of highly abundant lysine side chains within histones would eliminate their positive charge and thus their interaction with the negatively charged DNA phosphate backbone. This might "loosen" their interaction with DNA render them less able obstacles for an elongating RNA polymerase II. Certainly, histone acetylation has since been observed to influence nucleosome dynamics in vitro (Ferreira et al., 2007a; Neumann et al., 2009), as well as be required for efficient in vitro transcription of chromatin templates (Kundu et al., 2000). Perhaps most convincing is the observation that targeting a histone acetyltransferase to a gene of interest leads to its activation and that this activation is dependent on an intact catalytic site (Hilton et al., 2015).

Considering lysine methylation, it is not immediately evident how such a modification might affect chromatin. Lysine side chains can also be mono di or trimethylated. Like lysine acetylation, it is a small modification, however, methylated lysine residues retain their charge, and, thus, electrostatic interactions with DNA are not drastically altered. Methylated lysine residues largely exert their effects by serving as stable binding epitopes, within the context of their surrounding amino acid sequence, for regulatory proteins. Many chromatin-associated proteins possess highly conserved domains that can specifically recognize various lysine methylation states—differentiating between the presence of 0, 1, 2, or 3 methyl groups. A variety of such domains exist allowing for the ability to interpret a large number of methylation ‘signals’ on chromatin (Musselman et al., 2012; Patel, 2016). As a canonical example, trimethylation of the 9th lysine residue in histone H3 (H3K9me3) plays a very well characterized role in the formation of transcriptionally silent constitutive heterochromatin found at centromeres (Maison and Almouzni, 2004). The H3K9me3 modification is deposited by the Suv39h1 methyltransferase (Rea et al., 2000), and recognized by heterochromatin protein 1 (HP1) via its N-terminal chromodomain (Bannister et al., 2001; Jacobs et al., 2001). Oligomerization of has been shown to be important for gene silencing in vivo (Canzio et al., 2011), a function disrupted by mutation of key binding residues in its chromodomain, supporting a role for H3K9me3 recognition in heterochromatin
formation (Jacobs et al., 2001; Platero et al., 1995). Targeting of HP1 in vivo has also been shown to induce local chromatin condensation. Altogether, this results in the current view that H3K9me3 marks constitutive heterochromatin, serving as an anchor for HP1 monomers, which then oligomerize to compact, and, in cooperation with a host of other nuclear factors, silence these regions of the genome (Maison and Almouzni, 2004).

As exemplified above, there are two major ways that histone modifications participate in the specific regulation of underlying nucleic acid sequences. “Cis” effects can be described as those that directly affect the biophysical nature of chromatin, while “trans” effects are considered those that rely on the recruitment of biological activity through a specific interaction between a functional protein and a histone modification. Cis-regulation of chromatin occurs through adjustment of histone-histone interactions, histone-DNA interactions, and interactions between nucleosomes. At the mononucleosome level, modifications lying under the DNA gyres can have major effects on nucleosome dynamics if they significantly alter the electrostatic properties of amino acids that interact with DNA. For example, acetylation of the ε-amino group on lysine side chains eliminates its positive charge, and phosphorylation of serine and threonine residues introduces a negative charge. Both types of modifications can drastically affect how modified amino acids interact with the phosphate backbone of DNA. One effect of such modifications is increased ‘DNA breathing’ (Anderson and Widom, 2000; Li et al., 2005), an equilibrium process of continual unwrapping and re-wrapping of DNA where it enters and leaves the nucleosome (Anderson et al., 2001; Brehove et al., 2015; Neumann et al., 2009; North et al., 2011). Such an event can facilitate invasion of nucleosomal DNA by transcription factors to engage their cognate oligonucleotide recognitions site, which, theoretically, would contribute to downstream transcriptional activity (Li and Widom, 2004; Tims et al., 2011). Histone acetylation can also affect nucleosome mobility (i.e. their ability to slide and occupy different translational positions on DNA) (Ferreira et al., 2007a). At the level of the chromatin fiber, certain nucleosome modifications have been shown to affect interactions between nucleosomes, affecting chromatin compaction. The most well-known example is H4K16ac-mediated disruption of the engagement of the H4 tail with the H2A-H2B acidic patch of neighboring nucleosomes, preventing fiber folding (Luger et al.,
1997; Shogren-Knaak et al., 2006; Wilkins et al., 2014). Indeed, H4K16ac localization correlates with chromatin accessibility in vivo (Bell et al., 2010). Other modifications such as H2BK120ub have also been shown to prevent chromatin condensation in vitro (Debelouchina et al., 2017; Fierz et al., 2011). Consistent with this observation, H2BK120ub functions in transcriptional activation, where DNA must be accessible for the transcription machinery (Kim et al., 2005).

Trans-regulation of chromatin occurs through histone modification-mediated recruitment of proteins that support structural changes or elicit various downstream biological functions such as gene transcription. Chromatin regulatory factors can anchor themselves to the genome by recognizing histone PTMs, achieved through specific interactions between the modification in question and a conserved recognition ('reader') domain (Musselman et al., 2012; Patel, 2016; Ruthenburg et al., 2007). The presence of such domains in chromatin-associated proteins is widespread, indicative of a direct role of histone PTM interpretation in their biological function. Single polypeptides can contain multiple reader domains, allowing interpretation of patterns of multiple marks. Potential for such 'multivalent' modes of recognition can also be achieved by multi-subunit protein complexes if more than one subunit possesses histone recognition capabilities (Ruthenburg et al., 2011; Ruthenburg et al., 2007). Such concepts lend support to the ‘Histone Code Hypothesis’ (Strahl and Allis, 2000), the idea that histone modifications, individual or in concert, form a 'vocabulary' of signals on chromatin, resulting in a language to be interpreted by various regulatory proteins that manage the genome. Though the idea that histone modifications form a strict code is debated (Henikoff and Shilatifard, 2011), their direct role in genome regulation cannot be argued. Development of methods to decipher each of their individual and combined functions will be essential to resolve this issue (Muller and Muir, 2015; Vora et al., 2016).

Reader domains that specifically recognize unmodified lysine, methyllysine, methylarginine, acetyllysine, phosphoserine, phosphothreonine, and crotonyllysine have been reported (Musselman et al., 2012; Zhao et al., 2017), and a wealth of structural biology studies have been critical to describe the nature of their specific modes of recognition. Trans regulatory effects on chromatin are typically observed when reader domains are coupled to some sort of
activity—either in the same polypeptide or through participation in a protein complex. This is not strictly always the case as described in the previous example involving HP1, which can exert effects on chromatin structure through H3K9me3 recognition and self-oligomerization (Canzio et al., 2011). A classic example of this mode of trans regulation (coupling of PTM recognition to enzymatic activity) is ‘spreading’ of the H3K9me2/3 mark by the Suv39h1 methyltransferase. This enzyme was first discovered in a genetic screen *Drosophila melanogaster* as a suppressor of position-effect variegation (PEV) (Tschiersch et al., 1994)—the dependence of gene activity on stochastic adoption of the silencing character of proximal pericentric heterochromatin due to induced genome rearrangement (e.g. X-ray damage) or transposition (Elgin and Reuter, 2013). Thus, the suppressive characteristics of heterochromatin are able to ‘spread’ into what is typically a euchromatic neighboring gene. Observation of the penetrance of the PEV phenotype (e.g. eye color), dependent on secondary mutations in the genome, can therefore be used to identify suppressors and enhancers of the phenotype (contributors to heterochromatin maintenance), which is how the *Drosophila melanogaster* homolog of Suv39h1, SU(VAR)3-9, was discovered (Tschiersch et al., 1994). The role of the H3K9me2/3 modification in heterochromatin formation is underscored by the fact that an enzymatically hyperactive variant of SU(VAR)3-9 enhances PEV (Ebert et al., 2004). Spreading of heterochromatin is initiated by a positive-feedback loop that involves continuous recognition and deposition of the H3K9me2/3 mark by Suv39h1 (Al-Sady et al., 2013; Muller et al., 2016). Recognition of the methyl mark is mediated by Suv39h1’s chromodomain (Al-Sady et al., 2013) and its installation carried out by its SET methyltransferase domain (Muller et al., 2016; Rea et al., 2000). It thus interprets the very mark that it deposits. Finally, introduction of H3K9me2/3 can then recruit HP1 to facilitate downstream silencing of affected chromatin (Grewal and Moazed, 2003; Lachner et al., 2001; Talbert and Henikoff, 2006).

Proteins that interpret, install, and remove histone PTMs are commonly referred to as histone ‘readers’, ‘writers’, and ‘erasers’, respectively. Each of these functionalities exist to act on many different modifications to render the chromatin landscape highly dynamic (Badeaux and Shi, 2013). Proteins that ‘read’ what they ‘write’ are generally thought serve to maintain essential chromatin states through a positive feedback mechanism (Margueron and Reinberg, 2010).
Proteins that ‘read’ and either ‘write’ or ‘erase’ a heterotypic modification undergo what is known as ‘histone modification cross-talk’ (Musselman et al., 2012; Suganuma and Workman, 2008). Both examples demonstrate that histone modifications can be functionally linked to reinforce chromosomal processes such as activation or repression of gene transcription.

Chromatin regulation by histone PTMs presents itself as an exceedingly complicated network composed of many interwoven layers. Untangling the hierarchical organization of these layers is, thus, essential to understanding eukaryotic genome function. Considering that the presence of many histone modifications correlate extremely well with functional genomic elements suggests a role in either establishment or maintenance of the status of the element they occupy (Bannister and Kouzarides, 2011). Indeed, mutations in proteins involved in the interpretation, installation and removal of histone PTMs are found in many developmental and disease-related disorders, including cancer. Mutations in histones themselves have also been discovered in certain tumor types and underscore the importance of such modifiable residues in nuclear biology (Behjati et al., 2013; Lewis and Allis, 2013; Maze et al., 2014; Schwartzentruber et al., 2012; Wu et al., 2012; Zhao et al., 2016). How such changes in amino acid composition in these cases disrupt histone PTM signaling systems and downstream process is currently a very pressing area of research (Lewis et al., 2013; Lu et al., 2016; Papillon-Cavanagh et al., 2017).

Unsurprisingly, the number of epigenetic regulators as drug targets is continually growing. Development of and search for pharmacological-based strategies to target histone reading, writing, and erasing capabilities (Audia and Campbell, 2016; Baylin and Jones, 2016; Brown et al., 2014) emphasizes the importance of these functionalities in chromosomal operations.

1.3 ATP-dependent Chromatin Remodeling Enzymes

1.3.1 Introduction

Organization of eukaryotic genomes into chromatin creates an inherent barrier to vital cellular processes by sequestering DNA binding sites. Correspondingly, chromosomal
transactions such as gene transcription require alterations in chromatin structure so that necessary trans-acting factors can interact with DNA to properly execute their roles. Such alterations depend on numerous chromatin-associated proteins as well as interactions between nucleosomes. While DNA residing between or outside of nucleosomes remains, in large part, accessible, nucleosomal DNA is unrecognizable by many DNA binding proteins due to its curvature and direct proximity to the octameric histone protein core. Thus, nucleosomes compete with DNA binding proteins for genomic access. Unbridled interpretation of genetic information hence requires the ability to maneuver and/or disassemble nucleosomes, overcoming an energy barrier created by numerous contacts between histone proteins and DNA. Conversely, chromatin assembly and placement of nucleosomes may be used to silence genetic material. ATP-dependent chromatin remodelers are a class of proteins that use ATP-hydrolysis to carry out these operations, their main functions consisting of chromatin access, assembly, and editing (incorporation and removal of histone sequence variants) (Clapier et al., 2017). In this way, they gate entry to the genome, playing critical roles every aspect of genome management (e.g. gene transcription, DNA replication, and DNA damage repair), processes that inevitably involve chromatin assembly and disassembly (i.e. addition, positioning, removal, or rearrangement of histone octamers along DNA) (Figure 1.6).
Figure 1.6 Functions of ATP-dependent Chromatin Remodelers. Chromatin remodelers carry out three basic functions—assembly, access, and editing. Regularly spaced nucleosome arrays are assembled by chromatin remodelers (left) and can be altered (right) to create access to underlying DNA sequences or diversified in function through incorporation of histone variants. Arrows indicate reversibility of these processes and the resulting dynamic nature of chromatin structure.

Early studies investigating genome accessibility revealed that active regulatory elements were found to be in more ‘open’ chromatin configurations relative to their inactive counterparts as indicated by mild nuclease digestion (Becker and Horz, 2002; Elgin, 1981). The development of the indirect end-labeling method (Wu et al., 1979), without question, revolutionized the study of genome structure as it allowed for dynamics of chromatin accessibility to be probed at individual genomic loci. Reik et al., 1991, used this technique to observe rapid and reversible nucleosome
removal over an enhancer upon inducible gene transcription though addition and removal of a signaling hormone (Reik et al., 1991). This study catalyzed the search for proteins involved in such a process. Around the same time, the SWI/SNF family of proteins were discovered as broad transcriptional activators in yeast (Peterson and Herskowitz, 1992), but how they might perform this function was at the time unclear. Several key discoveries in yeast led to the establishment of the family of proteins known today as ATP-dependent chromatin remodeling enzymes. The first was that SWI/SNF family protein SNF2/SWI2 contained a conserved helicase domain that had DNA-dependent ATPase activity and was necessary for its role in transcriptional activation (Laurent et al., 1993). The second was connecting its SWI/SNF protein function to chromatin itself. This involved the identification of mutations in histone genes that acted as suppressors of mutations in SWI/SNF proteins (Hirschhorn et al., 1992). Such genetic evidence suggested a functional interaction between these two protein families, and perhaps most importantly, altered chromatin structure at the yeast SUC2 promoter (requiring SWI/SNF proteins for transcription) was contingent on the presence of wild-type SNF2 and SNF5. The final steps in elucidating the basis of SWI/SNF function in gene activation required its biochemical purification. Purified yeast SWI/SNF components were characterized in vitro to be bona fide ATP-dependent chromatin remodeling enzymes by enabling transcriptional activator access to nucleosomal DNA through nucleosome disruption (Cote et al., 1994; Kwon et al., 1994).

Over two decades of in vitro characterization and in vivo functional studies have led to the classification of four families of ATP-dependent chromatin remodeling enzymes (Clapier et al., 2017). They are broadly divided based on similarities and/or differences within a core catalytic ATPase subunit: SWI/SNF (for SWItch/Sucrose Non-Fermentable), ISWI (for Imitation SWItch), CHD (for Chromodomain Helicase DNA-binding), and INO80 (Table 1). ISWI and CHD remodelers carry out nucleosome sliding (nucleosome repositioning without dismantling) and chromatin assembly. SWI/SNF remodelers are typically associated with octamer disassembly or ejection (though they can also slide nucleosomes). INO80 remodelers, depending on the subtype, can slide nucleosomes, but are most well-known for their nucleosome editing capabilities. The DNA translocase region in each remodeling ATPase is flanked by unique domains that regulate
ATPase activity, interact with chromatin, and organize assembly of higher-order protein complexes. These cis-regulatory domains (in addition to a characteristic insertion in the translocation domain of INO80 remodelers) and accessory (non-ATPase) protein partners that are modularly assembled around the common catalytic subunit confer unique properties to the ATPase and affect the overall outcome or efficiency of the remodeling reaction (Clapier et al., 2017).

<table>
<thead>
<tr>
<th>Remodeling family</th>
<th>SWI/SNF</th>
<th>CHD</th>
<th>INO80</th>
<th>ISWI</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATPase</td>
<td>BRG1, BRM</td>
<td>CHD1, CHD2, CHD3, CHD4, CHD5, CHD6, CHD7, CHD8, CHD9</td>
<td>INO80, SRCAP, p400</td>
<td>SNF2h, SNF2L</td>
</tr>
<tr>
<td>Complexes formed</td>
<td>BAF (BRG1/BRM), nBAF (BRG1/BRM), sBAF (BRG1/BRM), sBAF (BRG1), PBAF (BRG1)</td>
<td>NuRD (CHD4)</td>
<td>INO80 (INO80), SRCAP (SRCAP), TRRAP/TIP60 (p400)</td>
<td>ACF (SNF2h), CHRAC (SNF2h), WICH (SNF2h), Nuc (SNF2h), RSF (SNF2h), NURF (SNF2L), CERF (SNF2L)</td>
</tr>
</tbody>
</table>

**Table 1 ATP-dependent Chromatin Remodeling Families in Humans.** Table displaying four main remodeling families in humans, the ATPases that comprise them, and the complexes that these ATPases participate in.

Each remodeler’s output (i.e. assembly, access, or editing), however, is the result of a universal enzymatic property—DNA translocation—that lies at the heart of their ability to break histone DNA contacts (Saha et al., 2002; Whitehouse et al., 2003). Remodeler ATPases are part of the Snf2 helicase family, within the SF2 super family of helicases (Durr et al., 2006). Snf2 enzymes, however, do not behave like helicases in the sense that they use the energy of ATP-hydrolysis to traverse duplex DNA, without undergoing strand separation. Most knowledge regarding the mechanism of DNA translocation is drawn from work studying monomeric DNA helicases, which track along the phosphate backbone of one strand of the double helix (Singleton et al., 2007). The core translocase motor is split into two RecA-like lobes referred to as DExx (lobe 1) and HELICc (lobe 2), which together undergo a series of conformational changes throughout the ATP-hydrolysis cycle (Singleton et al., 2007). The result is sequential binding and
release of the two lobes and unidirectional movement along DNA—often referred to as ‘inchworming’ due to the small step size in number of base pairs (Singleton et al., 2007; Velankar et al., 1999). For ATP-dependent chromatin remodelers this step size has been measured to be 1-2 base pairs per molecule of ATP bound, hydrolyzed, and released (Deindl et al., 2013; Harada et al., 2016; Singleton et al., 2007; Sirinakis et al., 2011). The most detailed information on DNA translocation in the context of nucleosome remodeling comes from examination of ISWI remodelers by single-molecule fluorescence resonance energy transfer (FRET) (Deindl et al., 2013). Bound two helical turns from the dyad axis, the ATPase moves seven base pairs out of the exit site of the nucleosome without drawing in any additional DNA. At this point sufficient strain is generated to trigger DNA to be drawn into the nucleosome in three base pair increments, concomitant with three base pair increments exiting the nucleosome. Each of these increments can be further broken down into one base pair elementary steps. This process occurs until translocation stops (Figure 1.7). An important question is how the nucleosome is able to accommodate the absence of seven base pairs in the initial step. A rigid histone octamer would require significant unwinding of DNA, and perhaps this twist is relieved by diffusion around the octamer and out the exit side of the nucleosome in subsequent steps (Mueller-Planitz et al., 2013a). However, it was recently reported that ISWI ATPase, SNF2h, induces conformational in the histone octamer during remodeling and that such changes were necessary for nucleosome sliding to occur (Sinha et al., 2017). Such deformations are another possible mechanism to compensate for distortions in nucleic acid structure during DNA translocation. Though many molecular mechanisms remain to be elucidated, a core DNA translocation activity forms the foundation of the diversification of each remodeling family’s observed cellular functions.
29

Figure 1.7 Model for DNA Translocation in Nucleosome Remodeling by ISWI Enzymes.

Panel 1 shows a cartoon of the ISWI ATPase bound to the nucleosome with the octamer in light blue and the DNA in gray. The ATPase domain binds the nucleosome at superhelical loop (SHL) 2 and the HAND-SANT-SLIDE domain binds extranucleosomal DNA at SHL7. The dyad axis is also labeled. In the second panel, the ATPase domain hydrolyzes ATP to pump seven base pairs of DNA out of the nucleosome at the exit site, while drawing in no DNA at the entry site. This leaves minus seven base pairs of tension in the nucleosomal DNA (red) that must be accommodated. The ATPase then draws three base pairs of DNA into the nucleosome (panel 3), which is results in three base pairs of DNA being pumped out of the exit site (panel 4). This process repeats in 3 base pair increments until translocations ceases.

1.3.2 SWI/SNF Family

SWI/SNF remodelers were the first remodeling family to be discovered. They were initially characterized in yeast as antagonists to chromatin-mediated transcriptional repression through eviction of nucleosomes and facilitation of transcriptional activator function (Cote et al., 1994; Kwon et al., 1994; Peterson and Herskowitz, 1992). SWI/SNF complexes are typically thought of as ‘access’ remodelers due to their nucleosome disassembly properties. This activity is often used to clear transcription factor binding sites or maintain proper promoter chromatin architecture, viewed to permit productive engagement by the transcription machinery (Cote et al., 1994; Krietenstein et al., 2016; Kwon et al., 1994; Parnell et al., 2015; Reik et al., 1991; Whitehouse et al., 2007). However, recent reports describe an unexpected property of SWI/SNF in gene regulation, resolving the mechanism of known genetic antagonism between SWI/SNF and polycomb repressive complexes (PRCs; transcriptional inactivators) (Tamkun et al., 1992).
Surprisingly, it was discovered that the ATPase activity of SWI/SNF was required to evict PRC1 from chromatin, and changes in this activity through ATPase deletion or targeted localization of SWI/SNF subsequently led to changes in PRC1 occupancy and dysregulation of gene targets (Kadoch et al., 2017; Stanton et al., 2017).

In humans, there are two SWI/SNF family ATPase homogs, BRG1 and BRM, that serve as the catalytic subunits for so-called ‘BAF (Brg/Brahma-associated factors) complexes’ (i.e. human SWI/SNF complexes) (Kadoch and Crabtree, 2015). Treated as a distinct entity, the BAF complexes consist of 15 subunits. However, because they incorporate many developmental and tissue-specific subunits, the true diversity of their composition is unknown (Ho and Crabtree, 2010; Kadoch and Crabtree, 2015). In this sense, they demonstrate a major feature of chromatin remodelers—direction of biological specificity by combinatorial assembly of accessory protein subunits around a core ATPase. The importance of such modularity, as studied in yeast, is underscored by studies that demonstrate the genetic deletion of individual subunits in the complex can lead to formation of unique subcomplexes and distinct changes in gene expression (Dutta et al., 2017; Sen et al., 2017). Such investigations were inspired by the discovery that subunits in the BAF complex were mutated in >20% of all human cancers (Kadoch and Crabtree, 2015; Kadoch et al., 2013). Elucidating unique contributions individual subunits provide in regulating the core ATP-dependent motor and how this translates into biological activity is essential to study chromatin remodeling and its role in development and disease.

1.3.3 CHD Family

The CHD family of ATP-dependent chromatin remodelers includes 9 individual ATPases in humans, while yeast only have one (yChd1). Thus, the number of family members appears to increase with organismal complexity. CHD1 is one of the most well-studied chromatin remodeling enzymes in vitro and is known to act as a monomer in vivo (Tran et al., 2000). CHD1 is well-known to possess nucleosome sliding and spacing properties (Lusser et al., 2005; McKnight et al., 2011), which contribute to genome-wide nucleosome organization (Gkikopoulos et al., 2011;
However, as a whole and across species, CHD remodelers have been associated with all types of remodeler activity—sliding/assembly (Lusser et al., 2005; McKnight et al., 2011), promoter clearance (Murawska and Brehm, 2011), and editing (Konev et al., 2007). Aside from CHD1, the human CHD4 ATPase is likely the most well characterized due to its role in the NuRD (Nucleosome Remodeling and Histone Deacetylase) complex (Lai and Wade, 2011), aptly named as it has both nucleosome remodeling (from CHD4) and deacetylase activity (from HDAC1 and HDAC2). Other core subunits include MBD2/3, MTA1/2/3, and RbAp46/48. Like other chromatin remodeling complexes, combinatorial assembly is thought to direct specific biological roles (Lai and Wade, 2011). Consistent with the presence of deacetylase activity, the NuRD complex has gene repressive activity (Reynolds et al., 2012a; Reynolds et al., 2012b), and disruption of these properties has emerged as a contributor to certain cancers (Lai and Wade, 2011). The mechanistic contributions of the remaining members of the CHD family remodeling enzymes to nuclear biology are largely mysterious (Murawska and Brehm, 2011).

1.3.4 INO80 Family

INO80 family remodelers are primarily known for their nucleosome editing properties (the ability to introduce or remove histone sequence variants from the nucleosome). There are three ATPase homologs in humans: INO80, SRCAP, and p400, which contribute to the INO80, SRCAP, and TRRAP/Tip60 complexes, respectively (Clapier et al., 2017). Most of what is known about these complexes comes from studies on the yeast remodeling enzymes INO80 and SWR1. The nucleosome editing activity of INO80 remodelers specifically refers to the exchange of H2A-H2B dimers for those of a different histone composition. For example, the SWR1 complex exchanges canonical H2A-H2B dimers for those comprised of H2A.Z and H2B (Mizuguchi et al., 2004), a process that is known to be regulated by H3K56ac (Watanabe et al., 2013). The INO80 conducts the reciprocal reaction (Papamichos-Chronakis et al., 2011). In humans, the SRCAP and TRRAP/Tip60 complexes behave analogously to SWR1, while the INO80 possesses the same function as its yeast INO80 counterpart (Mizuguchi et al., 2004; Papamichos-Chronakis et
In addition, removal of H2A.X from chromatin is also dependent on the yeast INO80 enzyme (Papamichos-Chronakis et al., 2006; van Attikum et al., 2007).

INO80 family remodelers are associated with a variety of functions. H2A.Z dynamics, controlled by INO80 remodelers, influence gene transcription, consistent with its enrichment at the 5’ end of genes. Indeed, loss of H2A.Z by deletion in budding yeast leads to dysfunction in transcriptional programs (Marques et al., 2010). Additionally, the yeast INO80 complex has been shown to participate in the formation of proper promoter nucleosome architecture in vitro (Krietenstein et al., 2016), consistent with reports that it can also exhibit sliding and spacing activity and is found at these genomic locations (Udugama et al., 2011; Yen et al., 2012). Whether this activity is mutually exclusive with H2A.Z removal and how this dual functionality is regulated is not currently known. Supplementary to gene regulation, INO80 remodelers are also implicated in DNA damage repair and genomic integrity (Papamichos-Chronakis et al., 2006; van Attikum et al., 2004), a function also determined to be related to H2A variant dynamics during such events.

1.3.5 ISWI Family

The chromatin remodeling ATPase, ISWI, was originally discovered in Drosophila melanogaster due to its sequence similarity to the SWI2/SNF2 protein in yeast and Drosophila. In humans, there are two homologs of ISWI—SNF2h and SNF2L. Both possess a core DNA translocase domain surrounded several regulatory domains and C-terminal HAND-SANT-SLIDE (HSS) domain. Due to their relative simplicity (2-4 subunits total) ISWI family remodelers have long served as a valuable model for studying ATP dependent chromatin remodeling (Clapier et al., 2001; Deindl et al., 2013; He et al., 2008; Mueller-Planitz et al., 2013b; Yang et al., 2006). To date, ISWI remodelers have been characterized has having a “nucleosome sliding” activity, altering translational positions of histone octamers without their removal from (Langst et al., 1999). SNF2h forms a total of 5 complexes—the ACF complex (with ACF1), the CHRAC complex (with ACF1, CHRACH-15, and CHRAC-17), the WICH complex (with WSTF), the NoRC complex
(with TIP5), and the RSF complex (with RSF1). The SNF2L ATPase is known to form two different complexes—the NURF complex (with BPTF and RBAP46/48), and the CERF complex (with CECR2) (Erdel and Rippe, 2011). ISWI ATPases (along with the CHD1) are unique in the fact that they have been shown to require both extranucleosomal DNA and the H4 N-terminal tail for maximum remodeling activity (Clapier et al., 2001; Ferreira et al., 2007a; He et al., 2008). This regulation is facilitated by unique cis-regulatory sequences embedded within the ATPase itself. The C-terminal HSS domain, which binds extranucleosomal DNA, is required for nucleosome sliding. Additionally, the translocase domain is flanked two auto-inhibitory modules referred to as AutoN and NegC. AutoN functions to inhibit ATPase activity, while NegC acts to regulate coupling of ATPase activity to DNA translocation. The auto-inhibitory function of AutoN is relieved by the presence of a basic patch on the H4 tail (K18R17H16R19K20) (Clapier and Cairns, 2012; Clapier et al., 2001; Clapier et al., 2002; Yan et al., 2016), while negative regulation of coupling of ATP hydrolysis to movement of DNA by NegC ceases upon binding of the HSS domain to extranucleosomal DNA (Clapier and Cairns, 2012; Clapier et al., 2017).

ISWI family chromatin remodelers have been implicated in almost every aspect of genome regulation including transcription (Alcid and Tsukiyama, 2014; Parnell et al., 2015; Smolle et al., 2012; Sun et al., 2015; Whitehouse et al., 2007; Whitehouse and Tsukiyama, 2006), DNA replication (Bozhenok et al., 2002; Collins et al., 2002; Poot et al., 2004; Yadav and Whitehouse, 2016), DNA damage repair (Aydin et al., 2014), and chromatin structure (Corona et al., 2002). As with other chromatin remodelers, these ascribed functions correlate with specific subunit composition of ISWI complexes. In the context of the NoRC complex, mammalian ISWI ATPase SNF2h uses its binding partner, TIP5, to regulate rDNA transcription by binding to sequence specific DNA-binding transcription factor TTF-1 (Nemeth et al., 2004). However, when SNF2h resides in the WICH complex, with accessory subunit WSTF, it plays a role in maintaining chromatin structure during DNA replication (Collins et al., 2002). This occurs through targeting of the WICH complex to replication foci through an interaction between WSTF and PCNA. Additionally, many studies also connect interactions between specialized domains within remodeling complex subunits and histones or histone PTMs to properly localize remodelers and
remodeler function (not unique to ISWI remodelers). For example, recognition of H3K4me3 by a PHD finger in BPTF, a component of the NURF complex, is critical for proper chromatin association and regulation of *Xenopus* development. Indeed, this binding is crucial for recruitment of the SNF2L ATPase (Wysocka et al., 2006). Continuing to decipher how different accessory subunits communicate unique signaling inputs to a universal ATPase will be critical to fully understand remodeler function *in vivo*.

### 1.4 High-throughput Chromatin Biochemistry: Catching up with Epigenomics

Continual expansion of high-throughput DNA sequencing and proteomic methodologies have led to an explosion of data fueling the field we know as epigenomics (Encode Project Consortium, 2012; Huang et al., 2014; Janssen et al., 2017; Roadmap Epigenomics Consortium et al., 2015). The attachment of the '-seq' suffix to a seemingly endless list of experimental acronyms underscores the importance of next-generation DNA sequencing technologies in our current understanding of epigenetics. In fact, a PubMed search for '-seq' will show that the number of publications using this term from 2015-present is almost 1.5-times greater than the total number ever recorded before that time period. Such experimental techniques have had an impressive capability characterize aspects of chromatin and epigenetic regulation on a genome-wide scale. Dependent upon the development of highly specific antibodies, ChIP-seq furnishes the ability to localize chromatin associated proteins as well as histone modifications throughout the genome. Efforts employing ChIP-seq have led to the many histone PTM signatures now associated with many major genomic features such as gene promoters, enhancers, actively transcribed genes, and facultative and constitutive heterochromatin (Bannister and Kouzarides, 2011). MNase-seq, DNase-seq, FAIRE-seq, ATAC-seq and HiC have the capability to probe chromatin structure from the level of global individual nucleosome positions (MNase-seq) to identification of ‘open’ chromatin and DNA accessibility (DNase-seq, FAIRE-seq, ATAC-seq) to 3-dimensional genome organization (HiC) (Bell et al., 2011; Buenrostro et al., 2013; Lieberman-
Aiden et al., 2009). Sequencing technologies have also been developed to catalog DNA methylation (Yong et al., 2016).

The ENCODE project represents a massive tour de force characterization of the human genome that utilized many ‘-seq’ methods to generate immense amounts of data regarding functional genomic and epigenomic elements. While epigenomic studies afford vast amounts of information regarding the functional organization of eukaryotic chromosomes, including the distribution of histone modifications, popular techniques such as ChIP-seq, at best, provide correlative-type information on a genome-wide level. They do not reveal if and how histone PTMs regulate the genomic elements they are associated with. In this regard, biochemical approaches offer a powerful complement to epigenomic techniques by providing the means to explore potential causal relationships between histone PTMs and the regulation of the many proteins that act on chromatin. The sheer number and combinatorial complexity of histone modifications in vivo have driven the development of high-throughput methods to study their function. Initial efforts to characterize specificity of reader, writer and eraser specificities of chromatin effector proteins typically employed screening platforms using histone-derived peptides (e.g. peptide arrays) (Muller and Muir, 2015; Taverna et al., 2007). While highly instructive, it is difficult to replicate the structural intricacies of chromatin using peptides. Indeed, the use of histone peptides renders the study of recognition of multiple histones simultaneously impossible, and they do not incorporate the omnipresent nucleic acid component of chromatin—DNA. As many proteins possess enzymatic functions directed at nucleosomes that are mediated by the ability to sense the composition of their substrate, the use of DNA-barcode nucleosome libraries is one such strategy that overcomes these shortfalls, allowing biochemical analysis of chromatin effector proteins to be performed in a highly-parallelized manner (Nguyen et al., 2014). Combining the strengths of protein semi synthesis and next-generation DNA sequencing, such a technique offers the capability to test intact chromatin effectors on physiological chromatin substrates, whilst consuming minimal amounts of material.
1.4.1 Regulation of ATP-dependent Chromatin Remodeling Enzymes by Nucleosome Modifications

ATP-dependent chromatin remodelers regulate access to genetic information by contributing to nucleosome positions in vivo (Jiang and Pugh, 2009). Typically, individual members of a chromatin remodeling family possess the same core ATPase subunit (or homolog thereof), yet are able to perform functionally distinct roles within the nucleus. This specialization occurs through partnership with accessory protein subunits, which can offer additional protein binding partners or sensory inputs that direct chromatin remodeling activity to specific genomic loci. The underlying molecular mechanisms that control this functional specialization, however, remain largely unclear.

The current picture of nucleosome engagement by remodelers is painted by over a decade of studies employing primarily X-ray crystallography, cryo-EM, and proximity-based crosslinking (Chaban et al., 2008; Dang and Bartholomew, 2007; Grune et al., 2003; Liu et al., 2017; Nguyen et al., 2013; Racki et al., 2009; Yamada et al., 2011; Yan et al., 2016). Unfortunately, all high-resolution studies only investigate remodeler subdomains, and those employing full-length remodelers were acquired at too low resolution to yield atomic-level information. Thus, a complete and detailed picture of how intact remodelers integrate multiple molecular contacts to carry out nucleosome remodeling is still unclear. Most structural data come from examination of the ISWI family due to the experimental tractability associated with these enzymes. The ISWI ATPase occupies a position on the nucleosome two helical turns away from the dyad axis (near the activating epitope located on the H4 N-terminal tail), utilizing its ATP-dependent DNA translocase activity to track along a single strand of nucleosomal DNA (Dang and Bartholomew, 2007). It’s C-terminal HSS domain binds extranucleosomal DNA on the opposite side of the nucleosome (Dang and Bartholomew, 2007). Translocating DNA around the histone core, resulting in nucleosome sliding activity associated with ISWI enzymes, requires anchoring of the ATPase domain in a fixed position relative to the octamer. The histone binding domain mediating such an interaction currently remains elusive. Thus, ISWI remodelers engage
the octamer core, histone tails, nucleosomal DNA, and linker DNA, all which are subject to post-translational modification (Huang et al., 2014).

Chemical alterations to the nucleosomal substrate by covalent modification may provide a means to tune points of contact made by chromatin remodelers, either through enhancement or inhibition, to regulate the overall efficiency of the DNA translocation process. Though these effects have long been proposed, only a small number of studies have described how modifications to nucleosomes, which occur in massive complexity in vivo (Tan et al., 2011), affect remodeling efficacy (Chatterjee et al., 2015; Chatterjee et al., 2011; Ferreira et al., 2007a; Goldman et al., 2010; Neumann et al., 2009; North et al., 2011; Shogren-Knaak et al., 2006). In addition, it is not known how chromatin remodelers employ modular subunit assembly in their 'regulation by targeting epitopes' (Clapier et al., 2017), as chromatin modifications may serve as a means to localize their activity to specific genomic function. To date, there has been no systematic analysis of how remodeling activity is tuned by the presence of nucleosome modifications. Addressing this compelled the development of a DNA-barcoded mononucleosome library, tailored for studying chromatin remodeling, to profile the remodeling behavior (i.e. nucleosome sliding) of seven recombinantly produced human ISWI remodelers; namely the SNF2h ATPase and the ACF, CHRAC, WICH, NoRC, RSF and NURF complexes (Figure 1.8).
Figure 1.8 ISWI Chromatin Remodelers Profiled. Individual protein subunits of each complex are labeled. The SNF2h ATPase provides DNA translocase activity for the ACF, WICH, RSF, CHRAC, NoRC, and NURF complexes. The SNF2L ATPase provides DNA translocase activity for the NURF complex.

The resulting data show that accessory (non-ATPase) subunits of ISWI remodelers can distinguish between differentially modified nucleosomes, directing remodeling activity toward specific nucleosome substrates based on their modification state. Unexpectedly, the nucleosome acidic patch was discovered as necessary for maximum activity of all ISWI remodelers evaluated, and that this dependence also extends to CHD and SWI/SNF family remodelers, suggesting the acidic patch may be generally required for chromatin remodeling. Critically, remodeling activity was observed to be regulated by modifications neighboring the acidic patch, signifying it may act as a tunable interaction hotspot for ATP-dependent chromatin remodelers and, by extension, many other chromatin effectors that engage this region of the nucleosome surface (Armache et al., 2011; Barbera et al., 2006; Kato et al., 2013; Makde et al., 2010; McGinty et al., 2014; Morgan et al., 2016). The findings herein provide a foundation to understand how the discovered mechanisms of substrate recognition by ISWI remodelers occur in detail. This study also provides the impetus to validate such enzymatic regulatory mechanisms in vivo and how that modifies their nucleosome remodeling activity to regulate DNA-templated processes in the eukaryotic nucleus (Figure 1.9).
Figure 1.9 Regulation of ISWI Remodeling by Nucleosome Modifications. Nucleosome modifications (shapes, left) can act as positive, negative, or neutral inputs into ISWI nucleosome remodeling activity. This provides the potential for discovered regulatory outputs to modify remodeling activity \textit{in vivo}, regulating biological processes such as chromatin assembly/silencing (top right), gene transcription though transcription factor (red) access to its cognate binding site (purple; middle right), and gene transcription through promoter (green) occlusion by nucleosome occupancy (bottom right).
CHAPTER 2: HIGH-THROUGHPUT CHROMATIN REMODELING ASSAY DEVELOPMENT
AND NUCLEOSOME LIBRARY CONSTRUCTION
2.1 Introduction

A variety of laboratory methods have been developed to collectively give the ability to measure all three major ATP-dependent chromatin remodeling activities—nucleosome sliding, ejection, and editing. As exhibited by ISWI remodelers, nucleosome sliding is typically measured using either a restriction enzyme accessibility assay (REAA) or a nucleosome mobility gel-shift assay. REAAs rely on the availability of a restriction enzyme cleavage sequence, dependent on whether or not it is incorporated into nucleosomal DNA (He et al., 2008). On a timescale relevant to rates of chromatin remodeling, restriction enzymes are not able to cut their cognate recognition motif if it is occupied by a histone octamer. Juxtaposition of the double helix to histones and its adopted curvature upon nucleosome formation renders it essentially uncleavable. If the restriction enzyme site is incorporated into a strong nucleosome positioning sequence (e.g. the 601 sequence (Lowary and Widom, 1998)) such that all nucleosomes in a population uniformly occupy a single initial position, nucleosome movement onto an extranucleosomal DNA overhang away from the restriction site will expose it and allow cleavage by the restriction enzyme. The rate of their movement from that position (sliding) can then be accurately, however indirectly, quantified by measuring the rate of cleavage of nucleosomal DNA over time (Figure 2.1a).

Nucleosome sliding can be directly measured using an alternative method based on the observation that nucleosomes display different electrophoretic mobilities based on the position of the histone octamer on a DNA fragment (Eberharter et al., 2004). Nucleosomes that are positioned on the end of a DNA fragment display a higher mobility than those residing in the middle. Therefore, as with REAAs, this type of assay also requires a DNA fragment longer than what is minimally wrapped by the octamer (> 147 base pairs). Nucleosomes are initially directed to the end of the DNA fragment using a strong nucleosome positioning sequence and their movement away from this sequence is concomitant with the appearance of more slowly migrating bands over time as visualized by native gel electrophoresis. The C-terminal HSS domain of ISWI remodelers acts as a molecular ruler, sensing extranucleosomal DNA and directing nucleosome movement toward excess linker DNA. This process will happen repeatedly until the nucleosome
reaches an obstacle, such as a DNA binding protein, neighboring nucleosome, or the end of a DNA fragment. This property of ISWI remodelers renders them capable of organizing nucleosome arrays with a defined spacing between nucleosomes (Clapier et al., 2017), an observed function in vivo as it relates to chromatin assembly on newly replicated DNA (Yadav and Whitehouse, 2016). This tendency to slide nucleosomes toward ‘naked’ DNA, for example, nucleosome-free regions typically associated with active gene promoters, also likely competes with binding of the transcription machinery and underlies the transcriptionally repressive properties attributed to ISWI remodelers. As such, on sufficiently short oligonucleotide fragments, ISWI remodelers will equalize the amount of extranucleosomal DNA on each side of the histone octamer resulting in centering of the nucleosome. Since centered nucleosomes can be distinguished from their end-positioned counterparts by native gel electrophoresis, remodeling rates can be quantified by measuring movement of nucleosomes away from the end of a DNA fragment (Figure 2.1b).

Figure 2.1 Restriction Enzyme Accessibility and Gel Mobility Shift Assays for Chromatin Remodeler Analysis. a, Restriction enzyme accessibility assay measurement of chromatin remodeling uses cleavage of nucleosomal DNA as a proxy for chromatin remodeling. Time-points are collected over the course of a remodeling reaction, remodeling activity is quenched, and the sample is run on a polyacrylamide gel. The abundance of uncut DNA at each time point is measured via densitometry. A decay rate is calculated and assumed to be equivalent to that of nucleosome repositioning. b, Gel mobility shift assay measurement of chromatin remodeling directly measures nucleosome movement by separating differently positioned nucleosome species on a native polyacrylamide gel. Time-points are collected over the course of a remodeling reaction, remodeling activity is quenched, and the sample is run on a polyacrylamide gel. The
abundance of nucleosomes occupying the initial position is measured at each time point via densitometry. A decay rate equivalent to that of nucleosome repositioning is then calculated. In a and b, graphs are only examples of what densitometry data might look like when plotted and then fit to a 1-phase exponential decay to extract a rate constant of remodeling.

Sophisticated FRET-based methods have also been developed to measure nucleosome sliding at the single-molecule level and dissect DNA movement around the histone octamer by ISWI remodelers at single base-pair resolution (Deindl et al., 2013). It must also be noted that remodeling activity that results in eviction of nucleosomes can be measured by REAA-based or gel mobility-based experimental techniques as well (Clapier et al., 2016; Fan et al., 2003). Western blot-based assays have correspondingly been developed to measure nucleosome editing/dimer exchange activity of INO80 remodelers (Papamichos-Chronakis et al., 2011).

All knowledge about nucleosome remodeling activity that has accumulated to date has been generated by one or a variation of the assays described in the previous paragraph. While powerful, they lack the throughput to efficiently probe the number of potential substrates, with respect to nucleosome modifications, available to chromatin remodelers in the cell and determine how these modifications affect their activity. In addition, understanding how modular assembly of chromatin remodeling complexes affects their interpretation of such substrates can require the testing of many individual subunit combinations. Tackling these questions using currently available experimental techniques is tremendously daunting and points to the need to develop new methodologies to answer them. Some attempts have been made at multiplexing existing assays. For example, separate assembly of unmodified nucleosomes and nucleosomes containing H2A.Z on DNA linked to unique fluorophores allowed these nucleosomes to be mixed and analyzed simultaneously via a REAA as the results of each reaction could be uniquely visualized by their distinct fluorescence properties (Goldman et al., 2010). However, this method is limited to the number of available fluorophores with non-overlapping excitation and emission spectra. DNA-barcoded mononucleosome libraries benefit from the extremely high encoding power of DNA, with the potential to index sufficient numbers of substrates (tens of thousands or
more) and approach the combinatorial complexity of chromatin modifications in vivo (Nguyen et al., 2014). The remaining sections of this chapter will discuss purification of seven ISWI family remodelers with distinct subunit compositions and synthesis of a 115-member DNA-barcoded mononucleosome library capable of measuring nucleosome remodeling activity using and modified restriction enzyme accessibility assay.

2.2 Purification and Characterization of ISWI Family Chromatin Remodelers

Recombinant protein production techniques typically have high product yields when compared to purification of their native counterparts from wild-type cells. High-specificity and efficiency affinity tags for isolation can also be easily introduced via standard molecular cloning techniques rather than relying on superb quality antibodies or genetic manipulation recover the protein of interest from cell lysates. Baculovirus-based expression of eukaryotic proteins also have the benefit of being carried out in a eukaryotic cellular environment (Sf9 cells). In addition, the ability to infect a single Sf9 cell culture with multiple viruses containing genes of interest is well-suited for the study of chromatin remodelers as different viral isolates may simply be 'mixed and matched' to produce intact complexes of various subunit combinations. Accordingly, such a strategy was used to produce the SNF2h ATPase along with the ACF, CHRAC, WICH, NoRC, RSF, and NURF complexes in Sf9 cells. Unfortunately, remaining ISWI family members, the SNF2L ATPase and the CERF complex were not well-behaved in nucleosome remodeling assays post-purification, and thus were not pursued further.
**Figure 2.2 Nuclear Extraction of Chromatin Remodelers.** Example nuclear extract preparation procedure to enrich the ACF complex followed by SDS-PAGE and staining with Coomassie.

Infected cells were subjected hypotonic lysis and soluble nuclear extracts prepared, which led to a significant pre-enrichment of remodeler proteins (Figure 2.2). Subsequent anti-FLAG affinity isolation facilitated purification of remodelers to apparent homogeneity as observed by SDS-PAGE analysis and Coomassie staining. Their ATP-dependent nucleosome remodeling activity was confirmed using a restriction enzyme accessibility assay (REAA). Such high-quality enzyme preparations facilitate their biochemical characterization and allow attribution of distinct functions (e.g. the ability to interpret different modified nucleosome substrates) to individual subunit combinations. See Section 7.7 for details about remodeler purification.
Figure 2.3 Characterization of Recombinant ISWI Chromatin Remodelers. a, Purified chromatin remodelers were run on a 4-20% Mini-PROTEAN® TGX™ gel (Bio-Rad) for 35 min at 180 V. Proteins were stained with Coomassie. The composition of each remodeler/remodeling complex is depicted above each respective lane on the gel. Expected molecular weights: SNF2h: 122 kDa, ACF1: 179 kDa, CHRAC-15: 14.7 kDa, CHRAC-17: 16.9 kDa, WSTF: 171 kDa, TIP5: 208 kDa, RSF1: 164 kDa; migrates at higher apparent molecular weight, SNF2L: 121 kDa, BPTF: 338 kDa, RbAp46: 47.8 kDa. b, All remodelers display ATP-dependent nucleosome remodeling activity as detected by a restriction enzyme accessibility assay.

2.3 Design and Development of a High-Throughput Nucleosome Remodeling Assay

The construction of DNA-barcoded mononucleosome libraries for accelerated biochemical analysis of chromatin effectors has been previously described (Nguyen et al., 2014). It is a highly sensitive methodology dependent on massively parallelized next-generation DNA sequencing (NGS) technology that can deliver extremely large numbers of biochemical
measurements from a single experiment. This is allowed by a unique nucleotide “barcode” embedded within an individual nucleosomal DNA sequence describing the distinct chemical make-up that nucleosome. Such a strategy permits individual nucleosomes to be pooled together to form a “library” and thus for a single analysis to be carried out on multiple nucleosomes simultaneously. The results of a particular biochemical experiment are then measured via NGS as individual read counts and assigned to the behavior of particular nucleosomes by their unique barcode identifiers.

All nucleosome remodeling assays require the ability to distinguish unaffected nucleosomes from those that have been remodeled and the measurement of this conversion over time to calculate a remodeling rate. The two techniques characteristically employed to evaluate ISWI remodelers are the restriction enzyme accessibility assay and gel mobility shift assay (Figure 2.1). Each of these approaches relies on electrophoretic separation to detect chromatin remodeling. If either of these procedures are employed to analyze a nucleosome library there is no way to monitor individual nucleosome remodeling reactions via their distinct barcodes and thus the multiplexing capability that defines DNA-barcoded mononucleosome libraries cannot be utilized. However, the capacity of the REAA to ‘mark’ remodeled nucleosomes using an irreversible DNA cleavage event allows this particular assay to be adapted to and read out by NGS. As mentioned previously, in a REAA unmoved (uncut) and remodeled (cut) nucleosomes are differentiated by whether or not their associated DNA has been cut by a restriction enzyme. REAAs are performed identically in a nucleosome library format except that when a time point is taken and quenched, instead of visualizing cut and uncut DNA by electrophoresis, the uncut DNA is purified specifically amplified by PCR. This in-solution separation technique serves an analogous role to that of electrophoresis in the sense that specific amplification of uncut DNA essentially ‘separates’ it from cut DNA. This is accomplished in several ways. First, the design of primers that flank the restriction site embedded within nucleosomal DNA ensure only uncut DNA is amplified in a PCR reaction (cut DNA can only undergo primer extension). Second, the primers used in this PCR step are used to incorporate suitable NGS oligonucleotide adapters such that only full-length amplicons (uncut) can be sequenced. An additional multiplex barcode is also
introduced here so that individual time point and/or experimental samples can be distinguished from one another. Indeed, amplification must remain in the exponential phase to ensure that the relative abundances of all DNA fragments are maintained. The introduction of this PCR step also serves to significantly increase the sensitivity of nucleosome library-based analyses. A typical assay as developed herein can be carried out on a REAA-compatible DNA-barcoded nucleosome library as follows (1) remodeler of interest is incubated with or without ATP, an appropriate restriction enzyme, and the nucleosome library (2) time points are taken (in this case 1, 2.5, 5, 15, 30, and 60 minutes post-initiation) and quenched (3) DNA (both cut and uncut) is purified from each collected time point and PCR amplified to add NGS adapter sequences and multiplex barcodes (4) PCR reactions are then pooled, purified, and subjected to NGS for barcode decoding. Barcode decoding yields all information necessary to generate a plot analogous to that in Figure 2.1a from which a remodeling rate can be calculated. It generates the y-value, the abundance of uncut DNA, by providing read counts for each individual fragment; it generates the x-value, the time point, which is inferred from the multiplex identifier barcode; and it designates which nucleosome each DNA fragment belongs to by its unique nucleosome identifier barcode. The curve corresponding to the abundance of the uncut DNA fragment over time is then fit to a suitable exponential decay equation. Such a strategy is capable extracting remodeling rate information in the form of kinetic constants simultaneously for virtually as many unique nucleosome substrates as the DNA-barcoded library is large (Figure 2.4). In the case of this study this number is 115. See Sections 7.13.1, 7.13.7, 7.13.8, and 7.17-7.19 for detailed methods regarding remodeling assays using the nucleosome library.
2.4 DNA-barcoded Mononucleosome Library Construction

2.4.1 Preparation of Modified Histones Using Protein Chemistry

Access to native and site-specifically modified histone proteins is critical to the study of direct cause and effect relationships between histone PTMs and chromatin effector proteins and relies on sophisticated techniques in protein chemistry. There are several methods typically employed to generate site-specifically modified histones (Holt and Muir, 2015; Howard et al., 2015; Muller and Muir, 2015). The simplest requires mutation of the residue of interest to cysteine and its alkylation with various modification analogs such as methyl and acetyllysine (Howard et al., 2015; Simon et al., 2007). This approach, however, has the drawback of introducing a thioether into the aliphatic sidechains of such amino acids and there are cases where resulting effects have undesirable properties, such as hindered recognition by reader domains (Chen et al.,...
Another technique relies again on mutation of the residue to cysteine, however in this case it is chemically converted to dehydroalanine and reacted with a thiol-containing PTM analog (Chalker et al., 2012). Though this approach is also attractive in its facility and ability to modify whole proteins, it not only also results in thioether ‘scar’, but passage though the dehydroalanine intermediate abolishes the stereochemistry of the α-carbon of the amino acid being modified (Chalker et al., 2012). Though this method or variations thereof have seen improvements in replacement of thiol-ene chemistry with carbon-carbon bond formation, the products generated in these cases are still not chirally resolved (Wright et al., 2016; Yang et al., 2016). Amber suppression technology has the advantage of being able to incorporate a variety of unnatural and modified amino acids into proteins during polypeptide elongation on the ribosome (Dumas et al., 2015) resulting in fully native post-translationally modified proteins. It has been used to introduce modifications such as lysine acetylation, crotonylation, and methylation (Muller and Muir, 2015) as well as phosphoserine into histones (Lee et al., 2013). This is accomplished using an engineered aminoacyl-tRNA synthetase/tRNA pair. Replacing the codon at the site of interest with an amber stop codon and supplementing the growth medium with the modified amino acid allows its incorporation into ribosomal proteins (Figure 2.5a). However, low incorporation efficiencies dependent on the modified amino acid or its site of incorporation makes generation of polypeptides containing multiple modified residues (either homotypic or heterotypic) challenging. In addition, with respect to histones, N-terminal acetylation consistently found on histones H2A and H4 cannot be installed using this method. However, once an appropriate aminoacyl-tRNA synthetase/tRNA pair is generated, producing singly modified proteins is relatively straightforward.

Native chemical ligation (NCL) can overcome all drawbacks of previously mentioned approaches in the preparation of modified histone proteins (Dawson et al., 1994; Holt and Muir, 2015). The ligation chemistry involves a chemoselective reaction between an N-terminal cysteine residue and a C-terminal thioester to yield a native peptide bond at the site of ligation. The process is initiated by nucleophilic attack of the thioester by the cysteine thiol and transthioesterification in the presence of a thiol catalyst. This positions the α-amino group on the
cysteine residue in close proximity to the thioester, driving an S to N acyl shift. The resulting native amide bond is thermodynamically locked, covalently linking the two original fragments together (Figure 2.5b). The regioselective nature of the reaction between the cysteine and thioester allows reactant peptides to be unprotected and for the reaction to proceed under extremely mild conditions (i.e. compatible aqueous buffer, neutral pH, and room temperature). Of course, traceless ligation requires the presence of a cysteine at the ligation junction. This initially proved to be a major limitation for generation of native histone proteins, which, in their entirety only contain at most two cysteine residues in the H3.1 variant. Fortunately, methods to radically desulfurize cysteine to alanine, which is much more abundance in histones have been developed (Muller and Muir, 2015; Wan and Danishefsky, 2007).

![Diagram of Amber Suppression and Native Chemical Ligation Technologies](image)

**Figure 2.5 Amber Suppression and Native Chemical Ligation Technologies.** a, An engineered aminoacyl-tRNA synthetase/tRNA pair are used to incorporate a non-standard amino acid (red) into a polypeptide during ribosomal translation. After translation is completed, the polypeptide contains the modified amino acid. b, A two-step reaction scheme depicting the mechanism of native chemical ligation of two peptides.
Preparation of peptide fragments to be ligated by solid-phase peptide synthesis (SPPS) permits simple introduction of appropriate reactive moieties on each peptide fragment (i.e. an N-terminal cysteine and a C-terminal thioester) (Holt and Muir, 2015; Muller and Muir, 2015). SPPS also allows access a plethora of protein modifications through incorporation of appropriately protected modified amino acids during chain elongation (Holt and Muir, 2015; Muller and Muir, 2015). The presence of strong denaturants (e.g. guanidinium chloride) in a peptide ligation reaction also allows high reactant concentrations to be used (e.g. millimolar), which boosts reaction rate and efficiency. An important drawback to note is if tertiary structure in the final polypeptide product is desired and it cannot be refolded, this strategy cannot be used. Additionally, technical limitations of SPPS make preparation of individual peptides over approximately fifty amino acids very difficult. This makes synthetic polypeptides at a length reaching over one hundred amino acids from ligating two reactant peptides challenging to obtain. This is a problem for core histones as they range from approximately one hundred to one hundred and forty amino acids in length. Indeed, total synthesis of histones can be accomplished by sequential ligation strategies (Fang et al., 2011), but is laborious.

In many cases, length limitations imposed by SPPS can be easily relieved using protein semi synthesis. Protein semi synthesis combines SPPS with recombinant protein technology to create proteins that are part synthetic and part recombinant—a ‘semi synthetic’ protein (Holt and Muir, 2015). Ligation of a synthetic fragment to a much larger recombinant protein product (typically produced in E. coli) significantly increases the length of proteins that can be prepared. Protein semi synthesis employs NCL, however, how the appropriate reactive moieties (i.e. N-terminal cysteine and C-terminal thioester) are generated differs depending on where the synthetic fragment is to be introduced. If the fragment prepared by SPPS is to be attached to the N-terminus of the recombinant fragment, it will contain a chemically generated thioester for reaction with an N-terminal cysteine on the recombinant polypeptide. If the fragment prepared by SPPS is to be attached to the C-terminus of the recombinant fragment, a different approach is required. An expressed protein ligation strategy employing an intein fusion is used to generate the recombinant protein fragment with a C-terminal thioester (Muir et al., 1998). The synthetic
fragment can then be easily prepared by SPPS with an N-terminal reactive cysteine for NCL. Unfortunately, length restrictions of SPPS do limit incorporation of modified amino acids to within approximately fifty residues of each termini using two-piece ligations, but, as mentioned before, this can be overcome by sequential ligation and total protein synthesis (Fang et al., 2011). However, this route can be very laborious and is more dependent on the ability of longer synthetic fragments to refold. Fortunately, core histones are short enough to be accessible by three-piece (Casadio et al., 2013) or four-piece ligation schemes if needed, and, when denatured and combined in equimolar amounts, can easily be refolded into their octameric form (Dyer et al., 2004). The clear majority of well-characterized marks occur on the N-terminal tails of histones, typically within the first thirty to forty amino acids, and can therefore be obtained by a standard two-piece ligation.

Figure 2.6 Two-piece Ligation Schematic for the Preparation of an H3K4me3 Modified Histone. A recombinantly produced histone H3 fragment containing an N-terminal cysteine is ligated to a complementary H3 peptide prepared by solid-phase peptide synthesis. A trimethyllysine amino acid is appropriately incorporated into the peptide during chain elongation, and a C-terminal thioester is chemically generated. The peptide, corresponding to the N-terminal
portion of H3 undergoes native chemical ligation with the remaining recombinant fragment. Desulfurization renders the ligation traceless and yields a native full length H3 histone with the lysine at the fourth position trimethylated (H3K4me3).

To create the nucleosome library 81 unique histone proteins were prepared using a variety of techniques. Unmodified core histones (H2A, H2B, H3, and H4), histone variants (H3.3, H2A.Z, and H2A.X), and histone mutants were produced recombinantly in E. coli. Modified histones were generated by utilization of both amber suppression and protein semi synthesis techniques. Histone modifications present within the library include lysine acetylation (Kac), lysine mono, di, and trimethylation (Kme, Kme2, Kme3, respectively), arginine monomethylation as well as asymmetric and symmetric dimethylation (Rme, Rme2a and Rme2s, respectively), lysine ubiquitylation (Kub), serine and tyrosine phosphorylation (Sph and Yph, respectively), serine modified with N-acetylglucosamine (SglcNAc), and lysine crotonylation (Kcr).
Table 2.1 Histones Used to Construct the Nucleosome Library. All histones prepared for use in the nucleosome library are listed in the table above. Each column corresponds to subtypes of each of the four core histones and includes unmodified versions of each. Modifications include lysine acetylation (Kac), lysine mono, di, and trimethylation (Kme, Kme2, Kme3), arginine monomethylation as well as asymmetric and symmetric dimethylation (Rme, Rme2a and Rme2s, respectively), lysine ubiquitylation (Kub), serine and tyrosine phosphorylation (Sph and Yph, respectively), serine modified with N-acetylglucosamine (SglcNAc), and lysine crotonylation (Kcr).

Analytical reversed-phase HPLC (RP-HPLC) and electrospray ionization mass spectrometry (ESI-MS) were used to verify the purity and atomic composition of each histone. The majority of the histones used in library preparation were prepared specifically for this study, and their corresponding analytical RP-HPLC chromatograms and mass spectra can be found in Figure 2.7. Specifically, the H2AK119ub, H2BK120ub, H3K4me3, H3K9ac, H3K9me3, H3K14ac, H3K18ac, H3K18acK23ac, H3K23ac, H3K27ac, H3K27me3, H3KpolyAc, H3R42me2a, H4K5ac, H4K8ac, H4K12ac, H4K16ac, H4K20ac, H4K5,8,12,16,20ac (KpolyAc), H4R17A R19A, H4R45A.
H4K8ac, H4K12ac, H4K16ac, H4K20ac, H4KpolyAc, and all crotonylated histones were manufactured as part of previous studies and analytical information regarding their preparation is described therein (Casadio et al., 2013; Li et al., 2016; Nguyen et al., 2014; Xiong et al., 2016). Precise details describing the production of all histones can be found in Section 7.3-7.5.
Figure 2.7 Characterization of Histones by RP-HPLC/ESI-MS. (Left) Analytical C-18 RP-HPLC chromatograms (insets; gradients used are individually noted) and corresponding ESI-MS spectra of indicated histones prepared in this study. Corresponding deconvoluted spectra are shown on the right. Lysine acetylation: Kac; lysine mono, di, and trimethylation: Kme, Kme2 and Kme3, respectively; arginine mono, asymmetric and symmetric demethylation: Rme, Rme2a and Rme2s, respectively; serine and tyrosine phosphorylation: Sph and Yph, respectively; serine modified with N-acetylglucosamine: SglcNAc. TFA = trifluoroacetic acid adduct.

Design of the nucleosome library, and therefore histone selection, was in partially guided by the desire to tailor it to analysis of ATP-dependent chromatin remodeling enzymes. Mutation or modification of the basic patch in H4 to alanine is known to inhibit ISWI remodeling activity (Clapier and Cairns, 2012; Clapier et al., 2002; Shogren-Knaak et al., 2006) and corresponding histones (e.g. H4R17A,R19A, H4K16ac) were prepared to use as negative controls. The
presence of H2A.Z has been reported to enhance ISWI remodeling activity and, thus, this variant was prepared as a positive control. Many acetyl and methyl marks were also included as all ISWI family remodeling complexes have a subunit containing a bromodomain, PHD-finger, or both. Additionally, histones were prepared that contained modifications or mutations suspected to alter histone-DNA contacts. Such chemical alterations can change how these amino acids interact with DNA (Muthurajan et al., 2004; Tropberger and Schneider, 2013) and, thus, it was hypothesized, have the potential to reshape the energy landscape that must be overcome during chromatin remodeling as these interactions are broken and reformed during DNA translocation. Additionally, as the role of the nucleosome acidic patch (Luger et al., 1997) in chromatin recognition is continually increasing (McGinty and Tan, 2015), thus one nucleosome member was included with the acidic patch disrupted by mutation. Lastly, as the effects of recently discovered ‘oncohistones’ (Behjati et al., 2013; Lewis and Allis, 2013; Maze et al., 2014; Schwartzentruber et al., 2012; Wu et al., 2012; Zhao et al., 2016) on remodeling are unknown, therefore a series of histones containing disease relevant mutations were also included. Altogether, the entire histone collection represents modifications and mutations that cover major nucleosomal features including histones tails, the octamer core, and DNA (Figure 2.8).
Figure 2.8 A Diverse Library of Modified Nucleosomes. Diagram depicting all histone modifications, mutants, and variants present in the 115-member nucleosome library used in this study. Residues modified or mutated were mapped on to the nucleosome (PDB: 1KX5) in black using UCSF chimera. H2A (light yellow), H2B (light red), H3 (light blue), and H4 (light green) modification and mutation locations are indicated by boxes and lines. For clarity, connections are only shown to a single copy of each histone protein.

2.4.2 Preparation of Barcoded 601 DNA

A requisite for nucleosome library preparation is the ‘tagging’ of each individual nucleosome with a DNA barcode that corresponds to the unique chemical identity of the
nucleosome it comprises. This barcode is a short oligonucleotide sequence embedded within the DNA. Its length is arbitrary and only need be long enough to adopt as many different sequence permutations as the size of the library. In this case, a six base pair barcode, which has the capacity to encode $4^6$ (4,096) nucleosomes was used. This number is far greater than the number of nucleosomes that were actually prepared (115), and thus there is significant room for expansion using this barcoding scheme.

In a restriction enzyme accessibility assay, the cause of restriction enzyme cleavage must be limited, as much as possible, to a chromatin remodeling event. Accurate measurement of nucleosome remodeling rates by this method thus requires all nucleosomes occupy the same initial position on a DNA fragment. This was accomplished by using the Widom 601 strong nucleosome positioning sequence (Lowary and Widom, 1998) to precisely position nucleosomes.

The 601 sequence is a synthetic oligonucleotide discovered in a high-throughput screen for DNA sequences that have a high propensity to form nucleosomes. For this reason, it has become very widely used in the field of chromatin biochemistry. To enable restriction enzyme cleavage of nucleosomal DNA, a PstI site was introduced by mutation seventeen base pairs from one end of the 601 sequence (He et al., 2008). Addition of unique barcodes to the 601 by PCR is not cost-efficient due to expense of reagents and amount of DNA needed to assemble individual nucleosomes. Therefore, each piece of nucleosomal DNA was assembled from two pieces. The 601 sequence contains 5'-DraIII and 3'-BsaI overhangs and was prepared by digestion of a plasmid containing 16 copies of the desired sequence followed by purification by polyacrylamide gel electrophoresis. As this fragment is common to all nucleosomes, large-scale production and purification of the plasmid was carried out in DH5α cells as previously described to generate a large amount of the final nucleosomal DNA fragment (Dyer et al., 2004). Importantly, digestion with BsaI and DraIII generated non-palindromic overhangs on each side, which allows specific ligation of DNA fragments with corresponding ‘sticky ends’. Unique oligonucleotide barcodes were added to the universal 601 sequence in this way. Barcoded DNA was prepared by first hybridizing individual oligonucleotides. Formation of double-stranded DNA resulted in a sticky end
compatible with the DraIII site on the 601 sequence. These oligos were then individually ligated to the 601 and the products were analyzed by gel electrophoresis.

![Figure 2.9 Characterization of Barcoded 601 (BC-601) DNA. a, BC-601 DNA prepared for all 115 nucleosome library members as described in Section 7.8.1. Ligation products are 192 b.p. in size and were visualized by polyacrylamide gel electrophoresis (5% acrylamide, 0.5x TBE, 200 V, 40 minutes) and staining with SYBR® Safe DNA gel stain. A faint band corresponding to unligated 601 DNA (601) is slightly visible in certain cases. b, BC-601 DNA for nucleosome 99 (see Table 2.2) was CpG methylated by the M.SssI methyltransferase (NEB) according to the manufacturer’s instructions and characterized by digestion with the Rsal restriction enzyme, which is sensitive to CpG methylation, and PstI, which is not.

The barcoded DNA has several features that are important to discuss. The first is obvious as it now makes each piece of DNA unique due to the introduction of a six base pair nucleosome identifier barcode. The second is that it provides a forty five base pair overhang opposite the PstI site for nucleosomes to slide onto, conducive to restriction site exposure via chromatin remodeling (He et al., 2008). Lastly, it is used to partially introduce a forward adapter sequence such that the DNA fragment is eventually compatible with Illumina® high-throughput sequencing.

Detailed information on the preparation of barcoded DNA can be found in Section 7.8.1.
Figure 2.10 Barcoded 601 (BC-601) Nucleotide Sequence. The full nucleotide sequence for each BC-601 DNA is shown above. The 601 sequence is represented in red with the PstI site bold and underlined. The rest of the sequence corresponds to the forty-five base pair overhang—the DralII ligation junction is in black, the position of the six base pair nucleosome identifier barcode is in blue (NNNNNN), and the partial Illumina® forward adapter sequence is in green.

2.4.3 Octamer and Nucleosome Preparation and Library Characterization

A major advantage of the nucleosome library based approach is its sensitivity. Since the final readout of such assays is DNA sequencing, collected samples can be amplified by PCR. This significantly reduces the amount of starting material needed (Nguyen et al., 2014). This proves advantageous in situations where large numbers of nucleosomes must be prepared and precious material is limited. Nucleosome library based strategies thus reduce the amount of material needed in a typical preparation of histone octamers by about thirty-fold (Nguyen et al., 2014). This allowed octamers to be prepared using only 0.75 nmol (around 10 µg) of each of the four histones. Octamer formation was carried out by denaturing each of the four core histones in guanidinium chloride and mixing them together in equimolar amounts. Three dialysis steps of four hours each (with one step overnight) into 2 M NaCl were carried out to refold the histones into an octamer containing two copies each of the four histones (Dyer et al., 2004; Nguyen et al., 2014).

A major achievement in the manufacture of DNA-barcoded nucleosome libraries was the development of a nucleosome preparation method compatible with the use of crude (unpurified) octamer preparations directly in nucleosome assembly (Nguyen et al., 2014). This also factors in to the small amount of starting material needed because it negates an octamer purification step as well as nucleosome assembly ratio tests (Dyer et al., 2004), significantly lowering the amount of octamer needed to generate a quality nucleosome product. Typically, octamers need to be
purified away from excess H2A-H2B dimers and/or H3-H4 tetramers by gel filtration. In this method, crude octamers are mixed with 0.5 equivalents of BC-601 DNA and 0.5 equivalents of biotinylated mouse mammary tumor virus (bio-MMTV) DNA, a known weak nucleosome positioning sequence (Flaus and Richmond, 1998) and nucleosomes are generated by salt-gradient dialysis (Dyer et al., 2004; Nguyen et al., 2014). Nucleosome formation is thermodynamically favored on the 601 sequence while any non-productive histone-DNA interactions that resulting from absence of octamer purification are left to occur on the bio-MMTV DNA. The limiting amount of 601 also ensures that nucleosomes are exclusively formed on the 601 DNA. 'Excess' nucleosomes and other histone-DNA species that may form on the bio-MMTV DNA are removed by streptavidin pull-down, yielding a homogenous nucleosome preparation (Figure 2.11). A list of all nucleosomes can be found in Table 2.2.
Figure 2.11 Nucleosome Assembly Using a Crude Octamer Preparation. One of each of the four core histone types (light blue quarter circles) is mixed in equimolar amounts and dialyzed into 2 M NaCl to yield a crude octamer preparation. Crude octamers are then mixed with 0.5 equivalents of BC-601 DNA and bio-MMTV DNA and nucleosomes are formed by salt-gradient dialysis. Excess nucleosomes and other histone-DNA species that may form in the process are removed by streptavidin pull-down, which results in a pure and homogenous nucleosome preparation.
Table 2.2 Histone Composition of Nucleosome Library Members. Individual histone composition of each member of the nucleosome library. Modifications include lysine acetylation (Kac), lysine mono, di, and trimethylation (Kme, Kme2, Kme3), arginine mono, asymmetric and symmetric dimethylation (Rme, Rme2a and Rme2s, respectively), lysine ubiquitylation (Kub), serine and tyrosine phosphorylation (Sph and Yph, respectively), serine modified with N-acetylg glucosamine (SGlcNAc), and lysine crotonylation (Kcr). * = CpG methylation of nucleosomal DNA.

Individual nucleosomes were analyzed native polyacrylamide gel electrophoresis to ensure proper assembly (Figure 2.12a). All nucleosome preparations were then pooled to form the library, which was then reanalyzed by native polyacrylamide gel electrophoresis to ensure its final quality (Figure 2.12b). At the pooling step, two critical, non-nucleosomal, DNA fragments were included. The first is used as an internal reference to normalize relative abundances of DNA between samples (DNA Standard 1), and the second is used to ensure the amount of PstI restriction enzyme is not limiting under the experimental conditions being used (DNA Standard 2). Integrity of the nucleosome library was further assessed by incubating the library with an anti-H3K4me3 antibody and performing a pull-down experiment. DNA from the isolated material was subjected to DNA sequencing, and, satisfyingly, all nucleosomes bearing the H3K4me3 mark were efficiently isolated (Figure 2.12c). Detailed information regarding the preparation of octamers and nucleosomes as well as library characterization as well as explanation of the DNA standards can be found in Sections 7.9-7.11.
Figure 2.12 Analysis of the Quality and Integrity of the Nucleosome Library. Analysis of individual nucleosome preparations (a) and the final library after pooling of nucleosomes (b) by native gel electrophoresis and staining with ethidium bromide. c, Antibody pull-down of library members using an anti-H3K4me3 antibody. Every nucleosome member containing an H3K4me3 mark (red) was efficiently isolated relative to other library members (black). Notably, the antibody
was also able to pull down a nucleosome possessing solely the H3K4me2 mark (blue), indicating a lack of antibody specificity in this case. This experiment was performed once.

2.5 Summary

ATP-dependent chromatin remodelers absolutely require a nucleosome to operate, and therefore exemplify the necessity of employing high-throughput analytical approaches that use more physiological substrates in the study of chromatin biochemistry. Such assay formats not only have the potential to tackle the complexity of chromatin effector binding specificity, but also possess the ability to connect this information with a functional output – a relationship that is generally poorly understood chromatin biochemistry, including with respect to remodeling activity (Musselman et al., 2012; Stricker et al., 2017). ISWI remodeler complexes offer an attractive vehicle to explore this question due to the prevalence of reader domains within their accessory subunits and availability of quantitative in vitro assay formats for studying ATP-dependent nucleosome sliding activity. Although the expanded nucleosome library prepared in this study was used to profile ATP-dependent chromatin remodelers, it is also entirely compatible with the analysis of proteins that bind, modify, or remove modifications from nucleosomes.
CHAPTER 3: DIVERSE REGULATION OF ISWI REMODELERS BY NUCLEOSOME MODIFICATIONS
3.1 Data Summary

The nucleosome library and developed experimental workflow were used to analyze the remodeling behavior of the SNF2h ATPase and the ACF, CHRAC, WICH, NoRC, RSF and NURF complexes (see Sections 7.13.1, 7.13.7, 7.13.8, and 7.17-7.19 for detailed procedures as well as data analysis and processing). In total, this resulted in the collection of 784 individual observed remodeling rate measurements from 28,224 data points (Figure 3.1a). Notably, these experiments only consumed only a total of 21 pmol of nucleosome library, underscoring the sensitivity of this approach. Assays were performed with and without the addition of ATP, and, as expected, remodeling activity was negligible when ATP was not present (Figure 3.1b and Figure 3.2). DNA Standard 2 was efficiently cleaved by PstI in both the presence and absence of ATP, in accordance with the fact that it is not occupied by a nucleosome (Figure 3.1b and c and Figure 3.2).
Figure 3.1 Diverse Regulation of ISWI Remodelers by Nucleosome Modifications. a, Heatmap displaying ISWI remodeling data against the nucleosome library. Rows were sorted based on values for SNF2h (low to high). $k_{MN} = \text{nucleosome remodeling rate}$. $k_{\text{unmod.}} = \text{unmodified nucleosome remodeling rate}$. Values were capped at -4 and 4 for display purposes. b, Example decay curves depicting individual rates ($k_{MN}$) as in a. c, Rank-ordered remodeling rates for the ACF complex ($k_{MN}$) against the library. Dashed red line = $k_{\text{unmod.}}$. d, Relative remodeling rates as in a for select library members. All data are represented as the mean of experimental replicates (n=3). Error bars represent ± s.e.m. All histones are unmodified unless otherwise specified.
Figure 3.2 Nucleosome Remodeling Activity is Negligible in the Absence of ATP. Bar graphs showing individual DNA cleavage rates ($k_{\text{MN}}$, remodeling rates in the case of nucleosomes) from library remodeling experiments for each member of the library in the presence of the indicated chromatin remodeler with and without ATP. Rate values were rank ordered and displayed from low to high. The dashed red line represents the rate of remodeling of unmodified
nucleosomes. The related graphs for the ACF complex can be found in Figure 3.1c. Data are represented as the mean of experimental replicates ± s.e.m. (n = 3).

In reactions without ATP, the presence of a rapid decrease in the abundance of the uncut DNA fragment at the beginning of the reaction followed a sharp transition to a much slower second phase was observed in many cases. The initial sharp decrease was attributed to the cutting of non-nucleosomal or ‘free’ DNA present in the reaction, and the second phase to cutting of nucleosomal DNA by the PstI restriction enzyme in the absence of ATP (Figure 3.3 and see Section 7.19 for further discussion). Consistent with this hypothesis, such a phenomenon was much evident for nucleosomes that contained more free DNA in their preparations (Figure 3.3b). This therefore led us to fit the decrease in abundance of uncut DNA to a two-phase exponential decay equation, with the rate of the first phase fixed to the cutting of non-nucleosomal DNA in each particular reaction (provided by DNA standard 2). The rate of the second phase was equated to the rate of nucleosome remodeling. For consistency, this model was applied to reactions carried out in the presence of ATP as well.
Figure 3.3 Assessment of Modeling Nucleosome Library Remodeling as a 1- vs. 2-phase Exponential Decay. 

**a**, Rates calculated for ACF remodeling against select nucleosome library members by fitting to a 1-phase or 2-phase exponential decay equation. 

**b**, Gel images are the same as in Figure 2.12a. 

**c** and **d**, Curves used to calculate rates as in a for a H2A+H2B+H3+H4KpolyAc modified nucleosome (c) and an unmodified nucleosome (d) (nucleosomes 93 and 42 in Table 2.2, respectively). 

**e**, Mean rate values ($k_{obs} \text{ min}^{-1}$) calculated for
ACF remodeling against all nucleosome library members fit to either a 1-phase or 2-phase exponential decay equation. Nucleosomes remodeled at a faster rate when fit to a 1-phase vs. a 2-phase exponential decay equation are highlighted in red and labeled. f, Same as in e, but without highlighted nucleosomes. For a, c, and d data are represented as the mean of experimental replicates ± s.e.m. (n = 3).

Importantly, experimental conditions were used where the rate of cleavage of DNA Standard 2 (DNA not occupied by a histone octamer) was always faster than nucleosome remodeling (with and without ATP), ensuring that the amount of PstI present was not limiting. To ensure the validity of the experimental and data analysis workflow, several nucleosomes in the library were included as controls based on previously published literature (Figure 3.1d). As acetylation (H4K16ac) or mutation of the H4 N-terminal tail basic patch has been shown to reduce ISWI family remodeling activity (Clapier and Cairns, 2012; Clapier et al., 2002; Shogren-Knaak et al., 2006), H4K16ac and a H4R17A,R19A mutant nucleosomes were prepared. Additionally, a nucleosome containing the histone variant H2A.Z was included in the library, since nucleosomes containing this H2A variant are reported to ISWI remodeling activity (Goldman et al., 2010). Analysis of the behavior of each of these nucleosomes in a library experiment was in direct agreement with literature precedent (Figure 3.1d). The fact that relative trends are similar across all remodelers analyzed (at least those containing SNF2h) point an effect on the common ATPase subunit.

3.2 Accessory Subunits Direct ISWI Family Remodeling Activity Toward Unique Substrates

Inspection of the kinetic data clearly indicates that ISWI remodeling activity is sensitive to nucleosome modifications, with both stimulatory and inhibitory effects evident for all seven enzymes tested (Figure 3.1a).
Figure 3.4 PCA Analysis of Library Remodeling Data. Principal component (PC) analysis of library remodeling data. Percentages show the fractions of the variance accounted for by each PC. Individual nucleosomes are shown in light blue, and PC weight values for each remodeler are shown in either orange or black. Weights are scaled by a factor of 2 for visibility. PC1 vs. PC2 (left) and PC1 vs. PC3 (right) are plotted.

Principal component analysis (PCA) was performed to identify experimental trends that contributed to variability in remodeler activity toward different nucleosome substrates. The first principal component accounted for the large majority of variance (83%) in the data. Notably, all remodelers had very similar weight values for PC1, demonstrating that histone modifications that stimulate or impede remodeling tend to do so similarly over all the enzymes tested (Figure 3.4). Two non-mutually exclusive scenarios could explain this effect; either the remodelers share functional motifs (for example, in the common ATPase subunit) that are similarly sensitive to changes in the substrate structure, and/or certain modifications affect the stability of the nucleosome to make DNA translocation more or less difficult, i.e. by disrupting or augmenting histone-DNA contacts. To test the second possibility, thermal fluctuation/movement experiments (Flaus and Richmond, 1998) were performed, and indeed find a positive correlation between increases in remodeling rate and increased thermal nucleosome movement for two nucleosomes containing mutations at the nucleosome dyad axis that have relatively high contributions to PC1.
and relatively low contributions to PC2 and PC3 (Figure 3.5). An H3K14cr nucleosome, that had relatively high contributions to PC2 and PC3 was observed no such increase in thermal mobility (Figure 3.5). Additionally, there is literature precedent supporting a correlation between increased nucleosome dynamics with observed increases in nucleosome remodeling rate with certain modifications or mutants present in the nucleosome library: H3K56ac (Neumann et al, 2009); H3K122ac (Chatterjee et al. 2015); H3Y41ph (Brehove et al. 2015). The thermal fluctuation/movement experiments and previously published results support mechanistic interpretations of the library data.

Figure 3.5 Nucleosome Thermal Mobility Shift Experiments. a, Nucleosomes analyzed directly after assembly by salt gradient dialysis. b, Nucleosomes from a re-analyzed after incubation for 1 hour at either 37 °C or 47 °C. Mobile nucleosomes will shift and coalesce in to the most thermodynamically favorable position along the DNA fragment used in assembly. Experimental replicates are displayed (n = 3). Unmod. = unmodified. All histones are unmodified unless otherwise specified.
Figure 3.6 Specialization of ISWI Remodelers Toward Diverse Nucleosome Modifications.

**a**, Principal component analysis of library remodeling data. Nucleosomes = light blue; PC weight values for remodelers = orange. Weights are scaled by a factor of 2 for visibility. **b**, ISWI remodeling rates for select nucleosome substrates in the library. $k_{MN} = $ nucleosome remodeling rate. $k_{unmod.} = $ unmodified nucleosome remodeling rate. Values were capped at -4 and 4 for display purposes. All histones are unmodified unless otherwise specified. **c**, Single-site modifications mapped onto the nucleosome (PDB: 1KX5) and colored based on whether they had consistently positive (green), consistently negative (red), or variable (purple) effects on nucleosome remodeling activity across all ISWI remodelers analyzed.
The second and third principal components reveal how the activity of the remodelers varies in response to specific nucleosome substrates (Figure 3.6a). Interestingly, the behavior of the isolated SNF2h ATPase differs markedly from all complexes that incorporate the ATPase for DNA translocation activity, indicating the role that bound accessory subunits play in repurposing remodeling activity through unique interpretation of individual nucleosome modification states. We find that a subset of the nucleosome library drives differences in remodeler activity; the most prominent of which is H3K14ac, which activates only the NoRC complex (Figure 3.7, Table 3.1).

Figure 3.7 Subsets of the Nucleosome Library Drive Differences in Remodeler Activity.
PC2 vs. PC3 are plotted as in Figure 3.6a. Nucleosomes driving differences in remodeler activity were numbered as in Table 2.2 and grouped by their location in PC space.
Table 3.1 Nucleosome Library Members Contributing to Variation in Remodeler Behavior.

Table 3.1 contains nucleosomes grouped (by color) according to Figure 3.7. Per remodeler, values in each column correspond to log₂ values of the ratio of the rate of remodeling for a particular nucleosome to that of unmodified nucleosomes. Nucleosome modifications per histone are specified and nucleosomes are numbered as in Table 2.2.

Overall, the remodelers show unique preferences for nucleosome substrates, with the exception of CHRAC and ACF, which exhibit overlapping activity profiles. This result is consistent with the very similar composition of the two complexes, which are identical except for two small histone-fold proteins present within CHRAC (CHRAC-15 and CHRAC-17, Figure 2.3a).

Therefore, it does not appear that these two additional proteins contribute significantly to substrate preference within the scope of the nucleosome library.

It is also interesting to consider how the location of histone modifications on the nucleosome particle correlate with remodeler activity. Notably, histone modifications located on

<table>
<thead>
<tr>
<th>Nucleosome</th>
<th>Group</th>
<th>Source</th>
<th>CHRAC</th>
<th>WCH</th>
<th>NuRC</th>
<th>RSF</th>
<th>NuRF</th>
<th>H2A Modifications</th>
<th>H2B Modifications</th>
<th>H3 Modifications</th>
<th>H4 Modifications</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>1</td>
<td>-0.35</td>
<td>-0.06</td>
<td>-0.01</td>
<td>-1.58</td>
<td>-2.24</td>
<td>-2.54</td>
<td>-1.04</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>1</td>
<td>-0.22</td>
<td>-0.04</td>
<td>-0.02</td>
<td>-1.86</td>
<td>-2.39</td>
<td>-2.10</td>
<td>-0.66</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>109</td>
<td>1</td>
<td>-0.42</td>
<td>-0.15</td>
<td>-0.26</td>
<td>-0.41</td>
<td>-1.17</td>
<td>-1.99</td>
<td>-4.04</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>1</td>
<td>-0.32</td>
<td>0.35</td>
<td>0.27</td>
<td>-0.02</td>
<td>-2.05</td>
<td>-1.65</td>
<td>-1.51</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>110</td>
<td>1</td>
<td>-0.41</td>
<td>0.08</td>
<td>-0.03</td>
<td>-0.10</td>
<td>-1.72</td>
<td>-1.91</td>
<td>-1.62</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>2</td>
<td>-0.26</td>
<td>-0.17</td>
<td>-0.19</td>
<td>-1.10</td>
<td>-2.15</td>
<td>-1.15</td>
<td>-2.26</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>2</td>
<td>-0.70</td>
<td>-0.35</td>
<td>-0.37</td>
<td>-1.38</td>
<td>-1.51</td>
<td>-2.34</td>
<td>-1.08</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>2</td>
<td>-0.91</td>
<td>0.70</td>
<td>-0.78</td>
<td>-2.28</td>
<td>-2.17</td>
<td>-4.37</td>
<td>-2.06</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>2</td>
<td>-0.93</td>
<td>0.56</td>
<td>-0.59</td>
<td>-1.05</td>
<td>-0.28</td>
<td>-1.88</td>
<td>-2.29</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>2</td>
<td>-1.02</td>
<td>-1.38</td>
<td>-1.50</td>
<td>-1.50</td>
<td>-1.35</td>
<td>-2.73</td>
<td>-4.04</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>2</td>
<td>-1.26</td>
<td>-1.08</td>
<td>-1.29</td>
<td>-3.66</td>
<td>-2.71</td>
<td>-4.37</td>
<td>-4.04</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>31</td>
<td>2</td>
<td>-1.20</td>
<td>-0.84</td>
<td>-0.91</td>
<td>-1.83</td>
<td>-1.65</td>
<td>-3.45</td>
<td>-4.04</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>61</td>
<td>2</td>
<td>-0.40</td>
<td>0.17</td>
<td>-0.20</td>
<td>-1.52</td>
<td>-2.38</td>
<td>-1.68</td>
<td>-2.11</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>95</td>
<td>2</td>
<td>1.36</td>
<td>1.78</td>
<td>2.07</td>
<td>1.36</td>
<td>1.06</td>
<td>1.47</td>
<td>0.48</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>102</td>
<td>2</td>
<td>-0.33</td>
<td>0.17</td>
<td>0.20</td>
<td>-0.95</td>
<td>-0.71</td>
<td>-1.06</td>
<td>-2.12</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>103</td>
<td>2</td>
<td>-0.41</td>
<td>0.10</td>
<td>0.09</td>
<td>-1.33</td>
<td>-1.09</td>
<td>-1.59</td>
<td>-3.02</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>107</td>
<td>2</td>
<td>0.02</td>
<td>0.81</td>
<td>0.86</td>
<td>0.63</td>
<td>-0.61</td>
<td>-1.39</td>
<td>-1.57</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>28</td>
<td>3</td>
<td>0.35</td>
<td>0.72</td>
<td>0.62</td>
<td>-0.82</td>
<td>2.36</td>
<td>0.44</td>
<td>0.92</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>33</td>
<td>3</td>
<td>-0.79</td>
<td>-0.69</td>
<td>-0.95</td>
<td>-2.39</td>
<td>1.63</td>
<td>-1.90</td>
<td>-4.04</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>35</td>
<td>3</td>
<td>-1.06</td>
<td>-0.63</td>
<td>-0.85</td>
<td>-1.87</td>
<td>1.09</td>
<td>-1.50</td>
<td>-3.18</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>37</td>
<td>3</td>
<td>0.14</td>
<td>0.15</td>
<td>0.12</td>
<td>-0.38</td>
<td>2.29</td>
<td>0.17</td>
<td>0.28</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>93</td>
<td>3</td>
<td>-3.33</td>
<td>-1.15</td>
<td>-1.50</td>
<td>-2.45</td>
<td>0.70</td>
<td>-2.31</td>
<td>-4.04</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>94</td>
<td>3</td>
<td>-0.58</td>
<td>-0.15</td>
<td>-0.43</td>
<td>-1.46</td>
<td>1.54</td>
<td>0.62</td>
<td>0.20</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>43</td>
<td>4</td>
<td>4.37</td>
<td>4.63</td>
<td>5.01</td>
<td>3.86</td>
<td>-3.70</td>
<td>0.17</td>
<td>-0.04</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>48</td>
<td>4</td>
<td>-0.31</td>
<td>-0.49</td>
<td>-0.58</td>
<td>-0.12</td>
<td>-0.39</td>
<td>0.17</td>
<td>0.05</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>49</td>
<td>4</td>
<td>-0.16</td>
<td>-0.66</td>
<td>-0.48</td>
<td>-0.182</td>
<td>-0.432</td>
<td>0.013</td>
<td>0.018</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>4</td>
<td>-0.61</td>
<td>-0.41</td>
<td>-0.54</td>
<td>0.14</td>
<td>-0.36</td>
<td>0.31</td>
<td>0.08</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>56</td>
<td>4</td>
<td>0.5298</td>
<td>0.195</td>
<td>0.195</td>
<td>0.2518</td>
<td>0.2848</td>
<td>0.6902</td>
<td>1.235</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>76</td>
<td>4</td>
<td>-0.127</td>
<td>-0.454</td>
<td>-0.48</td>
<td>-0.077</td>
<td>-0.279</td>
<td>0.1999</td>
<td>-0.031</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>77</td>
<td>4</td>
<td>-0.128</td>
<td>-0.337</td>
<td>-0.436</td>
<td>-0.066</td>
<td>-0.273</td>
<td>0.1937</td>
<td>0.0142</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>82</td>
<td>4</td>
<td>-0.56</td>
<td>-1.37</td>
<td>-1.53</td>
<td>-1.01</td>
<td>-1.33</td>
<td>-0.85</td>
<td>-0.12</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>83</td>
<td>4</td>
<td>-0.07</td>
<td>-0.64</td>
<td>-0.73</td>
<td>-0.44</td>
<td>-0.62</td>
<td>-0.03</td>
<td>0.25</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>87</td>
<td>4</td>
<td>-0.07</td>
<td>-0.86</td>
<td>-0.96</td>
<td>-0.52</td>
<td>-0.96</td>
<td>-0.18</td>
<td>0.08</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>97</td>
<td>4</td>
<td>-3.94</td>
<td>4.11</td>
<td>4.66</td>
<td>-0.37</td>
<td>-0.70</td>
<td>-4.37</td>
<td>-4.04</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>98</td>
<td>4</td>
<td>-4.37</td>
<td>5.90</td>
<td>5.13</td>
<td>-4.30</td>
<td>-3.70</td>
<td>-4.37</td>
<td>-4.04</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>4</td>
<td>-0.67</td>
<td>-1.33</td>
<td>-1.45</td>
<td>-0.87</td>
<td>-1.76</td>
<td>-0.59</td>
<td>-0.53</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>104</td>
<td>4</td>
<td>-3.81</td>
<td>-3.78</td>
<td>-4.50</td>
<td>-0.35</td>
<td>-3.70</td>
<td>-4.37</td>
<td>-4.04</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>54</td>
<td>5</td>
<td>-0.21</td>
<td>-0.70</td>
<td>-0.97</td>
<td>-0.60</td>
<td>-1.67</td>
<td>-0.46</td>
<td>-0.26</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>51</td>
<td>5</td>
<td>-0.088</td>
<td>-1.551</td>
<td>-1.467</td>
<td>-2.192</td>
<td>-2.174</td>
<td>-0.867</td>
<td>-3.29</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>62</td>
<td>5</td>
<td>-0.137</td>
<td>-0.431</td>
<td>-0.456</td>
<td>-0.349</td>
<td>-1.361</td>
<td>-1.542</td>
<td>-0.589</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

85
more accessible regions of the nucleosome (e.g. on the flexible tails) show more variability in their effects across different remodelers as compared to those residing under the DNA, which show consistent increases in remodeling activity (Figure 3.6c). Indeed, there are many individual cases where a specific nucleosome type had a broad range of effects across different remodelers analyzed (Figure 3.6b), further demonstrating that the unique subunit compositions of each complex play a direct role in determining nucleosome substrate preference.

3.3 Histone Binding Properties Do Not Necessarily Correlate with Remodeling Activity

Correlations between chromatin remodeling activity and binding preferences of reader domains present within the ISWI remodelers were also examined. This analysis proved particularly enlightening for the NoRC, WICH, and NURF complexes, all of which have reader domains within their accessory subunits whose binding preferences have been studied. In the case of the NoRC complex, we find an excellent correlation between remodeling activity and the known binding preferences of the associated reader domains (Figure 3.8a). Thus, nucleosomes containing H3K14ac and unmodified H3K4, the respective preferred ligands for the bromodomain and PHD-finger in the TIP5 subunit (Tallant et al., 2015), were the best substrates for the remodeler. Interestingly, crotonylation of H3K14 (H3K14cr) led to a decrease in nucleosome remodeling. Thus, the minor change from H3K14ac to H3K14cr results in a striking switch in enzymatic activity, the former activates while the latter inhibits.

The correlation between substrate preference and reader domain binding activity was less pronounced for the WICH remodeler complex. In this case, the binding properties of the WSTF bromodomain did not always translate to improved nucleosome remodeling (Figure 3.8b). For example, the bromodomain binds strongly to H3K4, K9, K14, K18, K23, K27, K36, K37, K56, K64, K115, K122 acetylated peptides (Filippakopoulos et al., 2012), but only a subset of these lead to stimulation of remodeling activity by WICH (principally, the H3K56ac and H3K122ac marks) and in the case of H3K4ac we observed inhibition of activity.
Finally, in the case of the NURF complex, an inverse correlation between expected binding preferences and remodeling activity was observed. BPTF, the largest subunit of this remodeler, contains a tandem PHD-finger bromodomain module near its C-terminus that is known to engage nucleosomes through a bi-valent mechanism involving H3K4me3 (PHD-finger) and H4K16ac (bromodomain) (Ruthenburg et al., 2011). Surprisingly, we find that nucleosomes containing these marks were remodeled less efficiently than unmodified nucleosomes (Figure 3.8c). In considering the origins for such discrepancies, it is worth noting that binding studies on chromatin factors typically employ isolated reader domains and/or histone-derived peptides (Muller and Muir, 2015; Taverna et al., 2007). These simplified systems, while often informative, are not subject to the same structural and steric constraints present in a more physiological setting. Indeed, taken as a whole, these data highlight the advantages of performing biochemical measurements with intact chromatin effectors and nucleosomal substrates.
Figure 3.8 Remodeling of Nucleosomes Containing Modifications Preferred by Histone Recognition Domains. Library remodeling data generated by the NoRC (a), WICH (b), and NURF (c) complexes for nucleosomes containing residues known to interact with histone binding modules in accessory subunits of each complex (NoRC: TIP5; WICH: WSTF; NURF: BPTF).
Literature binding specificities are displayed in corresponding tables on the right. Bar graphs display \( \log_2 \) values of the rate of remodeling of individual nucleosome library members (\( k_{MN} \)) relative to unmodified nucleosomes (\( k_{unmod} \)). Data are represented as the ratio of the mean of experimental replicates ± s.e.m. (\( n = 3 \)). Note that H3KpolyAc includes the H3K14ac modification (Table 2.2). All histones are unmodified unless otherwise specified. BRD = bromodomain, PHD = PHD-finger, PHD-BRD = tandem PHD-finger-bromodomain module. ND = not determined.

3.4 Histone-DNA Contacts Modulate Remodeling Activity

At the most fundamental level, ATP-dependent chromatin remodeling enzymes function to break contacts between histones and DNA by utilizing the energy released from the hydrolysis of ATP. Remarkably, in the case of DNA translocation by ISWI enzymes, these contacts must be broken and reformed while leaving the nucleosome intact (Deindl et al., 2013; Langst et al., 1999). Many histone residues that either contact, or lie directly under, the phosphate backbone are also post-translationally modified (Tropberger and Schneider, 2013). Such chemical alterations have the potential to reshape the energy landscape that must be traversed during chromatin remodeling (Tropberger and Schneider, 2013). With this in mind, the nucleosome library included a variety of histone mutations and PTMs that were considered likely to disrupt histone-DNA interactions (Figure 3.9a and b). Specifically, two key locations on the octamer surface were targeted: the DNA entry-exit site, and the dyad axis. As expected, mutation of residues in these two regions, in general, led to an increase in remodeling activity across all remodelers tested (Figure 3.9b). Note, these stimulatory effects could be overridden by modifications to the basic patch of the H4 tail, a region known to be essential for ISWI ATPase activity (Clapier et al., 2002). Importantly, several PTMs within these regions also led to a stimulation of remodeling activity (Figure 3.9a). In particular, modifications present at or near the entry-exit site of the nucleosome that result in significant changes in the electrostatic properties of the modified residue were generally observed to have the strongest effect (e.g. H3Y41ph and H3K56ac). With some exceptions, all remodelers reacted similarly to the presence of these
PTMs. This points to general structural effects that these modifications may have on the nucleosomes, rather than specific interactions with individual remodelers. Experimental evidence to support this is discussed in Section 3.2 and displayed in Figure 3.5

**Figure 3.9 Alteration of Histone-DNA Contacts Affects Remodeling Activity.** a, Modified histone residues in the nucleosome library that lie under the DNA (tan) are highlighted on the nucleosome (PDB: 1KX5) in red. $k_{MN}$ = nucleosome remodeling rate. $k_{unmod} =$ unmodified nucleosome remodeling rate. PTMs are numbered and labeled on the nucleosome structure. Values were capped at -2 and 2 for display purposes. b, Histone mutants present in the nucleosome library that lie under the DNA (tan) are highlighted on the nucleosome (PDB: 1KX5) in red. The heatmap is displayed as in a. Locations of each mutation are individually labeled on the nucleosome structure. Values were capped at -3 and 3 for display purposes. All histones are unmodified unless otherwise specified.
3.5 Summary

In this study, the activity of seven ISWI remodelers against a diverse set of nucleosome modifications were profiled. Critical to this endeavor was the ability to adapt a DNA-barcoded nucleosome library technology to measure the rate of nucleosome sliding by remodelers. This provided quantitative biochemical information (rate constants) while consuming extremely small amounts (picomoles) of precious nucleosome material. From this dataset, important conclusions could be drawn about how histone modifications affect the activity of ISWI family remodelers through differential communication with non-catalytic subunits, with a key insight being that accessory subunits of ISWI remodelers can help a single remodeling ATPase interpret differentially modified nucleosomes.

The data presented provide strong evidence in support of the idea that combinatorial assembly of chromatin remodeling subunits drives functional specialization (Narlikar et al., 2013). For example, while ISWI remodelers may have a universal nucleosome sliding mechanism, distinct subunit compositions may help dedicate this activity to specific genomic loci that are defined by unique histone modification landscapes, thus contributing to the various cellular roles already described for different ISWI complexes. In this regard, histone modifications could function as either a positive or a negative signal. For example, the presence of H3K14ac, which is known to interact with a bromodomain in the NoRC complex subunit TIP5 (Tallant et al., 2015), was discovered to significantly enhance NoRC remodeling. By contrast, crotonylation of this same residue impedes NoRC remodeling, as do modifications on H3K4. Additional biochemical and biophysical studies will be needed to determine whether these regulatory effects (as well as the many other relationships revealed by our data) are driven by simple binding phenomena, or reflect allosteric-type processes. In this sense, the generated dataset exists as a resource that should drive the design of future biochemical and cell-based studies geared towards further understanding ISWI regulation.
CHAPTER 4: THE ACIDIC PATCH AS A DYNAMICALLY REGULATABLE NUCLEOSOME INTERACTION HOTSPOT
4.1 The Nucleosome Acidic Patch is Critical for Nucleosome Remodeling and is Tunable by Histone PTMs

The so-called ‘acidic patch’ on the nucleosome surface has emerged as a key binding hotspot for chromatin effector proteins (McGinty and Tan, 2015). This unique surface feature is created by a cluster of acidic residues located within the core regions of histones H2A and H2B that together create a negatively charged cleft on each face of the nucleosome disc (Figure 4.1a). The central importance of the acidic patch in chromatin recognition is revealed by high-resolution structural studies of chromatin factors bound to the nucleosome core (Armache et al., 2011; Barbera et al., 2006; Kato et al., 2013; Makde et al., 2010; McGinty et al., 2014; Morgan et al., 2016). Indeed, in nearly every structure solved to date, the trans-acting factor engages the nucleosome by inserting an arginine side-chain into a negatively charged pocket defined by a trio of residues in H2A (E61, D90 and D92). This specific mode of engagement has been dubbed the ‘arginine anchor’ (McGinty and Tan, 2015). Given the importance of the acidic patch in chromatin biochemistry, a nucleosome in which the acidic patch was disrupted through mutation was included in the library (Figure 2.8, Table 2.2). Analysis of the library dataset revealed that this nucleosome significantly impeded the ability of all ISWI family remodelers to reposition nucleosomes (Figure 4.1a).
Figure 4.1 The Nucleosome Acidic Patch is Crucial for Remodeling and Regulatable by Histone PTMs. 

**Histone PTMs.**

a. Coulombic surface rendering of the nucleosome (PDB: 1KX5, far right). (Far left) Effect of acidic patch modifications and mutations on ISWI remodeling activity. $k_{unmod.}$ = unmodified nucleosome remodeling rate. Mutations and PTM locations are individually numbered on the nucleosome structure (middle panel) with decimals indicating multiple changes per nucleosome ($7.1 = H2AE56A$, $7.2 = H2BE113A$, $7.3 = H2AE61A$, $7.4 = H2AE64A$, $7.5 = H2AD90A$, $7.6 = H2AE92A$, $7.7 = H2BE105A$, $7.8 = H2AE91A$; $2.1 = N94D$ ($H2A \rightarrow H2A.Z$), $2.2 = K95S$ ($H2A \rightarrow H2A.Z$). Values were capped at -4 and 4 for display purposes. b. E61A, D90A, and E92A mutations in H2A reduce remodeling activity of the ACF complex, CHD4, and BRG1 as read out by a gel-based restriction enzyme accessibility assay (corresponding example replicates shown on right). For a and b all histones are unmodified unless otherwise specified. For b data are represented as the mean of experimental replicates (n=3). Error bars represent ± s.e.m.
In follow-up studies involving the analysis of individual mutant nucleosomes, this effect was narrowed down to the afore-mentioned arginine-binding pocket (Figure 4.1b). These results were also confirmed using a gel mobility shift remodeling assay (Figure 4.2a), which provides an orthogonal and direct readout of remodeling, independent of restriction enzyme cleavage (see Section 2.1 and Figure 2.1 for more information about this type of assay and Section 7.16 for experimental details). Further implicating this region, remodeling of unmodified nucleosomes by the ACF complex was inhibited by addition of the latency-associated nuclear antigen (LANA) peptide from the Kaposi sarcoma-associated herpes virus, which is known to bind the acidic patch using the arginine anchor paradigm (Barbera et al., 2006) (Figure 4.2b).

**Figure 4.2 A Gel Mobility Shift Assay and Inhibition by the LANA Peptide Confirm the Importance of the Acidic Patch.** **a,** Activity of the ACF complex on unmodified and acidic patch mutant nucleosomes. **b,** Remodeling of unmodified nucleosomes is inhibited by the presence of the LANA peptide when compared to a LANA peptide with key binding residues mutated (LRS to AAA). Gel images of example replicates used to generate densitometry measurements in each subpanel are shown above respective graphs. **a** employs a nucleosome repositioning electrophoretic mobility shift assay. **b** employs a restriction enzyme accessibility assay. All histones are unmodified unless otherwise specified. All data are represented as the mean of experimental replicates ± s.e.m. (n = 3).
Given the universal translocase activity possessed by all ATP-dependent chromatin remodelers, in addition to the many other similarities they share (Clapier et al., 2017), it is curious to consider that sensitivity to the acidic patch might also extend to members of other remodeling families. To test this ATPases from the CHD family (CHD4) and SWI/SNF family (BRG1) of remodelers were produced using a baculovirus expression vector system, purified, and confirmed to have ATP-dependent nucleosome remodeling activity (Figure 4.3a, b). Interestingly, both enzymes displayed reduced activity on nucleosomes containing mutations in the acidic patch, suggesting that it may be generally required for chromatin remodeling (Figure 4.1b).

Figure 4.3 Characterization of Recombinant BRG1 and CHD4 Chromatin Remodelers. a, Purified remodelers were run on a 4-20% Mini-PROTEAN® TGX™ gel (Bio-Rad) and run for 35 min at 180 V. Proteins were stained with Coomassie. Expected molecular weights: BRG1: 185 kDa, CHD4: 218 kDa. b, BRG1 and CHD4 remodelers display ATP-dependent nucleosome remodeling activity as detected by a restriction enzyme accessibility assay.

The number of proteins known to bind and utilize the acidic patch in their interaction with chromatin is continually rising (McGinty and Tan, 2015). Nucleosomes themselves act as ligands for the acidic patch; the N-terminus of histone H4 binds the acidic patch in nearby nucleosomes to regulate higher order structure (Luger et al., 1997). Notably, this interaction is disrupted by
acetylation of H4K16, i.e. it is regulated through a PTM (Shogren-Knaak et al., 2006). A variation on this mode of regulation is seen in the case of the high mobility group nucleosomal protein, HMGN2, which also engages the nucleosome acidic patch via a basic motif (Kato et al., 2011). In this case, phosphorylation of serine residues proximal to this motif leads to release of the protein, likely through electrostatic repulsion (Prymakowska-Bosak et al., 2001). Given these examples, it is intriguing to speculate that a reciprocal-type relationship might exist, that is where modifications in the neighborhood of the acidic patch alter the binding of ligands, perhaps in a differential manner. Previous studies hint at this possibility, namely; the histone variant, H2A.Z, which increases the size of the nucleosome acidic patch, enhances remodeling by ISWI remodelers (Goldman et al., 2010), and H2BS112GlcNAc, a modification that is juxtaposed to the acidic patch, stimulates ubiquitylation of histone H2B by the Bre1 E3 ligase (Fujiki et al., 2011).

Analysis of the library dataset reveals that ISWI remodeling activity is indeed affected by a range of diverse histone PTMs located near the acidic patch (Figure 4.1a). Notably, both stimulation and inhibition of nucleosome sliding activity was observed in response to the PTMs (Figure 4.1a). The activity profile of each remodeler was similar across this substrate set, suggesting that a common sensing mechanism was operational across the ISWI remodeler family. It is anticipated such a mechanism is mediated by the ATPase subunit, since it is the only shared component of the complexes. Consistent with this assignment, the isolated SNF2h ATPase subunit was also found to require an intact acidic patch for activity (Figure 4.1a). The acidic patch may act through stabilizing a SNF2h-nucleosome interaction. Currently, ISWI ATPases are known to contact the nucleosome at several locations. The translocase domain binds nucleosomal DNA two helical turns from the dyad axis at SHL2 (Dang and Bartholomew, 2007) and makes additional contacts with the H4 tail (Yan et al., 2016), which is required for activity (Clapier et al., 2002). The C-terminal HSS domain is known to bind extranucleosomal DNA on the opposite side of the nucleosome. Furthermore, all chromatin remodeling ATPases are hypothesized to have a ‘histone-binding domain’ (HBD) that anchors the ATPase to the core region of the histone octamer (Clapier et al., 2017). Such an anchor point would be required for translocation of DNA relative to the histone octamer. Otherwise, the ATPase would simply track
along the double helix without breaking necessary histone-DNA contacts to carry out nucleosome remodeling. For ISWI remodelers, the location of the HBD and where it contacts the nucleosome is unknown. Perhaps, the acidic patch mediates this interaction. To confirm this, follow-up studies combining cross-linking and mass spectrometry are currently being carried out by a Post-Doctoral Researcher in the Muir lab, Hai Dao. Considering known modes of ISWI regulation by nucleosomal epitopes (Clapier and Cairns, 2012), possible allosteric mechanisms can also be envisioned. The H4 tail is required to relieve autoinhibition of the ISWI ATPase, such that it can efficiently carry out nucleosome sliding (Clapier and Cairns, 2012). Perhaps the acidic patch collaborates in this operation, or induces conformation changes in the ATPase that facilitate remodeling by an entirely different means.

Based on the above, the nucleosome library platform was used to investigate the possibility that certain PTMs or histone variants may act as ‘selectivity filters’ for the acidic patch, granting access to subsets of chromatin effectors. The ability of modifications in the unstructured tails of histones to be mimicked using peptide has led to their extensive characterization relative to those that in this histone core regions. The reason for this is that modified nucleosome substrates, which are substantially more difficult to prepare, are required to mimic important aspects of such modifications. For example, native protein secondary structure, or the involvement of neighboring histones or DNA may play a role in chromatin effector recognition of certain PTMs. These scenarios cannot be recapitulated using peptides. As a result, histone tails are well-known nucleosomal regulatory ‘hotspots’. The extent to which the core (‘non-tail’) regions of the histone octamer can serve analogous roles is currently unknown. The high-throughput and physiological nature of DNA-barcoded mononucleosome libraries provide a foundation to rapidly screen the biochemical regulatory effects of such modification.

In addition to ISWI remodelers, modifications were observed to influence CHD4, RCC1, and Sir3, all factors that are demonstrated in this study, or have previously been demonstrated (Armache et al., 2011; Makde et al., 2010), to engage the acidic patch (Figure 4.4a). Each of these factors is sensitive to PTMs within this region of the nucleosome, as predicted by the developed model (Figure 4.4b), and has its own unique characteristic signature of sensitivity. In
addition to those near the acidic patch, library experiments employing CHD4, RCC1, and Sir3 also generated a wealth of biochemical insights related to modifications located throughout the nucleosome that should spur follow-up studies unrelated to ISWI remodeling.

Potential downstream biological effects of acidic patch PTMs could be analogous to the diverse regulatory outputs elicited by modification of histone tails. For example, the N-terminal tail of H3 can be decorated with a range of PTMs that are associated with a variety of, sometimes opposing, biological processes (Bannister and Kouzarides, 2011). *In vivo* validation will be required to determine if this regulatory paradigm extends to the acidic patch. Taking current knowledge into consideration, this is the first biochemical evidence that PTMs in and around the acidic patch can influence chromatin engagement.

**Figure 4.4 Diverse Biochemical Regulation of Acidic Patch Interactors by Nucleosome Modifications.** a, Acidic patch modifications differentially affect remodeling activity and binding of chromatin factors relative to unmodified nucleosomes. b, Model for how histone modifications (yellow triangle, red square) proximal to the acidic patch (pink) might differentially regulate the binding and hence function of chromatin factors. For a all histones are unmodified unless otherwise specified, and data are represented as the mean of experimental replicates (n=3). Error bars represent ± s.e.m.
4.2 Summary

Compared to PTMs within histone N-terminal tails, which have been extensively studied with respect to chromatin regulation, those residing within the core of the histone octamer are much more enigmatic (Tropberger and Schneider, 2013). Thus, it is proposed the acidic patch is able to act as a compositionally dynamic and regulatable interaction hotspot for chromatin effector proteins (Figure 4.3b). By extension, perturbation of this region of the nucleosome either through aberrant PTM installation or mutation could lead to dysregulation of epigenetic processes. Indeed, a mutation in the acidic patch (H2AE56Q) has recently been implicated in human uterine and ovarian carcinomas (Zhao et al., 2016). Follow-up studies should include deciphering the precise mechanisms by which acidic patch modifications (or mutations) influence chromatin effector activity (e.g. binding or remodeling) and connecting these results to biological outputs. This will go hand in hand with development of highly specific ChIP-grade antibodies against core histone marks. Generation of such antibodies can be difficult as histone-derived peptide antigens commonly used cannot replicate the structural environment (e.g. protein secondary structure or neighboring histones or DNA) experienced by such modifications as they exist in chromatin in vivo. Nonetheless, the ability to demonstrate genomic co-localization between the mark and chromatin effector in question will provide support for functional interplay between the two. Lastly, cleverly designed experiments or newly developed assays will be required to uncover how biochemical regulatory mechanisms, generated by the library screen, serve to carry out associated DNA-templated processes in cells.
CHAPTER 5: A DNA-BARCODED NUCLEOSOME LIBRARY AS A GENERAL PLATFORM FOR HIGH-THROUGHPUT CHROMATIN BIOCHEMISTRY—ONGOING PROJECTS
5.1 Introduction

The nucleosome library developed herein serves as a general platform for multiple avenues of chromatin biochemical investigation beyond those of ATP-dependent chromatin remodeling and nucleosome binding (Figure 5.1). For example, this technology has also been applied to study PTM recognition by histone reader domains. The ability to analyze large number of substrates simultaneously was proven to useful to efficiently study the specificity of bromodomains (Nguyen et al., 2014), which tend to recognize sites of acetyllysine somewhat promiscuously (Filippakopoulos et al., 2012). Additionally, nucleosome libraries have been used to study effects of functional crosstalk between histone PTMs using purified writer enzymes in vitro as well as activity present within nuclear lysates (Nguyen et al., 2014). In this case, the activity of writer enzymes is analyzed in response to pre-existing nucleosome modifications by monitoring the efficiency with which the writer is able to modify a nucleosome. Nucleosomes that become modified are then immunoprecipitated using an antibody that recognizes the mark installed. Histone PTMs that tune enzymatic activity (e.g. through affinity or allostery) can be discovered in this way. Studies of nucleosome binding essentially only depend on a quality preparation of the bait protein of interest and the capacity to perform a pull-down experiment to yield biochemical information about bait-nucleosome interactions. On the other hand, analysis of writer enzymes is absolutely dependent on the efficient and specific isolation of modified substrates. In the vast majority of cases this relies on the availability of high quality antibodies that recognize the mark of interest. As the readout for such an experiment is solely sequencing the DNA of the nucleosomes isolated, immunoprecipitation of nucleosomes due to off-target antibody recognition can yield false positives. Thus, antibody-based library experiments must be approached with caution and proper controls put in place to ensure quality results are obtained. In this sense, however, nucleosome libraries can serve as excellent systems to assess the reliability of ChIP-grade antibodies on physiological substrates (Nguyen et al., 2014). Major cataloging efforts in this arena have already been put forth using modified histone peptide arrays and reinforce awareness of important issues such as the difficulty some antibodies have in...
distinguishing different lysine methylation states (Rothbart et al., 2015). However, studying antibody recognition on peptide substrates cannot replicate important scenarios encountered in the context of cellular chromatin, for example, when the mark of interest is embedded within nucleosome secondary/tertiary structure or lies under/adjacent to DNA. Nucleosomes can much better recapitulate the circumstances of antibody-chromatin interactions that occur during ChIP experiments. The remaining sections of this chapter will give an overview of key results obtained from ongoing collaborative projects, which showcase the versatility of the developed nucleosome library to study a range of different chromatin regulatory proteins.
Figure 5.1 Demonstrated Applications of DNA- Barcoded Nucleosome Libraries to Study Chromatin Effector Proteins. To date, DNA-barcoded nucleosome libraries (top, center) have been utilized in the study nucleosome binding/PTM recognition (left, middle), histone writer enzymes/functional cross-talk between PTMs (center, middle), and ATP-dependent chromatin...
remodelers (right, middle). All library-based experiments rely on high-throughput DNA sequencing to generate biochemical information corresponding to interactions between the chromatin effector in question and each individual nucleosome substrate (center, bottom).

5.2 b-JIM, a Novel PWPD-domain Containing Protein Functionally Regulates the Chromosomal Kinase JIL-1 in Drosophila melanogaster—with Catherine Regnard, Peter Becker lab, LMU Munich

JIL-1 is an essential nuclear protein kinase in Drosophila that undergoes autophosphorylation and phosphorylates histone H3 at serine 10 in vitro (Jin et al., 1999; Wang et al., 2001). JIL-1 localizes to interband (euchromatic) regions of Hoechst stained polytene chromosomes, showing very little overlap with banded (DNA dense) regions, suggesting it plays a role in gene transcription via its histone phosphorylation activity (Jin et al., 1999). Consistent with this observation, although JIL-1 is observed to associate with chromosomes throughout the cell cycle it is enriched approximately two-fold on the hypertranscribed Drosophila male X-chromosome, indicating involvement in dosage compensation. It is thought that JIL-1 may partially facilitate transcription by affecting chromatin structure. In addition to its general localization to euchromatin, ectopic recruitment of JIL-1 leads to local chromatin decondensation, and these effects are dependent on JIL-1 kinase activity (Bao et al., 2008). However, the precise role of JIL-1 in transcription remains unclear. Consistent with its previously reported effects on chromatin structure, JIL-1 mutant alleles have been found to be suppressors of position effect variegation (Ebert et al., 2004). JIL-1 deletion also leads to a genome-wide redistribution of H3K9me2 (installed by the Su(var)3-9 methyltransferase) and HP1, which become enriched on both male and female X chromosomes (Zhang et al., 2006). Together, these data imply JIL-1 could play a role in maintaining transcriptionally active regions of the genome by preventing Su(var)3-9/HP1-mediated heterochromatin formation. The role of JIL-1 H3S10ph kinase activity in activating gene expression in response to external stimuli such as heat shock remains controversial (Cai et al., 2008; Ivaldi et al., 2007; Karam et al., 2010).
Catherine Regnard, a scientist in the Becker lab, is interested in JIL-1 specifically as it relates to the process of X chromosome dosage compensation in *Drosophila*, a mechanism to increase gene transcription approximately two-fold on the male X chromosome to match the collective transcription of two X chromosomes in females (Straub and Becker, 2007). High-resolution ChIP-seq studies have shown that JIL-1 has a very strong co-localization with H3K36me3 and use of this histone mark as a targeting motif has been proposed (Regnard et al., 2011). Examination of JIL-1 protein domain organization yielded no evidence of any conserved histone reader domains, leading to the hypothesis JIL-1 is indirectly targeted to H3K36me3 modified chromatin by a protein interacting partner. Interestingly, a protein of approximately 60 kDa quantitatively co-purified with JIL-1 when it was immunoprecipitated from *Drosophila* nuclear extracts, suggesting formation of a complex. This previously uncharacterized protein has been named b-JIM (for JIL-1 Modulator). The ability of these to proteins to form a complex was confirmed by co-expression of FLAG-tagged JIL-1 and b-JIM in Sf21 cells. FLAG affinity purification of JIL-1 also led to the purification of roughly stoichiometric amounts of b-JIM (Figure 5.2). This is further supported by the fact that these to protein show similar localization patterns *in vivo* (Figure 5.3), and was also confirmed by high-resolution ChIP-seq experiments that also revealed a co-occurrence of these two proteins with H3K36me3 (experiments performed by Catherine Regnard, data not shown).
Figure 5.2 JIL-1 and b-JIM Physically Interact to Form a Protein Complex. FLAG-tagged JIL-1 and b-JIM were co-expressed in Sf21 cells. Anti-FLAG immunoprecipitation led to co-purification of b-JIM indicating it forms a complex with JIL-1. (Starting from the left) Lane 1 shows the Sf21 cell lysate. Lane 2 shows the flow-through after incubating the lysate with anti-FLAG affinity beads. Lane 3 shows the beads after eluting bound protein with FLAG peptide. Lane 4 shows bound protein that was eluted from the anti-FLAG affinity beads. All samples were run on an 8% SDS-PAGE gel and stained with Coomassie. Figure courtesy of Catherine Regnard.
Figure 5.3 b-JIM and JIL-1 Co-localize In Vivo. *Drosophila* polytene chromosomes fluorescently stained with DAPI (for DNA) and antibodies directed against b-JIM and JIL-1 show highly similar banding patterns for the two proteins, suggesting co-localization in vivo. Figure courtesy of Catharine Regnard.

How might JIL-1 be targeted to H3K36me3? Interestingly, b-JIM contains a PWPP domain (after the Pro-Trp-Trp-Pro motif). PWPP domains present in other proteins have previously been reported to recognize the H3K36me3 modification (Musselman et al., 2012; Patel, 2016). This prompted an investigation to uncover any histone binding specificity elicited by the b-JIM PWPP domain, while maintaining the overall hypothesis that it mediates local phosphorylation of H3 serine 10 by anchoring the JIL-1 kinase to chromatin. The developed nucleosome library served as an excellent tool to rapidly screen a large number of potential b-JIM nucleosome binding substrates in a massively parallel manner. Pull-down experiments using recombinant b-JIM revealed a high specificity for H3K36me3 nucleosomes, displaying an enrichment of approximately over forty-fold relative their unmodified counterparts (Figure 5.4a). Based on sequence conservation with previously described PWPP domains, a mutant protein was created by replacing key residues in the putative histone binding pocket. A purified recombinant version of this b-JIM mutant lacked the ability to enrich H3K36me3 nucleosomes in a library context, indicating a direct role for the b-JIM PWPP domain in H3K36me3 recognition (Figure 5.4b). As a control, pull-down experiments using the library were also performed with *Drosophila* MSL3, which has a chromodomain known to bind H3K36me3 (Sural et al., 2008). Indeed, this led to a highly specific enrichment of H3K36me3 nucleosomes (approximately...
seventeen-fold) relative to unmodified nucleosomes (Figure 5.4c). The b-JIM:JIL-1 complex also exhibited a preference for binding 
H3K36me3 nucleosomes, in this case resulting in an enrichment of over seventy-fold relative to unmodified nucleosomes (Figure 5.4d). The complex containing the b-JIM PWWP mutant displayed no such enrichment (Figure 5.4e). Therefore, the specific recognition of H3K36me3 by the PWWP domain of b-JIM also occurs in the context of the b-JIM:JIL-1 complex.
Figure 5.4 Pull-down Experiments Using the Nucleosome Library. Results from pull-down experiments against the nucleosome library using recombinant b-JIM (a), b-JIM PWWP mutant (b), MSL3 (c), b-JIM:JIL-1 complex (d), and b-JIM:JIL-1 PWWP mutant complex. Results included all nucleosome library members and are rank-ordered from high to low. Data were normalized to unmodified nucleosome 42 (see Table 2.2). Data corresponding to the H3K36me3 nucleosome is colored in red. Data are a result of experimental replicates (n=3). Error bars represent ± s.e.m.

These results inspired experiments to probe the existence of a positive cross-talk between H3K36me3 and H3S10ph in the presence of the b-JIM:JIL-1 kinase complex. In other words, relative to unmodified nucleosomes, are H3K36me3 nucleosomes better substrates for JIL-1 kinase activity? This was accomplished by incubating the nucleosome library with either the b-JIM:JIL-1 or the b-JIM:JIL-1 PWWP mutant complex in the presence of ATP and allowing
phosphorylation to occur. Substrates that underwent H3S10 phosphorylation were then immunoprecipitated using an anti-H3S10ph antibody. The relative enrichment of each substrate thus correlates with the extent of phosphorylation. Therefore, if recognition of H3K36me3 by the b-JIM PWWP domain recruits bound JIL-1 to drive H3S10ph, the H3K36me3 nucleosome should be enriched over unmodified nucleosomes after H3S10ph antibody (Figure 5.5).

Figure 5.5 A High-throughput In Vitro Kinase Assay Using the Nucleosome Library. The nucleosome library (left) is incubated with the b-JIM:JIL-1 kinase complex (wildtype and PWWP mutant). De novo phosphorylated nucleosomes (center) are isolated via immunoprecipitation with an anti-H3S10ph antibody, and the DNA from this enriched sample is subjected to high-throughput DNA sequencing. Barcode decoding of the resultant sequencing data provides information regarding which nucleosome library members are preferred or undesirable substrates for the kinase complex. H3K36me3 is represented by a green star and H3S10ph is represented by a yellow triangle.

Indeed, the results of the kinase assays demonstrate that out of all members of the nucleosome library, the H3K36me3 nucleosome is the preferred substrate for the b-JIM:JIL-1 complex (Figure 5.6a). This preference was eliminated in the context of the b-JIM:JIL-1 PWWP mutant (Figure 5.6b). Additionally, a pre-modified H3S10ph nucleosome present in the library served as a positive control, ensuring that nucleosomes containing this modification were able to be efficiently isolated.
Figure 5.6 H3K36me3 are Preferred Substrates at H3S10 for the b-JIM-JIL1 Kinase Complex. Kinase assays against the nucleosome library were performed with the b-JIM:JIL-1 (a) and b-JIM:JIL-1 PWWP mutant (b) complexes. Results included all nucleosome library members and are rank-ordered from high to low. Data were normalized to unmodified nucleosome 42 (see Table 2.2). Data corresponding to the H3K36me3 nucleosome is colored in red and data for the pre-modified H3S10ph nucleosome is colored in green. Data are a result of experimental replicates (n=3). Error bars represent ± s.e.m.

The observed preference for phosphorylation by the b-JIM:JIL-1 complex is thus confirmed to be dependent on an intact b-JIM PWWP domain. Enhancement of the b-JIM:JIL-1 kinase activity by the H3K36me3 modification has also been established using conventional radioactive ATP assays on individual nucleosome substrates (personal communication with Catherine Regnard, data not shown). Collectively, these results provide direct biochemical support for the recruitment of the JIL-1 kinase to H3K36me3 modified chromatin by b-JIM to establish/maintain H310ph decorated chromatin domains. Of course, these results are a small piece to a much larger collaborative puzzle surrounding the biology of the newly discovered b-JIM protein and its role in modulating the JIL-1 kinase collectively in gene transcription and dosage compensation in *Drosophila melanogaster*. 
Ubiquitylation of H2B on lysine 120 (H2BK120ub) is widely regarded as a transcriptionally active chromatin mark. This 76-residue protein is installed by the E2 ligase, UBE2A/B, and the heterodimeric RING-type E3 ligase, hRNF20/40, which interact directly with transcription cofactors (Yao et al., 2015). H2BK120ub is further implicated in facilitating gene transcription through its chromatin decompaction properties (Debelouchina et al., 2017; Fierz et al., 2011) and ability to stimulate H3K4 and H3K79 methylation, marks localized to actively transcribed genes (Wang et al., 2008), by the methyltransferases SET1 and Dot1L, respectively (Kim et al., 2009; McGinty et al., 2008; Zhou et al., 2016). The capacity to ‘loosen’ chromatin structure (Debelouchina et al., 2017; Fierz et al., 2011) has also implicated H2BK120ub to function during DNA repair, allowing the responsible enzymatic machinery access and fix lesions in the double-helix (Schwertman et al., 2016). There is also ongoing work in the Muir lab studying how post-translational modification of ubiquitin itself (termed ‘second-tier’ modifications) can further diversify its function once present on chromatin.

Compared to what is known about downstream signaling events (Fuchs et al., 2014; Kim et al., 2009; McGinty et al., 2008; Minsky et al., 2008; Ng et al., 2002; Sun and Allis, 2002) and structural effects (Debelouchina et al., 2017; Fierz et al., 2011) that result from the presence of H2BK120ub, there is little insight into the regulation of its installation. Because the responsible enzymatic machinery acts on nucleosomal substrates it was hypothesized that chromatin modifications may act as inputs to tune H2BK120ub deposition. In order to test this, Felix Wojcik, a post-doctoral fellow in the Muir lab, employed the nucleosome library to study how variations in the nucleosome substrate affect the activity of the UBE2A:RNF20/40 ubiquitylation machinery using a cleverly developed in vitro ubiquitylation assay. This experimental setup utilizes recombinant UBE2A, the RNF20/40 heterodimer, and human E1 activating enzyme. Detection of modified nucleosomes is reliant upon the use of an N-terminally HA-tagged form of ubiquitin such that de novo ubiquitylated substrates can be immunoprecipitated using an anti-HA antibody. DNA
from the isolated sample is subjected to high-throughput sequencing and the abundance of each nucleosome barcode quantified. This number is then compared to that of unmodified nucleosomes to calculate the efficiency in which the nucleosome in question can be ubiquitylated and how any pre-existing modifications might affect this enzymatic process (Figure 5.7).

Figure 5.7 A High-throughput In Vitro Ubiquitylation Assay Using the Nucleosome Library.

The nucleosome library (left) is incubated with the recombinant ubiquitin ligase complex comprised of UBE2A, the RNF20/40 heterodimer, and the E1 activating enzyme in the presence of ubiquitin fused to an affinity tag (in this case, an N-terminal HA tag). De novo ubiquitylated nucleosomes (center) are isolated via an affinity tag-based based pull-down, and the DNA from this enriched sample is subjected to high-throughput DNA sequencing. Barcode decoding of the resultant sequencing data provides information regarding which nucleosome library members are preferred or undesirable substrates for the ubiquitin ligase complex.

The results of the library screen revealed that the UBE2A:RNF20/40 complex is widely sensitive to nucleosome modifications (Figure 5.8). Stimulation and reduction of activity were both observed as well as modifications that had little or no effect on de novo ubiquitylation. Gratifyingly, several controls served to substantiate the resulting dataset. As expected, substrates already bearing the H2BK120ub mark were poor substrates for de novo ubiquitylation. In addition, the H2BS112GlcNAc modification led to a stimulation in activity, consistent with previously published biochemical results (Fujiki et al., 2011). Interestingly, nucleosomes that had either modifications or mutations near the entry/exit site also led to an increase in activity.
Figure 5.8 Analysis of UBE2A:RNF20/40 Enzymatic Activity Against the Nucleosome Library. UBE2A:RNF20/40 ubiquitylation activity analyzed against the nucleosome library following the experimental strategy outlined in Figure 5.6.

Unexpectedly, ubiquitylation was reduced on nucleosomes acetylated on their N-terminal tails, with the largest effects observed when acetylation was localized to K13 and K15. A negative cross-talk was also observed between UBE2A:RNF20/40 and nucleosomes containing the variant H2A.Z. Interestingly, H2A.Z contains a valine in place of K15, and a K15V mutation in H2A led to a decrease in activity similar to that of the H2AK15ac modification (personal communication with Felix Wojcik, data not shown). The reciprocal experiment resulted in derepression of de novo ubiquitylation on H2A.Z nucleosomes. Together these results suggest the importance of the positive charge conveyed by the lysine residue at this position in regulating the activity of the UBE2A:RNF20/40 complex. The importance of this residue in regulating H2BK120ub was also confirmed in a cellular context, taking advantage the fact that glutamine can mimic acetyllysine, which offers a straightforward genetic approach to test the effects of the H2AK15ac modification (personal communication with Felix Wojcik, data not shown). FLAG-tagged H2AK15Q mutant and wild-type H2A histones were transiently expressed in HEK293 cells, chromatin was isolated and digested to mononucleosomes. Immunoprecipitation of FLAG-tagged nucleosomes demonstrated a lower level of H2BK120ub incorporation in those that incorporated H2AK15Q when compared to those having wild-type H2A, demonstrating this mode of biochemical regulation is operative in
cells (personal communication with Felix Wojcik), data not shown). Overall, these results point to
the N-terminal region of H2A (and H2A.Z) as being important for regulation of H2BK120ub.

The question of how reduction in incorporation of H2BK120ub might serve to tune
biological processes led the investigation toward the negative cross-talk between this modification
and H2A.Z. Specifically, the co-localization of H2BK120ub and H2A.Z around transcription start
sites was examined using pre-existing ChIP-seq datasets, with the hypothesis that the genomic
organization of these two modifications on the chromatin template may provide insight into
potential biological regulatory mechanisms at play. This analysis revealed a clear anti-correlation
between H2A.Z and H2BK120ub, with H2A.Z typically found adjacent to gene promoters and
H2BK120ub extending into gene bodies (Figure 5.8). This finding provides evidence supporting
the presence of a negative biochemical cross-talk between these modifications, perhaps in the
sense that H2A.Z limits the incorporation of H2BK120ub at gene promoters. H2BK120ub also
showed significant overlap with H3K79me2, which is consistent with the well-characterized
positive biochemical crosstalk between these two modifications (Figure 5.9) (McGinty et al.,
2008). It is possible that H2BK120ub modification interferes with binding of the transcription
machinery, and the presence of H2A.Z plays a role in establishing the organization of
H2BK120ub over actively transcribed gene bodies. Perhaps H2A acetylation serves as a similar
signal to shut down H2BK120ub in more dynamic processes.
Figure 5.9 ChIP-seq Analysis Proximal to Transcription Start Sites Supports Negative Biochemical Cross-talk between H2A.Z and H2BK120ub. Available ChIP-seq databases were mined to reveal modification patterns surrounding transcription start sites. The H3K4me3 modification (yellow) is a hallmark of active gene promoters. The modification patterns of H2A.Z (red) and H2BK120ub (blue) appear to be mutually exclusive, while the modification patterns of H2BK120ub and H3K79me2 show strong overlap. Figure courtesy of Felix Wojcik and Leslie Beh.

5.4 Histone Acetylation Regulates ADP-ribosylation in DNA Damage—with Glen Liszczak and Katharine Diehl, Tom Muir lab, Princeton University

PARP1 is the founding member of a family of proteins known as poly(ADP-ribose) polymerases (PARPs). PARP1 catalyzes the addition of negatively charged ADP-ribose units to itself and other proteins via the ADP donor NAD⁺. ADP-ribose can be attached as a monomer or as either linear or branched multi-unit polymers—typically on aspartate, glutamate, and lysine residues (Ray Chaudhuri and Nussenzweig, 2017). Individual ADP-ribose or poly(ADP-ribose) (PAR) units can exert regulatory effects by altering protein-protein interactions or serving as a
scaffold for proteins to bind to (Krietsch et al., 2013; Teloni and Altmeyer, 2016). PARP1 is composed of three principal domains—an N-terminal DNA binding domain, a central automodification domain, and a C-terminal catalytic domain, all of which contribute to its primary function in the DNA damage response (DDR) (Ray Chaudhuri and Nussenzweig, 2017). Paramount to its role is the PARylation of itself, histones, and non-histone proteins (Gagne et al., 2008; Isabelle et al., 2010; Jungmichel et al., 2013).

Considering the role of PARP1 in chromatin regulation through the DDR and modification of histone proteins, post-doctoral researchers in the Muir lab, Glen Liszczak and Katharine Diehl, hypothesized that PARP1 activity may be targeted or regulated by signals embedded within its chromatin substrate. This prompted use of the nucleosome library to assess if and how the binding and enzymatic activity of PARP is affected by nucleosome modifications contained within. While binding assays can be carried out via standard pull-down procedures using an affinity tag on recombinant PARP1, analysis of enzymatic activity relies on the use of biotinylated NAD⁺ to enrich de novo modified nucleosome substrates (Figure 5.10) (Zhang, 1997) and the presence of HPF1 to stimulate histone modification activity (Gibbs-Seymour et al., 2016).

**Figure 5.10 A High-throughput In Vitro ADP-ribosylation Assay Using the Nucleosome Library.** The nucleosome library (left) is incubated with recombinant PARP1 and HPF1. De novo ADP-ribosylated nucleosomes (center) are isolated via a streptavidin based pull-down, and the DNA from this enriched sample is subjected to high-throughput DNA sequencing. Barcode decoding of the resultant sequencing data provides information regarding which nucleosome library members are preferred or undesirable substrates for PARP1 enzymatic activity.
Effects on PARP1 binding affinity caused by nucleosome modifications on PARP1 affinity were small (approximately two-fold or less) within the scope of the substrate library (Figure 5.11a). On the other hand, twenty-one nucleosome library members displayed a two to six-fold decrease in their ability to be modified by PARP1 (Figure 5.11b). Interestingly, serine has recently been discovered as modifiable by ADP-ribose (Bonfiglio et al., 2017; Leidecker et al., 2016), and many of these newly characterized modification sites follow a ‘KS’ motif, where the modified serine residue is preceded by a lysine, suggesting importance in PARP1 installation of ADP-ribose on nucleosomes. Both serine 9 and serine 28 follow lysine residues in histone H3 and were reported to be modified (Leidecker et al., 2016). All 15 nucleosomes in the library modified at H3K9 all showed a greater than two-fold decrease in their ability to be modified by PARP1 (Figure 5.11b). The H3K27ac modified nucleosome also resisted modification, however, to a lesser extent. This suggests S10 and S28 are the major sites of modification and native lysine sidechains at K9 and K27 are requisite for efficient PARP1 targeting. Follow-up mutational analysis on individual nucleosomes confirmed S10 and S28 as the major sites of modification on H3, with S10 being the preferred site (Figure 5.12).
Figure 5.11 Analysis of PARP1 Binding and Enzymatic Activity Against the Nucleosome Library

a, PARP1 binding analyzed using pull-down experiments against the nucleosome library. b, PARP1 ADP-ribosylation activity analyzed against the nucleosome library following the experimental strategy outlined in Figure 5.9.
Figure 5.12 S10 and S28 on Histone H3 are the Major Sites of ADP-ribosylation by PARP1.

ADP-ribosylation assays were carried out on individual unmodified (wildtype) or H3S10A,S28A mutant nucleosomes in the presence of PARP1 and HPF1. Anti-pan-ADP-ribose binding reagent (MABE1016, EMD Millipore) was used for detection.

Biological follow-up will include the investigation of the role of SIRT6 in the activation of PARP1 during the DNA damage response (Mao et al., 2011). Of major interest is the H3K9 histone deacetylase activity of SIRT6 (Michishita et al., 2008). The discoveries generated by the nucleosome library based experiments provide the impetus to investigate a potential mechanism for PARP1 activation—that deacetylation of H3K9 by SIRT6 is a prerequisite for ADP-ribosylation of H3S10 and procession of the DNA damage response.
5.5 Summary

All three experimental examples outlined in this chapter demonstrate investigation of enzymatic activities never before analyzed by a nucleosome library based-approach, namely phosphorylation, ubiquitylation, and ADP-ribosylation. Each of these methods relied on a highly efficient means with which to enrich de novo modified substrates from the rest of the library, as off-target enrichment can lead to the appearance of false-positives in the resultant dataset. In the case of the b-JIM:JIL-1 complex, a highly specific anti-H310ph antibody was required in an immunoprecipitation step. To analyze the UBE2A:RNF20/40 heterodimer, installation of an N-terminal HA affinity tag into ubiquitin was crucial to permit use of highly specific anti-HA antibodies for isolation of ubiquitylated nucleosomes. In certain cases, strategies to employ high quality antibodies are unavailable. Examination of ADP-ribosylation activity by PARP1 demonstrates one way to get around this through use of a modified NAD+ with a biotin handle for enrichment of ADP-ribosylated nucleosomes using streptavidin. In other cases, such methods are currently unavailable, and analysis of chromatin effector proteins that fall into this category will be difficult. For example, many anti-methyllysine antibodies show cross-reactivity between the various modification states (Kme1, Kme2, Kme3), which makes specific enrichment of a particular modification state in the presence of others difficult. This is compounded by the fact that many methyltransferases iteratively methylate their substrates which can result in the simultaneous existence of multiple methylation states. For example, NSD2 can both mono and dimethylate H3K36 (Kuo et al., 2011). Distinguishing between these two modification states has important clinical implications as only the dimethyl mark has been linked to the role of NSD2 as an oncoprotein (Kuo et al., 2011). Therefore, analysis of NSD2 dimethylation activity requires a highly specific ChIP anti-H3K36me2 antibody with, ideally, no cross-reactivity with the mono or trimethyllation states at this residue. If such an antibody is unavailable significant assay optimization may be required. Pushing the enzymatic modification to a single methylation state such that there is only a single mark for the antibody to recognize is one potential solution. Analysis of total activity, rather than a sole methylation state, can also be achieved using a
mixture of antibodies such that all modified substrates are isolated together. Substrates of interest can be further assessed by individual methyltransferase assays and the extent of modification clarified by mass spectrometry. At its foundation, a DNA-barcoded nucleosome library-based approach is a high-throughput screening tool, and the identification of false positives and negatives with such experimental strategies are not uncommon. It is up to the researcher to minimize these cases through proper design and understanding of their experimental system. In many cases, significant upfront development may be required before certain types of chromatin regulatory activities can be addressed in a nucleosome library format.

Sections 5.2-5.4 demonstrate that, in moving forward with this technology, separating library construction from platform development and application will significantly accelerate both mechanistic and biological discovery from library-based experiments. It can take years to generate a library, and as evidenced by these three projects (which did not participate in preparation of the library), they explored cell-based work and/or mechanistic biochemistry that wasn’t reached in the analysis of chromatin remodelers within the typical tenure of a Ph. D student (who also had to design and prepare the nucleosome library in addition to applying it experimentally). Dedicated personnel for library synthesis will allow significant reduction in the bias of library experiments by allowing uninterrupted focus on histone preparation to increase library diversity. These same researchers would also concentrate on library design and construction, and their resultant familiarity with the platform will permit efficiency in fielding collaborative requests from other laboratories in the field. When use of nucleosome library is the starting point, this allows more time for engagement in assay development, application, and detailed biochemical and biological follow-up. This is important as, in the author’s opinion, the most impressive accomplishments yet to come will not be demonstration of the technology but the new discoveries it will make that can be validated as bonafide biological phenomena.
CHAPTER 6: PERSPECTIVES
6.1 Introduction

The purpose of the DNA-barcoded nucleosome library constructed was to serve as a screening tool to accelerate discovery of novel aspects of chromatin regulation with respect to nucleosome modifications. In a single study, such analyses generate far too much information to exhaustively follow up every measurement fully. Of the 784 rate nucleosome remodeling rate constants generated by examination of the ISWI family of chromatin remodelers, only a small number that were considered to be the 'most interesting' were investigated further. Even then, the level of mechanistic detail these follow-up experiments involved was quite rudimentary. Therefore, many questions left to be answered regarding precisely how certain PTMs exert their observed effects, both biochemically and, potentially, in vivo. This, of course, includes all library-based measurements made regarding the remodeling activity of CHD4 and the nucleosome binding properties of RCC1 and the BAH domain of Sir3. The remaining sections of this chapter will discuss follow-up experiments founded upon the library screen that were not able to be addressed due to time constraints. Ideas concerning key future developments and applications of DNA-barcoded nucleosome libraries will also be considered.

6.2 ISWI Remodeling Dataset—Other Considerations

Positive and negative effects on chromatin remodeling by nucleosome modifications can arise from multiple sources. For example, they can affect the on or off rate of the remodeler (affinity), DNA translocation efficiency (activity), or the ability of a remodeler to release its substrate (turnover). The mechanistic basis for such outcomes was left unexplored in this study, and thus provides motivation for related follow-up studies.

To begin to address these questions, series of pull-down experiments using the nucleosome library should be performed, generating a binding profile for each remodeler to gain an understanding of how nucleosome modifications affect their relative affinities for each substrate. It will then be useful to plot this data against all the remodeling rate data for each
substrate to identify any correlations between the two datasets. Any positive correlations found in such an analysis would point to remodeler binding effects as being the driving force behind the observed changes in nucleosome sliding activity (i.e. increased remodeling rate correlates with increased affinity or decreased remodeling rate correlates with decreased affinity. Classical biochemical approaches (e.g. electrophoretic mobility shift assays) could then be used to determine dissociation constants ($K_d$) for nucleosomes of interest. Nucleosomes that increase or decrease remodeler activity without altering its affinity could potentially act through allosteric means or perhaps their effects are realized through structural changes to the nucleosome that make DNA translocation more or less difficult (e.g. by altering histone-DNA contacts). Lastly, increased remodeler affinity may anticorrelate with activity. In this case, it is possible that tightly bound remodelers have difficulty releasing their nucleosomal substrate, resulting in reduced efficiency of the remodeling reaction.

Even without having performed these experiments (which were deemed beyond the scope of this thesis), careful analysis of the ISWI remodeling dataset points to several scenarios where a positive correlation between affinity and activity may describe changes in remodeler activity. As discussed in Section 3.3, the TIP5 subunit of the NoRC complex contains a tandem PHD-finger bromodomain module that has been previously reported to display a preference for the H3K4me0K14ac modification state of the N-terminal tail of H3 (Tallant et al., 2015). The interaction with H3K4me0 is mediated by the PHD finger, while the bromodomain binds H3K14ac. Because this configuration of H3 shows a positive correlation only for NoRC activity, perhaps an increase in affinity, mediated by these modifications, is responsible for increased remodeling activity of this enzyme. This should be confirmed by the strategy outlined above and also by rationally designed mutations in the PHD-finger and bromodomain of TIP5 to address whether or not these reader domains play a direct role. In this case, an increase in remodeler binding would be expected to concomitantly increase its processivity, and, thus, overall remodeling rate.

An additional situation where this mechanism might be at play concerns the WICH complex, and is purely based upon the author's own speculation and not previously published information. Interestingly, all nucleosomes modified at H3K4, as well as the H3R2me nucleosome
show a decrease in activity (Figure 6.1a). This suggests that the presence of an unmodified lysine at that position is important for engagement of nucleosomes by the WICH complex, and the fact that this trend is not observed for SNF2h alone (Figure 6.1b) points to the WSTF subunit as being responsible for this effect. Primary sequence analysis of the WSTF PHD-finger reveals that it does not possess the conserved residues needed to form the ‘aromatic cage’ typically present in reader domains that recognize higher lysine methylation states (Sanchez and Zhou, 2011). Therefore, it is reasonable to postulate that recognition of H3K4me0 by the WSTF PHD-finger plays a role in chromatin remodeling, and that modification of this residue disrupts this interaction to decrease WICH affinity for such nucleosomes. Such a decrease in binding would be expected to result in the observed decreases in remodeling on H3K4 modified nucleosomes, and, again, could be explored using the experimental workflow described above. Pending these results, studies on the binding preference of the isolated WSTF PHD-finger for a series of H3K4 modified histone peptides would also be informative.
Figure 6.1 Unmodified H3K4 is Important for WICH Remodeling. Library remodeling data for H3R2me and all H3K4 modified nucleosomes corresponding to experiments with the WICH complex (a) and the SNF2h ATPase (b). $k_{\text{MN}}$ = nucleosome remodeling rate. $k_{\text{unmod.}}$ = unmodified nucleosome remodeling rate. All data are represented as the mean of experimental replicates (n=3). Error bars represent ± s.e.m. All histones are unmodified unless otherwise specified.
Nucleosomes modifications that result in increases or decreases in remodeling activity without altering the affinity of the remodeler are likely exert their effects in one of two ways. First, they may exert their effects allosterically, inducing conformational changes in the remodeler more or less conducive to remodeling. Such a mechanism has already been described, where the ISWI ATPase recognizes the unmodified H4 tail to relieve an autoinhibitory state (Clapier and Cairns, 2012). As this study and others have shown, modification of the H4 tail can affect this switch (Shogren-Knaak et al., 2006). Second, modifications that introduce changes in nucleosome structure may also affect efficiency of the universal DNA translocation mechanism. It has been previously demonstrated that the octamer undergoes conformational distortions during remodeling by ISWI remodelers (Sinha et al., 2017). Therefore, it is entirely plausible to speculate that modifications on the core regions of the histone octamer that disturb local histone structure may modify the ability of the nucleosome to undergo these transitions.

Perhaps a more intuitive manner by which histone modifications may affect DNA translocation is by perturbing histone-DNA contacts (discussed in Section 3.4). By lowering the number of interactions that must be broken and reformed by the remodeler, this class of modification is hypothesized to reduce the energy barrier to remodeling. Many of these modifications have already been biochemically or biophysically characterized to affect interactions between histones and DNA (Bowman and Poirier, 2015; Ferreira et al., 2007b; Muthurajan et al., 2004), and, thus, their overall effect on nucleosome remodeling supports this hypothesis. However, whether they change how ISWI remodelers slide nucleosomes is unknown. The sequence of DNA translocation steps taken by of ISWI remodelers has been worked out in great detail (Deindl et al., 2013) and is discussed in Section 1.3.1 and Figure 1.7. ISWI remodelers must first draw seven base pairs of DNA out of the nucleosome, which is thought to generate the tension required to begin draw DNA from the other side of the nucleosome in three base pair steps, pumping it around the octamer. If these step sizes are dictated by a precise energy landscape on the octamer surface it is possible they could be changed by PTMs that alter this landscape. On the other hand, perhaps step sizes do not change, and the lower energy requirements for DNA translocation to occur simply make individual steps easier to complete,
reducing the dwell time between each step. Pauses between steps are significantly longer than the translocation steps, and thus play a greater role in limiting the overall rate (Deindl et al., 2013). A study employing the single-molecule FRET technology developed in the laboratory of Xiaowei Zhuang to examine DNA translocation by ISWI remodelers containing modifications or mutations (e.g. H3Y41ph, H3K56ac, H3T118H, H4R45A) at various locations under the DNA (e.g. entry-exit site or dyad axis) could prove informative.

An anti-correlation between remodeler affinity and activity is also possible and is one potential scenario that surrounds the NURF complex discussed in Section 3.3 and Figure 3.7c. BPTF, the largest subunit in the NURF complex, possesses a PHD-finger and bromodomain and that recognize H3K4me3 and H4K16ac, respectively, in a bivalent manner (Ruthenburg et al., 2011). The library dataset revealed that nucleosomes bearing these modifications were poor substrates for NURF remodeling relative to their unmodified counterparts. This result is counterintuitive as it had initially been thought that dual recognition of these marks served as a signal to recruit NURF remodeling activity to chromatin (Ruthenburg et al., 2011). Inhibition of NURF remodeling by H3K4me3 and H4K16ac would make sense if recognition of these PTMs boosted NURF complex affinity such that it had difficult releasing its nucleosome substrate following DNA translocation. Assuming such effects on enzyme turnover, the overall rate of individual H3K4me3:H4K16ac nucleosome remodeling would not be expected to change, but the rate at which a nucleosome population as a whole was remodeled would be expected to be slower when compared to unmodified nucleosomes. This could be tested using single-molecule FRET strategies mentioned above (Deindl et al., 2013).

6.3 Considerations Regarding Streamlined Library Preparation

Given that virtually all components of a DNA-barcoded nucleosome library must be individually prepared by hand, library construction demands an enormous amount of time spent at the bench. Therefore, the opportunities to automate or streamline experimental procedures would be welcomed at any step of assembly (e.g. preparation of barcoded DNA, histones, octamers,
nucleosomes, or assay work up). Development of experimental strategies to parallelize even small parts of these stages could have significant effects in accelerating library synthesis. This is, in part, due to the fact library construction requires an immense amount of organization and focus on sample handling for efficient sample tracking. An opportunity therefore exists for development of an automated platform (e.g. a sample handling robot perhaps interfaced with a microfluidics device) in library assembly. Such an engineering innovation would, allow the investigator more time to focus on experimental design and application.

There several instances where the approach to obtaining barcoded DNA could benefit from the use of automation. Individual barcoded oligos could be purchased, solubilized, and stored in 96- or 384-well plates. As described in section 2.4.2, perhaps oligo pairing for hybridization, ligation to the 601 sequence, and parallel sample purification could all be automated. Commercial liquid-handling robots can perform plate-based nucleic acid purifications using bead-based strategies as well as prepare samples for standard molecular biological assays such as PCR, suggesting the potential for barcoded DNA-to-601 DNA ligations. Purified samples could be accurately quantified using a plate-based UV spectrophotometer for subsequent use in nucleosome assemblies.

Octamer and nucleosome assembly using a liquid-handling robot may be difficult for several reasons. Octamer assembly begins by unfolding individual histones in buffered 6 M guanidine hydrochloride (GndHCl). The high concentration of GndHCl and dissolved histone renders the solution very viscous and its pipetting requires extreme care in human hands. Whether a robot would be able to accurately handle samples at this step is unknown and should be tested empirically. Octamer refolding requires combination of core histones in equimolar amounts. A lack of precision at this step could result in less efficient octamer formation. As octamers used in library construction are unpurified, this can lead to subsequent nucleosome 'underassembly'. The reason for this is that when octamer concentrations are calculated, total histone content is quantified by UV spectroscopy, and the relative ratios of octamer:tetramer:dimer in the mixture are unknown. If there is not enough octamer present to saturate the 601 DNA in nucleosome assembly, there will be a larger than desirable amount of
free DNA left over. The presence of free DNA can be problematic in certain types of library-based assays. While there are ways around it, it affects nucleosome remodeling assays dependent on restriction enzyme accessibility as discussed in Section 3.1. In addition, pull-down assays can be affected if the chromatin effector being analyzed binds free DNA. In this type of binding experiment, the amount of the barcoded DNA isolated by the protein of interest is thus a combination of nucleosomal DNA and free DNA. Therefore, relative binding affinities nucleosomes that have a significant amount of free DNA may be difficult to calculate. The inclusion of uniquely barcoded non-nucleosomal DNA in nucleosome libraries can provide information about how effectively free DNA is being bound and tell the researcher if experimental conditions need to be adjusted to reduce these effects. Increasing salt concentration or the addition of divalent metal cations such as Mg$^{2+}$ are several ways to accomplish this. To avoid nucleosome undersassembly, individual ratio tests could be performed for several octamer preparations and the results extrapolated to the rest of the octamers in the assembly. However, carrying out ratio tests for each individual member would undermine the high-throughput nature of this workflow. For the same reasons described above, liquid-handling robots may not be able to accurately prepare nucleosome assembly mixtures due to the high viscosity of octamer solutions—which contain 50% glycerol. Even if this step must be carried out ‘offline’, removal of biotinylated MMTV DNA used in nucleosome assembly (see Section 2.4.3 and Figure 2.11) via streptavidin pull-down would likely benefit from bead-handling capabilities of certain robotic systems. In addition, many of the assay work-up steps involving nucleic acid purification and PCR preparation for multiplex barcoding could also be automated. In the meantime, moving any steps mentioned above to multi-well plates and use of a multi-channel pipette and ‘traditional manpower’ rather than an expensive robotic system would be a noteworthy advancement.

One major advantage of nucleosome library-based approaches is their sensitivity and, thus, the small amount of material required to prepare them. Therefore, consideration of dead-volume requirements present in liquid-handling systems and the amount of extra material needed to satisfy those requirements would be a significant consideration and potential drawback.

Advances in automation of microfluidic device technology could provide an alternative solution.
As the fields of microfluidics and chromatin biochemistry require very diverse sets of skills, such a project will have the highest chance of success a collaboration between experts, ensuring device fabrication meets experimental needs and vice versa. Reflecting on the fact that the nucleosome library prepared in this study contains a larger number of substrates than some histone peptide libraries analyzed is quite impressive (Muller and Muir, 2015). However, if library sizes are to rival the diversity of larger peptide-based studies on the scale of multiple hundreds or even one thousand, parallelization and automation will almost certainly need to be implemented.

The current library utilized 81 uniquely modified histone proteins to generate 110 distinct nucleosomes and encompasses the most well-studied histone PTMs. However, it is not exhaustive in its coverage of chromatin modification states in vivo. For example, many nucleosome members only contained a single modification. Histone PTM analysis by mass spectrometry demonstrates that many histones contain multiple modifications simultaneously (Young et al., 2009). Numerous ChIP-seq studies also report co-localization of many different histone marks at certain functional genomic elements, providing the potential for them to co-exist on the same nucleosome (Bannister and Kouzarides, 2011). In future iterations of the nucleosome library, in vivo-relevant combinatorially modified histones should be prepared as well nucleosomes that ‘represent’ various genomic features. For example, a nucleosome located at an active gene promoter may contain both H3K4me3 and H2A.Z, while a nucleosome located in an active gene body may contain H3K79me2, H3K36me3, and H2BK120ub. Library screens would benefit from having nucleosomes that represent such major regions of the genome, for example, constitutive heterochromatin, facultative heterochromatin, telomeric chromatin, pericentric heterochromatin, or chromatin at gene enhancers. How a protein behaves in a comprehensive chromatin environment, rather than in response to a single PTM, could provide more detailed insight into its regulation in cells. A library larger than several hundred members would likely be able to capture this.
6.4 DNA-barcoded Mononucleosome Libraries—Untapped Applications

6.4.1 Profiling Histone Eraser Activity

To date, chromatin remodelers, histone readers, and histone writers have been studied using DNA-barcoded nucleosome libraries (this study and Nguyen et al., 2014). Eraser activity, displayed by enzymes that remove covalent modifications from histones, remains to be explored. The library prepared herein, was initially designed to contain a very large number of acetyl marks as, a Post-Doctoral Researcher in the lab, Glen Liszczak had planned to probe the deacetylase activity of the NuRD complex (Lai and Wade, 2011). As these experiments have yet to be carried out, their design and how nucleosome libraries might be used to generally profile eraser activities will be discussed.

Measurement of histone deacetylase activity requires the ability to detect nucleosome substrates on which deacetylation has occurred. This could potentially be accomplished by performing a typical pull-down experiment with the library using a pan-acetyl antibody to create a reference data set (Figure 6.2). In a separate experiment, the library would be incubated with a histone deacetylase, deacetylation will be allowed to occur, and an additional pull-down experiment using the same pan-acetyl antibody subsequently performed. After immunoprecipitation, nucleosomes that are deacetylated will appear depleted when compared to the reference dataset. In theory, this workflow could be applied to other eraser activities, as long as suitable antibodies are available. This type of experiment, of course, applies to scenarios where the amino acids targeted by the eraser enzyme are unknown. Site-specific antibodies could also be used if examination of activity at a particular residue is desired. Experiments of this type, however, would require careful planning in library composition for the specific enzyme in question. In order to be informative, the target modification should be combined in the presence of a large variety of other PTMs. Therefore, designing libraries applicable to many different site-specific eraser activities would need to be very large in size.
Figure 6.2 A High-throughput *In Vitro* Deacetylase Assay Using the Nucleosome Library.

The nucleosome library (left) is incubated with a recombinant histone deacetylase. After deacetylation occurs (center), nucleosomes that remain acetylated are isolated via a pan-acetyl antibody-based pull-down, and the DNA from this enriched sample is subjected to high-throughput DNA sequencing. Barcode decoding of the resultant sequencing data and comparison of the data to an identical experiment where deacetylation has not occurred provides information regarding which nucleosome library members are targets of deacetylase activity.

### 6.4.2 Profiling H2A.Z Incorporation by the hSRCAP Complex

As discussed in section 1.3.4, the hSRCAP complex catalyzes the exchange of H2A-H2B dimers for those containing H2A.Z (Mizuguchi et al., 2004; Papamichos-Chronakis et al., 2011). The fact that the modification introduced to the nucleosome is a protein makes the ability to isolate nucleosomes using a peptide affinity tag (analogous to the strategy outlined in Section 5.3 for *in vitro* ubiquitylation) very attractive. H3K56ac is already known to alter the substrate specificity of the yeast homolog of hSRCAP, SWR1, by reducing H2A.Z incorporation efficiency (Watanabe et al., 2013). However, the extent to which H2A.Z incorporation is regulated by other modifications is unclear. In theory, this could easily be explored using the nucleosome library prepared in this study (Figure 6.3). Incubation of the library in the presence of the hSRCAP complex, ATP, and excess H2A.Z-H2B dimers with HA-tagged H2A.Z would allow efficient isolation of nucleosomes that have undergone dimer exchange via immunoprecipitation with anti-HA antibodies. Rates of H2A.Z incorporation could be calculated for each individual substrate.
over a time course and compared to determine how pre-existing nucleosome modifications influence hSRCAP activity. Such a dataset could provide insight into possible chromatin-mediated mechanisms of regulating H2A.Z incorporation across the genome.

Figure 6.3 A High-throughput In Vitro H2A.Z Incorporation Assay Using the Nucleosome Library. The nucleosome library is incubated with recombinant hSRCAP complex and H2A.Z-H2B dimers with HA-tagged H2A.Z in the presence of ATP (left). After dimer exchange occurs (center), nucleosomes that with newly incorporated H2A.Z are isolated HA-tag immunoprecipitation, and the DNA from this enriched sample is subjected to high-throughput DNA sequencing. Barcode decoding of the resultant sequencing data provides information regarding which nucleosome library members are targets for H2A.Z incorporation by the hSRCAP complex.

6.5 Expanding DNA-barcoded Chromatin Library Substrate Complexity

The ability to rival the scope of peptide-based studies using DNA-barcoded nucleosome libraries represents a significant achievement in the field of chromatin biochemistry. In many ways, nucleosomes are far superior than histone-derived peptides in the study of chromatin effector proteins. They represent the most fundamental unit of chromatin—a complete histone octamer composed of two copies each of histones H2A, H2B, H3, and H4 in complex with DNA. In this way, they are able replicate the many structural and steric constraints imposed by chromatinization of eukaryotic genomes. As a consequence, aspects of chromatin regulatory proteins that involve recognition of histone secondary structure, multiple histones, nucleosomal
DNA, extranucleosomal DNA, or any of these features simultaneously can be probed using nucleosomes. Chromatin remodelers, by their very nature, require a nucleosome as their substrate, and exemplify the advantages provided by using a more physiological substrate.

However, chromatin, in a biological sense, is indeed represented by repeating array of nucleosomes, and thus generation of libraries using multi-nucleosome substrates should be considered. The need for such substrates is underscored by existence of multi-megadalton (MDa) complexes known to operate on chromatin. For example, the mammalian BAF complex has a molecular weight of 2-MDa, approximately twelve-times that of the nucleosome. This size is consistent with the fact that its calculated average genomic footprint is approximately two to four kilobases (Kadoch and Crabtree, 2015). Using a two-hundred base pair nucleosome repeat length, this corresponds to arrays of 10-20 nucleosomes. Although BAF complexes can remodel mononucleosome substrates (Wang et al., 1996), the influence of adjacent nucleosomes must be taken into account to more closely replicate what occurs in vivo. In the case of many chromatin effector proteins, it is possible that signaling inputs from multiple nucleosomal neighbors could easily override those provided by the nucleosome being acted upon. These regulatory aspects are impossible to dissect using the mononucleosome library developed in this study.

Preparation of multi-nucleosome substrates must retain as many aspects of streamlined assembly present for mononucleosomes as possible. This includes requiring small material amounts and not having to individually purify octamers or perform ratio tests. Therefore, the difficulty of preparation will likely be the limiting factor in the size of multi-nucleosome substrate used. At present, twelve-nucleosome arrays can only be manufactured using purified histone octamers, and DNA:octamer ratios must be individually optimized for each preparation (Debelouchina et al., 2017). As many more pieces must come together to yield a homogenous final product, substrates of this scale are thus currently out of reach for high-throughput assembly. Therefore, it would be informative to apply the high-throughput assembly strategy developed for mononucleosomes (see Section 2.4.3 and Figure 2.11) to smaller multinucleosome substrates such as di, tri, or tetranucleosome arrays. If these substrates can be prepared using an identical procedure, or one requiring minimal optimization, it would represent an important step
forward in DNA-barcodechromatin library technology. Of course, increasing the physiological nature of DNA-barcodechromatin libraries introduces new technical considerations, a significant one being potential formation of higher-order chromatin structures. Taking this into account is especially important using assays that require Mg$^{2+}$ (e.g. kinase or restriction enzyme accessibility chromatin remodeling assays), as the ability of this divalent metal cation to induce chromatin compaction is well-documented (Shogren-Knaak et al., 2006). Nonetheless, libraries of this type will allow chromatin biochemistries to be explored on templates more representative of their counterparts in vivo.

6.6 Summary

As outlined in this chapter, this entire body of work provides a starting point for in-depth follow-up studies, its major focus to survey of the interplay between chromatin modifications and remodeling activity. The discoveries made point to exciting new avenues of investigation related to ISWI remodelers. Principally, the data not only establish that chromatin modifications can affect remodeling rates, but that accessory subunits within the complexes mediate these effects. Furthermore, the arginine-anchor region of the acidic patch was discovered to be essential for chromatin remodeling and that modifications in and around this 'hotspot' can regulate activity – a finding shown to extend to other chromatin binding factors. Following up detailed mechanistic aspects of these findings will be a key focus of future efforts in the Muir laboratory, and, presumably, others as well.

It must be stressed that it is currently not possible to study the direct biochemical relationships underlying the regulation of chromatin remodeling in a cell, which absolutely necessitates quantitative biochemical studies of this type. Much remains to be explored with respect to development and application of DNA-barcodechromatin libraries. Increasing substrate number and diversity to challenge that present in vivo will require major advances in process development and likely involve the incorporation robotic automation. In addition, expanding substrate complexity will be important to uncover chromatin regulatory mechanisms
that cannot be addressed using mononucleosomes. With these platforms in place, it is ultimately up to the imagination of innovative researchers to develop novel applications of DNA-barcoded chromatin libraries, in hopes to provide both a launch pad for new discoveries and a foundation of mechanistic understanding for the complex phenomena observed in cell-based studies.
CHAPTER 7: METHODS
7.1 Analytical Methods

Analytical reversed-phase HPLC (RP-HPLC) was performed on an Agilent 1100 series instrument employing a Vydac C18 column (5 μm, 4 × 150 mm) at a flow rate of 1 mL/min. Semi-preparative scale purifications were performed on an Agilent 1100 series instrument employing a Vydac C18 semipreparative column (12 μm, 10 mm × 250 mm) at a flow rate of 4 mL/min. Preparative scale purifications were conducted on a Waters DeltaPrep 4000 system equipped with a Waters 486 tunable detector (Waters). A Vydac C18 preparative column (15–20 μm, 20 × 250 mm) was used at a flow rate of 20 mL/min. 0.1% TFA in water (HPLC solvent A) and 90% acetonitrile, 0.1% TFA in water (HPLC solvent B) were used as the mobile phases in all RP-HPLC analyses and purifications. ESI-MS analysis was conducted on a MicrOTOF-Q II ESI-Qq-TOF mass spectrometer (Bruker Daltonics).

7.2 Production and Purification of RCC1 and Sir3 Constructs

Full-length RCC1 (Uniprot ID: P18754) with an N-terminal 6xHis-SUMO tag and a C-terminal GST tag and the BAH domain of S. cerevisiae Sir3 (Armache et al., 2011) with a C-terminal FLAG-6xHis tag were subcloned into bacterial expression plasmids (pET, Novagen) for protein production in E. coli. In brief, BL21 Rosetta™ (DE3) cells were transfected with expression plasmids and grown in LB medium at 37 °C until reaching an OD$_{600}$ of 0.6. For the 6xHis-SUMO-RCC1-GST construct, protein expression was induced by the addition of 0.6 mM IPTG overnight at 18 °C. For the Sir3-BAH-FLAG-6xHis construct, protein expression was induced by the addition of 0.6 mM IPTG for 1h at 37 °C. Cells were harvested by centrifugation at 4,000 x g for 10 minutes at 4 °C, and cell pellets were washed twice with 10 mL of cold PBS per L of culture. Cells were lysed in 10 mL of lysis buffer (40 mM Tris, 1 M NaCl, 5 mM β-mercaptoethanol, 10% glycerol, pH 7.7 at 4 °C) per L of culture by passage through an EmulsiFlex-C3 homogenizer (Avestin). Each protein was initially purified by Ni-NTA affinity chromatography. The 6xHis-SUMO tag was cleaved from the RCC1 construct by Ulp1 protease,
and removed by reverse Ni-NTA affinity chromatography. The RCC1-GST and Sir3-BAH-FLAG-6xHis constructs were further purified over Superdex 200 10/300 GL and Superdex 75 10/300 GL gel filtration columns (GE Healthcare) equilibrated with gel filtration buffer (50 mM Tris, 100 mM NaCl, 2 mM tris(2-carboxyethyl)phosphine, 10% glycerol), respectively, using an AKTA FPLC system (GE Healthcare) equipped with a P-920 pump and UPC-900 monitor prior to use in subsequent experiments.

7.3 Production and Purification of Recombinant Histones

Unmodified recombinant human histones (H2A, Uniprot ID: Q6FI13; H2B, Uniprot ID: O60814; H3C96A, C110A, Uniprot ID: P68431; H4, Uniprot ID: P62805), histone variants, and histone mutants were produced in and purified from E. coli. In brief, BL21 Rosetta™ (DE3) cells were transfected with histone expression plasmids (pET, Novagen) and grown in LB medium at 37 °C until reaching an OD_{600} of 0.6. Protein expression was induced by the addition of 0.6 mM IPTG for 2-3 hours at 37 °C. Cells were harvested by centrifugation at 4,000 x g for 10 minutes at 4 °C, and cell pellets were washed twice with 10 mL of cold PBS per L of culture. Cell pellets were resuspended in 10 mL of cold lysis buffer (50 mM Tris, 100 mM NaCl, 1 mM EDTA, 1 mM 2-mercaptoethanol, pH 7.6 at 4 °C) with cOmplete™, EDTA-free Protease Inhibitor Cocktail (Sigma-Aldrich) per L of culture and homogenized by passage several times through an 18-gauge needle. Cells were lysed via sonication and the resulting suspension was centrifuged at 30,000 x g for 30 minutes at 4 °C. The inclusion body pellet was washed twice with cold lysis buffer containing 1% Triton-X 100 and once without detergent. Inclusion body pellets were resuspended in 10 mL of inclusion body resuspension buffer (6 M guanidine hydrochloride, 20 mM Tris, 1 mM EDTA, 100 mM NaCl, 1 mM DTT, pH 7.5 at 4 °C) per L of culture, and mutated at 4 °C for 2 hours. Resuspensions were then centrifuged at 30,000 x g for 30 minutes at 4 °C, and the supernatants were transferred to 3.5 kDa MW cutoff dialysis tubing and dialyzed overnight in 2 L of low salt urea buffer (7 M urea, 10 mM Tris, 1 mM EDTA, 100 mM NaCl, 1 mM DTT, pH 7.5 at 4 °C). Resuspensions were dialyzed an additional 2 hours prior to loading onto a HiTrap SP HP 5
mL column (GE Healthcare) equilibrated with low salt urea buffer (filtered and degassed). The
histones were purified using a gradient over 20 column volumes from low salt urea buffer to high
salt urea buffer (7 M urea, 10 mM Tris, 1 mM EDTA, 1 M NaCl, 1 mM DTT, pH 7.5 at 4 °C) on an
AKTA FPLC system from GE Healthcare equipped with a P-920 pump and UPC-900 monitor.
Histones were further purified using preparative C-18 RP-HPLC. Purified histones were analyzed
by analytical RP-HPLC and ESI-MS (Figure 2.7).

7.4 Production of Modified Histones by Protein Semisynthesis

H2AK119ub, H2BK120ub, H3K4me3, H3K9ac, H3K9me3, H3K14ac, H3K18ac,
H3K18acK23ac, H3K23ac, H3K27ac, H3K27me3, H3KpolyAc, H4K5ac, H4K8ac, H4K12ac,
H4K16ac, H4K20ac, and H4KpolyAc histones were prepared by semisynthesis as described
previously (Nguyen et al., 2014). The semisynthesis of the H3R42me2a histone was conducted
as described previously (Casadio et al., 2013).

7.4.1 Preparation of H2AK5ac, H2AK9ac, H2AK13ac, H2AK15ac, and H2AKpolyAc Histones

Proteins were assembled from two pieces: A recombinant fragment was generated
encompassing residues 21-129 of histone H2A with an A21C mutation fused to an N-terminal
6xHis-SUMO tag. This sequence was inserted into a pET expression plasmid and confirmed by
gene sequencing. BL21 Rosetta™ (DE3) cells were transfected with the plasmid and grown in LB
medium at 37 °C until reaching an OD600 of 0.6. Protein expression was induced by the addition of
0.6 mM IPTG for 2-3 hours. Cells were harvested, lysed, and inclusion bodies were prepared
identically as for recombinant histones in this study. The 6xHis-SUMO-H2A(21-129)A21C protein
was purified by Ni-NTA affinity chromatography. The sample was dialyzed for 2 hours in 2 M
urea, 50 mM Tris, 200 mM NaCl, 1 mM DTT, pH 7.4 at 4 °C before being moved to an otherwise
identical buffer containing 1.5 M urea. Ulp1 protease was added to cleave off the 6xHis-SUMO
and the sample was dialyzed overnight at 4 °C. The resulting H2A(21-129)A21C fragment was
purified using preparative C-18 RP-HPLC. The final product was characterized by ESI-MS and analytical C-18 RP-HPLC.

H2AK5ac, H2AK9ac, H2AK13ac, H2AK15ac, and H2AKpolyAc peptides corresponding to residues 1-20 with acetylated lysine residues at the indicated positions were synthesized as follows: approximately 500-750 mg of Trityl-OH resin (ChemMatrix) resuspended with 10 mL of dichloromethane (DCM) and chlorinated by reaction with 200 µL of SOCl₂ on a shaker overnight. The resin was washed thoroughly with DCM and dimethylformamide (DMF) followed by an additional wash with 5% (v/v) N,N-diisopropylethylamine (DIEA). 500 mg of chlorinated resin was combined with 4-fold excess (relative to resin loading) of hydrazine monohydrate in 2 mL DMF along with an 8-fold excess of DIEA. The reaction was stirred on ice for 1 hour and subsequently washed with DMF prior to repeating the reaction with hydrazine/DIEA in DMF to ensure quantitative loading. To ensure all reactive sites on the resin had been consumed, after 1 hour, 500 µL of methanol was added to the second hydrazine reaction and allowed to stir for an additional 30 minutes.

The peptides were synthesized on a CEM Discover Microwave Peptide Synthesizer using Fmoc chemistry. Fmoc-acetyllysine was used to incorporate acetyllysine residues at the appropriate positions. For a 0.25 mmol-scale reaction the resin was first swelled in DMF. In between, the resin was washed with DMF and DCM. Each amino acid was dissolved in DMF (200 mM) and introduced to the resin after the deprotection step. DIEA (2 M) in N-methylpyrrolidone (NMP) and HOBt (500 mM)/HBTU (500 mM) in DMF were added along with the amino acid to perform each coupling (twice). The resin was then removed from the synthesizer without deprotecting the final Fmoc group for stability in storage. The N-terminus of each peptide was manually deprotected with 20% piperidine and acetylated with acetic anhydride. Peptides were cleaved from the resin as C-terminal acyl hydrazides in 95% TFA, 2.5% triisopropylsilane (TIS), and 2.5% water and purified by preparative C-18 RP-HPLC. The final products were characterized by ESI-MS and analytical C-18 RP-HPLC.

Full-length histones were prepared by traceless native chemical ligation of acetylated peptides and the recombinant fragment as previously described (Nguyen et al., 2014) and purified
by semi-preparative C-18 RP-HPLC. Final products were characterized by analytical C-18 RP-HPLC and ESI-MS (Figure 2.7).

7.4.2 Preparation of the H2AK118ac Histone

A recombinant fragment corresponding to residues 1-112 of H2A was fused to a fused Npu DnaE intein with a C-terminal 6xHis tag. This protein was produced in E. coli and purified by Ni-NTA affinity chromatography. An α-thioester was installed in the histone fragment by thiolysis (via 2-mercaptoethanesulfonate) of the corresponding intein fusion similar to methods previously described (McGinty et al., 2008). An H2AK118ac peptide (113-129, A113C) was synthesized as previously described in Section 7.4.1, with minor modifications. The peptide was synthesized on Wang resin that was purchased pre-charged with the C-terminal amino acid (Boc-protected lysine). Acetylation of the N-terminal amine with acetic anhydride was omitted after final Fmoc deprotection, and the peptides were cleaved (95% TFA, 2.5% triisopropylsilane (TIS), and 2.5%) as C-terminal carboxylates. Full-length histones were prepared by traceless native chemical ligation of acetylated peptide and the recombinant fragment as previously described (Nguyen et al., 2014) and subsequently purified by semi-preparative C-18 RP-HPLC. Final products were characterized by analytical C-18 RP-HPLC and ESI-MS (Figure 2.7).

7.4.3 Preparation of H2BK5ac, H2BK11ac, H2BK12ac, H2BK15ac, H2BK16ac, and H2BK20ac Histones

Proteins were assembled from two pieces: A recombinant fragment was generated encompassing residues 21-125 of histone H2B with an A21C mutation fused to an N-terminal 6xHis-SUMO tag. This sequence was inserted into a pET expression plasmid and confirmed by gene sequencing. The recombinant H2B(21-125)A21C fragment was produced and purified analogous to that described for the H2A(21-129)A21C fragment in Section 7.4.1. The final product was characterized by ESI-MS and analytical C-18 RP-HPLC.
H2BK5ac, H2BK11ac, H2BK12ac, H2BK15ac, H2BK16ac, and H2BK20ac peptides corresponding to residues 1-20 with acetylated lysine residues at the indicated positions were synthesized analogous to that described in Section 7.4.1, omitting acetylation of the N-terminal amine with acetic anhydride. Full-length histones were prepared by traceless native chemical ligation of acetylated peptides and the recombinant fragment as previously described (Nguyen et al., 2014) and were purified by semi-preparative C-18 RP-HPLC. Final products were characterized by analytical C-18 RP-HPLC and ESI-MS (Figure 2.7).

7.4.4 Preparation of H2BK108ac, H2BK116ac, H2BK120ac, H2BK125ac Histones

H2BK108ac, H2BK116ac, H2BK120ac, and H2BK125ac histones were produced in a similar manner to H2AK118ac as described Section 7.4.2. A recombinant fragment corresponding to residues 1-106 of H2B containing an α-thioester was prepared via thiolysis of fused Npu DnaE intein fusion for ligation to H2BK108ac(107-125, A107C) and H2BK116ac(107-125, A107C) peptides. Additionally, a recombinant fragment corresponding to residues 1-116 of H2B containing an α-thioester was prepared via thiolysis of a fused Npu DnaE intein fusion for ligation to H2BK120ac(117-125, A107C) and H2BK125ac(117-125, A117C) peptides. Full-length histones were prepared by traceless native chemical ligation of acetylated peptides and the recombinant fragment as previously described (Nguyen et al., 2014) and were subsequently purified by semi-preparative C-18 RP-HPLC. Final products were characterized by analytical C-18 RP-HPLC and ESI-MS (Figure 2.7).

7.4.5 Preparation of Phosphorylated H3 Histones

The H3H41ph histone was prepared using a previously described 3-piece traceless ligation strategy (Casadio et al., 2013) with minor modifications. Notably, the middle peptide fragment (residues 29-46, A29C) was synthesized with inclusion of Fmoc-Tyr(PO(OBzl)OH)-OH, which was used to introduce a phosphotyrosine residue at position 41 of H3. The H3S10ph
histone was prepared using a 2-piece traceless ligation strategy analogous to that described previously for ligation of synthetic peptide fragments corresponding to residues 1-14 of H3 to a recombinant H3 fragment corresponding to residues 15-135 (A15C, C96A, C110A) (Nguyen et al., 2014). The N-terminal peptide fragment (residues 1-14) was synthesized with inclusion of Fmoc-Ser(PO(OBzl)OH)-OH to introduce a phosphoserine residue at position 10 of H3. Final products were characterized by analytical C-18 RP-HPLC and ESI-MS (Figure 2.7).

7.4.6 Preparation of H3K4me1 and H3K4me2 Histones

H3K4me1 and H3K4me2 histones were prepared using a 2-piece traceless ligation strategy analogous to that described previously for native chemical ligation of synthetic peptide fragments corresponding to residues 1-28 of H3 to a recombinant H3 fragment corresponding to residues 29-135 (A29C, C96A, C110A) (Nguyen et al., 2014). The N-terminal peptide fragments (residues 1-28) were synthesized using Fmoc chemistry as previously described utilizing an on-resin alkylation strategy to generate mono and dimethylated lysine residues at position 4 of H3 (Biron et al., 2006; Brown et al., 2014). Final products were characterized by analytical C-18 RP-HPLC and ESI-MS (Figure 2.7).

7.4.7 Preparation of H3R2 Methylated Histones

H3R2me and H3R2me2a histones were prepared using a 2-piece traceless ligation strategy analogous to that described previously for ligation of synthetic peptide fragments corresponding to residues 1-14 of H3 to a recombinant H3 fragment corresponding to residues 15-135 (A15C, C96A, C110A) (Nguyen et al., 2014). The N-terminal peptide fragment (residues 1-14) was synthesized with inclusion of either Fmoc-Arg(Me,Pbf)-OH or Fmoc-ADMA(Pbf)-OH to introduce either a monomethyl or asymmetric dimethyl arginine residue at position 2 of H3. Final products were characterized by analytical C-18 RP-HPLC and ESI-MS (Figure 2.7).
7.4.8 Preparation of the H3K36me3 Histone

The H3K36me3 histone was prepared with a 3-piece ligation strategy analogous to that previously described (Casadio et al., 2013) with minor modifications. Notably, the full-length sequence corresponded to H3 and contained C96A, C110A mutations, and the middle peptide fragment (residues 29-46, A29C) was synthesized with inclusion of Fmoc-Lys(Me$_3$)-OH, which was used to introduce a trimethyllysine residue at position 36 of H3. The final product was characterized by analytical C-18 RP-HPLC and ESI-MS (Figure 2.7).

7.4.9 Preparation of H4R3 Methylated Histones

Synthetic H4 peptides were assembled as alkyl-thioesters (TAMPAL) (Hackeng et al., 1999) on solid phase. Up to residue 4, amino acids were incorporated using Boc-protected amino acids and in situ neutralization protocols. Side-chain protected amino acids (H4(1-14)R3Me and R3Me2s) were coupled using standard HBTU/DIPEA procedures; subsequent couplings were carried out using Fmoc amino acids, thus maintaining acid sensitive side-chain protecting groups (e.g. Boc or Pbf). To reduce Fmoc-deprotection cocktail exposure time, the terminal residue could be coupled with Boc-protection. Final Fmoc deprotection was carried out in 1-methylpyrrolidine (25 % v/v), hexamethylene imine (2 % v/v), HOBt (2 % w/v), in NMP-DMSO (1:1); (2) DBU (1 % v/v), HOBt (1 % w/v) in DMF.

The H4(1-14)R3Me2a thioester peptide was synthesized, using resin from above, that was split prior to coupling methyl arginine residues. Following Boc-deprotection, neutralization of the TFA salt (5 % (v/v) DIPEA-DMF), and thorough washing with DMF, the resin was washed with 5 % (w/v) HOBT-DMF, ensuring a slightly acidic environment necessary to maintain arginine side-chain protonation. Subsequent base-free activation of Fmoc-ADMA, in the presence of peptidyl-resin, was achieved using 5 eq of Fmoc-ADMA-OH, 5 eq of N,N-diisopropylcarbodiimide (DIC) and 5.5 eq of HOBT in 10 % DMF-DCM, with excess HOBT used to maintain side-chain protonation. Fmoc-deprotection was subsequently mediated using 1-methylpyrrolidine (25 % v/v),
hexamethylene imine (2 % v/v), HOBt (2 % w/v), in NMP-DMSO (1:1); (Li et al., 1998) (2) DBU (1 % v/v), HOBt (1 % w/v) in DMF, and the described coupling process repeated for subsequent amino acids, with neutralization omitted following Fmoc-deprotection.

Peptides were cleaved from the resin using liquid HF. HF cleavage (10mL HF/g resin) was performed for 1.5 h at 0-4 °C with 5 % (v/v) p-cresol as a scavenger. Following HF cleavage, the HF was removed under reduced pressure, peptides were precipitated in ice-cold ether, filtered, dissolved in 40% (v/v) aqueous MeCN containing 0.1% (v/v) TFA, and lyophilized.

Peptides were purified by preparative C-18 RP-HPLC and final products were characterized by analytical C-18 RP-HPLC and ESI-MS.

To generate the remaining recombinant portion of H4 suitable for native chemical ligation, the H4(1-14) tail was fused to the hexahistidine-tagged, thrombin cleavable truncated H4 protein bearing an A15C mutation – insertion of the His-tag and thrombin cleavage sequence into the H4 protein between residues 14 and 15 was found to improve expression levels. This approach yielded protein as inclusion bodies, which after Ni²⁺-affinity chromatography, was subjected to on-column cleavage with thrombin. By incorporating a mutated thrombin cleavage site (LVPRC), the desired N-terminal cysteine containing H4 fragment could be obtained by extracting the cleaved protein from the column using 8 M urea. The truncated histones were then further purified by preparative C-18 RP-HPLC.

Native chemical ligation was performed in 0.1 M sodium phosphate pH 7.5, 6 M guanidine-HCl, 0.1 M NaCl, 60 mM TCEP, 30 mM 4-mercaptophenylacetic acid (MPAA) at 37 °C. Reactions were performed at approximately 3 mM peptide concentration in siliconized 1.5 mL microcentrifuge tubes. Significantly lower yields were obtained when standard polypropylene tubes were used. Ligation products were purified, desulfurized, and purified analogous to that previously described (Nguyen et al., 2014). Final protein products were characterized by analytical C-18 RP-HPLC and ESI-MS (Figure 2.7).
7.4.10 Preparation of H4K77ac and H4K79ac Histones

H4K77ac and H4K79ac histones were produced similarly as described in Section 7.4.2. A recombinant fragment corresponding to residues 1-75 of H4 containing an α-thioester was prepared via thiolysis of a fused Npu DnaE intein fusion for ligation to H4K77ac(76-102, A76C) and H4K79ac(76-102, A76C) peptides. Full-length histones were prepared by traceless native chemical ligation of acetylated peptides and the recombinant fragment analogous to that previously described (Nguyen et al., 2014) and were subsequently purified by semi-preparative C-18 RP-HPLC. Final products were characterized by analytical C-18 RP-HPLC and ESI-MS (Figure 2.7).

7.4.11 Preparation of H2BS112GlcNAc

Fmoc-Ser(β-D-GlcNAc(Ac)₃)-OH was synthesized as previously described (using aglycone Fmoc-Ser(OH)-OBn and glycosyl donor GlcNTroc(Ac)₃-Br) (Mitchell et al., 2001) for use in solid-phase peptide synthesis. Standard Fmoc solid-phase peptide synthesis was used to access the glycopeptide H2B(107-125, S112GlcNAc, A107C). After successful peptide assembly, the peracetylated glycopeptide was deprotected on resin by hydrazinolysis (Mitchell et al., 2001), cleaved from the resin, and purified by preparative C18 RP-HPLC. The H2B(1-106) α-thioester was obtained by expression and thiolysis of a recombinant intein fusion fragment of H2B, H2B(1-106)GyrA-His₆ as previously described (McGinty et al., 2008). A one pot ligation/desulfurization procedure (Thompson et al., 2014) was used to obtain the final glycosylated H2BS112GlcNAc histone which was purified by semi-preparative C-18 RP-HPLC. The final product was characterized by analytical C-18 RP-HPLC and ESI-MS (Figure 2.7).
7.5 Production and Purification of Modified Histones by Amber Suppression

All crotonylated histones were produced using amber suppression in *E. coli* as previously described (Li et al., 2016; Xiong et al., 2016).

7.5.1 Preparation of H3K4ac, H3K36ac, H3K37ac, H3K56ac, H3K64ac, H3K79ac, H3K115ac, H3K122ac, and H4K91ac Histones

H3K4ac, H3K36ac, H3K37ac, H3K56ac, H3K64ac, H3K79ac, H3K115ac, H3K122ac histones were prepared by fusing a fused Npu DnaE intein to the C-terminus of H3 (H3C96A, C110A) or H4 and produced using amber suppression in *E. coli* as previously described (Batjargal et al., 2015; Neumann et al., 2009). Following purification, all proteins were characterized by analytical C-18 RP-HPLC and ESI-MS (Figure 2.7).

7.6 Synthesis of LANA Peptides

Two LANA peptides (Barbera et al., 2006) corresponding to residues 2-22 with a norleucine at position 6 substituted for methionine were prepared. The first contained the otherwise native sequence, and the second had residues 8-10 (LRS) mutated to alanine, which is known to hinder nucleosome binding (Barbera et al., 2006). Peptides were synthesized on a CEM Discover Microwave Peptide Synthesizer using Fmoc chemistry and a ChemMatrix Rink amide resin (0.47 mmol/g). Following chain assembly, peptides were cleaved from the resin using 95% TFA, 2.5% TIS, and 2.5% water and purified by C-18 RP-HPLC. Final products were characterized by analytical C-18 RP-HPLC and ESI-MS (Figure 7.1).
Figure 7.1 Characterization of LANA Peptides. Analytical C-18 RP-HPLC chromatograms (insets; gradients used are individually noted) and corresponding ESI-MS spectra of indicated peptides are shown.

7.7 Production of Chromatin Remodelers in Sf9 Cells

All chromatin remodelers used in this study were produced in Sf9 cells using a baculovirus expression vector system. The SNF2h (Uniprot ID: O60264), ACF1 (Uniprot ID: Q9NRL2), WSTF (Uniprot ID: Q9UIG0), CHRAC-15 (Uniprot ID: Q9NRG0), and CHRAC-17 (Uniprot ID: Q9NRF9) coding sequences were a gift from R. Kingston (Harvard University). The RSF1 (Uniprot ID: Q96T23) coding sequence was a gift from D. Reinberg (New York University). The BPTF (Uniprot ID: Q12830, Isoform 4) coding sequence was a gift from C. D. Allis (Rockefeller University). The SNFL, TIP5 (Uniprot ID: Q9UIF9, Isoform 1), CHD4 (Uniprot ID: Q14839), and RbAp46 (Uniprot ID: Q16576) coding sequences were purchased from Open Biosystems. The SNF2L coding sequence was purchased as an inactive splice variant (Barak et al., 2004) with an insertion that was deleted using standard molecular cloning procedures to yield the active variant (Uniprot ID: P28370, Isoform 2). pFastBac1-Flag-BRG1 was a gift from Robert Kingston (Addgene plasmid # 1957) (Phelan et al., 1999). All coding sequences were confirmed by Sanger sequencing. All coding sequences were subcloned into vectors compatible with bacmid generation by standard restriction enzyme cloning methods or Gibson Assembly® (NEB) with FLAG affinity tags introduced suitable for purification of all desired proteins and protein complexes (flag = N- or C-terminal FLAG affinity tag): pFastBac1-SNF2h-flag, pACEBac1-SNF2h, pACEBac1-ACF1-flag, pACEBac1-WSTF-flag, pACEBac1-CHRAC-15, pACEBac1-
CHRAC-17, pACEBac1-TIP5-flag, pACEBac1-flag-RSF1, pFastBac1-flag-BPTF, pACEBac1-RbAp46, pACEBac1-SNF2L, and pACEBac1-CHD4-flag. For bacmid generation, SNF2h, ACF1-flag, WSTF-flag, CHRACH-15, CHRAC-17, TIP5-flag, flag-RSF1, CHD4-flag constructs were transfected into DH10MultiBacTurbo™ E. coli competent cells, and bacmids were produced per manufacturer’s instructions (Geneva Biotech). SNF2h-flag, flag-BPTF, SNF2L, RbAp46, and flag-BRG1 constructs were transfected into DH10bac E. coli competent cells, and bacmids were produced per manufacturer’s instructions (Bac-to-Bac® Baculovirus Expression System, ThermoFisher Scientific). To generate virus for protein production, bacmid transfection into Sf9 cells was carried out in 6-well plates. All transfection, viral amplification, and infection steps were performed in a sterile hood. Typically, 1-10 μg of bacmid was transfected into 1 x 10^6 attached Sf9 cells according to manufacturer’s instructions (Bac-to-Bac® Baculovirus Expression System, ThermoFisher Scientific). After transfection, cells were overlaid with 2 mL fresh medium (Sf-900™III SFM, Thermo Fisher Scientific) and incubated at 27 °C for 72 hours in the dark. The supernatant was collected for virus amplification and cleared by centrifugation to generate the P1 virus. 2% FBS (v/v) was added. Between uses, all viral stocks were stored at 4 °C in the dark. Subsequent steps were carried out in medium with penicillin/streptomycin. To generate the P2 virus, 800 μL P1 virus was added to 20 mL of Sf9 cells in a sterile flask at 2 x 10^6 cells/mL. Cells were grown at 27 °C in suspension culture in the dark until they reached 40% viability as monitored by Trypan Blue staining. The culture supernatant was then collected and cleared by centrifugation, and 2% FBS (v/v) was added. To generate the P3 virus (used for protein production), 200 μL P2 virus was added to 50 mL of Sf9 cells in a sterile flask at 2 x 10^6 cells/mL. Cells were grown at 27 °C in suspension culture in the dark until they reached 40% viability as monitored by Trypan Blue staining. The culture supernatant was then collected and cleared by centrifugation, and 2% FBS (v/v) was added. During virus amplification Sf9 cell density was kept around 2 x 10^6 cells/mL, diluting if needed, until growth arrested and viability dropped. Aside from the monomeric SNF2h (SNF2h-flag), CHD4 (CHD4-flag), and BRG1 (flag-BRG1) ATPases, individual viruses were directly combined with Sf9 cell cultures to generate the ACF (ACF1-flag, SNF2h), CHRAC (ACF1-flag, SNF2h, CHRACH-15, CHRAC-17), WICH (WSTF-flag, SNF2h),
NoRC (TIP5-flag, SNF2h), RSF (flag-RSF1, SNF2h), and NURF (flag-BPTF, SNF2L, RbAp46) complexes. Production of CHD4 and the ACF, CHRAC, WICH, NoRC, and RSF complexes was carried out by adding a 1:1000 dilution of each P3 virus to Sf9 suspension cultures at 2 x 10^6 cells/mL. Cells were harvested by centrifugation after 72 hours at 27 °C in the dark. Production of SNF2h, BRG1, and the NURF complex was carried out by adding a 1:100 dilution of P3 virus to Sf9 suspension cultures at 2 x 10^6 cells/mL. Cells were harvested by centrifugation after 48 hours at 27 °C in the dark. All proteins and protein complexes were purified by first washing harvested cell pellets with 10 mL of cold TBS (50 mM Tris, 150 mM NaCl, pH 7.5 at 4 °C) per 50 mL of culture. Washed pellets were resuspended in 400 µL of cold nuclear extraction buffer A (buffer A; 10 mM HEPES, pH 7.9, 10 mM KCl, 0.1 mM EDTA, 0.1 mM EGTA) with 1 mM DTT, cOmplete™, EDTA-free Protease Inhibitor Cocktail (Sigma-Aldrich), and 0.5 mM PMSF per 10 mL of suspension culture and swelled on ice for 15 minutes. 25 µL of 10% IGEPAL CA-630 per 400 µL of buffer A were then added, and the cells were vortexed briefly for lysis. Nuclei were spun down at 4 °C for 30 seconds at 17,000 x g and the supernatant was removed. The nuclei were washed once with 400 µL of cold buffer A (with 1 mM DTT, cOmplete™, EDTA-free Protease Inhibitor Cocktail, and 0.5 mM PMSF) per 10 mL of suspension culture. The resulting nuclear pellet was resuspended in 50 µL of cold nuclear extraction buffer B (buffer B, 20 mM HEPES, pH 7.9, 400 mM KCl, 1 mM EDTA, 1 mM EGTA, 10% glycerol) with 1 mM DTT, cOmplete™, EDTA-free Protease Inhibitor Cocktail, and 0.5 mM PMSF per 10 mL of suspension culture. Nuclear pellets were then incubated with end-over-end rotation for 15 minutes at 4 °C. Nuclear debris was then spun down at 4 °C for 10 minutes at 17,000 x g. The supernatant (nuclear extract) was removed for FLAG affinity purification of remodelers. 1.5 µL (3 µL of 50% slurry) of ANTI-FLAG® M2 beads (Sigma-Aldrich) per mL of suspension culture were used for purification. First, beads were washed twice with 10 bead volumes of BC-100 buffer (20 mM HEPES, 100 mM KCl, 0.2 mM EDTA, 10% glycerol, 1 mM DTT, 0.2 mM PMSF, pH 7.9). The nuclear extract was diluted 2-fold with BC-0 buffer (BC-100 buffer without KCl) and centrifuged at ≥ 4,000 x g for 10 minutes at 4 °C. The supernatant was removed from any precipitation, added directly to the washed ANTI-FLAG® M2 beads, and incubated for 1 hour at 4 °C with end-over-end rotation. The beads were
centrifuged for 5 minutes at 1000 x g and 4 °C, and the supernatant was removed. The beads were washed once with 10 bead volumes of BC-100 buffer, once with 10 bead volumes of BC-300 buffer (BC-100 buffer with 300 mM KCl), and once with 10 beads volumes of BC-100 buffer. Bound protein was eluted by incubation with one bead volume of 0.25 mg/mL FLAG peptide (Sigma-Aldrich) in BC-100 buffer for 20 minutes at 4 °C. This process was repeated twice. Pure elutions were combined and concentrated using appropriate MW cut-off Vivaspin 500 centrifugal filter units (Vivaproducts) if needed. Concentrations of remodelers were determined using BSA standards and SDS-PAGE with Coomassie blue staining referencing the intensity of the ATPase subunit for each complex. Following purification, remodelers were aliquotted, flash-frozen in liquid nitrogen and stored at -80 °C until use. Purity of ISWI family remodelers (Figure 2.3a) as well as BRG1 and CHD4 (Figure 4.3a) was assessed by SDS-PAGE analysis and Coomassie blue staining. Additionally, ATP-dependent nucleosome remodeling activity on unmodified nucleosomes was verified for ISWI family members (Figure 2.3b) as well as BRG1 and CHD4 (Figure 4.3b) using a restriction enzyme accessibility assay. All assays were performed in 50 µL reactions with 10 nM unmodified nucleosomes with under conditions described in Section 7.15. Time points were taken at 1 minute and 60 minutes, quenched, and deproteinized as described in Section 7.13.1.

. Samples were directly run on a 5% polyacrylamide gel (0.5x TBE, 200 V, 40 minutes). Staining was performed with SYBR® Safe DNA gel stain, and gels were imaged on a Typhoon scanner (GE Healthcare). Each remodeler was analyzed in the presence and absence of 2 mM ATP. In all cases, the appearance of a lower band at 60 minutes that is dependent on the presence of ATP (indicative of remodeling) is visible. Concentrations of remodelers used: SNF2h: 200 nM, ACF: 2 nM, CHRAC: 20 nM, WICH: 50 nM, NoRC: 20 nM, RSF: 10 nM, NURF: 50 nM, BRG1: 370 nM, CHD4: 200 nM.
7.8 DNA Preparation

7.8.1 Barcoded 601 (BC-601) DNA Preparation

A PstI restriction site was introduced into a 147 b.p. 601 sequence via site-directed mutagenesis to detect nucleosome remodeling. The 601 sequence contains 5’-DraIII and 3’-Bsal overhangs and was prepared by digestion of a plasmid containing 16 copies of the desired sequence; full sequence of one repetitive unit:

5’- GTGACAGGATGTATATATCTGACACGTGCTGGAGACTAGGGAGTAATCCCCTTGGCGGT TAAAACGCAGGGGACAGCGCGCTACGTCGTGTTTTAACGGGTGTAGAGCTGTCTACGACCA ATTGAGCGGCTGCAGCACCAGGGATTCTCCAG – 3’

601 sequence – bold
PstI site – underlined (CTGCAG)

The plasmid was produced in and purified from DH5α cells as previously described (Dyer et al., 2004). Approximately 20 mg of plasmid was obtained from 6 L of bacterial culture. 10 mg of the plasmid was then digested with Bsal and DraIII (total DNA concentration of 1 mg/mL in the presence of 500 U/mL of each restriction enzyme) at 37 °C overnight. Digested DNA was purified by phenol chloroform extraction and concentrated by ethanol precipitation. The pellet was resuspended in roughly 1 mL of TE buffer (50 mM Tris, 0.1 mM EDTA, pH 7.5) and sucrose was added to a final concentration of 15%. The desired fragment was then purified by polyacrylamide gel electrophoresis (5% polyacrylamide gel, 0.5x TBE buffer) using a preparative cell (Bio-Rad). A peristaltic pump in-line with a Foxy R1 fraction collector (Teledyne Isco) was used to collect eluting fractions every 45 seconds at a flow rate of 1 mL/minute. Fractions containing the desired fragment were pooled and concentrated by ethanol precipitation. The pellet was resuspended in water yielding approximately 1-2 mg of 601 DNA. Unique nucleosome identifier barcodes were
then individually ligated to 601 DNA fragment using the non-palindromic 5’-DralI site. Individual oligo pairs were purchased from Integrated DNA Technologies. Each oligo was dissolved in TE buffer to a concentration of roughly 100 µM. For hybridization, equimolar amounts of oligo pairs (top and bottom strand) were mixed at a concentration of 10 µM each oligo in a volume of 500 µL. Samples were heated on a heat block for 5 minutes at 95 °C, and were subsequently left to cool slowly by placing the heat block on a lab benchtop at room temperature for 1 hour. Individual oligos contained the following sequence:

Top strand – 5’ – CTCTTTCCCTACACGACGCTCTTCCGATCTNNNNNNCACC GC – 3’
Bottom strand – 5’ – GTGNNNNNNAGATCGGAAGAGCGTCGTG TAGGGAAGAA – 3’

Partial Illumina forward adapter sequence – bold
Unique hexanucleotide nucleosome identifier barcode – NNNNNN

Relative to the top strand, hybridization resulted in double-stranded BC DNA with a 3’ overhang compatible with ligation to the 5’-DralI overhang on the purified 601 sequence and a 3’-AA dinucleotide overhang to prevent blunt-end ligation of hybridized oligos. In a typical ligation reaction 0.41 µM of 601 DNA was combined with 1.2 equivalents of double stranded BC DNA and incubated at 37 °C with 0.1 U/µL T4 polynucleotide kinase (NEB) in a total volume of 100 µL in 1x T4 ligase buffer (NEB) for 1 hour. Subsequently 5 U/µL T4 DNA ligase were added, and the mixture was incubated for 1 hour at room temperature. Ligation reactions were monitored by native polyacrylamide gel electrophoresis (5% polyacrylamide gel, 0.5x TBE, 200 V, 40 minutes) and staining with SYBR® Safe DNA gel stain (ThermoFisher Scientific; Figure 2.9a). The final product was purified using a Qiagen PCR purification kit and quantified by UV spectroscopy at 260 nm. The final product resulted in a 192 b.p DNA fragment composed of the 147 b.p. 601 sequence and a 45 b.p. overhang containing a unique hexanucleotide barcode and a partial Illumina forward adapter sequence:
5’ –
CTCTTTCCCTACACGAGCTCTTCCGATCTNNNNNNCACCGCGTGCACAGATGTATATATCT
GACACGTGCCTGGAGACTAGGGAGTAAATCCCTTTGGCGGTTAAAAACCGGCGGGGACAGCG
CGTACGTGCCTTTAAGCGGTGC TAGAGCTGTCTACGACCAATTGAGCGGCTGCAGCACCG
GGATTCTCCAG – 3’

601 sequence – bold
PstI site – underlined (CTGCAG)
Unique hexanucleotide nucleosome identifier barcode – NNNNNN
The partial Illumina forward adapter sequence is the 30 b.p. 5’ of the hexanucleotide barcode.

CpG methylated nucleosomal DNA was prepared as described in Figure 2.9b.

7.8.2 Biotinylated MMTV ‘Buffer’ DNA

Biotinylated MMTV ‘buffer’ DNA was prepared by PCR using a MMTV DNA template
(Flaus and Richmond, 1998) with the following primers purchased from Integrated DNA Technologies:

Forward – 5’ – biotin-TATCACTTGCAACAGTCCTAACATTCACCTC – 3’ (HPLC purified)
Reverse – 5’ – ATCCAAAAAACGTGCGCCAGTCGG – 3’

The PCR product was purified using a Qiagen PCR purification kit. The final product was
quantified by UV spectroscopy at 260 nm and stored at -20 °C.
7.8.3 192 b.p. DNA Fragment Used in Follow-up Studies on Individual Nucleosomes

601 DNA used in non-library remodeling assays on individual nucleosomes was generated via PCR. The 147 b.p. 601 sequence containing a PstI site was inserted into a KS BlueWhite screening vector using standard restriction enzyme techniques. Primers were designed to amplify the 601 sequence with a 45 b.p. overhang generated from the surrounding vector sequence.

Forward – 5' - GGC CGC TCT AGA ACT AGT – 3'
Reverse – 5' - CTG GAG AAT CCC GGT – 3'

DNA fragment generated:

GGCCGCTCTAGAAGTGGATCCGATATCGCTTTCACCACGTGACAGGATGTATATATC
TGACACGTGCCCTGAGAATGGAGAATCCTCTTGGCGTTAAACACGCCGCGGACAGC
GCGTACGTGCCTTTAAGCGGTGCTAGAGCTGTCTACGACCAATTGAGCGCGGCTGCAGCACC
GGGATTCTCCAG

601 sequence – bold
PstI site – underlined (CTGCAG)

7.9 Octamer Formation for the Nucleosome Library

Octamers containing desired histone compositions were assembled as previously described (Nguyen et al., 2014) with minor modifications. In brief, histones were dissolved in histone unfolding buffer (6 M guanidine hydrochloride, 20 mM Tris, 5 mM DTT, pH 7.5 at 4 °C) and combined in equimolar ratios (0.75 nmol each of desired version of histones H2A, H2B, H3,
and H4). The total histone concentration was adjusted to 1 mg/mL, and the mixtures were placed in Slide-A-Lyzer MINI dialysis devices (3.5 kDa MW cutoff, ThermoFisher Scientific) and dialyzed at 4 °C against 3 x 400 mL of octamer refolding buffer (2 M NaCl, 10 mM Tris, 0.5 mM EDTA, 1 mM DTT, pH 7.8 at 4 °C) for at least 4 hours for each step, with one dialysis step overnight. The mixtures were then transferred to clean microcentrifuge tubes and spun down at 17,000 x g for 5 minutes at 4 °C to remove any precipitate. Supernatants were transferred to fresh microcentrifuge tubes, and 50% (v/v) glycerol was added. Octamer concentrations were measured by UV spectroscopy at 280 nm, and stored at -20 °C until use in nucleosome assembly.

7.10 Nucleosome Reconstitution for the Nucleosome Library

Nucleosomes were assembled as previously described (Nguyen et al., 2014) with minor modifications. In brief, in a typical nucleosome assembly, a crude histone octamer preparation (50 pmol) was combined with the appropriate BC-601 DNA (25 pmol) and biotinylated MMTV ‘buffer’ DNA (25 pmol) in 70 µL of octamer refolding buffer (2 M NaCl, 10 mM Tris, 0.5 mM EDTA, 1 mM DTT, pH 7.8 at 4 °C). Mixtures were placed in Slide-A-Lyzer MINI dialysis devices (3.5 kDa MW cutoff, ThermoFisher Scientific) and dialyzed at 4 °C against 200 mL nucleosome assembly start buffer (10 mM Tris, 1.4 M KCl, 0.1 mM EDTA, 1 mM DTT, pH 7.8 at 4 °C) for 1 hour at 4 °C. Subsequently, 330 mL of nucleosome assembly end buffer (10 mM Tris, 10 mM KCl, 0.1 mM EDTA, 1 mM DTT, pH 7.8 at 4 °C) were added at a rate of 1 mL/minute using a peristaltic pump, followed by two final dialysis steps against nucleosome assembly end buffer (1 hour and overnight). The dialysis mixture was transferred to a microcentrifuge tube, and removal of biotinylated MMTV ‘buffer’ DNA-associated side products was accomplished by affinity depletion using 140 µL of MyOne streptavidin T1 coated Dynabead slurry (ThermoFisher Scientific) for 1 hour at room temperature. The unbound material in the supernatant was transferred to a clean microcentrifuge tube, and any precipitate was removed by centrifugation. Final nucleosome preparations were quantified by UV spectroscopy at 260 nm. The quality of individual nucleosomes was assessed by native polyacrylamide gel electrophoresis (5% acrylamide gel,
0.5x TBE, 200 V, 40 minutes), followed by ethidium bromide staining. This resulted in a main band migrating around 500 b.p. (variable depending on the nucleosome variant; Figure 2.12a) To form the library, nucleosome preparations were directly combined and concentrated using Vivaspin 500 centrifugal filter units (10 kDa MW cutoff, Vivaproducts). The final nucleosome library was analyzed by native gel electrophoresis as previously described (Figure 2.12b). For storage, 20% (v/v) glycerol was added resulting in a final library concentration of approximately 1.5 µM. Library aliquots were flash frozen in liquid nitrogen, and stored at -80 °C until being thawed before use.

7.11 Analysis of Nucleosome Library Integrity by Antibody Pull-down

We assessed nucleosome library integrity by antibody pull-down against a specific histone mark. The nucleosome library (12 fmol of each member) was combined with a modification specific antibody (anti-H3K4me3, Abcam: ab8580; 15 µg/mL final concentration) in 100 µL of antibody binding buffer (20 mM Tris, 50 mM NaCl, 5 mM EDTA, 0.1% TWEEN® 20) and incubated at room temperature for 1 hour. The volume was then brought to a final volume of 200 µL with antibody binding buffer. 10 µL of Pierce Protein G agarose slurry (previously washed with antibody binding buffer) were then added and the mixture was incubated for an additional hour at room temperature to allow antibody binding to the Protein G beads. The beads were then washed 4 times with antibody binding buffer and incubated with 100 µL of elution buffer (100 mM Tris, pH 7.5, 10 mM EDTA, 1% SDS, 10 mM, 2-mercaptoethanol, 8 U/mL Proteinase K) for 90 minutes at 50 °C. DNA was purified using a Qiagen PCR purification kit, and was subsequently quantified using a Qubit high-sensitivity dsDNA quantification kit. DNA was diluted with water to a final concentration of approximately 2 pg/µL. The sample was PCR amplified according to conditions described in Section 7.13.7. An input sample was prepared by directly mixing an equivalent amount of nucleosome library with elution buffer followed by identical sample processing. Sequencing was performed as described in the section 'Illumina sequencing'.
Experimental samples were normalized against the input, and, within the experiment, the indicated variant (Figure 2.12c).

7.12 Nucleosome Thermal Mobility Shift Experiments

Nucleosome thermal mobility shift experiments were performed similarly to that previously described (Flaus and Richmond, 1998). Octamers and nucleosomes were prepared similarly to that previously described (Dyer et al., 2004) with minor modifications. Nucleosomes were assembled via salt gradient dialysis on a DNA fragment that was PCR amplified from a pTF vector to generate a 283 b.p. fragment containing a centrally positioned MMTV sequence (NucA) (Flaus and Richmond, 1998) with 64 b.p. of flanking DNA on each side.

Forward: 5’- ATT TAT TAT GCA TTT AGA ATA AAT TTT GTG TCG CCC TTG – 3’
Reverse: 5’- CAG TCG AAA GAC TGG GCC TTT C – 3’

DNA fragment generated:

5’ –
ATTTTATGCAATTTAGAAATATTGTTGTCGCCCTTGTGCTGAGGTACCAGATCTGATAT
CACTTGCAACACGTCTAACATTCACCTTCTTTGTGTGTTTGTGTTGCTGCTGCCATCCCGTCTCC
GCTCGTCATTTATCCTCTTTCCAGAGGGTCCCCGGCCAGACCCCGGGCGACCCTGGGT
GGCCGACTGCGGCCACAGTTTTTTTGATGATCGGATCCGTCATCGAAGGGCGACACCC
CTAATTAGCCCCCCGGCGAAGGGCGACAGCTTCTCGACTG – 3’

MMTV sequence – bold

Approximately 1 pmol of each nucleosome type were incubated in assay buffer (50 mM Tris, 150 mM NaCl, pH 7.5) for 1 hour at either 37 °C or 47 °C. After incubation, 5 µL of 50%
sucrose were added to each sample and they were run on a 5% TBE gel in 0.5x TBE buffer for 40 minutes at 200V. Nucleosomes were visualized by staining with SYBR® Gold Nucleic Acid Gel Stain (ThermoFisher Scientific). Unmodified and H3K14cr nucleosomes are not expected to be as mobile as H3T118H and H4R45A nucleosomes, which contain mutations that affect interactions between histones and DNA.

7.13 Experiments Employing the Nucleosome Library

7.13.1 Remodeling Assays Using the Nucleosome Library

Remodeling assays employed a restriction enzyme accessibility (REA)-based strategy and were performed similarly to those previously described (He et al., 2008). In each remodeling reaction, 10 nM of library was used along with varying amounts of remodeling enzyme depending on the baseline activity of each protein preparation. Typically, an amount of remodeler that resulted in approximately 50% of unmodified nucleosomes remodeled in 1 hour was used (200 nM SNF2h, 2 nM ACF, 5 nM CHRAC, 10 nM WICH, 5 nM NoRC, 40 nM RSF, and 50 nM NURF). 50 µL remodeling assays were carried out in REA buffer (12 mM HEPES, pH 7.9, 4 mM Tris, pH 7.5, 60 mM KCl, 10 mM MgCl₂, 10% glycerol, and 0.02% (v/v) IGEPAL CA-630) with or without 2 mM ATP and in the presence of 2 U/µL PstI restriction enzyme (NEB). Reactions were pre-incubated for 10 minutes at 30 °C prior to initiation by the addition of nucleosomes. Note that ATP (or an equivalent volume of REA buffer in the case of reactions without ATP) was not added until approximately 2.5 minutes prior to initiation. After initiation, remodeling assays were carried out for 1 hour at 30 °C. 6 µL time points were taken at 1, 2.5, 5, 15, 30, and 60 minutes and each was quenched with 9 µL of quench buffer (10% glycerol, 70 mM EDTA, pH 8, 20 mM Tris, pH 7.7, 2% SDS, 0.2 mg/mL bromophenol blue). Samples were deproteinized with 30 U/mL Proteinase K (NEB) for 1 hour at 37 °C. DNA was purified using a Qiagen PCR purification kit, eluting in 50 µL of TE buffer (50 mM Tris, 0.1 mM EDTA, pH 7.5). For each enzyme, reactions with and without ATP were performed in triplicate. With 6 time points taken per reaction, this generated 36
samples per enzyme analyzed. Samples were quantified using a Qubit high-sensitivity dsDNA quantification kit. DNA was diluted with water to a final concentration of approximately 2 pg/µL. The sample was then PCR amplified as per conditions described in the Section 7.13.7.

7.13.2 Pull-down Experiments with the RCC1-GST Construct using the Nucleosome Library

10 µL of Pierce Glutathione Agarose slurry were washed 4 times with 100 µL of binding buffer (20 mM Tris, 50 mM NaCl, 5 mM EDTA 0.1% TWEEN® 20, pH 7.5). The beads were resuspended in 100 µL of binding buffer after the final wash and 1.5 pmol of RCC1-GST were added. Binding to the beads was allowed to occur for 1.5 hours at room temperature with end-over-end rotation. After incubation, the beads were washed again 4 times with 100 µL of binding buffer. The beads were resuspended in 100 µL of binding buffer after the final wash and combined with the nucleosome library (12 fmol of each member). After incubation with end-over-end rotation for 2 hours at 4 °C, the beads were washed again 4 times with 100 µL of binding buffer. The beads were then directly resuspended in 100 µL of DNA elution buffer (100 mM Tris, 10 mM EDTA, 1% SDS, 10 mM β-mercaptoethanol, 200 μg/mL Proteinase K) and incubated for 1.5 hours at 50 °C. At this point an input sample for normalization was prepared by taking a small amount of nucleosome library, resuspending it in 100 µL of DNA elution buffer, and incubating for 1.5 hours at 50 °C. DNA from all samples was purified using a Qiagen PCR purification kit, eluting in 50 µL of TE buffer (50 mM Tris, 0.1 mM EDTA, pH 7.5). Samples were quantified using a Qubit high-sensitivity dsDNA quantification kit. DNA was diluted with water to a final concentration of approximately 2 pg/µL. The sample was then PCR amplified as per conditions described in Section 7.13.7.
7.13.3 Pull-down Experiments with the Sir3-BAH-FLAG-6xHis Construct Using the Nucleosome Library

10 µL of ANTI-FLAG® M2 (Sigma-Aldrich) slurry were washed 4 times with 100 µL of binding buffer (20 mM Tris, 50 mM NaCl, 5 mM EDTA 0.1% TWEEN® 20, pH 7.5). The beads were resuspended in 50 µL of binding buffer after the final wash and 40 pmol of Sir3-BAH-FLAG-6xHis were added. Binding to the beads was allowed to occur for 1 hour at room temperature with end-over-end rotation. After incubation, the beads were washed again 4 times with 50 µL of binding buffer. The beads were resuspended in 40 µL of binding buffer after the final wash and combined with the nucleosome library at a final concentration of 40 nM. After incubation with end-over-end rotation for 1 hour at room temperature, the beads were washed again 4 times with 50 µL of binding buffer. The beads were then directly resuspended in 100 µL of DNA elution buffer (100 mM Tris, 10 mM EDTA, 1% SDS, 10 mM β-mercaptoethanol, 200 µg/mL Proteinase K) and incubated for 1.5 hours at 50 °C. At this point an input sample for normalization was prepared by taking a small amount of nucleosome library, resuspending it in 100 µL of DNA elution buffer, and incubating for 1.5 hours at 50 °C. DNA from all samples was purified using a Qiagen PCR purification kit, eluting in 50 µL of TE buffer (50 mM Tris, 0.1 mM EDTA, pH 7.5). Samples were quantified using a Qubit high-sensitivity dsDNA quantification kit. DNA was diluted with water to a final concentration of approximately 2 pg/µL. The sample was then PCR amplified as per conditions described in the Section 7.13.7.

7.13.4 Experimental Procedures Corresponding to Section 5.2 'b-JIM, a Novel PWWP-domain Containing Protein Functionally Regulates the Chromosomal Kinase JIL-1 in Drosophila melanogaster'

All recombinant proteins and protein complexes were prepared in Sf21 cells by Catherine Regnard. Individual proteins (b-JIM, b-JIM PWWP mutant, and MSL3) are FLAG tagged. Protein b-JIM:JIL-1 complexes contain a FLAG tag on the b-JIM subunit. Pull-down experiments using
the nucleosome library with b-JIM, b-JIM PWWP mutant, and MSL3 were performed using M2 α-FLAG slurry. 10 µL of slurry were first washed with binding buffer (20 mM Tris-HCl, pH 7.5, 50 mM NaCl, 1 mM MgCl₂, 0.1% TWEEN). Beads were resuspended in 100 µL of binding buffer and 1.5 pmol of each recombinant protein were added to the beads. Binding to the beads was allowed to occur for 1.5 hours at room temperature with end-over-end rotation. Following incubation, the beads were washed four times with 100 µL of binding buffer, and, after the last wash, resuspended in 200 µL of binding buffer. The nucleosome library (12 fmol each member) was then added to the protein-bound beads and incubated at 4 °C for 2 hours with end-over-end rotation. Post-incubation, the beads were washed 4 times with 200 µL of binding buffer. After the last wash, 100 µL of DNA elution buffer (100 mM Tris, pH 7.5, 10 mM EDTA, 1% SDS, 10 mM β-mercaptoethanol, 8 U/mL proteinase K) were added, and the samples were incubated for 1.5 hours at 50 °C. DNA was then purified using a Qiagen PCR purification kit.

Kinase assays carried out with the b-JIM:JIL-1 and b-JIM:JIL-1 PWWP mutant complexes were carried out at 30 °C in a volume of 20 µL with nucleosome library (12 fmol each member) in kinase buffer (binding buffer with 0.5 mM ATP). After pre-incubation for 5 minutes at 30 °C, assays were initiated by the addition of 0.15 pmol of each enzyme complex and lasted 40 minutes. Assays were quenched with 5 mM EDTA. Each assay sample (20 µL) was then added to H3S10ph antibody (Abcam ab5176)-bound Protein G beads (20 µL in kinase buffer) that were prepared as follows. 10 µL of Pierce Protein G slurry (per sample) were washed 4 times with 100 µL of kinase buffer and resuspended in 20 µL of kinase buffer after the last wash. 0.75 pmol of antibody were added, and binding to the beads was allowed to occur for 1 hour at room temperature with end-over-end rotation. After 1 hour, the beads were washed 4 times with 20 µL of kinase buffer and resuspended in 20 µL of kinase buffer. Kinase reactions were mixed and incubated with antibody-bound beads for 1 hour at room temperature with end over end rotation. After 1 hour, the beads were washed 4 times with kinase buffer. Following the last wash, the beads were resuspended in 100 µL of DNA elution buffer and incubated for 1.5 hours at 50 °C. DNA was then purified using a Qiagen PCR purification kit.
All purified assay samples were individually barcoded and subjected to high-throughput DNA sequencing as described in Sections 7.13.7 and 7.13.8 except that sequencing was performed in an Illumina MiSeq instrument with a 150 or 300 nucleotide single-end read option. Data was processed exactly as described in Section 7.18.

7.13.5 Experimental Procedures Corresponding to Section 5.3 ‘Functional Crosstalk Between Histone H2B Ubiquitylation and H2A Modifications and Variants’—Performed by Felix Wojcik

For ubiquitylation assays, 5 pmol of mononucleosome library, 100 ng of hE1, 200 ng of His-tagged UBE2A, 250 ng of RNF20/40, and 1 µg of HA-tagged ubiquitin were incubated in 20 µL reaction buffer (50 mM Tris, pH 7.9, 5 mM MgCl₂, 2 mM NaF, 0.4 mM DTT and 4 mM ATP) at 37°C for 4 hours. Afterward, the reaction was quenched by addition of N-methylmaleimide (final concentration 3 mM) for 15 minutes at room temperature. 80 µL of binding buffer (25mM Tris, pH 7.5; 150 mM NaCl, 0.1%BSA, 0.1% NP-40, 10% Glycerol, 1 mM DTT) were then added.

For pull-down experiments 1 pmol of ubiquitylated nucleosome library (20 µL of the quenched solution) was mixed with 20 pmol of unmodified mononucleosomes in 120 µL binding buffer and incubated overnight at 4°C with anti-HA functionalized magnetic beads (containing 12 µg anti-HA antibody and 60 µL Invitrogen Dynabeads™ Protein G slurry). The beads were then washed with 3 times with 1 mL binding buffer. The beads were then directly resuspended in 100 µL of DNA elution buffer (100 mM Tris, 10 mM EDTA, 1% SDS, 10 mM β-mercaptoethanol, 8 U/mL Proteinase K) and incubated for 1.5 hours at 50°C. The resulting DNA was purified using a Qiagen PCR purification kit, eluting in 30 µL of TE buffer (50 mM Tris, 0.1 mM EDTA, pH 7.5). Samples were quantified using a Qubit high-sensitivity dsDNA quantification kit. DNA was diluted with water to a final concentration of approximately 2 pg/µL.

All purified assay samples were individually barcoded and subjected to high-throughput DNA sequencing as described in Sections 7.13.7 and 7.13.8 except that sequencing was
performed in an Illumina MiSeq instrument with a 150 or 300 nucleotide single-end read option. Data was processed exactly as described in Section 7.18.

7.13.6 Experimental Procedures Corresponding to 5.4 ‘Histone Acetylation Regulates ADP-ribosylation in DNA Damage’—Performed by Glen Liszczak and Katharine Diehl

ADP-ribosylation of the nucleosome library by PARP1 was carried out in 10 µL reaction buffer (50 mM Tris, pH 7.5, 20 mM NaCl, 2 mM MgCl₂, 2 mM DTT, 100 µM biotinylated NAD⁺) with 100 nM PARP1, 2 µM HPF1, and 100 nM nucleosome library for 30 minutes at 30 °C. Reactions were quenched by addition of 390 µL of wash buffer (50 mM Tris, pH 7.5, 100 mM NaCl, 2 mM MgCl₂, 1% Triton-X 100, 2 mM DTT, 1 µM olaparib PARP inhibitor). 300 µL magnetic streptavidin bead slurry (NEB) were then equilibrated with wash buffer. After equilibration, the beads were pelleted, supernatant removed, and reactions were directly added. Samples were incubated for 1 hour at room temperature with end-over-end rotation for isolation of ADP-ribosylated nucleosome by streptavidin pull-down. Samples were then washed 9 times with 1 mL of wash buffer. After the last wash the beads were resuspended in 200 µL of DNA elution buffer (100 mM Tris, 10 mM EDTA, 1% SDS, 10 mM β-mercaptoethanol, 8 U/mL Proteinase K) and incubated for 1.5 hours at 50°C. The resulting DNA was purified using a Qiagen PCR purification kit, eluting in 50 µL of TE buffer (50 mM Tris, 0.1 mM EDTA, pH 7.5). Samples were quantified using a Qubit high-sensitivity dsDNA quantification kit. DNA was diluted with water to a final concentration of approximately 2 pg/µL.

PARP1 binding studies using the nucleosome library were performed in 100 µL of binding buffer (50 mM Tris, pH 7.5, 150 mM NaCl, 1 mM MgCl₂, 0.1% Triton-X 100, 1 mM DTT) with 10 nM nucleosome library, 10 nM PARP1-flag, and 2 µM HPF1. Samples were incubated for 1 hour at 4 °C with end-over-end rotation. 20 µL of equilibrated ANTI-FLAG® M2 beads (Sigma-Aldrich) were added and the samples were incubated again for 1.5 hours at 4 °C. The beads were subsequently washed 5 times with 100 µL of binding buffer, followed by resuspension in 200 µL.
of DNA elution buffer and incubation for 1.5 hours at 50 °C. The resulting DNA was purified using a Qiagen PCR purification kit, eluting in 50 µL of TE buffer (50 mM Tris, 0.1 mM EDTA, pH 7.5). Samples were quantified using a Qubit high-sensitivity dsDNA quantification kit. DNA was diluted with water to a final concentration of approximately 2 pg/µL.

All purified assay samples were individually barcoded and subjected to high-throughput DNA sequencing as described in Sections 7.13.7 and 7.13.8 except that sequencing was performed in an Illumina MiSeq instrument with a 150 or 300 nucleotide single-end read option. Data was processed exactly as described in Section 7.18.

7.13.7 Addition of Multiplex Barcodes and Illumina Forward and Reverse Adapter Sequences via PCR

Approximately 10 pg of DNA from each purified experimental sample collected from remodeling assays or binding experiments using the nucleosome library were added to a PCR reaction to be amplified using the following primer pair:

Forward primer

5’ – AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACG – 3’

Reverse primer

5’ – CAAGCAGAAGACGGCATACGAGATXXXXXCTGGAGAATCCCGGTG – 3’

Unique hexanucleotide multiplexing barcode – XXXXXX
Specific amplification of full-length (un-cut) nucleosomal DNA at each assay time point was carried out using a Phusion® High-Fidelity PCR kit from NEB and primers flanking the PstI restriction cut site (1 mM dNTPs, 10 pg DNA, 0.5 µM each primer, 0.02 U/µL polymerase) with the following amplification conditions: Step 1: 98 °C for 30 seconds, Step 2: 98 °C for 10 seconds, Step 3: 47 °C for 15 seconds, and Step 4: 72 °C for 8 seconds. Steps 2-4 were repeated for a total of 15 cycles followed by a final extension step at 72 °C for 7 minutes. To ensure reactions remained in the exponential phase and relative abundances of DNA fragments were maintained, qPCR analysis was performed using 10 pg of a chosen remodeling reaction time point using SYBR Green I Dye for detection on an ABI 7900 quantitative PCR instrument. Identical conditions were used as previously described except a 10 second extension time was used (instead of 8 seconds) due to the instrument requirements. A mean C_t value of 13.9 was determined (compare to 15 cycles used for amplification of DNA prior to Illumina sequencing). The following amplicon was generated that is compatible with Illumina sequencing:

5’ –
AATGATACGCGACCCGAGATCTACACTTTTCCTTTAGACGCTCTTTCCGATCTNNNN
_NNCACCGGCGTGACAGGATGTATATATCTGACACGTGCCTGGAGACTAGGGAGTAATCC
TTGGCGGTTAAACGCAGGGGGACAGCGCGTACGCGCTGTTAAACGGGTGCTAGAGCTGCT
ACGACCAATTGAGCGGCTGCAGACCAGGGATTCCCGAGXXXXXATCTCGTATGCGGCTTT
CTGCTTG – 3’

601 sequence – bold
PstI site – underlined (CTGCAG)
Unique hexanucleotide nucleosome identifier barcode – NNNNNN
Unique hexanucleotide multiplexing barcode – XXXXX

The full Illumina forward adapter sequence is 5’ of the unique hexanucleotide nucleosome identifier barcode (NNNNNN). The full Illumina reverse adapter sequence is 3’ of the
unique hexanucleotide multiplexing barcode (XXXXXX). Samples containing unique multiplexing barcodes could be pooled directly after PCR and purified using a Qiagen PCR purification kit in one pot. After purification, samples were subjected to high-throughput DNA sequencing as described in Section 7.13.8. In each case, the correct amplicon size (250 b.p.) was verified by agarose gel electrophoresis and staining with EtBr.

7.13.8 Illumina Sequencing

Single-end sequencing of barcoded DNA libraries (starting from the forward adapter and covering the unique nucleosome identifier barcode) was performed by the Lewis Sigler Institute for Integrative Genomics Sequencing Core Facility at Princeton University on an Illumina HiSeq 2500 with read length of 67 b.p. Due to the significant sequence homogeneity present in samples generated from remodeling experiments, these libraries were diluted with a PhiX control library to enable sequencing. A custom sequencing primer was used to sequence the index read (5’ – CAATTGAGCGGCTGCAGCACCGGGATTCTCCAG – 3’). This primer annealed to the 601 sequence and the 8 b.p. index read covered the unique multiplexing barcode.

7.14 Nucleosome Remodeling Experiments in the Presence of LANA Peptides

Remodeling experiments in the presence of LANA and LANA mutant (LRS to AAA) peptides and rate calculations were performed as described in Section 7.15. LANA peptides were pre-incubated with unmodified nucleosomes for 10 minutes on ice before being used in remodeling assays at a final concentration of 10 µM peptide.
7.15 Nucleosome Remodeling Rate Analysis for Validation Experiments on Individual Nucleosomes Using a Restriction Enzyme Accessibility Assay

All remodeling experiments validating results obtained from the nucleosome library were performed on nucleosomes assembled with a 192 b.p. fragment described in the Section 7.8.3. Remodeling assays were performed as described in Section 7.13.1 with some modifications. Experiments involving the NURF complex were carried out using nucleosomes at 10 nM and NURF at 50 nM. Experiments involving the ACF complex were performed using nucleosomes at 10 nM and ACF at 2 nM. Experiments involving CHD4 were carried out using nucleosomes at 10 nM and CHD4 at 5 nM. Experiments involving BRG1 were carried out using nucleosomes at 10 nM and BRG1 at 200 nM. All reactions were carried out in the presence of 2 mM ATP. Time points were taken, quenched, and deproteinized as described in Section 7.13.1. Samples were directly run on a 5% polyacrylamide gel (0.5x TBE, 200 V, 40 minutes). Staining was performed with SYBR® Safe DNA gel stain and gels were imaged on a Typhoon scanner (GE Healthcare). Densitometry measurements were performed using Image Studio Lite (LI-COR). Rates were determined using GraphPad Prism by fitting to a single exponential decay equation (see Figure 4.2b, Figure 3.3, and Section 7.19).

7.16 Nucleosome Remodeling Assays and Rate Analysis for Validation of Acidic Patch Mutant Nucleosomes (H2A E61A, D90A, E92) Using an Electrophoretic Mobility Shift Nucleosome Repositioning Assay

Octamers and nucleosomes were prepared similarly to previously described (Dyer et al., 2004) with minor modifications. Nucleosomes were assembled via salt gradient dialysis on a 227 b.p. DNA fragment that was prepared similarly as in Section 7.8.3. In this case, primers were designed to amplify the 601 sequence with a 80 b.p. overhang generated from the surrounding vector sequence.
Forward – 5’ - CAC TAT AGG GCG AAT TGG AG – 3’
Reverse – 5’ - CTG GAG AAT CCC GGT – 3’

DNA fragment generated:

CACTATAGGGCGAATTGGAGCTCCACCAGCGGCTTGCGCCGCTCTAGAACTAGTGATCCG
ATATCGCTGTACCCCGCTGACAGGATGTATATATCTGACACGTGCCTGGGAGACTAGGGA
GTAATCCCCCTTGGCGGTAAAACGCGGGGCCAGCGCGCTACGTCGCTTATAAGCGGTGCTA
GAGCTGTCTACGAACATCGGACGCTGACGCACCGGGATTCTCCAG

601 sequence – bold

50 µL remodeling assays were carried out in assay buffer (12 mM HEPES, pH 7.9, 4 mM Tris, pH 7.5, 60 mM KCl, 10 mM MgCl₂, 10% glycerol, and 0.02% (v/v) IGEPAL CA-630) with 30 nM ACF complex and 90 nM nucleosomes. Reactions were pre-incubated for 10 minutes at 30 °C prior to initiation by the addition of nucleosomes. Note that ATP (final concentration equivalent to 2 mM) was not added until approximately 2.5 minutes prior to initiation. After initiation, remodeling assays were carried out for 1 hour at 30 °C. 6 µL time points were taken at 1, 2.5, 5, 15, 30, and 60 minutes and each was quenched by addition of 6 µL of quench buffer (assay buffer with 800 ng/µL sheared salmon sperm DNA (ThermoFisher Scientific)) and placement on ice. Quenched assay samples were directly run on a 5% TBE gel in 0.5x TBE buffer for 40 minutes at 200V. Nucleosomes were visualized by staining with SYBR® Gold Nucleic Acid Gel Stain (ThermoFisher Scientific). Densitometry measurements were performed to quantify the movement of nucleosomes away from their initial position using Image Studio Lite (LI-COR). Rates were determined using GraphPad Prism by fitting to a single exponential decay equation (Figure 4.2b).
7.17 Processing of Sequencing Data for Chromatin Remodeling Experiments Using the Nucleosome Library

As previously noted, each remodeling experiment (per enzyme) generated 36 samples (triplicate reactions with and without ATP with 6 time points collected per reaction). Individual remodeling experiments were processed as follows: 67 b.p. single-end and 8 b.p. index reads were imported into Galaxy (Princeton University installation). Corresponding read pairs (single-end and index) generated from the same fragment were joined to generate single nucleotide sequences that contained both a unique nucleosome identifier barcode and related unique multiplexing barcode. All reads were then split into separate FASTQ files (36 in total) based on their unique multiplexing barcodes using Galaxy’s Barcode Splitter tool. Each FASTQ file was exported from Galaxy, and occurrences of each unique nucleosome identifier barcode within each file were counted using a custom R script. This resulted in individual data points composed of a distinct number of read counts associated with a unique nucleosome identifier barcode and a unique multiplexing barcode.

Also, included in the nucleosome library was a non-nucleosomal DNA fragment identical to BC-601 DNA except that it did not contain a PstI site (DNA Standard 1). It was carried with the sample from the initiation of remodeling experiments through Illumina sequencing, and used as an internal reference to normalize relative abundances of DNA between multiplexed samples. After normalization, sample read counts were organized by time point/experimental condition (with or without ATP) and nucleosome type. At this point, nucleosomes 71, 78, and 113 were removed from further analysis as their unique nucleosome identifier barcodes created an additional PstI site with the surrounding nucleotide sequence.
7.18 Processing of Sequencing Data for Binding Experiments Using the Nucleosome Library

Barcode sorting was performed as described in Section 7.17. To account for variation in abundance of individual library members, raw read counts for each nucleosome were first normalized to corresponding values from the input sample. Note that data from a single input sample was used for processing all pull-down experiments. To better visualize how nucleosome modifications or mutations affect binding, these values were then normalized to one of the unmodified nucleosomes in the nucleosome library (nucleosome 42, see Table 2.2) to calculate relative affinity measurements for all nucleosome library members.

7.19 Nucleosome Remodeling Rate Analysis for Library Experiments

For each enzyme, remodeling rate constants per nucleosome were calculated for experimental triplicates in GraphPad Prism. Also, included in the nucleosome library was a non-nucleosomal DNA fragment identical to BC-601 DNA (DNA Standard 2). Nucleosome remodeling data were fit to a 2-phase exponential decay equation (see below). The first phase ($k_{\text{fast}}$) was fixed as the rate of cutting of DNA Standard 2 in each experiment (fit to 1-phase exponential decay equation). The y-intercept was allowed to vary and the plateau was set to zero. Fitting to this model allowed the second phase ($k_{\text{slow}}$) to be determined as the rate of nucleosome remodeling (this parameter was left unconstrained) (see Figure 3.3 and below). Additionally, to estimate the minimum rate of remodeling capable of being determined by our analysis methods we averaged the bottom 1% of rates that had 95% confidence interval non-overlapping with zero and calculated the associated error. Any values below this threshold were set equal to it. This included cases where $k_{\text{slow}}$ was extremely slow (the curve very flat) and determined to be slightly negative. The remodeling rates of unmodified nucleosomes in the library were then averaged and the associated error was calculated. Log$_2$(fold-change) values relative to the unmodified nucleosome rate average and associated errors were then computed for each nucleosome in the
library. Singular value decomposition of the nucleosome library data was calculated for PCA with the Python module matplotlib.mlab (v. 1.4.3).

Fitting the remodeling kinetic data to a 2-phase decay was necessary to account for the presence of free BC-601 DNA for some members in the library – in particular nucleosomes containing polyacetylated H3 and H4 (Figure 3.3b). Use of a 1-phase exponential decay led to an overestimation of the rate constants (and a poorer fit) for several library members, due to the contribution of the free DNA which is rapidly cut by the PstI restriction enzyme, irrespective of remodeling activity (Figure 3a, c). Fitting the data to the 2-phase exponential led to a better fit for those library members. The first phase accounts for the free DNA, whereas the slower second phase reports on the remodeling rate of the nucleosome. Note, the vast majority of library members had minimal free BC-601 DNA and could be analyzed equally well using 1- or 2-phase decays (Figure 3e, f). For consistency, we used the 2-phase decay fit throughout. Contributions of free DNA to rate measurements were generally not thought to be dependent on systematic irreproducibility in the nucleosome assembly procedures used in this study, but typically on the relative affinities of certain modified histone octamers for nucleosomal DNA as observed by gel electrophoresis. For example, polyacetylated nucleosomes tended to show such behavior (Figure 3b). We therefore propose that the amount of free DNA in a nucleosome preparation must be carefully considered when fitting nucleosome remodeling data generated by a restriction enzyme accessibility assay to a 1-phase exponential decay equation. In the case, where an internal standard to measure the rate of cutting of non-nucleosomal DNA is present, (DNA standard 2 in our high-throughput nucleosome remodeling assay) data may be better modeled by a 2-phase exponential decay equation to aid in accounting for the presence of any free DNA. This proves advantageous in situations where large numbers of nucleosomes must be prepared and precious material is limited or if nucleosome assembly reactions prove difficult to optimize.


