ROLE OF LINKER DNA LENGTH AND LINKER HISTONE OCCUPANCY IN CHROMATIN COMPACtion

A Thesis in
Biochemistry and Molecular Biology

by

Michael B. Schubert

© 2015 Michael B. Schubert

Submitted in Partial Fulfillment of the Requirements for the Degree of Master of Science

August 2015
The thesis of Michael B. Schubert was reviewed and approved* by the following:

Sergei A. Grigoryev  
Professor of Biochemistry and Molecular Biology  
Thesis Advisor

Laura Carrel  
Associate Professor of Biochemistry and Molecular Biology

Gregory S. Yochum  
Assistant Professor of Biochemistry and Molecular Biology

James R. Broach  
Chair of Biochemistry and Molecular Biology

*Signatures are on file in the Graduate School.
ABSTRACT

Chromatin condensation underlies many important biological processes such as cell differentiation, cell division and senescence. In this work, reconstituted nucleosomal arrays with various DNA linkers were constructed and used in combination with electron microscopy and analytical ultracentrifugation to investigate the internal organization of condensed chromatin in vitro and the effects of linker histone levels on chromatin folding. The internucleosomal interaction pattern within reconstituted chromatin is shown to closely resemble that obtained with chromatin fixed in situ, suggesting that the reconstituted arrays correctly recapitulate the nucleosome folding in living cells. A negative correlation between nucleosome repeat length (NRL) and chromatin folding was consistently observed for chromatin arrays with parallel nucleosome interaction faces. For arrays with shorter NRLs, chromatin folding showed a strong rotational periodicity, whereas folding in arrays with longer NRLs was unaffected by internucleosomal rotational changes. Our data is also consistent with results of mesoscopic modeling predicting that in condensed chromatin, nucleosome association with linker histone that triggers linker DNA crossing (“stem” formation) tends to maintain the two-start zigzag topology despite generation of metastable interdigitated forms without a distinct fiber diameter. Furthermore, in chromatin arrays with sub-saturated ratios of linker histone, a tendency toward stronger long-range nucleosome interactions is observed, similar to that revealed in condensed metaphase chromosomes in situ. Taken together, these observations suggest that condensed chromatin preserves zigzag features and is compacted via processes that promote formation of multiple loops and bundles of the nucleosome zigzag chain, rather than longitudinal compaction into regular 30 nm fibers. Such looped chromatin forms could be modulated by the ratio of linker histone and histone charge modifications, thereby altering global and local chromatin architecture.
TABLE OF CONTENTS

List of Tables .............................................................................................................................. vi
List of Figures .............................................................................................................................. vii
List of Abbreviations ................................................................................................................... viii
Acknowledgments ....................................................................................................................... x

Chapter 1. Introduction to chromatin higher-order structure.................................................... 1
  1.1 Overview of chromatin higher-order structure ................................................................. 1
     1.1.1 Primary structure ....................................................................................................... 2
     1.1.2 Secondary structure ................................................................................................. 3
     1.1.3 Tertiary and quaternary structure ............................................................................. 5
  1.2 Regulation of chromatin higher-order structure .............................................................. 6
     1.2.1 Linker DNA ............................................................................................................ 6
     1.2.2 Linker histone occupancy ....................................................................................... 7
     1.2.3 Linker histone phosphorylation .............................................................................. 8
     1.2.4 Core histone tail interactions and modifications ..................................................... 9
     1.2.5 Core histone variants .............................................................................................. 10
     1.2.6 Non-histone proteins .............................................................................................. 11
  1.3 Conclusions ....................................................................................................................... 12

Chapter 2. Materials and methods for chromatin structural experiments................................. 14
  2.1 Generation of DNA constructs ......................................................................................... 14
     2.1.1 PCR modification of existing monomer constructs ................................................. 15
     2.1.2 Multiple cloning of monomers into 12-mers ........................................................... 17
  2.2 Isolation of core histones ................................................................................................. 18
     2.2.1 Isolation of erythrocyte nuclei .............................................................................. 19
     2.2.2 Isolation of core histones from erythrocyte nuclei ................................................. 19
  2.3 Reconstitution of chromatin arrays .................................................................................. 20
     2.3.1 Core array reconstitution ....................................................................................... 20
     2.3.2 Linker array reconstitution .................................................................................... 22
  2.4 Electrophoretic techniques .............................................................................................. 23
  2.5 Magnesium chloride self-association assays .................................................................... 24
  2.6 Analytical ultracentrifuge experiments .......................................................................... 25
  2.7 Electron microscopy experiments .................................................................................... 26
     2.7.1 Standard transmission electron microscopy ........................................................... 26
     2.7.2 Electron microscopy-assisted nucleosome interaction capture ............................... 27
     2.7.3 Cryo-electron microscopy ...................................................................................... 28
  2.8 Materials .......................................................................................................................... 29
     2.8.1 Reagents ................................................................................................................ 29
     2.8.2 Equipment .............................................................................................................. 30

Chapter 3. Rotational and translational nucleosome settings in chromatin structure ................. 32
  3.1 Rationale: nucleosome positioning and chromatin structure ........................................ 32
  3.2 Results ............................................................................................................................... 33
LIST OF TABLES

Table 2.1: Linker DNA sequences of constructs used in the assembly of biochemically defined reconstituted nucleosome arrays .................................................................................................................................. 14-15
Table 2.2: PCR primers used to generate constructs used in the assembly of biochemically defined reconstituted nucleosome arrays .................................................................................................................................. 15-16
LIST OF FIGURES

Figure 1.1: Diagram of interdigitation in a chromatin array .................................................. 6
Figure 2.1: Assembly of multimeric DNA templates for reconstitution into chromatin arrays ... 18
Figure 2.2: SDS-PAGE of purified core histones ......................................................................... 20
Figure 2.3: Agarose gel electrophoresis of core histone arrays .................................................. 21
Figure 2.4: Agarose gel electrophoresis and SDS-PAGE of linker histone arrays ....................... 23
Figure 2.5: Representative sucrose density gradient purification of a 188 bp NRL array .............. 23
Figure 2.6: Representative magnesium chloride self-association assay of a 200 bp NRL array... 25
Figure 3.1: Comparison of relative c(S) and van Holde-Weischet plots for the display of sedimentation coefficient (S) values .................................................................................... 34
Figure 3.2: Sedimentation coefficient (S) values of full-turn NRL chromatin arrays in relation to cation concentration ........................................................................................................ 35
Figure 3.3: Sedimentation coefficient (S) values of short, partial-turn chromatin arrays in relation to cation concentration ........................................................................................................ 36
Figure 3.4: Sedimentation coefficient (S) values of long, partial-turn chromatin arrays in relation to cation concentration ........................................................................................................ 37
Figure 3.5: Magnesium chloride self-association assay quantifications from chromatin arrays with full DNA helical turns .................................................................................................. 38
Figure 4.1: Sequence alignment and comparison between five linker histone variants ............ 43
Figure 4.2: Stages of native linker histone isolation from HeLa cells ........................................ 45
Figure 4.3. Comparison of chromatin folding at different NRLs with and without stoichiometric loading of human linker histone H1<sup>0</sup> ......................................................................................................... 47
Figure 4.4. Comparison of chromatin folding without and with stoichiometric loading of human linker histone H1<sup>0</sup> ......................................................................................................... 48
Figure 4.5. Comparison of chromatin self-association without and with stoichiometric loading of human linker histone H1<sup>0</sup> ......................................................................................................... 49
Figure 4.6. Diagram of nucleosome interactions ........................................................................ 50
Figure 4.7. Quantification of nucleosome interactions observed by EMANIC in arrays with 188 and 207 bp NRLs .................................................................................................................. 50
Figure 4.8. Overall compaction and self-association in linker histone-containing chromatin arrays with sub-stoichiometric ratios of H1<sup>0</sup> ......................................................................................................... 51
Figure 4.9. Increase in chromatin folding with increase in linker histone-to-nucleosome ratio in a 188 bp NRL array ........................................................................................................ 52
Figure 4.10. Increase in chromatin looping with increase in linker histone-to-nucleosome ratio in a 188 bp NRL array ........................................................................................................ 54
Figure 4.11. Diagrams of the core nucleosome and linker DNA in open and stem formations ... 56
Figure 5.1. Cryo-electron microscopy of linker histone arrays .................................................. 64
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ac</td>
<td>acetylation</td>
</tr>
<tr>
<td>bp</td>
<td>base pairs</td>
</tr>
<tr>
<td>CENP-A</td>
<td>centromere protein A</td>
</tr>
<tr>
<td>cryo-EM</td>
<td>cryo-electron microscopy</td>
</tr>
<tr>
<td>DMSO</td>
<td>dimethyl sulfoxide</td>
</tr>
<tr>
<td>DNA</td>
<td>deoxyribonucleic acid</td>
</tr>
<tr>
<td>DNP</td>
<td>deoxynucleoprotein</td>
</tr>
<tr>
<td>EDTA</td>
<td>ethylenediaminetetraacetic acid</td>
</tr>
<tr>
<td>E. coli</td>
<td>Escherichia coli</td>
</tr>
<tr>
<td>EM</td>
<td>electron microscopy</td>
</tr>
<tr>
<td>EMANIC</td>
<td>EM-assisted nucleosome interaction capture</td>
</tr>
<tr>
<td>G1</td>
<td>gap 1 phase of cell cycle</td>
</tr>
<tr>
<td>G2</td>
<td>gap 2 phase of cell cycle</td>
</tr>
<tr>
<td>H1</td>
<td>linker histone protein H1</td>
</tr>
<tr>
<td>H1.2</td>
<td>linker histone protein H1 variant H1.2</td>
</tr>
<tr>
<td>H1.4</td>
<td>linker histone protein H1 variant H1.4</td>
</tr>
<tr>
<td>H1^0</td>
<td>linker histone protein H1 variant H1^0</td>
</tr>
<tr>
<td>H2A</td>
<td>core histone protein H2A</td>
</tr>
<tr>
<td>H2B</td>
<td>core histone protein H2B</td>
</tr>
<tr>
<td>H3</td>
<td>core histone protein H3</td>
</tr>
<tr>
<td>H4</td>
<td>core histone protein H4</td>
</tr>
<tr>
<td>H5</td>
<td>linker histone protein H5</td>
</tr>
<tr>
<td>HE</td>
<td>HEPES-EDTA (buffer)</td>
</tr>
<tr>
<td>HEPES</td>
<td>4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (buffer)</td>
</tr>
<tr>
<td>HMG</td>
<td>high mobility group</td>
</tr>
<tr>
<td>HNE</td>
<td>HEPES-sodium-EDTA (buffer)</td>
</tr>
<tr>
<td>HP1</td>
<td>heterochromatin protein 1</td>
</tr>
<tr>
<td>I</td>
<td>interphase</td>
</tr>
<tr>
<td>LB</td>
<td>lysogeny broth</td>
</tr>
<tr>
<td>LH</td>
<td>linker histone</td>
</tr>
<tr>
<td>M</td>
<td>mitosis</td>
</tr>
<tr>
<td>Me</td>
<td>methylation</td>
</tr>
<tr>
<td>MENT</td>
<td>myeloid and erythroid nuclear termination stage-specific protein</td>
</tr>
<tr>
<td>MWCO</td>
<td>molecular weight cut-off</td>
</tr>
<tr>
<td>NP-40</td>
<td>Tergitol-type nonyl phenoxypolyethoxylethanol</td>
</tr>
<tr>
<td>NRL</td>
<td>nucleosome repeat length</td>
</tr>
<tr>
<td>OD</td>
<td>optical density</td>
</tr>
<tr>
<td>PBS</td>
<td>phosphate-buffered saline (buffer)</td>
</tr>
<tr>
<td>Acronym</td>
<td>Full Form</td>
</tr>
<tr>
<td>---------</td>
<td>-----------</td>
</tr>
<tr>
<td>PCR</td>
<td>polymerase chain reaction</td>
</tr>
<tr>
<td>PMSF</td>
<td>phenylmethylsulfonyl fluoride</td>
</tr>
<tr>
<td>PTA</td>
<td>phosphotungstic acid (phosphotungstate)</td>
</tr>
<tr>
<td>qPCR</td>
<td>quantitative polymerase chain reaction</td>
</tr>
<tr>
<td>RPM</td>
<td>revolutions per minute</td>
</tr>
<tr>
<td>RSB</td>
<td>reticulocyte standard buffer</td>
</tr>
<tr>
<td>S</td>
<td>synthesis phase of cell cycle</td>
</tr>
<tr>
<td>S value</td>
<td>sedimentation coefficient (or S)</td>
</tr>
<tr>
<td>SDS</td>
<td>sodium dodecyl sulfate</td>
</tr>
<tr>
<td>SDS-PAGE</td>
<td>sodium dodecyl sulfate polyacrylamide gel electrophoresis</td>
</tr>
<tr>
<td>TAD</td>
<td>topologically associating domain</td>
</tr>
<tr>
<td>TAE</td>
<td>Tris-acetate-EDTA (buffer)</td>
</tr>
<tr>
<td>TE</td>
<td>Tris-EDTA (buffer)</td>
</tr>
<tr>
<td>TEM</td>
<td>transmission electron microscopy</td>
</tr>
<tr>
<td>UA</td>
<td>uranyl acetate</td>
</tr>
</tbody>
</table>
I would like to thank all members of the Grigoryev laboratory, past and present, for their assistance with this project. In particular, I am grateful to Sergei, whose patient mentorship taught me more about chromatin, and about science as a whole, than I knew there was to know when I arrived in his lab. Sergei’s hard-won familiarity with the literature and seemingly endless ability to trace experimental mysteries back to their roots will most likely always serve as my template for the consummate scientist. I also owe my thanks to former lab member Sarah Correll, who taught me many of the techniques I needed for this work and remained positive even in the face of puzzling or disappointing results. Also in the lab when I arrived were Misha Gomberg and Sangita Chakraborty, both of whom were tolerant of my initial confusion and helped to make the laboratory a welcoming place to begin my graduate studies. As time passed, old graduate students were replaced with new, and current lab members Jenna Buckwalter and Abigail Harris Becker provided support and entertainment through departmental renovations, lab moves and equipment failures. Many hands make light work in the laboratory, and the assistance of students Tamreen Khan, Elena Klabukova, Beth Blaisse and Valentyna Kostiuk was very welcome during long hours of summer experiments.

I also owe my thanks to the members of my thesis committee, Laura Carrel and Greg Yochum, and to the faculty mentors who helped to make my graduate school experience as smooth as possible – Colin Barnstable, Tom Spratt, Ira Ropson and Sarah Bronson. Their concern for me and my studies made a sometimes frustrating path easier, and ensured that my experiments and presentations were always done to the best of my ability. When I had questions or needed assistance, they welcomed me and shared advice I couldn’t have gotten anywhere else. Not only these faculty members, but the entire Department of Biochemistry and Molecular Biology, helped me over the course of my graduate career – from writing letters to washing dishes, none of this work could have been accomplished without them.

My laboratory was lucky enough to have collaborators across America who contributed to our research. Tamar Schlick and Toni Luque at New York University provided us with modeling data, even altering their models to fit our exacting experimental specifications. Chris Woodcock at the University of Massachusetts, Amherst, provided the facilities we needed to prepare carbon-coated mica for application to our electron microscopy grids. Jeff Hayes and his laboratory at the University of Rochester provided us with an engineered human histone H1° plasmid for transfection, as well as with advice and encouragement on numerous occasions.

Finally, I owe a great debt to my family and friends, who encouraged and reassured me no matter how bright or dim my own outlook. I am grateful for the assistance of my friends at the graduate school – Richard Gill, Heather Young, Jessica Wagner, Nicole Briley, Lindsay Horvath, Sarah Arnold-Croop, and my entire entering-year cohort – and also for the unwavering support of my father Peter, my mother Susan, my brother Alex, and my partner, Giles Crawford.
Chapter 1

Introduction to chromatin higher-order structure

1.1 Overview of chromatin higher-order structure

Eukaryotic chromatin consists of DNA wrapped around a nucleosome core comprised of eight histone proteins (1). Nucleosomes are connected by linker DNA, which varies in length and sequence dependent on such factors as species (2) and tissue type (3) and which, in turn, affects the higher-order structure of chromatin. At the nucleosomal level, known as the primary structure of chromatin, the DNA forms a “beads on a string” configuration (4) that is then further compacted into secondary and tertiary structures.

Secondary chromatin structure, the formation of a fiber with diameter ~30 nm, can be compacted in two different nucleosome stacking conformations. In one, the “classical” ~30 nm fiber, there are approximately six nucleosomes per 11 nm; in the “heterochromatic” fiber, there are 12-15 nucleosomes per 11 nm (5). The former is the state most often observed in in vitro chromatin reconstitutes and reflects a largely unfolded state of euchromatin in the nucleus (6). The fiber itself exhibits a zigzag conformation with nucleosome cores at the periphery and linker DNA in the center (7). The greater packing density of the latter state is also seen in natural chromatin from various species (8) and in different regions of the compacted chromosome (9).

Tertiary chromatin structure involves interactions between chromatin fibers, resulting in assemblages that are denser by virtue of interdigitation (10) or possess diameters of greater than ~30 nm (11). These structures form in vitro both within and between single reconstituted chromatin arrays (12), indicating that they are involved in short- and long-range chromatin
interactions \textit{in situ}. Tertiary structure is seen to form irregularly in naturally-condensed heterochromatin of terminally differentiated cells (13), as well as in the mitotic/meiotic chromosome (14). During metaphase, in particular, tertiary and even higher-level interactions are seen in the chromosomes (15), though such structures can also be seen during other stages of the cell cycle (16).

1.1.1 Primary structure

Core histones are those around which the DNA wraps to form its first level of structure. There are four types – H2A, H2B, H3 and H4 – which together form an octamer known as the nucleosome core. The structure of this core histone octamer was solved by X-ray crystallography in the 1980s (17-19) and reached 1.3 Å by 1991 (20), showing that it is a tripartite assembly centered on an H3/H4 tetramer, with an H2A/H2B dimer on each side for a total diameter of ~65 Å. All four types of core histone contain the histone fold domain, a protein structure consisting of three $\alpha$-helices connected by loops; though all four folding patterns are similar, histones H2B, H3 and H4 resemble one another most closely, while H2A shows the greatest difference (21) (see Table 1.1). At the hydrophobic interfaces where the histones dimerize, sequence divergence is limited and dimerizations between certain histone pairs – for instance, H2A with H2B – are preferred. The dimerizations themselves are formed by the head-to-tail association of two histone fold domains, resulting in a curved outer surface where the positively charged histones interact with negatively charged nucleosomal DNA.

The nucleosome consists of the core histone octamer and the 145-147 bp of DNA that wraps around it. This structure, resolved to 2.8 Å in 1997 (22), has a molecular mass of ~206 kDa and repeats throughout the eukaryotic genome, allowing packaging and assembly into higher-order structures. Both the nucleosomal components themselves (for instance, the
sequence of the DNA) and factors from their environment (for instance, chemical groups modifying the histones, or ionic strength of the solution in which nucleosomes are contained) can affect their physical properties and, in turn, those of the structures they form (23). Each individual histone dimer of the octamer is associated with 27-28 bp of DNA, meaning that when the 4 bp linkers between dimers are included, histone fold domains account for the organization of 121 bp of DNA. The remainder comes from a backbone segment that extends out from each dimer for 34-36 bp, though these segments overlap with one another to yield the final 145-147 bp total (22).

The final component of the nucleosome is the DNA of the entry/exit and linker regions. In 1995, this DNA was determined by cryo-EM and orientation analysis to have an entry/exit angle of $33 \pm 19^\circ$ in the absence of linker histone (24); the DNA “arms” do not cross at the entry/exit site due to significant bending required to fit the linker DNA paths. Beyond the entry/exit region, linker DNA becomes a significant determinant of chromatin conformation, an effect dependent on many factors including sequence, length and external architectural factors reported on in this work.

1.1.2 Secondary structure

Once the primary structure of chromatin is established, it can further compact into secondary structures. The most widely known of these is the “30 nm fiber,” a structure that places the nucleosomes at the periphery with the linker DNA in the center. The precise conformation of the fiber is still poorly understood. It was first proposed as a solenoidal (one-start helix) construct wherein the primary structure supercoils to form the filaments (25), but shortly after this initial suggestion, an alternative zigzag (two-start helix) structure was suggested (26,27). Though several variations on these structures were later proposed, the true nature of the
~30 nm fiber remained a mystery until Schalch et al. (28) published the structure of a tetranucleosome array at a resolution of 9 Å in 2004. This study showed the formation of a truncated two-start helix with linker DNA in a zigzag conformation incompatible with a solenoidal structure. Subsequent investigations agree with this result (29,30). However, despite the seeming resolution of the conflict, some debate still exists – Robinson et al. were able to produce ~30 nm fibers in vitro whose dimensions were inconsistent with simple two-start models (31), and thus proposed that the presence of linker histone might regulate the difference between the looser (5-7 nucleosomes per 11 nm) zigzag and tighter (11-15 nucleosomes per 11 nm) solenoid conformations (32).

Even the existence of the ~30 nm fiber in living cells is a well-known controversy. First identified in the interphase nuclei of Triturus newts (33), these fibers have been observed in the nuclei of a variety of terminally differentiated cells (34-37); however, examination of proliferating cells has revealed no similar structures. It is possible that, in the course of chromatin higher-order structural changes that occur during mitosis (38,39), the fiber is lost in favor of more condensed conformations such as the 100 and 250 nm fibers (40,41). While it is alternatively possible that the structure is disrupted by sample preparation, a range of techniques (40,42-46) have failed to provide evidence of the ~30 nm fiber and instead offer evidence supporting higher-order organization into a “molten globule” (47) or other non-fiber structure. For this reason, much of the body of work analyzing chromatin secondary structure has been conducted either in vitro using reconstituted chromatin arrays, or in situ in cells sharing particular characteristics – low levels of transcription, low levels of non-histone architectural proteins, high levels of linker histone, and high nucleosome repeat lengths (NRLs) – in which chromatin fibers can be observed in situ. The most commonly used of these cell types is the
avian erythrocyte, though vertebrate retinal cells and echinoderm sperm also exhibit fiber formation.

1.1.3 Tertiary and quaternary structure

Chromatin conformation beyond secondary structures involves the interactions of those structures with one another. However, doubt over the existence of the ~30 nm fiber means that the existence of a rigid hierarchy of higher-order structures is also called into question. That tertiary structures exist is known from observations of chromosomes in situ (40,43,48), but their nature is not yet fully understood. Self-association of nucleosomes may lead to fiber-like structures with diameters greater than ~30 nm, or to dense conformations that result from the interdigitation of fibers with one another (see Figure 1.1).

Figure 1.1. Diagram of interdigitation in a chromatin array. Nucleosomes in each half of the array are indicated by colored circles (pale yellow or green). Linker histones are indicated by black circles. Nucleosomes can be seen interdigitating with one another at the center of the loop structure.
Tertiary structure forms *in vitro* both within reconstituted arrays of sufficient length (generally a minimum of 12 nucleosomes in a single array) and between arrays in solution, indicating that this level of compaction is involved in short- and long-range chromatin interactions *in situ*. The structures can also be seen in the condensed chromatin of terminally differentiated cells (13), as well as throughout the cell cycle in native mitotic and meiotic chromosomes (14-16). A study by Shopland *et al.* demonstrated that the higher-order structure of chromatin in the chromosome reflects the functional association of genes in the nucleus, showing that these areas tend to cluster together in the interior of the nucleus, while gene deserts are relegated to its periphery (49). Complex structures accommodated multiple functions that may be mechanistically connected, without relying on factors such as active transcription or proximity to a common nuclear site. Thus, the study concluded that chromatin tertiary and higher-order structure is dynamic and related to both gene sequence and function.

### 1.2 Regulation of chromatin higher-order structure

The higher-order compaction of chromatin is controlled by a variety of factors. Chief among them are the linker DNA length and sequence (which can control nucleosome rotational and translational settings), linker histone presence and occupancy, post-translational modification of the core and linker histones, and the presence and occupancy of non-histone architectural proteins.

#### 1.2.1 Linker DNA contributions

Linker DNA is a major contributor to the secondary structure and overall compaction of chromatin. “Short” linkers (<~40 bp) hold the nucleosome at a relatively fixed angle, as the
linker is not long enough to permit a high degree of internucleosomal flexibility. Thus, in chromatin whose NRL completes a full DNA helical turn between nucleosomes – that is, any NRL whose length is approximately \( x + 10.5n \), where \( x \) is the number of base pairs in the nucleosome positioning sequence (usually \( \sim 146 \)) \( 10.5 \) is the number of nucleotides required to complete a 360° turn (as each nucleotide contributes \( \sim 36° \) of rotation), and \( n \) is the number of completed turns – the nucleosomes are held in approximately parallel orientations to one another and are thus capable of complete cation-induced folding without contribution from any additional factors. In contrast, “long” linkers (>\( \sim 40 \) bp) permit flexibility between neighboring nucleosomes, meaning that even in chromatin whose NRL completes a full DNA helical turn, neighboring nucleosomes may not always lie parallel to one another. Additionally, the negative charge of such a long DNA linker is capable of overwhelming the positive charge of the nucleosome core, leading to repulsion and incomplete chromatin folding unless assisted by additional factors (50,51).

1.2.2 Linker histone occupancy

Chromatin may incorporate linker histones (H1 or H5) during compaction. Dependent on nucleosome repeat length, in terms of its effect on both rotational and translational setting of the core nucleosome, linker histone may be necessary for full compaction. In cases where the repeat length is long (\( \sim 188 \) bp or more) or does not result in a complete helical turn between nucleosomes, linker histone is required to mediate the transition from open to folded chromatin (52).

In the heterochromatic fiber, increased compaction occurs in the presence of linker histone. Addition of linker histone H5, which exhibits functionally similar behavior to H1 (53), has been shown to increase compaction of chromatin arrays with a 207 base pair nucleosome
repeat length *in vitro* in a size-independent manner. A similar effect is seen with the addition of linker histone to arrays with altered rotational settings, demonstrated in arrays with a 172 bp \((x + 10.5n + 5)\) repeat length. Chromatin possessing short nucleosome repeat lengths following the \(x + 10.5n\) pattern to yield complete DNA helical turns between nucleosomes is capable of condensing fully without the contribution of linker histone; however, chromatin that does not fulfil both requirements is not capable of full condensation without linker histone (52).

Naturally occurring linker histone may occur at a ratio of one or more molecules per nucleosome in terminally differentiated cells such as human glia or avian erythrocytes (54). However, in proliferating cells, it is frequently seen at a ratio of less than one molecule per nucleosome. When chromatin in such a cell is compacted, as in the metaphase chromosome, it reaches a high degree of condensation and coiling; thus, it is reasonable to hypothesize that changes to linker histone occupancy affects the capability of chromatin to form higher-order structures and reach full compaction.

1.2.3 *Linker histone phosphorylation*

The phosphorylation state of linker histone H1 changes during the cell cycle along with the overall compaction of chromatin. Levels of H1 phosphorylation are lowest in G1, when chromatin is most open, and increase steadily throughout the S and G2 phases to peak during mitosis (55). The degree to which H1 is phosphorylated is thus directly proportional to the degree of higher-order structure exhibited by chromatin. For numerous variants of H1, including those of the seven major mammalian variants relevant to studies performed in different cell types, the number and location of phosphorylation sites on the linker histone are known for each stage of the cell cycle. H1 phosphorylation during the cell cycle is site-specific and divided into *I sites*, which are phosphorylated even during interphase, and *M sites*, which are phosphorylated
only during mitosis. On H1.2, the I site is serine\textsuperscript{173} and M sites are threonine\textsuperscript{31}, threonine\textsuperscript{146} and threonine\textsuperscript{154}. On H1.4, I sites are serine\textsuperscript{172} and serine\textsuperscript{187} and M sites are threonine\textsuperscript{18}, threonine\textsuperscript{27} and, as on H1.2, threonine\textsuperscript{146} and threonine\textsuperscript{154} (56).

1.2.4 Core histone tail interactions and modifications

The positively charged N-terminal tail domains of core histones are key components of their interactivity. These tails, along with the C-terminal tail domain of H2A, are known to interact with one another to create and stabilize chromatin secondary (57) and tertiary (58) structures. The histone tails can mediate a variety of internucleosomal interactions; however, key among them is the interaction between the H4 tail and the H2A/H2B acidic patch. Other significant interactions that contribute to overall chromatin structure are the binding of the H3 tail to linker DNA, thus contributing to the negation of its charge (59), and the binding of the H2A C-terminal tail to linker histone (60).

The core histone tails are susceptible to a range of post-translational modifications. These include acetylation, phosphorylation and methylation, as well as less common changes like deimination. Negatively charged acetyl groups neutralize the positive charge of the lysines to which they bond, destabilizing chromatin architecture; these modifications may also be present in the globular core of the nucleosome and act as transcriptional coactivators (61). Phosphorylation, also a negatively charged modification, functions in a similar fashion, albeit more dynamically than acetylation. Most, but not all, core histone phosphorylation takes place on the histone N-terminal tail (62). Unlike acetylation and phosphorylation, methyl groups do not alter the charge of the histone protein; however, they do affect other characteristics of the target lysine or arginine residue, such as hydrophobicity. These modifications contribute to the recruitment of chromatin-regulating factors and can serve as markers for either active
(H3K4Me3, H3K36Me3) or repressed (H3K9Me2/3, H3K27Me3) genes (63). Deimination converts a positively charged arginine residue to a neutral citrulline, decreases methylation (64), and can induce rapid decondensation of the associated chromatin (65). The dynamic nature of many of these modifications indicates a situational role in their regulation of chromatin structure.

1.2.5 Core histone variants

In addition to the canonical histone proteins, there exist a number of variants that affect chromatin structure and stability. The major variants include those of histone H2A (including macroH2A, H2A.Bbd, H2A.W, H2A.X and H2A.Z) and histone H3 (H3.3 and CENP-A). MacroH2A and H2A.W both contribute to chromatin condensation; the former is well-known as a transcription repressor and also reduces histone acetylation in a form of indirect repression (66), whereas the latter is a heterochromatin marker critical for chromatin condensation (67). Both H2A.Bbd and H2A.X maintain open chromatin and active transcription. H2A.Bbd is particularly associated with rapid transcription because, having only ~50% identity with canonical H2A, it lacks the C-terminal domain, acidic patch, part of the docking domain and many lysine residues, rendering it unable to effectively compact chromatin (66). H2A.X becomes fully effective upon phosphorylation of its C-terminal tail (commonly signified by the term γH2A.X), which occurs at the sites of DNA double-strand breaks. This modification destabilizes the nucleosome and prevents histone H1 binding, maintaining an open chromatin environment in which DNA repair can take place (68). Finally, H2A.Z, a variant with ~60% identity to the canonical histone, influences internucleosomal interactions through its enhanced acidic patch. While this variant promotes chromatin folding, it may simultaneously inhibit self-association (69), possibly by packing nucleosomes too closely in parallel to permit higher-order interactions. Histone H3.3 has a highly conserved role as a marker of enhancers and sites of
nucleosome displacement in transcriptionally active chromatin, where it significantly impairs folding to prime genes for transcription (70,71). Centromeric H3, known as CENP-A in mammalian cells, is a specialized variant of H3 that defines the unique chromatin structure of the centromere and allows it to self-propagate. In all cases where variants deviate significantly from the canonical, they play a role in the regulation of chromatin structure and function.

1.2.6 Non-histone proteins

But histone variants are not the only proteins with an effect on structure. Many non-histone proteins can alter the architecture of chromatin, some more notably than others. HP1 (heterochromatin protein 1), for instance, can interact with core histones H3 and H4 (72) and shares functions with linker histone H1 (73), suggesting that the two proteins compete for binding and contribution to the condensation of heterochromatin. MENT (myeloid and erythroid nuclear termination stage-specific protein), which replaces HP1 in terminally differentiated non-mammalian cells such as erythrocytes and granulocytes, also mediates higher-order chromatin compaction (45).

The HMG (high mobility group) proteins are a superfamily of chromatin architectural factors with a wide range of functions. They can be separated into three families – HMG A, HMGB and HMG N – all of which share chromatin structural influences. HMG proteins compete with histone H1 for binding (74), thereby preventing chromatin compaction and allowing access to other proteins, such as transcription factors (75-77). Another protein family with a global effect are the polycomb group complexes; rather than compete for binding as HMG proteins do, however, these complexes function as repressors by post-translational modification, (78), broadly targeting and introducing histone H3 methylation marks (79). Yet another group of proteins, the sirtuins, also repress chromatin through chemical modification, acting as
deacetylases of histones, as well as other proteins. Members of the sirtuin family contribute to the deacetylation of H4K16Ac and H3K9Ac (both modifications that inhibit chromatin fiber folding), thus enabling the formation of condensed heterochromatin (80). Sirtuins are also known to deacetylate H3K18Ac and H1.4K26Ac while promoting heterochromatin marks such as H3K9Me3 and H4K20Me1, leading to an overall repressive and condensing effect.

1.3 Conclusions

A wide variety of factors can affect chromatin structure – from the most basic sequences that make up the DNA and histone components of the nucleosome to complex interactions between post-translational modifications and external protein factors. When examining any individual contributor, it is important not only to understand its function, but also the potential interplay between it and other factors influencing the overall compaction of chromatin. Some affect only secondary or only tertiary structure, while some affect both, and still more can enhance or compensate for the effects of others. Although many of these contributors have been investigated, few at this point are fully characterized and still fewer are understood in the wider context of chromatin in vivo.

The goal of this work is to contribute to closing the significant gaps in understanding overall chromatin structure. Biochemically defined systems of reconstituted chromatin were designed and constructed using engineered plasmid DNA and native avian erythrocyte core histone proteins. These chromatin arrays were then used to assess the contribution of linker DNA length to nucleosome translational and rotational settings, and the effect of those settings on chromatin folding and self-association. Human linker histone H1 was also investigated in an
array with a 188 bp NRL designed to mimic the chromatin found in many human tissues. Initially, the contribution of H1 presence to chromatin compaction was examined, but subsequently, the effect of differing levels of occupancy (that is, varying the ratio of H1 molecules to nucleosomes) was also explored as a key contributor to changing degrees of condensation throughout the cycle of a proliferating cell. It is hoped that the experiments in this thesis will contribute to a framework of understanding and further investigating the nature of changes to chromatin structure throughout the life of a cell.
Chapter 2

Materials and methods for chromatin structural experiments

Procedures for construction of DNA templates for mononucleosomes and multimeric oligonucleosome arrays, their reconstitution with core and linker histones by salt dialysis, characterization by micrococcal nuclease mapping, agarose gel electrophoresis, restriction enzyme protection, analytical ultracentrifugation, and electron microscopy to verify the correct number and positioning of the nucleosome cores are as described below.

2.1 Generation of DNA constructs

DNA templates were designed using Widom clone 601 DNA (81), which positions nucleosomes with single-nucleotide precision, and adding specific lengths of linker DNA (see Table 2.1). Templates were first designed as monomers with the length of linker DNA altered through primer modification PCR. Regardless of linker DNA length, all monomeric templates were designed with XbaI restriction enzyme digestion sites at the 5’ end and SpeI and SphI sites at the 3’ end. Monomeric templates were ligated into the pUC19 vector (82), transformed into *Escherichia coli* DH5α (Invitrogen, #18265-017), and grown in the presence of carbenicillin.

<table>
<thead>
<tr>
<th>NRL</th>
<th>DNA Linkers</th>
</tr>
</thead>
<tbody>
<tr>
<td>165</td>
<td>AGATATC ATCCTGTGCAGCAGT</td>
</tr>
<tr>
<td>167</td>
<td>AGATATC ATCCTGTGCATGACT</td>
</tr>
<tr>
<td>169</td>
<td>AGATATC ATCCTGTGCATCTGACT</td>
</tr>
</tbody>
</table>
Table 2.1. Linker DNA sequences of constructs used in the assembly of biochemically defined reconstituted nucleosome arrays.

<table>
<thead>
<tr>
<th>Name</th>
<th>Purpose</th>
<th>Length (bp)</th>
<th>Sequence</th>
<th>Direction</th>
<th>Melting Temperature (°C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MHS001</td>
<td>Expand plasmid with 167 bp NRL to 172 bp NRL</td>
<td>27</td>
<td>ATC TCT AGA GAA TAT CCC GCC CTG GAG</td>
<td>F</td>
<td>61.3</td>
</tr>
<tr>
<td>MHS002</td>
<td>Expand plasmid with 167 bp NRL to 172 bp NRL</td>
<td>29</td>
<td>GGA GCA TGC ATG ACT AGT TAC ATG CAC AG</td>
<td>R</td>
<td>61.5</td>
</tr>
<tr>
<td>MHS003</td>
<td>Reduce plasmid with 207 bp NRL to 188 bp NRL</td>
<td>30</td>
<td>TTC TCT AGA CTA TAC GCG GCC GCC CTG GAG</td>
<td>F</td>
<td>67.1</td>
</tr>
</tbody>
</table>

2.1.1 PCR modification of existing monomer constructs

Polymerase chain reaction (PCR) modification of monomer constructs was achieved by selecting an existing template close in size to the desired construct and designing primers appropriately to add or remove the requisite number of base pairs (see Table 2.2).
<table>
<thead>
<tr>
<th>MHS004</th>
<th>Reduce plasmid with 207 bp NRL to 188 bp NRL</th>
<th>32</th>
<th>GGC GCA TGC CGC GAC TAG TAT GAA TTC GGA TC</th>
<th>R</th>
<th>67.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>MHS005</td>
<td>Alter plasmid with 167 bp NRL to 165 or 169 bp NRL</td>
<td>27</td>
<td>TCC TCT AGA TAT CCC GCC CTG GAG AAT</td>
<td>F</td>
<td>61.3</td>
</tr>
<tr>
<td>MHS006</td>
<td>Reduce plasmid with 167 bp NRL to 165 bp NRL</td>
<td>28</td>
<td>TGC ATG CAT GAC TAG TCG CAC AGG ATG T</td>
<td>R</td>
<td>61.4</td>
</tr>
<tr>
<td>MHS007</td>
<td>Sequencing plasmid with 200 bp NRL 3’ to 5’ across pUC18 MCS</td>
<td>20</td>
<td>TTT ATG CTT CCG GCT CGT AT</td>
<td>R</td>
<td>49.7</td>
</tr>
<tr>
<td>MHS008</td>
<td>Expand plasmid with 167 bp NRL to 169 bp NRL</td>
<td>28</td>
<td>GCA TGC ATG ACT AGT CAG ATG CAC AGG A</td>
<td>R</td>
<td>61.4</td>
</tr>
<tr>
<td>MHS009</td>
<td>Expand plasmid with 172 bp NRL to 174 bp NRL</td>
<td>26</td>
<td>TCT AGA GAT GAT ATC CCG CCC TGG AG</td>
<td>F</td>
<td>61.1</td>
</tr>
</tbody>
</table>

Table 2.2. PCR primers used to generate constructs used in the assembly of biochemically defined reconstituted nucleosome arrays.

The alterations were conducted on a region of the linker DNA that would have little effect on the functionality of the final construct, ensuring that the clone 601 sequence, restriction sites, and notable features were unaffected. Primers were designed to have sufficient length and homology both before and after the base-pair alteration to ensure successful binding, and were tested in silico using OligoCalc (www.basic.northwestern.edu/biotools/oligocalc.html) (83).

PCR conditions were determined by calculating the appropriate annealing and extending temperatures based on primer sequences, then further altering conditions based on trial and error. Parameters that were modified to produce the highest degree of amplification included annealing and extending temperatures, number of amplification cycles, presence and concentration of magnesium chloride (MgCl₂, a necessary polymerase co-factor), presence and concentration of dimethyl sulfoxide (DMSO, an enhancing agent that improves amplification of GC-rich content and reduces formation of secondary structures), and amount of template DNA used. Typical PCR
conditions were 30 cycles at a denaturing temperature of 94°C, an annealing temperature of 55°C, and an extending temperature of 72°C, followed by storage at 4°C. Recipes contained a minimum of 0.5 mM MgCl₂, but in the case of constructs more difficult to amplify, might use up to 3 mM MgCl₂ or 10% DMSO and be amplified for up to 35 cycles. “Hot-start” PCR and higher annealing temperatures (up to 62°C) were attempted for some especially challenging constructs.

Following a successful PCR, the ends of new constructs were removed by restriction digestion using XbaI and SphI. Next, the product was purified using a Promega Wizard SV Gel and PCR Clean-Up System (Promega, #A9281) and ligated into a pUC19 vector at 16°C overnight.

2.1.2 Multiple cloning of monomers into 12-mers

Expansion into a 12-mer template was achieved through a stepwise approach in which monomer plasmids were digested using either SpeI and SphI to yield vectors containing monomers, or XbaI and SphI to yield monomeric inserts (see Figure 2.1). Upon ligation of monomeric insert to the vector already containing one copy of the nucleosome positioning template, the SpeI sites of the vectors and the XbaI sites of the inserts were both abolished. This process was repeated until constructs existed as 12-mers.
Figure 2.1. Assembly of multimeric DNA templates for reconstitution into chromatin arrays. Stepwise protocol of digestion, insertion and ligation was repeated until the desired number of repeats of the monomer was attained.

Once the positioning templates were constructed in pUC19 plasmids, they were transformed into *E. coli* and grown in ~2 L of lysogeny broth (LB)-Lennox medium (Fisher, #BP1427) supplemented with carbenicillin; medium was distributed between 3 to 6 flasks containing no more than one-third flask volume of medium each. Plasmids were isolated using modified large-scale alkaline lysis procedures (84). All inserts were subsequently digested from the vector backbone using the restriction endonucleases XbaI and HindIII. The vector backbone was digested into smaller fragments (1120, 862, 692, and 19 bp) using DraI to serve as a carrier DNA (see reconstitution procedure). The only exceptions were 200x12, which used XbaI and KpnI, and 177x12, which used EcoRV at each end to excise the template.

2.2 Isolation of core histones
2.2.1 Isolation of erythrocyte nuclei

Fresh chicken whole blood was obtained from the Bell and Evans (Fredericksburg, PA, USA) chicken processing plant. Blood was stored on ice in a 1% sodium citrate solution in water for transportation, with no air remaining in the bottle to prevent clotting. After transportation, blood was immediately filtered through cheesecloth to remove clots and other unwanted debris, then separated into 50 mL aliquots. As soon as possible, aliquots were processed to isolate the erythrocyte nuclei.

Each nuclear isolation procedure was conducted entirely on ice or at 4°C. Individual aliquots were centrifuged and the undesirable fractions aspirated to remove serum and non-erythrocyte cells; erythrocytes were then pelleted, washed in 40 mL of PBS-1% sodium citrate, resuspended and washed in 40 mL of RSB buffer, resuspended in 40 mL of RSB-0.5% NP-40, and homogenized using a 40 mL Dounce tissue grinder (Sigma-Aldrich, #D9188). Subsequent pellet washes were carried out in 40 mL of RSB-0.5% NP-40-150 mM NaCl-1 mM PMSF, in order to prevent protease activity and remove unwanted heme and globin proteins. Once all such factors were removed (determined by the disappearance of the red color from the nuclear pellet), nuclei could be used immediately or stored in RSB-0.5% NP-40-1 mM PMSF-50% glycerol at -20°C for several weeks.

2.2.2 Isolation of core histones from erythrocyte nuclei

Core histone octamers were isolated from chicken erythrocyte nuclei by the selective removal of histones through ultracentrifugation under increasing ionic strength as described (7), and then loaded on a ~10 mL 5-30% sucrose gradient in 2 M NaCl-10 mM HEPES-0.1 mM EDTA buffer. Sucrose gradients were spun on the ultracentrifuge in a Beckman SW 41Ti
swinging bucket rotor, 4°C, 38 000 RPM, for 48 h. After centrifugation, the gradient was divided into 1 mL fractions that were analyzed by SDS–PAGE (see Figure 2.2). Fractions containing equimolar amounts of core histone proteins were dialyzed against at least 200 mL of buffer (10 mM HEPES, 2 M NaCl, 0.1 mM EDTA) for ~36 h to remove sucrose and stored at -10°C.

Figure 2.2. SDS-PAGE of purified core histones. Fractions that feature a high concentration of core histone proteins and equal amounts of each individual histone can be used for chromatin array reconstitution.

2.3 Reconstitution of chromatin arrays

2.3.1 Core array reconstitution

Reconstituted 12-mer (and 24-mer) arrays were assembled using DNA templates (described above) and chicken erythrocyte histones. Reconstitutions were performed essentially as described in (85), but in the presence of carrier DNA (86) to ensure proper loading. The carrier DNA was created through digestion of pUC19 vector backbone by DraI, and used at a template:carrier ratio of 2:1. Core histones and purified DNA were combined in a final mixture containing 2 M NaCl and 1 mM PMSF. Samples were dialyzed in 3500 MWCO membranes against 200 mL of buffer (2 M NaCl, 10 mM HEPES, 0.2 mM EDTA, 0.1% NP-40, 5 mM β-
mercaptoethanol, pH=7.5) for 2 h. The salt was then lowered to 1.5 M NaCl by addition of 66.67 mL of salt-free buffer (10 mM HEPES, 0.2 mM EDTA, 0.1% NP-40, 5 mM β-mercaptoethanol, pH=7.5) and dialyzed for 2 h, followed by dialysis against 1 M NaCl (achieved by adding 133.33 mL of salt-free buffer) for 3 h, and a 0.75 M NaCl dialysis (achieved by adding 133.33 mL of salt-free buffer) for 3 h (all other buffer component concentrations were kept constant). The salt concentration was then decreased to 0.5 M NaCl by adding 266.67 mL of salt-free buffer and dialyzed for ~18 h. Reconstitutes were next dialyzed against 200 mL of a buffer containing 5 mM NaCl, 10 mM HEPES, 0.2 mM EDTA, 0.1% NP-40, 5 mM β-mercaptoethanol, pH=7.5, for 3 h. Then the buffer was changed to 200 mL of one lacking NP-40 (all other buffer components remained the same) and dialyzed for an additional 3 h to remove NP-40. The reconstituted core arrays were then quantified by spectrophotometry (A260 nm) measurements, run on 1% Type IV agarose (DNP) gels to ensure proper histone loading (see Figure 2.3), and assayed through SDS-PAGE to ensure integrity of core histones.

![Figure 2.3. Agarose gel electrophoresis of core histone arrays. 1% Type IV agarose in HE buffer. Lanes: 1) Reference 188 bp NRL array, 2) 1 kb DNA ladder (New England BioLabs, #N3232), 3-4) underloaded arrays, 5) correctly loaded array, 6-7) overloaded arrays.](image-url)
2.3.2 **Linker array reconstitution**

In instances where linker histone was added to the chromatin arrays, additional reconstitutions were performed by mixing reconstituted core arrays with human linker histone H1\(^0\) at or below a molar ratio of one molecule per nucleosome. Linker reconstitutes were dialyzed against 200-250 mL of 0.5 M salt buffer (500 mM NaCl-0.1 mM PMSF- 0.025% NP-40) using 3500 MWCO membranes for 2 h at 4°C, followed by dialysis against 200-250 mL of low-salt buffer (5 mM NaCl-10 mM HEPES-0.1 mM EDTA-0.025% NP-40, pH=7.5) for ~18 h, followed by dialysis against 200-250 mL of “standard buffer” (5 mM NaCl-10 mM HEPES-0.1 mM EDTA, pH=7.5) for 4 h to remove NP-40. The linker arrays were then removed from dialysis and stored at 4°C after the addition of 0.1 mM PMSF.

Carrier DNA was removed from reconstituted arrays through purification on a ~10 mL 5-30% sucrose gradient in 10 mM Tris-1 mM EDTA-0.5 mM PMSF buffer. Gradients were centrifuged at 35 000 RPM, 4°C, for 8 h on the ultracentrifuge in the Beckman SW 41Ti swinging bucket rotor. After centrifugation, samples were divided into ~1 mL fractions and analyzed on a 1% agarose gel (see Figure 2.4), and fractions containing sufficient concentration of array with no carrier DNA (see Figure 2.5) were dialyzed against at least 200 mL of 10 mM HEPES-5 mM NaCl-0.1 mM EDTA, pH=7.5 buffer for ~36 h to remove sucrose. Samples were analyzed by SDS-PAGE to confirm histone integrity.
2.4 Electrophoretic techniques

Low electro-endosmosis agarose (Lonza SeaKem LE agarose, #50002) gels of 1% in Tris-acetate-EDTA (TAE) buffer (BioRad, #161-0743) were used to visualize histone-free DNA.
Gels were run for 40 minutes at 80 V. Medium electro-endo-osmosis Type IV agarose (Sigma-Aldrich, #A3643) was used to analyze oligonucleosome reconstitutes by deoxynucleoprotein (DNP) electrophoresis. For these gels, Type IV agarose was dissolved in HE buffer (20 mM HEPES, 0.1 mM EDTA) at a concentration of 1%. These gels were run at 80 V for 70-80 minutes and stained with ethidium bromide. Type IV agarose (Sigma-Aldrich, #A3643) gels were also used to analyze mononucleosomes and free DNA resulting from test restriction digestions of oligonucleosome reconstitutes. For these gels, Type IV agarose was dissolved in TAE buffer at 1% concentration. The gels were run at 80 V for 70-80 minutes. All agarose gels were poured at an approximate volume of 25 mL for 7 cm gels and 50 mL for 10 cm gels. Gels to be imaged were stained with ethidium bromide (~1 μg/mL) for 20 minutes before visualizing.

Polyacrylamide gels were run throughout the experiments with nucleosome arrays to assay the stability and concentration of histones. These gels were typically 15% acrylamide with 1.5 mm thickness. For gels where band quantification was required, the NIH program Image J was used (http://rsbweb.nih.gov/ij/). Gel bands were marked using the rectangular selection tool and, for each lane, a background reading was taken from the area surrounding the band and subtracted from the measured band intensity. In the case of magnesium precipitation assays, relative band intensities were averaged from a minimum of three gels.

2.5 Magnesium chloride self-association assays

The extent of self-associated material was assayed through selective precipitation in magnesium (58). Arrays with A260 nm=0.5 were incubated with increasing concentrations of MgCl₂ (Sigma-Aldrich, #M1028-1 ml) for 20 minutes on ice. Incubation tubes contained a
volume of array equivalent to 5 µL at 1 OD, a volume of MgCl₂ yielding the appropriate concentration, and ultra-pure water to a total volume of 10 µL per tube. After incubation, the reactions were centrifuged at 12 000 RPM, 4°C, for 10 minutes. Supernatants were removed and loaded on a 1% agarose gel, while pellets were resuspended in 6 µL of water, then had 2 µL of 25% glycerol-50 mM EDTA-0.5% SDS added, and were loaded on a 1% agarose gel. The percentage of DNA in the supernatant was determined by DNA band quantification using Image J software. The concentration of magnesium at which 50% of the reconstituted chromatin was precipitated was reported (see Figure 2.6).

**Figure 2.6. Representative magnesium chloride self-association assay of a 200 bp NRL array.** 1% Type I agarose in TAE buffer. Each lane is representative of one 1 mM incremental increase in concentration of MgCl₂ from 0 to 6 mM. Percent array in supernatant was calculated as a fraction of total band intensity in supernatant at 0 mM MgCl₂. The point at which 50% array remains in supernatant is the mid-transition point.

*2.6 Analytical ultracentrifuge experiments*
The extent of nucleosome array folding was assayed through sedimentation velocity experiments on the analytical ultracentrifuge (Beckman Optima XL-A). Chromatin array samples at a volume of 400 µL, with A260 nm=0.5 OD or higher, were run in 10 mM HEPES-0.1 mM EDTA, pH=7.5 buffer under varying ionic strength conditions (5 mM, 60 mM, 150 mM NaCl and 0.6 mM, 1 mM, 2 mM MgCl$_2$). Cells were balanced with 405 µL of a reference solution containing 100 mM of NaCl in 10 mM HEPES-0.1 mM EDTA, pH=7.5 buffer. Scan data were collected using ProteomeLab XL-A (Beckman Coulter) at a wavelength of 260 nm, while samples were spun at 20 000 RPM at 20°C for sufficient time to achieve 21 successive readings for each sample. Data analysis from the analytical ultracentrifuge runs was carried out using the continuous c(S) distribution model of the SEDFIT software (www.analyticalultracentrifugation.com) (88). Boundary analysis was also conducted by the method of van Holde and Weischet (89) using the program Ultrascan II (www.ultrascan2.uthscsa.edu/).

2.7 Electron microscope experiments

2.7.1 Standard transmission electron microscopy

Electron microscopy (EM) samples were made with ~25 µL of reconstituted arrays with A260 nm between 0.5-2.5 OD. Arrays were incubated at various salt conditions (5 mM NaCl, 150 mM NaCl, or 1 mM MgCl$_2$), and 0.1% glutaraldehyde at 4°C for ~5 h, followed by ~12 h dialysis against 200 mL of buffer (10 mM HEPES, 5 mM NaCl, 0.1 mM EDTA, pH=7.5) at 4°C using 10 000 MWCO membranes. Specimens were diluted 10-20 times with 50 mM NaCl and applied to carbon-coated and glow-discharged (20 mA for 20 s) copper EM grids (Electron
Microscopy Sciences, #T1000-Cu). Grids were stained with 0.04% uranyl acetate (UA) for negative staining and with 1% phosphotungstic acid (PTA) for positive staining. Dark-field images were obtained and digitally recorded using a JEM-1400 electron microscope (JEOL USA, #JEM-1400Plus) at 120 keV with an SC1000 ORIUS 832 11-megapixel CCD camera (Gatan).

Staining with UA was accomplished by the “droplet” method after adhering sample in an appropriately diluted 50 µL droplet to the carbon grid for 5 min. Grids were rinsed 3 times for 30 s each in 50 µL of magnesium acetate, then stained for 30 s in 50 µL of 0.2% UA. After staining, grids were rinsed 3 times for 30 s each in 50 µL of ultra-pure water. Finally, stained grids were dried under desk lamp illumination in a low-traffic area for a minimum of 5 min. To stain with PTA using the “droplet” method, sample in an appropriately diluted 50 µL droplet was adhered to the carbon grid for 5 min, after which grids were rinsed for 20 s in 100 mL of 0.05 mM sodium borate-0.4% Photoflo 200-water, pH=9.0. After rinsing, grids were stained for 30 s in a 50 µL droplet of 1% PTA in 75% ethanol, followed by rinsing in 50 µL droplets of 95% and 100% ethanol for 30 s each. As with UA staining, grids were dried for at least 5 min under desk lamp illumination before proceeding with electron microscopy imaging.

2.7.2 Electron microscopy-assisted nucleosome interaction capture

Chromatin arrays were prepared for electron microscopy-assisted nucleosome interaction capture (EMANIC) imaging by formaldehyde crosslinking. Approximately 30 µL at 1-2 OD of reconstituted nucleosome array, stored in 10 mM HEPES-0.1 mM EDTA, pH=7.5, plus appropriate salt concentrations to induce chromatin folding if necessary, was used per sample. Each array was incubated on ice for 15 minutes, then equilibrated to room temperature. 0.1% formaldehyde was added to each tube and incubated at room temperature for exactly 5 minutes,
after which the reaction was stopped using 40 mM glycine, pH=8.0. The crosslinked chromatin arrays were dialyzed in 200 mL of a solution of very low ionic strength (5 mM Na-borate, pH=9.0) for 4 hours. The unfolded nucleosome arrays were fixed with 0.1% glutaraldehyde for 16 h by addition to the dialysis solution.

For transmission electron microscopy of the partially crosslinked nucleosome arrays, positive staining with UA was employed along with dark-field mode imaging, a highly sensitive staining technique (90) that was modified for EMANIC by optimizing the thickness of the ultra-thin carbon coating (2-4 nm), mesh size of EM grids (1000 mesh), stain concentrations, and grid treatments and washing regimens. EMANIC analysis of in vitro crosslinked arrays was conducted essentially as described in (7). Dark-field EM images were obtained and digitally recorded at 120 keV using a JEOL JEM-1400 electron microscope with an SC1000 ORIUS 11-megapixel CCD camera. Images were collected at 20K, 25K, 30K, 42K, 52K or 67K magnification. For each sample, nucleosomal arrays containing 12 distinguishable nucleosome cores were selected to score internucleosomal interactions. Individual nucleosomes were overlaid with masks of nucleosome-size disks that were scaled to a diameter of 110 Å for each given magnification, and centered over each nucleosome and connected with lines tracing the underlying linker DNA. Internucleosomal interactions were scored as positive if the scaled nucleosome disks contacted each other. Standard deviations and standard error of mean were obtained from at least three EM experiments and at least two independently crosslinked samples; $p$ values represent probability associated with a Student’s two-sample unequal variance $t$ test with a two-tailed distribution.

2.7.3 Cryo-electron microscopy
Reconstituted chromatin array samples for cryo-EM were not fixed or stained in any way. Instead, samples were vitrified in solution at -196°C using liquid nitrogen to preserve their native three-dimensional structures, then imaged using a JEM-2100 electron microscope with a lanthanum hexaboride filament (JEOL, #JEM-2100 LaB6) and single-particle reconstruction analysis.

2.8 Materials

2.8.1 Reagents

100G), Igepal CA-630 (NP-40, #I3021-100mL), 1 M magnesium chloride (#M1028-1mL), and polyethylene glycol (PEG) (#P-2139). From Thermo Scientific (thermoscientific.com): PMSF (#36978). From VWR (www.vwr.com): Tris (#BH4500-1KGP) and type I agarose (Lonza SeaKem LE agarose, #50002).

2.8.2 Equipment

The equipment used for PCR was a Bio-Rad S1000 thermal cycler (#184-2000) with a dual 48/48 fast reaction module (#184-0148) and a 96-well fast reaction module (#184-0196). For agarose gel electrophoresis, the following Bio-Rad horizontal systems were used: Mini-Sub Cell GT systems (#170-4486, #170-4467) and Wide Mini-Sub Cell GT systems (#170-4485, #170-4469). For acrylamide gel electrophoresis, the Bio-Rad Mini-Protean 3 Cell system (#165-3301) was used; where Western blotting was conducted, the Bio-Rad Mini-Trans-Blot Cell system (#170-3930) was used.

For column filtration, Amersham/Pharmacia Hi-Trap SP cation exchange columns (#17-1151-01) were used. For small-volume sample dialysis (<0.1 mL), the following Thermo Scientific Slide-A-Lyzer MINI dialysis devices were used: 0.1 mL 3.5K MWCO (#PI-69552), 0.1 mL 10K MWCO (#PI-69572), floats (#PI-69588). For moderate-volume sample dialysis (0.1-0.5 mL), Thermo Scientific Slide-A-Lyzer 0.1 to 0.5 ML 3.5K MWCO dialysis cassettes (#PI-66333) were used. For large-volume sample dialysis (>0.5 mL), Spectrum Labs’ Spectra/Por 3 standard-grade 3.5K MWCO dialysis tubing (#132720) was used. For small-volume sample concentration, EMD Millipore’s Amicon Utra-0.5 mL centrifugal filters (#UFC5010) were used. For large-volume sample concentration, Pall Corporation’s Microsep Advance centrifugal devices with Omega membrane (#MCP010C41) were used. For small-volume phenol/chloroform extraction, 5 PRIME’s Phase Lock Gel Heavy 2 mL tubes
(#2302830) were used. For large-volume phenol/chloroform extraction, 5 PRIME’s Phase Lock Gel Heavy 15 mL tubes (#2302850) were used.

For general centrifugation, Eppendorf’s Centrifuge 5810 R (#5811 000.010) was used with fixed-angle rotor F-34-6-38 (#5804 727.002) and swinging bucket rotor A-4-62 (#5810 709.008). For preparative ultracentrifugation, Beckman Coulter’s Optima LE-80K (#365668) was used with fixed-angle rotor Type 70 Ti (#337922) and swinging bucket rotor Sw 41 Ti (#331362) with bucket set (#333790). For analytical ultracentrifugation, Beckman Coulter’s Optima XL-A was used with analytical rotor An-60 Ti (#361964). For electron microscopy, the following JEOL instruments were used: JEM-1400Plus with SC1000 ORIUS 832 11-megapixel camera, JEM-2100 LaB6.
Chapter 3

Rotational and translational nucleosome settings in chromatin higher-order structure

Portions of the data in this chapter were included in the paper published by Correll et al. (52). Specifically, data contained in Figures 3.2, 3.3, 3.4 and 3.5 of this work were also included in the aforementioned paper. Authors were Sarah Correll, Michael Schubert and Sergei Grigoryev. Of the work shown in this chapter, Michael Schubert was responsible for designing and constructing the 165, 169, 172, 174 and 188 bp monomers and the 167±2 trimer, and for expanding all of those constructs as well as the 200 bp monomer into 12-mer arrays. This author also isolated core histone, reconstituted and characterized the 165, 167±2, 169, 172, 188 and 200 bp 12-mer arrays and analyzed the data resulting from all experiments. Additional data for Figure 3.1 of this chapter were generated by Michael Schubert. Further information is included in the Appendix of this work.

3.1 Rationale: nucleosome positioning and chromatin structure

The positioning of nucleosomes is a key determinant of the structure of compact chromatin. Nucleosomes interact largely on the basis of positively charged N-terminal tails of the core histones, as well as the flexible C-terminal tail of H2A and other features on the nucleosome surface (22,91,92). In order for these features to contribute to nucleosome interactions, each nucleosome must be appropriately positioned – a factor affected by the nucleosomal DNA and, in particular, the linker DNA.
Prior experiments have shown that chromatin arrays with different NRLs have different folding requirements; for instance, arrays with short NRLs (167 bp) are capable of full compaction without contribution from environmental factors, whereas arrays with long NRLs (>188 bp) require contribution from linker histone (50), presumably in order to neutralize the additional negative charge of the longer linker DNA. In the absence of linker histone or other architectural factors, chromatin arrays with long NRLs are less compact than those with shorter NRLs under equivalent conditions (51). To examine the effect of precise changes in NRL on chromatin compaction, biochemically defined nucleosome arrays with NRLs from 167 (common in lower eukaryotes and actively transcribed tissues) to 209 (common in tissues of terminally differentiated systems) base pairs (94) were assembled in order to achieve specific variations in translational settings.

However, NRL does not determine only the translational setting of the nucleosome (that is, its distance from its nearest neighbor). Each additional nucleotide in the NRL adds ~36° of rotation to the system, meaning that it also affects each nucleosome’s angle in relation to its nearest neighbor (known as rotational setting). For this reason, arrays that affect only translational nucleosome settings occur at intervals of ~10.5 bp (360°) – in the case of this chapter, NRLs of 167, 177, 188 and 209. In order to observe the effect of the nucleosome rotational setting on chromatin compaction, arrays with NRLs featuring a partial (<360°) rotation of the nucleosome, including 165, 169, 172 and 174 bp, were examined.

3.2 Results

3.2.1 Chromatin folding as a function of translational setting
Experiments were conducted to determine the difference in chromatin secondary structure as a result of changes to the NRL. DNA templates were designed and purified, core histones were purified, and reconstituted chromatin arrays assembled as described in chapter 2, with NRLs of 167, 177, 188 or 207 bp to ensure fixed rotational setting. Folding was examined by sedimentation velocity experiments conducted under various sets of identical conditions. Figure 3.1 shows a comparison between relative c(S) and van Holde-Weischet plots, demonstrating that data for substances with a single sedimentation coefficient can be displayed with equal accuracy using either method, but that when multiple arrays or folding states are present in a sample, only the relative c(S) plot is capable of accurately displaying each sedimentation coefficient.

**Figure 3.1. Comparison of relative c(S) and van Holde-Weischet plots for the display of sedimentation coefficient (S) values.** Top: 167 bp NRL array reconstitute (52). Bottom: 167 and 207 bp NRL array co-reconstitute (52). Both plots are equally capable of displaying single sedimentation coefficients; however, when multiple substances with different S values are present in a sample, only the relative c(S) plot is capable of displaying each one separately.
Figure 3.2 shows that with increasing concentration of mono- or divalent cation (sodium or magnesium chloride), the sedimentation differences, and thus the degree of compaction, between arrays with different NRLs become more pronounced.

![Figure 3.2. Sedimentation coefficient (S) values of full-turn NRL chromatin arrays in relation to cation concentration.](image)

**Figure 3.2**. Sedimentation coefficient (S) values of full-turn NRL chromatin arrays in relation to cation concentration. At low cation concentrations (5 mM NaCl), all arrays show a similar degree of compaction; however, at higher cation concentrations, shorter NRLs show significant compaction while longer NRLs show much less. Data report average S values from at least 3 experiments; error bars represent the 95% confidence levels of the c(S) analysis. This data contributed to research paper (52).

### 3.2.2 Chromatin folding as a function of rotational setting in short-linker arrays

Having established the effect of translational setting alterations with fixed rotational settings, experiments were next conducted to determine the effect of nucleosome rotational setting on chromatin secondary structure. As before, DNA templates were designed and purified, core histones were purified, and reconstituted chromatin arrays assembled as described in chapter 2, this time with NRLs of 165, 169, 172, 200 and 205 bp. (A 174 bp NRL was also assembled, but no conclusive results were obtained.) Folding was examined by sedimentation velocity experiments conducted under various sets of identical conditions.
Figure 3.3 shows that with increasing concentration of mono- or divalent cation (sodium or magnesium chloride), sedimentation increases in those arrays whose NRLs most closely resemble a full-turn array – in this case, arrays with 165 and 169 bp NRLs, each only ~2 bp away from the classic 167 bp full-turn NRL. In the array with a 172 bp NRL, on the other hand, compaction increases far less with increasing cation concentration; the difference in S value between low-salt (5 mM NaCl) and high-compaction (2 mM MgCl₂) conditions is only 8 S, in contrast to the other two arrays, in which the difference is 20-21 S.

![Figure 3.3. Sedimentation coefficient (S) values of short, partial-turn chromatin arrays in relation to cation concentration.](image)

**Figure 3.3. Sedimentation coefficient (S) values of short, partial-turn chromatin arrays in relation to cation concentration.** At low cation concentrations (5 mM NaCl), all arrays show a similar degree of compaction; however, at higher cation concentrations, arrays with NRLs closer to a full 360° helical turn show significant compaction, while those with NRLs resulting in approximately a half-turn show much less. Data report average S values from at least 3 experiments; error bars represent the 95% confidence levels of the c(S) analysis. This data contributed to research paper (52).

3.2.3 Chromatin folding as a function of rotational setting in long-linker arrays
The effect of long, 200 and 205 bp NRLs on chromatin structure was also examined. For this analysis, sedimentation velocity experiments were conducted under various sets of identical conditions.

Figure 3.4 shows that with increasing concentration of mono- or divalent cation (sodium or magnesium chloride), sedimentation increases in those arrays whose NRLs most closely resemble a full-turn array – in this case, arrays with 200 bp NRLs, only ~1-2 bp away from a 198-199 bp full-turn NRL. In the array with a 205 bp NRL, on the other hand, compaction increases somewhat less with increasing cation concentration; the difference in S value between low-salt (5 mM NaCl) and high-compaction (2 mM MgCl₂) conditions is only 9 S. In general, in arrays with longer NRLs, the difference between those close to a full turn and those far from it is less pronounced than in arrays with shorter NRLs.

Figure 3.4. Sedimentation coefficient (S) values of long, partial-turn chromatin arrays in relation to cation concentration. At low cation concentrations (5 and 60 mM NaCl), both arrays show a similar degree of compaction, whereas at higher cation concentrations there is a significant difference between arrays with NRLs closer to a full 360° helical and those resulting in approximately a half-turn show less. In arrays with long (> 188 bp) NRLs, the difference is less pronounced than in those with short NRLs. Data report average S values from at least 3
3.2.4 Chromatin self-association as a function of nucleosome repeat length

Magnesium chloride self-association assays were conducted to determine the difference in chromatin tertiary structure as a result of changes to the NRL. Biochemically defined, reconstituted chromatin arrays were constructed as described above, with NRLs ranging from 165 to 209 bp. The assays were conducted as described in chapter 2, with magnesium chloride (MgCl$_2$) concentrations ranging from 0 to 6 mM in increments of 1 mM. The concentration of MgCl$_2$ at which 50% of the chromatin precipitated out of solution was recorded. For additional specificity, further assays were conducted to better determine this point, known as the “mid-transition point.” These assays used increments of 0.5 mM MgCl$_2$.

Experiments conducted on arrays with a full 360° DNA helical turn showed little difference in self-association behavior regardless of NRL (see Figure 3.5).
**Figure 3.5.** Magnesium chloride self-association assay quantifications from chromatin arrays with full DNA helical turns. Data are displayed as percent array remaining in supernatant. All mid-transition points occur between ~2.5 and 3.5 mM MgCl₂. Data report the average from at least 3 experiments; error bars represent the standard deviation of the 50% precipitation point. There are no statistically significant differences between the mid-transition points of rotationally fixed arrays. This data contributed to research paper (52).

When expanded to include arrays whose NRLs resulted in partial DNA helical turns, the results showed slightly more variation, but remained within a 2.0 mM range of MgCl₂ concentrations.

### 3.3 Discussion

The results presented in this chapter indicate that when the rotational setting of the nucleosome is fixed, there is a negative correlation between NRL and chromatin compaction. Sedimentation coefficient (S) values ranged from 38 S for a 167 bp array to 52 S for a 207 bp array at full (2 mM MgCl₂) compaction. This effect is likely due in part to the difference in charge, with shorter NRLs having significantly less negative DNA backbone charge requiring neutralization by the histones (81). It is known that nucleosome arrays with NRLs above ~200 bp require additional architectural factors in order to reach full compaction (50), and these may be responsible for neutralizing the additional negative charges of long NRLs. The decrease in compaction is also likely to be due in part to the increased length of the NRL’s providing greater flexibility for the array, so that the nucleosomes do not come as readily into the necessary contact – namely, the parallel orientation required for interaction between the histone H4 N-terminal tail and the histone H2A/H2B acidic patch (28,91,92).

When the rotational setting of the nucleosome is varied, however, the correlation between NRL and chromatin compaction becomes nonlinear. This is because nucleosomes that are fixed
in angles other than parallel to one another at short linker lengths, as seen in arrays like the one featuring a 172 bp NRL, will not be able to make the essential interaction required for compaction. The difference in folding between arrays with 167, 172 and 177 bp NRLs is striking and clearly indicates the importance of nucleosome rotational setting in arrays with short NRLs.

When arrays with longer NRLs, such as those at 200 and 205 bp, have rotational variation, the effect is not as pronounced. It is likely that the same flexibility preventing parallel nucleosomes from coming into contact as readily in long-NRL arrays as in short also increases the likelihood that those nucleosomes will come into contact in long-NRL arrays with rotational variation, especially if the degree of rotation is minimal. The increased negative charges of these NRLs also contributes to the decreased overall effect, as chromatin is less compact regardless of the nucleosome rotational state, thus resulting in a less significant difference between nucleosomes with and without rotational variation.

It appears, then, that the NRL directs chromatin into one of two types of higher-order structure: charge-dependent folding based largely on the length of the DNA backbone and heavily influenced by nucleosome rotational settings, or architectural factor-dependent folding based largely on the addition of linker histone and other proteins that fold chromatin without dependence on rotational settings. The former can be seen in reconstituted arrays with short NRLs; the latter, in arrays with long NRLs.

Neither the rotational nor the translational setting of chromatin arrays appears to impact their ability to self-associate and form tertiary or higher structures. It is therefore likely that the majority of chromatin compaction under these conditions is as a result of secondary structure. However, folding alone is insufficient to explain the extent of chromatin compaction in vivo. It is likely that the lack of observable effect on chromatin self-association in these experiments is due
to the fact that all of the arrays tested possessed a maximum of 12 nucleosomes and were unable to undergo significant self-association without the assistance of external architectural factors.
Chapter 4

Linker histone presence and occupancy in chromatin higher-order structure

4.1 Rationale: linker histone and chromatin structure

Both *in vitro* and *in vivo*, linker histone has demonstrated an effect on chromatin compaction. The contribution of linker histone to chromatin higher-order structure was proposed in 1995 by Blank and Becker, who used a cytoplasmic extract from *Drosophila* embryos to assemble regularly spaced nucleosomes onto plasmid DNA with an NRL of ~185 bp (95). They observed that, among other factors, linker histone was capable of altering the NRL in a gradual way, rather than in direct proportion to binding by the globular domain. *In vivo*, this observation is borne out by the fact that in actively proliferating cells, where the NRL tends to be shorter, the amount of linker histone present in the cell is correspondingly decreased (54,94), whereas when cells reach terminal differentiation and acquire a longer NRL, the amount of linker histone also increases (84).

Prior experiments have shown that linker histone has an effect on the formation of chromatin structures (34,79). However, this effect has never been fully characterized, so it is not yet known at what NRL linker histone is required for full compaction; it is also not known whether linker histone has the ability to overcome unfavorable chromatin characteristics such as an NRL that results in a half-helical turn.

Many *in vitro* experiments involving linker histone have been conducted using avian H5, which closely resembles the human linker histone variant H10. However, in many human tissues, including the commonly used HeLa cell line, the most common linker histone variants are H1.2
and H1.4 (85), which occur in approximately equal amounts. Figure 4.1 shows the alignment of
and similarity between five linker histone variants. Thus, experiments conducted using human
variants H1.2 and H1.4, or structurally similar linker histones (including mouse histone H1c or
rat H1.4), may provide more accurate insight into the behavior of linker histone-containing
chromatin in human cells \textit{in vivo}. For this reason, experiments were conducted to determine the
best way of isolating human linker histone variants from HeLa cell cultures.

\begin{verbatim}
Figure 4.1. Sequence alignment and comparison between five linker histone variants. Avian
linker histone H5, human histone H10, mouse histone H1c (identical to human histone H1.2),
human histone H1.4 and rat histone H1.4 are compared. Human and rat histone H1.4 bear the
closest resemblance to one another, while histone H1.2/H1c is similar to both. Histones H5 and
H10 are similar to one another, but bear little resemblance to the other three variants. Blue
sequences represent the N-terminal domains of each protein, while red sequences represent the
central globular domains and black sequences represent the C-terminal domains.
\end{verbatim}
Furthermore, *in vitro* experiments involving linker histone have typically been conducted using a “stoichiometric” ratio (that is, one molecule of linker histone per nucleosome core). In nature, however, various tissues in a diverse set of organisms exhibit a “sub-stoichiometric” loading (that is, less than one molecule of linker histone per nucleosome core). Wild-type liver, thymus and embryonic stem cell tissue in mice exhibit sub-stoichiometric linker histone loading, as do yeast cells and human neurons (83). It is significant to note that all of these tissues are composed of proliferating cells. Terminally differentiated cells exhibit either a stoichiometric (as seen in human glia) or even a super-stoichiometric (as seen in chicken erythrocyte chromatin) ratio of linker histone. This is initially paradoxical, as one would expect mitotic cells with highly condensed chromatin to exhibit high levels of linker histone; however, the affinity of linker histone to the mitotic chromosome is actually less than to interphase chromatin (86). In fact, mitotic chromosomes are capable of condensation even in the absence of linker histone (87).

In order to observe the effects of linker histone presence and occupancy on chromatin compaction, arrays with a 188 bp NRL mimicking that of many human tissues including the well-characterized HeLa cell line (88) were examined both with and without human linker histone H1\(^0\). In the case of arrays assembled with linker histone, the effects of both stoichiometric and sub-stoichiometric ratios were observed.

### 4.2 Results

#### 4.2.1 Linker histone isolation from HeLa cells

Linker histones H1 and H5 have previously been isolated from terminally differentiated cells, including the same chicken erythrocytes from which core histones were obtained (89).
Histone H1 has also previously been isolated from human HeLa cells. In this case, numerous native linker histone isolation procedures were conducted on HeLa cell cultures. For each procedure, 10 cell culture surfaces – either 100 mm dishes or 25 cm$^2$ flasks – were seeded with HeLa cells and grown to confluence. Linker histone isolation methods attempted include those previously used to purify chromatin architectural proteins (89) as well as those described by other research groups for the isolation of linker histones from HeLa and other cell cultures (50,89,91,92). Modifications to these methods included additional buffer washes, changes to ionic concentrations, and addition of ultracentrifugation steps to separate linker histones from other proteins with similar properties in solution. More commonly used acid extraction techniques were avoided wherever possible due to the potential negative effects of the acid on histone tails and post-translational modification. A representative outcome of these attempts is shown in Figure 4.2.

![Image](image_url)

**Figure 4.2. Stages of native linker histone isolation from HeLa cells.** 18% polyacrylamide 1.5 mm gel with electrophoresis at 11-22 mA, stained with Coomassie Brilliant Blue R-250. a) Crude nuclear isolate after buffer washing and centrifugation to remove small proteins and peptide fragments. b) Nuclear isolate after further purification. c) Supernatant after high-salt precipitation of non-histone proteins, with remaining linker histone, if any, too dilute for effective staining.

From these extraction attempts, it appears that native linker histone does not exist in large enough quantities in mammalian cells to be successfully isolated from small tissue culture
volumes without risking damage to the histone tail. To isolate native linker histone from mammalian cells, large tissue culture volumes may be necessary; otherwise, future researchers may benefit from overexpressing and subsequently isolating recombinant linker histone. Due to a repeatedly low yield of human linker histone from HeLa cell culture, further linker histone experiments were conducted with commercially available human linker histone H1\textsuperscript{0} (New England BioLabs, #M2501).

4.2.2 Effect of linker histone presence on chromatin compaction

The addition (as described in Chapter 2) of a stoichiometric ratio of linker histone to previously reconstituted chromatin arrays has an effect on their overall compaction (both folding and self-association) that ranges from negligible to highly significant. When linker histone is added to an array with a short NRL whose nucleosomes are oriented parallel to one another, such as the 167 bp array, it has little discernible effect on chromatin folding because these arrays have no need of contributions from external factors to fully compact (see Figure 4.3). When it is added to an array with a longer NRL, such as the 207 bp array, or one whose nucleosomes are not in parallel orientation, such as the 172 bp array, it has a much greater effect on overall folding because these arrays are unable to fully compact without assistance. Thus, stoichiometric linker histone compacts chromatin with longer DNA linkers to a much greater extent than chromatin with shorter DNA linkers (see Figure 4.4).
Figure 4.3. Comparison of chromatin folding at different NRLs with and without stoichiometric loading of human linker histone H1\(^0\). Folding is measured by sedimentation velocity experiments on the analytical ultracentrifuge and expressed as a sedimentation coefficient. The addition of linker histone affects the folding of the array with 167 bp NRL very little, the array with 188 bp NRL markedly, and the array with 207 bp NRL very markedly (top). Linker histone also increases the compaction of affected arrays regardless of ionic strength of solution (bottom). Data report average S values from at least 3 experiments; error bars represent the 95% confidence levels of the c(S) analysis.
Figure 4.4. Comparison of chromatin folding without (top) and with (bottom) stoichiometric loading of human linker histone H10. Folding is measured by sedimentation velocity experiments on the analytical ultracentrifuge and expressed as a sedimentation coefficient. Without the contribution of stoichiometric linker histone, the array with intermediate (188 bp) NRL folds to a greater extent than the array with long (207 bp) NRL under all tested conditions. With the contribution of linker histone, the cation-induced folding of the long-NRL array surpasses that of the short-NRL array. Data report average S values from at least 3 experiments; error bars represent the standard deviation between S values.

Unlike a simple alteration of nucleosome translational or rotational settings, the addition of linker histone has an effect not only on chromatin secondary structure, but on tertiary structure as well. Adding approximately one molecule of human histone H10 per nucleosome to arrays resulted in a significant decrease in the concentration of magnesium chloride required to
precipitate 50% of the array out of solution (see Figure 4.5), indicating an increase in chromatin self-association.

**Figure 4.5. Comparison of chromatin self-association without and with human linker histone H1\(^0\).** Numbers represent the range of MgCl\(_2\) concentrations, from 0.5 to 2 mM. Top panels represent array in supernatant, whereas bottom panels show array present in resuspended pellet. a) “Core” array with 188 bp NRL and no histone H1\(^0\), showing only a small percentage of precipitated array at 2 mM MgCl\(_2\). Core arrays typically exhibit a mid-transition point of ~3.5 mM MgCl\(_2\). b) “Linker” array with 188 bp NRL and stoichiometric loading of histone H1\(^0\), showing a mid-transition point of ~1.5 mM MgCl\(_2\), a much lower concentration than is required to precipitate 50% of the core array.

Examining the interactions between nucleosomes with the aid of electron microscopy (electron microscopy-assisted nucleosome interaction capture, EMANIC) reveals predominantly non-interacting nucleosomes (“monos”) or nucleosomes that interact with their nearest \((i\pm1)\) or next-nearest \((i\pm2)\) neighbors (see Figure 4.6).
Figure 4.6. Diagram of nucleosome interactions. From left to right: “mono” (no interaction), $i\pm1$ (nearest-neighbor interaction), and $i\pm2$ (next-nearest neighbor interaction), known as “doubles.”

These types of interactions are typical of chromatin folding, wherein nearby nucleosomes contact one another, whereas higher-order interactions ($i\pm3$ and above) are typical of “looping,” or self-association, wherein more distant nucleosomes interact. Interactions at $i\pm3$ and above are observed at low levels in core arrays, but do not increase significantly in proportion with the addition of stoichiometric linker histone (see Figure 4.7).
Figure 4.7. Quantification of nucleosome interactions observed by EMANIC in arrays with 188 and 207 bp NRLs. In both cases, interactions without linker histone and with the addition of stoichiometric human linker histone H1^0 are predominantly i±1 and i±2 (‘doubles’), characteristic of chromatin folding. The addition of linker histone increases total interactions (shown by a decrease in ‘monos’), but no statistically significant increase in looping. Data report average counts from at least 3 experiments with 80 to 540 particles counted per experiment; error bars represent the standard deviation between counts.

4.2.3 Effect of linker histone occupancy on chromatin compaction

Adding linker histone in ratios other than one molecule per nucleosome results in increased compaction correspondent to the amount of linker histone added. A ‘sub-stoichiometric’ ratio (in this case, 0.6 molecules of histone H1^0 per nucleosome) results in greater overall compaction than core arrays, but less than arrays with a stoichiometric ratio of linker histone. A ‘super-stoichiometric’ ratio (in this case, 1.2 molecules histone H1^0 per nucleosome) results in stronger self-association than core, sub-stoichiometric or stoichiometric arrays (see Figure 4.8).
**Figure 4.8. Overall compaction and self-association in linker histone-containing chromatin arrays with sub-stoichiometric ratios of H1\textsuperscript{0}.** a) Gel shift assay comparing arrays with core (0), sub-stoichiometric (0.6) and super-stoichiometric ratios of H1\textsuperscript{0}. b) Magnesium chloride self-association assay on core array. c) Self-association assay on sub-stoichiometric linker array. d) Self-association assay on super-stoichiometric linker array.

To better understand the nature of this increased self-association and overall compaction, sedimentation velocity and EMANIC experiments were conducted on human tissue-like intermediate (188 bp) NRL arrays featuring a range of sub-stoichiometric linker histone ratios. Sedimentation velocity experiments revealed a steady increase in folding correspondent to the increase in linker histone-to-nucleosome ratio (see Figure 4.9).

**Figure 4.9. Increase in chromatin folding with increase in linker histone-to-nucleosome ratio in a 188 bp NRL array.** Sedimentation velocity experiments were conducted at physiologically isotonic cation concentration (150 mM NaCl). Data report average S values from at least 3 experiments; error bars represent the standard deviation between S values.

EMANIC experiments allowed the visualization of short- and long-range interactions between nucleosomes in arrays with approximately 0, 0.2, 0.8 and 1.0 molecules of human H1\textsuperscript{0} per nucleosome. Upon scoring, these experiments revealed a significant decrease in the number
of non-interacting nucleosomes per array, with a corresponding increase in “loops,” or interactions at the $i\pm3$ or higher level indicative of chromatin looping and self-association (see Figure 4.10). The array with the most significant increase in looping interactions contained 0.2 molecules of H1$^0$ per nucleosome, although arrays with 0.8 H1$^0$ molecules per nucleosome also showed an increase in looping over core or stoichiometric arrays. Arrays with 0.1 H1$^0$ molecules per nucleosome showed a greater degree of looping than core arrays, but this increase was not significantly different to that of arrays containing stoichiometric ratios of linker histone.

Thus we show that sub-stoichiometric association of H1, while having only a small effect on chromatin compaction, notably increased the looping interactions between distal nucleosomes.
Figure 4.10. Increase in chromatin looping with increase in linker histone-to-nucleosome ratio in a 188 bp NRL array. Each array was fixed and decondensed for electron microscopy (top left). Arrays with sub-stoichiometric ratios of linker histone showed fewer non-interacting nucleosomes (“monos”) and more long-range interactions (“loops,” or interactions at the i±3 or higher level) than core arrays (top right; bottom). This effect was most pronounced in arrays with 0.2 molecules of H1⁰ per nucleosome. Data report average counts from at least 3 experiments with 80 to 540 particles counted per experiment; error bars represent the standard deviation between counts.

4.3 Discussion
In all cases, attempts to isolate native human linker histone H1 variants from HeLa tissue culture were unsuccessful. While some procedures based on previous research (87,93,96) were capable of retaining histone H1 in the final protein solution, these procedures also consistently retained non-histone proteins and other factors that made H1 purification or subsequent use impossible. Procedures that were capable of removing these non-histone factors also removed most or all of the H1, resulting in a solution too dilute and of too low a volume to allow experimental use of the linker histone. This inability to successfully isolate histone H1 from HeLa cultures stems partly from the fact that the protocols used avoided acid extraction and other techniques likely to damage chemical modifications to the histone, or the histone itself (97). However, the bulk of the problem arises from degree – that is, while investigators who were able to successfully isolate human histone H1 from HeLa cells used high-volume tissue culture (109), a maximum of 20 culture vessels were used at a time for the previously stated isolation attempts. For situations when culturing large volumes of HeLa cells is impractical, the engineered overexpression of a desired variant as in (98) should be considered in order to obtain the necessary amount of linker histone.

The results presented in this chapter indicate that linker histone plays a significant role in the compaction of chromatin arrays that are otherwise unable to fully compact – for instance, arrays with partial helical turns (such as the 172 bp array) or those with long linkers (such as the 207 bp array). The former cannot fully compact without the contribution of an external factor because the nucleosomes are oriented at non-parallel angles to one another and, due to the short linker length, lack the flexibility to reach the necessary positions for interaction. The former cannot compact without external factors because the length of the linkers introduces excessive negative charge, as well as flexibility, into the system. Because arrays with full helical turns and
short linkers (such as the 167 bp array) compact so efficiently without assistance, the addition of linker histone has little effect on these arrays’ overall compaction, whereas its effect on the compaction of arrays with long linkers or partial helical turns is significant and can be observed under all cationic conditions. References for this research are provided by the previous work of others (29,99), who have conducted structural work on arrays with 187 and 207 bp NRLs.

To assist in chromatin compaction, the linker histone binds at the DNA entry/exit site of the nucleosome and alters the conformation of the DNA at the site. Under the influence of linker histone, the DNA forms a “stem” motif (see Figure 4.11) (100). The closing of the DNA entry/exit site by inducing the formation of the stem, along with the neutralization of the negative charge on the linker DNA, allows H1 to increase the degree of chromatin folding in arrays that are unable to fold fully without external assistance.

Figure 4.11. Diagrams of the core nucleosome and linker DNA in open and stem formations. The open (left) conformation of the DNA entry/exit site arises in the absence of linker histone, whereas the stem (right) conformation is stabilized by the C-terminal domain of the linker histone (101).

Arrays with sub-stoichiometric linker histone show an increase in self-association as well as folding, in much the same manner as the metaphase chromosome of a mitotic cell (102), which shows a higher degree of condensation than any other point in the cell cycle (103).
likely due to long-range interactions in which the folded nucleosomes interdigitate, creating thicker fibers and other condensed structures.

However, linker histone occupancy is not the sole determining factor of chromatin condensation; the phosphorylation state of H1 also changes during the cell cycle. Levels of H1 phosphorylation are lowest in G1, when chromatin is most open, and peak during mitosis, when it is condensed (55). The degree of H1 phosphorylation is directly proportional to the degree of higher-order structure exhibited by chromatin. In the case of HeLa cells whose most common H1 variants are H1.2 and H1.4 (56), phosphorylation during the cell cycle is site-specific and does not occur at all instances of the (S/T)P(K/A)K consensus sequence. Rather, phosphorylation sites are divided into I sites (S<sub>173</sub> in H1.2; S<sub>172</sub> and S<sub>187</sub> in H1.4) phosphorylated even during interphase and M sites (T<sup>31</sup>, T<sup>146</sup> and T<sup>154</sup> in H1.2; T<sup>18</sup>, T<sup>27</sup>, T<sup>146</sup> and T<sup>154</sup> in H1.4) phosphorylated only during mitosis (56). The effect of this cyclical phosphorylation and condensation on cell cycle regulation remains incompletely understood, though it is hypothesized that H1 phosphorylation reduces secondary chromatin structure while increasing tertiary structure to allow chromatin accessibility during cell cycle progression.

The results obtained from H1<sup>0</sup> experimentation indicate that linker histone is required for full compaction of chromatin with either partial DNA helical turns or long NRLs. They also indicate that, in situations where linker histone is necessary, maximum chromatin compaction is achieved at sub-stoichiometric (less than one molecule per nucleosome) ratios, a result that closely adheres to the linker histone occupancy ratios observed in vivo in proliferating cells. In arrays with sub-stoichiometric ratios of linker histone, a greater proportion of condensation is achieved by chromatin “looping,” or long-range (i≥3 or greater) interactions between nucleosomes than in arrays with core or stoichiometric linker histone ratios – allowing increased
interdigitation and overall compaction. Future experiments should investigate the relationship between linker histone occupancy and chemical modification, especially phosphorylation, throughout the cell cycle.
Chapter 5

The future of chromatin structural studies

The results of the investigations presented in this work indicate that both the linker DNA and the linker histone have a significant role to play in the overall condensation of chromatin. The length of the linker DNA can affect both the rotational and translational settings of individual nucleosomes, as well as the stability of their internucleosomal angles, and thereby affect their ability to interact with one another and generate higher-order structures of chromatin. The presence and occupancy of linker histone also affects higher-order structure formation. In chromatin where linker histone is present, chromatin compacts to a greater degree due to the formation of linker DNA stem motifs that position the nucleosomes for interaction. When linker histone is present at sub-stoichiometric levels, as seen in many proliferating tissues, chromatin compaction is increased even further; this phenomenon results from the additional flexibility of regions with lower levels of linker histone occupancy, which allow the chromatin to interdigitate and form denser structures.

This is, however, only the beginning of studies into chromatin compaction and the factors affecting it. To obtain a more complete picture, it will be vital to consider both the length and the sequence of linker DNA, modifications of the linker histone as well as presence and occupancy, and the interplay of all of these factors in solution and, ultimately, in living cells.

5.1 Role of linker DNA length in chromatin dynamics, energies and topology
Linker DNA linker length affects not only the angles of the nucleosomes, but also other characteristics of the chromatin. Single-molecule force spectroscopy on long (15-mer 197 bp NRL or 30-mer 167 bp NRL) arrays was conducted by Meng et al. (104) to better understand the folding and unfolding of single chromatin fibers. By treating each nucleosome individually, the authors were able to resolve a metastable nucleosome conformation, quantify the compositional heterogeneity of individual fibers, and gain an understanding of each fiber’s unfolding mechanisms. From this, they discovered that fibers with a 197 bp NRL unfold differently from those with a 167 bp NRL. Additionally, 167 bp NRL arrays show approximately half the maximum extension at rupture force as 197 bp NRL arrays (7 nm versus 13 nm), and possess approximately four times the stiffness. These data are consistent with a zigzag folding conformation of the 167 bp NRL array and a solenoidal structure of the 197 bp NRL array – folding mechanisms that, because of the high level of reproducibility between experiments, appear well defined. It is not yet clear how the folding of these arrays would be affected by rotational alterations to the nucleosomes, or by variable linker lengths such as the 167±2 bp NRL construct previously mentioned in this work. Other experiments still to be conducted include an investigation of the effect of additional structural factors such as histone post-translational modifications might have on chromatin folding and dynamics.

Norouzi and Zhurkin (105) also revealed a topological difference between chromatin arrays with NRLs of the 10n (full helical turn) and 10n+5 (half helical turn) forms in arrays with NRLs ranging from 160 to 184 bp. These observations are consistent with previous experimental results, and so the authors made several predictions. They predict that the full and half helical turn fibers will show differences in flexibility, degree of DNA supercoiling, and level of transcription with relation to local NRL in the yeast genome, where NRL varies. Analysis of
biochemically reconstituted chromatin arrays of different long NRLs, including the 188 bp NRL array and other arrays with full and partial helical turns, will result in a more complete picture of chromatin topology at all NRLs.

5.2 Role of linker histone modifications in chromatin fiber relaxation

A recent paper by Lopez et al. (106) investigated the effects of partial phosphorylation of the linker histone on chromatin higher-order structure. This paper revealed a relaxation of the ~30 nm fiber as measured by sedimentation rate through a sucrose density gradient, and an increase in overall accessibility as measured by a technique involving micrococcal nuclease digestion and subsequent qPCR. Chromatin aggregation was induced by high concentrations of divalent cation (MgCl$_2$) and measured by dynamic light scattering. The authors observed that chromatin containing partially phosphorylated linker histone had impaired aggregation, indicating the onset of chromatin remodeling for greater chromatin accessibility during cell proliferation.

To make their analyses, Lopez et al. used chromatin samples digested from chicken erythrocyte and human embryonic kidney cell nuclei; these were partially phosphorylated ex vivo with CDK2-cyclinA kinase, which is specific to the linker histones in soluble chromatin. Investigations into other linker histone modifications has also historically made use of either soluble or in situ chromatin. Aubin, et al., for instance, studied the effect of H1 ADP-ribosylation on chromatin structure (107). They discovered that hyper(ADP-ribosyl)ation of histone H1 led to local relaxation of the chromatin fiber without loss of the histone, likely in order to facilitate DNA damage repair. This work, like the previous one, was conducted on chromatin isolated
from tissue (in this case, rat pancreatic cells); however, linker histone modifications have also been examined in cell culture – as in the case of the citrullination studies conducted by Christophorou et al. (108). Like phosphorylation and ADP-ribosylation, citrullination of histone H1 residue was shown to be correlated with chromatin decondensation.

However, examining higher-order structure using *in vitro* reconstituted chromatin arrays like those discussed in this thesis, with linker histone in controlled modification states, will allow future researchers to fully characterize the relationship between H1 modifications and chromatin condensation without contribution from other factors within or outside the cellular environment. This can be combined with an analysis of the relationship between such modifications and their effect on chromatin condensation to transcription factor binding and gene transcription levels, so that researchers may develop an understanding of the pathway between linker histone phosphorylation, chromatin structure and transcription during the cell cycle. Such an investigation will be particularly useful if conducted with a reconstituted chromatin array featuring a 188 bp NRL similar to those designed and constructed for the experiments in this work – but ideally containing histone proteins isolated from HeLa cells. It will be additionally useful if conducted using linker histone stoichiometry similar to that seen in proliferating human cells; that is, less than one molecule of H1 per nucleosome.

The knowledge gained from such an undertaking will permit the generation of a new model of chromatin higher-order structural interactions and the accessibility of condensed chromatin to binding factors; this, in turn, will yield insight into the genetic landscape of the cell throughout mitosis. With this information, it will be possible to further analyze changes in higher-order structure throughout the cell cycle, as well as investigating the accessibility and transcription of specific genes and regulatory regions.
5.3 Role of reconstituted chromatin arrays in *in situ* modeling

Biochemically reconstituted chromatin arrays are additionally useful for modeling *in situ* chromatin, so that researchers can analyze the behavior of chromatin in living cells. Arrays like this work’s 188 bp NRL are intended to mimic as closely as possible the attributes of native chromatin; however, the commonly used 12-mer array is too short to capture long-range interactions such as those that occur in metaphase chromosomes. While observing short-range interactions in the 12-mer, a longer 22-mer construct was also assembled, which is able to successfully recapitulate interactions in metaphase chromatin. The manuscript detailing the results of this work is currently in preparation (110). This and other long constructs (including 24-mer arrays with 167 and 207 bp NRLs, already assembled in the laboratory) can be used to investigate the effect of specific conditions and alterations on the condensation of chromatin *in situ*.

5.4 Role of cryo-EM tomography in analyzing DNA conformation in solution

Cryo-EM tomography is a method of observing the conformations of objects without the potential alterations induced by extensive sample preparation and fixation. In this case, chromatin arrays are vitrified in solution, then imaged by transmission electron microscopy in order to capture their behavior in that solution. This method was used to visualize the 12-mer 188 bp NRL array in solution, and will be used in future to gain a more complete understanding of chromatin conformation in solution under various conditions.
Figure 5.1. Cryo-electron microscopy of linker histone arrays. Representative cryo-EM images of a 12-mer 188 bp NRL array with stoichiometric histone H1$^0$, in 10 mM NaCl.

Recently, Song et al. (29) published a cryo-EM study of a 187 bp NRL 12- and 24-mer array, comparing its conformation to a similar array with a 177 bp NRL. Their three-dimensional cryo-EM structures revealed a two-start zigzag topology of the fiber, but the authors warn that this may not be true under all conditions – for instance, arrays with longer NRLs compacting with the assistance of linker histone and cation may show a solenoidal structure as seen by Meng et al. in 197 bp NRL arrays. The structures elucidated by Song et al. also indicated a previously unknown location of histone H1 on the chromatin fiber, resulting in asymmetric binding and interaction. This research permitted not only the generation of a cryo-EM map of the 187 bp NRL chromatin fiber, but also of a double-helix structural model for the array in the presence of histone H1.
5.5 Conclusions

In this work, new 12-mer chromatin arrays with various short and long NRLs featuring both full and partial helical turns were constructed. These arrays had different rotational and translational settings that permitted analysis of each attribute’s contribution to the overall folding and self-association of the chromatin fiber. The behavior of these arrays was observed under different environmental conditions (namely, concentration of mono- or divalent cation) and with and without the contribution of human linker histone H10, at ratios of greater than, equal to, and in particular less than one molecule per nucleosome. All of these factors are known to vary between tissues or between stages of the cell cycle; as such, all of the constructs and methods designed will allow the study of chromatin structure throughout the cell cycle to progress.
LITERATURE CITED


93. Khadake JR, Rao MR. DNA- and chromatin-condensing properties of rat testes H1a and H1t compared to those of rat liver H1bdec; H1t is a poor condenser of chromatin. Biochemistry. 1995; 34:15792.


APPENDIX


License #3653161315647 for use of this material in full for the purposes of thesis publication was granted by John Wiley and Sons on June 20, 2015.
This paper contains work by Sarah Correll, Michael Schubert and Sergei Grigoryev. Michael Schubert was responsible for designing and constructing the 165, 169, 172, 174 and 188 bp monomers mentioned in this paper, and for expanding all of those constructs as well as the 200 bp monomer into 12-mer arrays. This author also isolated core histone, reconstituted and characterized the 165, 169, 172, 188 and 200 bp 12-mer arrays and analyzed the resulting data.

Abstract

In eukaryotic cells, DNA is organized into arrays of repeated nucleosomes where the shorter nucleosome repeat length (NRL) types are associated with transcriptionally active chromatin. Here, we tested a hypothesis that systematic variations in the NRL influence nucleosome array folding into higher-order structures. For NRLs with fixed rotational settings, we observed a negative correlation between NRL and chromatin folding. Rotational variations within a range of longer NRLs (188 bp and above) typical of repressed chromatin in differentiated cells did not reveal any changes in chromatin folding. In sharp contrast, for the shorter NRL range of 165–177 bp, we observed a strong periodic dependence of chromatin folding upon the changes in linker DNA lengths, with the 172 bp repeat found in highly transcribed yeast chromatin imposing an unfolded state of the chromatin fiber that could be reversed by linker histone. Our results suggest that the NRL may direct chromatin higher-order structure into either a nucleosome position-dependent folding for short NRLs typical of transcribed genes or an architectural factor-dependent folding typical of longer NRLs prevailing in eukaryotic heterochromatin.
**Introduction**

In eukaryotic cells, the DNA is organized into chromatin structures. At the basic level of compaction, this involves wrapping \(\sim 147\) bp of DNA around an octamer of core histones to form the nucleosome core particle (Richmond and Davey, 2003). Nucleosome cores are connected by 10–70 bp of linker DNA forming nucleosome arrays (beads-on-a-string) that are further compacted with the aid of linker histone into higher-order chromosomal structures (Tremethick, 2007; Fussner et al, 2011; Li and Reinberg, 2011; Schlick et al, 2012).

One of the key determinants of chromatin structure is the nucleosome repeat length (NRL). The NRL includes the constant length of DNA wrapped around the histone octamer (\(\sim 147\) bp) and the variable length of linker DNA connecting adjacent nucleosome cores. Linker DNA varies in length between different tissues and in different organisms, producing a range of NRLs from the shortest 155 bp NRL found in fission yeast (Lantermann et al, 2010) to the longest measured NRL occurring in echinoderm sperm (\(\sim 240\) bp) (Athey et al, 1990). Linker histone levels in chromatin are correlated with the NRL (Pearson et al, 1984; Woodcock et al, 2006). There is also an inverse correlation between linker DNA length and gene activity, where actively transcribed genomes (e.g., yeast, embryonic stem cells, and tumour cells) tend to have shorter NRLs (\(\sim 160–189\) bp) and mature, transcriptionally inactive genomes (e.g. echinoderm sperm and chicken erythrocytes) have longer NRLs (\(\sim 190–240\) bp) (Thomas and Furber, 1976; Weintraub, 1978; Athey et al, 1990; Perisic et al, 2010). Biochemical analysis showed that the average NRL in yeast is quantized following the rule of \(10n+5\) bp per linker (Lohr, 1981); this
feature was also notable in the longer-NRL rat liver and chicken erythrocyte chromatin (Strauss and Prunell, 1983; Lohr, 1986). More recently, genome-wide studies of NRLs in yeast and human cells confirmed the quantized $10n+5$ bp spacing between yeast nucleosomes (Wang et al, 2008) and showed that, in human chromatin, active genes have a notably shorter repeat (178 bp) than repressed or heterochromatic non-coding sequences (206 bp) (Valouev et al, 2011).

While changes in the NRL were initially expected to alter the chromatin fiber diameter proportionally to NRL (Athey et al, 1990), more recent electron microscopic studies of linker histone-dependent chromatin compaction by Rhodes and her colleagues revealed two types of chromatin fiber distinguished by a stepwise increase in chromatin diameter from ~33 nm for chromatin with 177–207 bp linkers to ~42 nm for 217–227 bp linkers (Robinson et al, 2006). A third type of structure was observed for very short NRLs typical of yeast and neuronal cells (167 bp). Electron microscopic studies showed that, in contrast to the longer arrays, those with a short 167 bp NRL could fold into compact fibers without linker histone (Routh et al, 2008) in agreement with earlier biochemical observations of linker histone-independent compact folding of 167–177 bp NRL arrays (Dorigo et al, 2003; Shogren-Knaak et al, 2006). The folded 167 bp NRL arrays also had a distinct fiber diameter about 21–23 nm with a clear zigzag morphology (Routh et al, 2008) consistent with nucleosome stacking in the tetranucleosome X-ray crystal structure (Schalch et al, 2005).

Due to the helical nature of DNA, smaller alterations to the NRL in the range of several base pairs were predicted to have a strong effect on chromatin fiber folding (Widom, 1992; Woodcock et al, 1993; Leuba et al, 1994; Stehr et al, 2010). Here, we examine whether and how chromatin higher-order folding depends on the internucleosomal rotational setting. This parameter refers to linker DNA length variability under 10.5bp (one turn of the DNA double
helix) that is expected to change internucleosomal rotation angle by 36° per 1 bp of DNA length. It is distinct from the more conventional definition of rotational setting as the local orientation of the DNA helix on the histone surface (Jiang and Pugh, 2009), instead referring to the rotational settings relative to adjacent nucleosomes, which may be altered by DNA orientation at the surface and local variations in the linker length.

We show that for rotationally similar nucleosomes with different NRLs, there is a general negative correlation of chromatin compaction with DNA linker length. For long-linker arrays, rotational alterations do not affect chromatin folding. In a striking contrast, for shorter NRLs (165–177 bp), we observed a periodic dependence of chromatin fiber folding upon local rotational variations. Our data thus demonstrate that internucleosomal rotational settings are crucial in determining nucleosome packing of arrays with short linker DNA, while long-linker DNAs accommodate rotational variability without a substantial effect on chromatin folding. Our results suggest that internucleosomal rotational settings play a crucial role in the higher-order structure of chromatin with short NRLs such as those found in the yeast genome (Lohr, 1981; Wang et al, 2008) and in transcriptionally active chromatin of human cells (Valouev et al, 2011).

Results

Nucleosome array folding reflects the intrinsic properties of the underlying NRL

Reconstituted nucleosome arrays with various NRLs were assembled using the clone 601 sequence and histone octamers isolated from chicken erythrocyte nuclei. Since precise histone stoichiometry is crucial for proper folding, it was monitored through several independent techniques. First, to ensure that our arrays were not underloaded, the reconstituted arrays were
digested with restriction nucleases into mononucleosomes and assayed by agarose DNP electrophoresis in TAE buffer optimized for resolving mononucleosomes and free DNA. On these gels, the presence of free DNA monomer bands was indicative of underloaded material (arrow on Supplementary Figure S1A). Next, the reconstitutes were analyzed using agarose DNP gels in HEPES buffer optimized for oligonucleosome separation. On these gels, smearing of carrier DNA indicated that high-affinity (clone 601) templates had been saturated and excess histones had been bound to the low-affinity carrier DNA (Lane 2, Supplementary Figure S1B). These same gels also allowed us to visualize instances where nucleosome arrays were overloaded with histone octamers through the appearance of additional bands above the main array band (brackets on Supplementary Figure S1B). Based on the preliminary experiments, we reconstituted arrays at a larger scale with a typical yield of 100 μg DNA. Agarose DNP gels in parallel with analytical ultracentrifugation provide an efficient way to monitor histone loading in multiple arrays and prepare homogeneous samples for solution and imaging studies (Figure 1). In these experiments, higher mobility of the main electrophoretic band (black arrow, Figure 1A) and sedimentation coefficients S20,w below 28 S (Figure 1B, top panel) indicated underloading and were discarded. Overloaded samples, which showed additional dimer bands above the main array band (white arrows in Figure 1A), also had additional peaks during ultracentrifugation (white arrows in Figure 1B, middle panel) and were discarded. The optimally loaded arrays (grey arrows, Figure 1A and B) were examined by transmission electron microscopy to confirm full saturation of the nucleosome arrays with the majority of the arrays containing 12 nucleosomes by EM (Figure 1C and D).
Figure 1. Oligonucleosome reconstitution and characterization. (A) DNP type IV agarose gel in HE buffer, showing reconstituted nucleosome arrays, 169×12 arrays with different histone loadings. Lanes 1–4: 1—DNA molecular
weight markers, 2—underloaded (black arrow), 3—overloaded (white arrows), and 4—properly loaded (grey arrow) nucleosome arrays. (B) Distributions of sedimentation coefficients, c(S), for 169×12 arrays with different histone loadings: underloaded (top panel), overloaded (middle panel), and properly loaded (bottom panel) at 0.6 mM MgCl2. (C) Electron micrograph (uranyl acetate staining, dark-field imaging) of 207×12 core arrays (top panels) and 167×12 core arrays (bottom panels) fixed at 5 mM NaCl. (D) Histograms showing distribution of nucleosome arrays containing a certain number of nucleosomes per array calculated from several EM fields of 207×12 arrays (top panel) and 167×12 arrays (bottom panel).

The homogeneity and extent of nucleosome array folding for arrays with different NRLs were assayed by sedimentation velocity analysis using the continuous c(S) distribution model (Schuck, 2000) as well as by the boundary analysis of van Holde and Weischet, 1975. For highly homogeneous samples of 167×12 and 207×12 arrays, we observed that the main boundary positions in van Holde–Weischet plots are fully consistent with the main peaks in the c(S) distribution (cf. Figure 2A and B to Supplementary Figure S2A) showing a dramatic difference in sedimentation (36 versus 52 S) at 1 mM MgCl2. We also observed that c(S) analysis was very efficient in resolving sedimentation distribution in heterogeneous systems (cf. Figure 2C to Supplementary Figure S2B). Therefore, we applied the c(S) method for sedimentation analysis of all heterogeneous samples including those partially self-associated due to overloading or excessive ionic strength as well as intentionally heterogeneous samples throughout this work.
Figure 2. Nucleosome array folding depends upon the intrinsic properties of the underlying NRL. (A–C) Distribution of sedimentation coefficients, c(S), for 167×12 (A), 207×12 (B), and the 167/207×12 coreconstitute (C) at 1 mM MgCl2. (D–F) Electron micrographs of 167×12 (D), 207×12 (E), and the 167/207×12 coreconstitute (F) at 1 mM MgCl2 showing different degrees of compaction.
To independently verify that the observed differences in folding, such as those between arrays with NRLs of 167 and 207 (Figure 2A and B), were not due to variations in histone loading, chromatin arrays were reconstituted with an equimolar mixture of each of 167×12 and 207×12 DNA templates and core histones in the same dialysis bag. These coreconstituves were then assayed for folding in sedimentation velocity experiments on the analytical ultracentrifuge. The coreconstituves formed two distinct peaks, indicating two types of particles with different folding extents (Figure 2C), with each one of these double peaks corresponding to the single peak of the uniform array in Figure 2A and B. EM imaging of the 167×12/207×12 coreconstitute further confirms the presence of two distinct states of chromatin folding for 167×12 and 207×12 arrays (Figure 2D and E), as both compact and open arrays are seen in the coreconstituted sample (Figure 2F).

NRL with fixed rotational setting is negatively correlated with chromatin folding

We constructed and characterized a series of nucleosome core arrays in which the NRL was altered so that due to the 10.5 bp periodicity of DNA, the arrays were expected to maintain the same rotational settings. As such, reconstituted 12-mer chromatin arrays were assembled with NRLs of 167, 177, 188, or 209 bp. While the sedimentation peaks observed at 5 mM NaCl had very similar distributions of sedimentation coefficients around 30 S as expected for unfolded nucleosome arrays (Figure 3A), upon the induction of chromatin folding by 60–150 mM NaCl or 0.6–2 mM MgCl2, we observed a notable increase in sedimentation velocity for all samples (Figure 3B and C and Supplementary Table S3). Remarkably, the salt-dependent increase in S values was much stronger for shorter NRLs—about 20 S for 167×12 compared to a moderate increase of about 5 S for 209×12 at 1 mM MgCl2. This result is consistent with the strong salt-
dependent folding previously reported for 167 versus 197 (Routh et al, 2008) and 167 versus 207 (Grigoryev et al, 2009) arrays at the same ionic strength. Now we observed a clear negative correlation between the NRL and the degree of nucleosome array compaction for a range of monovalent and divalent cation concentration known to support compact folding of the nucleosome arrays.
Rotational variations do not alter folding for chromatin arrays with long NRL

While native chromatin species have NRLs differing ±2 to ±4 from the average repeat lengths (Strauss and Prunell, 1982; Widom, 1992), most experiments with reconstituted nucleosomal arrays employed nucleosomal DNA templates with uniform NRLs. To explore the importance of local variations in the internucleosomal rotations mimicking native chromatin for chromatin folding, we constructed and characterized 12-mer arrays (205–207–209)×4 and (205–209)×6, in which the NRL varied in a defined manner by either ±2 bp or ±4 bp, respectively. These arrays were expected to show structural variations reflecting local changes in the internucleosomal rotation angle (Woodcock et al, 1993). However, for neither of these reconstitutes did we observe any significant difference in chromatin folding between the uniform (207×12) and variable (±2 or ±4 bp) arrays at any of the ionic conditions tested (Table I, Supplementary Figure S3A).

<table>
<thead>
<tr>
<th></th>
<th>5 mM NaCl</th>
<th>60 mM NaCl</th>
<th>150 mM NaCl</th>
<th>0.6 mM MgCl2</th>
<th>1 mM MgCl2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Uniform −LH</td>
<td>29</td>
<td>36</td>
<td>36</td>
<td>35</td>
<td>36</td>
</tr>
<tr>
<td>±2 bp −LH</td>
<td>30</td>
<td>37</td>
<td>38</td>
<td>33</td>
<td>36</td>
</tr>
<tr>
<td>±4 bp −LH</td>
<td>30</td>
<td>36</td>
<td>39</td>
<td>35</td>
<td>36</td>
</tr>
<tr>
<td>Uniform +LH</td>
<td>38</td>
<td>53</td>
<td>54</td>
<td>57</td>
<td>60+</td>
</tr>
<tr>
<td>±2 bp +LH</td>
<td>36</td>
<td>55</td>
<td>56</td>
<td>60</td>
<td>60+</td>
</tr>
<tr>
<td>±4 bp +LH</td>
<td>37</td>
<td>52</td>
<td>55</td>
<td>56</td>
<td>58+</td>
</tr>
</tbody>
</table>
Table 1. Linker length variations do not affect chromatin folding for uniform and variable 12-mer arrays with long NRLs. Sedimentation coefficient main peak positions (average of three independent experiments) were determined from distributions of sedimentation coefficients, c(S), for 12-mer oligonucleosome core arrays (−LH) and linker arrays (+LH) with uniform (207 bp) and variable (207±2 and 207±4 bp) NRLs at various salt conditions as indicated.

Next, the folding of uniform and variable arrays was assayed in the presence of linker histone H5, which is known to be important for the complete folding of nucleosome arrays with 207 bp NRL (Carruthers et al, 1998). As expected, the addition of linker histone increased the sedimentation coefficients to ~55–60 S in the presence of MgCl2; however, the uniform and variable arrays again had similar sedimentation coefficients both in the presence and absence of linker histone H5 (Table I, Supplementary Figure S3B).

To confirm the absence of size dependence of nucleosome folding in our reconstituted arrays, the uniform and ±4-bp arrays were doubled in length to form 24-mer arrays, which were again assayed for folding. There were no significant differences in folding between the uniform and variable 24-mer arrays in either the core arrays or in those containing linker histone H5 (Supplementary Table S2). These results clearly demonstrate that the uniform and variable arrays with net 207 bp NRL are completely folded into the chromatin fiber in the presence of linker histone and that the degree of their folding is independent of internucleosomal rotational settings.

Internucleosomal rotational variations affect folding in chromatin arrays with short NRL.
Next, we asked whether uniform arrays with altered rotational settings would follow a similar pattern of increased NRL leading to decreased S values. To test this, we employed 12-mer core arrays with repeats of 165, 169, 172, 200, 205, and 207 bp, that is, NRL values intermediate between those explored above (167, 177, 188, and 209). Within each array, the rotational setting is constant, but is different relative to the other arrays because the NRL changes are not in increments of DNA helical periodicity (10.5 bp). If the rotational settings did not affect chromatin folding, then we expected the compaction of the nucleosome arrays with intermediate NRLs to fall between the values obtained for arrays with the nearest longer and the nearest shorter NRLs. Indeed, as seen in Figure 4A and B and Supplementary Table S3, we found that core arrays with NRL of 200 bp have similar folding to arrays with NRL of 188 bp (~38–40 S), and the 205-bp arrays were similar to 207 bp (~35–36 S).
Figure 4. NRL variations affect chromatin folding but not self-association. (A) Graphic plotting of main sedimentation coefficient peaks in c(S) distribution (average of three independent experiments) for 12-mer
oligonucleosome core arrays with varying NRL (165, 167, 169, 172, 177, 188, 200, 205, 207, and 209 bp) at 5 mM NaCl, 60 mM NaCl, or 150 mM NaCl. Student’s t-test P-values for significant differences between the data sets are shown over the brackets. (B) Graphic plotting of main sedimentation coefficient peaks in c(S) distribution (average of three independent experiments) for 12-mer oligonucleosome core arrays with varying NRL (165, 167, 169, 172, 177, 188, 200, 205, 207, and 209 bp) at 5 mM NaCl, 60 mM NaCl, or 150 mM NaCl. Student’s t-test P-values for significant differences between the data sets are shown over the brackets. (B) Graphic plotting of main sedimentation coefficient peaks in c(S) distribution (average of three independent experiments) for 12-mer oligonucleosome core arrays with varying NRL (165, 167, 169, 172, 177, 188, 200, 205, 207, and 209 bp) at 0.6 mM MgCl2, 1 mM MgCl2, or 2 mM MgCl2. (C) Histograms of the concentration of MgCl2 (average of three independent experiments) that results in 50% precipitation of material for arrays with varying NRL.

Surprisingly, in a sharp contrast to the apparently similar folding of the longer-NRL arrays, the shorter arrays displayed strong NRL-dependent structural deviations. Figure 4A and B shows that 172×12 array was dramatically unfolded, with sedimentation coefficient values of 41 S in 150 mM NaCl and 38 S in 1 mM MgCl2, compared to the 167×12 and 177×12 arrays, which have sedimentation coefficients at about 48–54 S under those conditions. The 165×12 and 169×12 arrays showed intermediate folding with S values distributed around 43–47 S in either 150 mM NaCl or 1 mM MgCl2. The NRL-dependent periodic changes of chromatin compaction were also observed for a wider range of Mg2+ concentrations—between 0.6 and 2 mM (Figure 4B). These data show that the unfolded short-NRL arrays (165 and 169) and, especially, the intermediate NRL arrays (188 and 200 bp) undergo additional compaction at 2 mM MgCl2. Sedimentation of neither the most folded or unfolded arrays with short NRLs (167, 172, and 177) nor the unfolded arrays with long NRLs was significantly affected by increased Mg2+, showing that the divalent cation may partially curb the difference between NRLs 165, 167, and 169, but cannot fold the most open arrays (unlike linker histone in experiments shown in Figure 5 below). These results, together with sedimentation analysis of the longer-NRL nucleosome arrays facilitated by linker histone (Table I), show that the internucleosomal rotational settings have a
strong periodical effect on chromatin folding for nucleosome core arrays with short NRLs (167–177 bp) but not for the long (>200 bp) NRLs. Our findings are consistent with recent modelling studies predicting a stronger counterion requirement for packing the longer and more negatively charged linkers (Perisic et al, 2010), and the linker DNA geometry being most important for folding of the nucleosome arrays with short NRL (see Discussion).
Figure 5. A 5-bp difference in linker DNA length destabilizes chromatin folding for short NRLs. (A–D) Distribution of sedimentation coefficients, c(S), at 1 mM MgCl₂ for separately analyzed 167×12 and 172×12 core arrays (A), 167/172×12 coreconstitute (B), 167×12 reconstituted with and without linker histone (C), and 172×12 reconstituted with and without linker histone (D). (E–G) Electron micrographs of 172×12 core arrays at 5 mM NaCl (E) and
Nucleosome array self-association does not depend upon the NRL

In addition to chromatin folding (secondary chromatin structure), chromatin fibers are compacted by self-association (tertiary chromatin structure) (Woodcock and Dimitrov, 2001; Luger and Hansen, 2005). To examine whether NRL variations affect chromatin tertiary structure, we monitored the extent of self-association for the various constructs differing in NRL by magnesium-dependent self-association assays (Schwarz et al, 1996). We observed that wide changes in NRL, including those that significantly affected chromatin folding, did not notably alter self-association rates, with most chromatin arrays having 50% precipitation between 3.0–3.5 mM MgCl2 (Figure 4C). It thus appears that nucleosome self-association is not affected by the broad range of NRLs, suggesting that neither general NRL nor its local rotational variations can notably affect formation of tertiary chromatin structures.

While chromatin precipitation by low-speed centrifugation is typically employed to determine chromatin self-association (Schwarz et al, 1996), we noticed that under certain conditions analytical ultracentrifugation can be successfully used to assay the initial steps of self-association in solution resulting in array dimerization (tertiary structures). The dimers appear as a ‘second peak’, running at 90 S (or above). This peak is most commonly observed on arrays with
NRLs ~207 bp in the presence of linker histone H5 and 1 mM MgCl2. Using this approach, we also did not observe any significant differences in Mg-dependent formation of the 90 S peak between the 207×12 uniform array and the variable ±2- or ±4-bp arrays (Supplementary Figure S3B, bottom), confirming the results of the precipitation assay.

Intrinsic properties of the short-NRL DNA templates determine nucleosome array folding

After observing periodic modulations in chromatin folding for the short-NRL arrays, we also wished to independently verify that these differences were not due to variations in histone loading. Therefore, we assembled nucleosome reconstitutes with a mix of 167×12 and 172×12 DNA templates. These coreconstitutes were then assayed for folding by sedimentation velocity experiments on the analytical ultracentrifuge and by transmission EM. The coreconstitutes formed two distinct peaks at 1 mM MgCl2, indicating two samples with different folding extent (Figure 5B), with each one of these double peaks corresponding to the single peak of the 167×12 or 172×12 uniform array (cf. panels A and B in Figure 5). EM imaging (42K magnification) of the 172×12 in the unfolded state (Figure 5E) as well as 172×12 and 167×12 arrays folded at 1 mM MgCl2 (Figure 5F–H) confirms the notably larger diameters of 172×12 arrays at 1 mM MgCl2. This difference is also observed in the EM images of platinum-shadowed samples obtained at 110K magnification (Figure 5, cf. panels I and J).

We next asked whether the addition of linker histone, which is required to stabilize compact folding (>50 S) of nucleosome arrays with longer NRLs (Carruthers et al, 1998; Routh et al, 2008), would also promote the folded state of the short-NRL arrays. In this set of experiments, we added linker histone H5 to 167×12 and 172×12 arrays and observed that H5
induced only minor increases in the already folded 167×12 (52–55 S, Figure 5C), consistent with previous findings by Routh et al (2008). Strikingly, adding linker histone caused a dramatic increase in 172×12 sedimentation from 38 to 55 S in 1 mM MgCl2 (Figure 5D). It thus appears that certain short NRLs can be completely folded by linker histone, suggesting that short NRLs following the 10n+5 rule impose an architectural factor-dependent folding on the overlaying chromatin fiber. In contrast, the short NRLs containing an integer number of linker DNA turns (such as 167×12 and 177×12) promote compact folding independently of the architectural proteins.

Discussion

Arrays of positioned nucleosomes with regular NRLs are widely used in the studies of chromatin higher-order structure. However, nucleosomes are mobile on native DNA (Flaus-Owen-Hughes, 2003) as most native DNA sequences do not bind or position histone octamers as tightly and precisely as they bind to the clone 601 sequence and other strong positioning templates (Lowary and Widom, 1998) utilized in most recent experiments. As such, the NRL is not uniform within native chromatin, but rather contains variations around well-defined average peaks with 10 bp intervals ±2 bp (Widom, 1992). Several models have directly suggested a dependence of chromatin higher-order structure on the internucleosomal rotational variations (Widom, 1992; Woodcock et al, 1993; Leuba et al, 1994; Stehr et al, 2010). This raises the question of whether the regular arrays fold into higher-order structures similar to native chromatin. Our experiments with uniform (207×12) and variable (207±2 and ±4 bp) nucleosome arrays clearly showed that these arrays exhibited similar compaction whether they were in
unfolded, partially folded, or completely folded states (Table I and Supplementary Figure S3). Thus, our results imply that nucleosome linkers are flexible enough to be independent from the rotational settings for the NRLs of around 200 bp so that previous biochemical experiments with regular repeats, as well as computative modelling conducted by us and other groups, adequately reproduce native chromatin states for long NRLs typical of differentiated tissues.

How does the compact chromatin fiber accommodate internucleosomal rotational variability? One explanation was provided by X-ray crystal studies showing that nucleosome core DNA contains four superhelical locations where the DNA major grooves may be stretched or compressed to incorporate 5 or 6 bp (Richmond and Davey, 2003; Makde et al, 2010) so that linker DNA length variations of 4 bp might be absorbed by nucleosome core. If this was indeed happening in solution, then we would expect that linker DNA length difference of 2–4 bp has no effect on chromatin folding. However, our experiments showing periodic dependence of the short-NRL array conformation upon linker DNA lengths (Figures 4 and 5) argue against absorbing the NRL variations by the core DNA as the ability to stretch and absorb linker variations should be the same for the short and long NRLs. On the other hand, our recent computative modelling and EM experiments showed that compact chromatin fibers packed with linker histones incorporate conformationally heterogeneous DNA linkers (Grigoryev et al, 2009). This conformational heterogeneity can readily explain the independence of chromatin folding on rotational settings for long nucleosomes.

The most striking result of our work is that the shortest NRLs confer a strong dependence of nucleosome folding upon small rotational variations. To further explore the mechanism of the different folding capabilities of the 167 and 172 NRLs, we constructed dinucleosome molecular models based on the nucleosome core structure taking 145 bp of clone 601 nucleosome (Makde
et al, 2010) pdb ref# 3MVD, to which we added either 22 or 27-bp linker DNA (from pdb ref#1ZBB). The models of the 167 and 172 dinucleosomes (Figure 6) show that the internucleosomal rotational setting change resulting from the 5 bp NRL difference leads to a striking contrast in the linker DNA spatial positions. While the DNA linkers in the 167 bp dinucleosomes are well separated in space (Figure 6A and B), the 172 bp dinucleosome brings its linkers to an immediate proximity if the linker DNA contains no structural deviations such as DNA bending and twisting (Figure 6D and E).
Figure 6. Molecular modelling shows the proposed effect on chromatin folding of the 5-bp difference in the NRL between the 167 bp and 172 bp repeats. (A, B) Molecular model of the dinucleosome with NRL 167 (X–Y and Z–Y planes are shown) was constructed using UCSF Chimera (http://plato.cgl.ucsf.edu/chimera/) based on 147 bp from the clone 601 nucleosome core structure PDB 3MVD for the nucleosome after removing RCC protein from the file. For adding linker DNA, we used a DNA fragment containing base pairs DG345I/DC3J through DC336I/DG12J from PDB 1ZBB and aligned multiple copies in three-dimensional space to get the 20-bp linker length and 167 bp NRL. (C) A flexible wire model for DNA geometry in a compact tetranucleosome with NRL 167. The dashed lines show the 30-nm fiber axis. The arrows show the points of contact that limit the longitudinal compaction of the fiber. (D–E) Molecular model of the dinucleosomes with NRL 207 based on 147 bp from the clone 601 and constructed as in panels A–B but adding 25-bp linker DNA to get 172 bp NRL. (F) A flexible wire model for DNA geometry in a compact tetranucleosome with NRL 172. The dashed line and the arrow indicate the fiber axis and the points of contact as in panel C.

Recent crystal study of the clone 601 nucleosome showed that its core constrains 145 bp of DNA (Makde et al, 2010) implying that in solution the core DNA may additionally incorporate 2 bp from the linker. In that case, the internucleosomal rotational angle would differ by 72° from the one shown on Figure 6. However, we found that a spatial model recapitulating the DNA path in the 167-bp NRL array and based on the clone 601 nucleosome core containing 145 bp DNA (Makde et al, 2010) allows this repeat to be easily folded in a conformation-promoting stacking of the nucleosome discs (Figure 6C). The predicted size and shape of the most compact 12-mer array with 167 bp NRL (Supplementary Figure S4A) are consistent with high-resolution EM images of the 167 NRL repeats (Routh et al, 2008) and Figure 5I as well as with the tetranucleosome crystal structure (Schalch et al, 2005). The agreement between our models and the experiments argues that the actual DNA conformation in the 167×12 nucleosome arrays is similar to that of its crystal form.
In striking contrast, for similar models but with a 172 bp repeat (Figure 6F and Supplementary Figure S4B), we noticed that the topology of the DNA chain does not allow the nucleosome discs to be brought into the close contacts needed for stacking. Specifically, the nucleosome linkers in the folded 172×4 structure comes to a very close proximity between themselves and limit the array's folding (arrow on Figure 6F) consistent with our high-resolution EM data clearly showing an incomplete, uneven packing of the 172 bp repeats (Figure 5J). This result agrees with recent modelling studies predicting a stiff and open conformation for a 173-bp NRL repeat (Perisic et al, 2010) and by Scipioni et al (2010) suggesting that excluded volume at the linker crossing and torsional energy are both limiting factors for close packing of nucleosomes. It also predicts strong peaks of torsional energy at 162 and 172 bp with a valley at 167 bp (Scipioni et al, 2010). The fact that linker histone promotes the folding of the 172-bp array more strongly than that of the 167-bp array (Figure 5C,D) while increased MgCl2 levels cannot fold the 172 repeat (Figure 4B) confirms that linker DNA geometry rather than electrostatic charge is the main obstacle to complete folding of the 172-bp arrays.

Earlier biochemical experiments (Lohr, 1981) and recent sequencing of positioned yeast nucleosomes (Wang et al, 2008) both showed that linker DNA is preferentially quantized with a linker length of 10n+5 bp, making the 162- and 172-bp NRLs predominant in yeast. According to our structural experiments, this indicates that the yeast genome evolved to maintain its chromatin in an open configuration similar to the one observed for the 172-bp NRL. This is consistent with the generally open structure of the yeast genome with very low levels of linker histone and about two nucleosomes per unit length found by in vivo mapping of chromatin interactions through crosslinking analysis (Dekker, 2008).
Our results clearly demonstrate that for a broad range of nucleosome arrays with different NRLs, only a subset of short nucleosome arrays with optimal NRLs such as 167 and 177 bp can completely fold at physiological ionic conditions (Figure 3B,C), consistent with previous observations for 167 (Routh et al, 2008; Grigoryev et al, 2009) and 177 bp (Dorigo et al, 2004; Shogren-Knaak et al, 2006) repeats. At first glance, this result seems surprising as the increase in DNA linker length usually accompanies chromatin condensation (Thomas and Furber, 1976; Weintraub, 1978; Athey et al, 1990; Perisic et al, 2010). However, there also exists a trend where differentiated cells with specialized gene expression patterns tend to have, in addition to longer NRLs, increased linker histone levels (Weintraub, 1978; Pearson et al, 1984; Woodcock et al, 2006). It is now clear that the addition of linker histone is capable of compensating for decreased folding of the long-NRL arrays (Routh et al, 2008; Grigoryev et al, 2009) as well as the shorter-NRL arrays with unfavorable nucleosomal rotations (Figure 5). We expect that identification of the NRLs inhibiting chromatin folding and constructing the set of native-like nucleosome array templates described here will provide valuable tools for experimental and modelling studies of chromatin higher-order structure and interactions with chromatin architectural factors.

Most recently, high-throughput genomic maps of nucleosome positions in the human genome revealed that transcriptionally active genes have a notably shorter repeat (178 bp) than repressed or heterochromatic non-coding sequences (206 bp) (Valouev et al, 2011). Our findings, together with the newly estimated NRLs of transcribed genes, provide an explanation for the paradoxical observations of relatively condensed conformation of transcriptionally active chromatin (Hu et al, 2009) despite lower levels of linker histone (Kamakaka and Thomas, 1990) as well as its shorter retention time on the nucleosomes (Misteli et al, 2000). Our data imply that the shorter 10n NRL close to the 178 bp in active genes would maintain a relatively compact
conformation independent of architectural factors. Moreover, we predict that the higher-order structure of 10n repeats that includes strong internucleosomal stacking (Figure 6C) would be especially receptive to regulation by histone H4-K16 acetylation that is known to inhibit internucleosomal stacking (Shogren-Knaak et al, 2006; Allahverdi et al, 2011).

Finally, the observation that the 172×12 array with unfavorable rotational settings requires linker histone (Figure 5) rather than MgCl2 (Figure 4B) for complete folding shows that architectural protein factors may be necessary for the folding of specific short NRLs following the 10n+5 rule. For example, the predictably unfavorable neuronal NRL of 162 bp (Pearson et al, 1984) seems to contain an anomalously high concentration of linker histone for this given repeat (see Figure 1 in (Woodcock et al, 2006)). Yeast chromatin may be folded at certain heterochromatic loci by special architectural proteins such as Sir3 (McBryant et al, 2008). It thus appears that specific NRLs may direct chromatin structure into either a charge-dependent folding typical of short 10n NRLs or, alternatively, into an architectural factor-dependent plastic folding typical of short 10n+5 repeats in yeast or longer NRLs in higher eukaryotes. There is little doubt that future genome-wide studies will continue to provide more examples of specific nucleosomal positioning whose structural significance could be revealed by systematic studies in biochemically defined reconstitution systems.

Materials and Methods

Nucleosome positioning templates and arrays

DNA templates were designed using clone 601 DNA (Lowary and Widom, 1998), which positions nucleosomes with single-nucleotide precision, and adding specific lengths of linker
DNA (Supplementary Table S1). Templates were first designed as monomers with the length of linker DNA altered through primer modification PCR. Regardless of linker DNA length, all monomeric templates were designed with XbaI restriction sites at the 5′-end and SpeI and SphI sites at the 3′-end. Monomeric templates were ligated into the pUC19 vector, transformed into Escherichia coli DH5α (Invitrogen, #18265-017), and grown in the presence of carbenicillin. Expansion into a 12-mer template was achieved through a stepwise approach in which monomer plasmids were digested using either SpeI and SphI to yield vectors containing monomers, or XbaI and SphI to yield monomeric inserts. Upon ligation of monomeric insert to the vector already containing one copy of the nucleosome positioning template, the SpeI sites of the vectors and the XbaI sites of the inserts were both abolished. This process was repeated until constructs existed as 12-mers.

Once the positioning templates were constructed in pUC19 plasmids, they were transformed into E.coli and grown in ~2 l of media supplemented with carbenicillin. Plasmids were isolated using large-scale alkaline lysis procedures (Birnboim and Doly, 1979). All our plasmids were subsequently digested from the vector backbone using the restriction endonucleases XbaI and HindIII. The vector backbone was digested into smaller fragments (1120, 862, 692, and 19 bp) using DraI to serve as a carrier DNA (see reconstitution procedure). The only exceptions were 200×12, which used XbaI and KpnI, and 177×12, which used EcoRV to excise the template.

**Histone octamer isolation and purification**

Core histone octamers were isolated from chicken erythrocyte nuclei by the selective removal of histones through ultracentrifugation under increasing ionic strength as described
(Grigoryev et al, 2009), and then loaded on a 5–30% sucrose gradient in 2 M NaCl–10 mM HEPES–0.1 mM EDTA buffer. Sucrose gradients were spun on the ultracentrifuge in a Beckman SW41Ti swinging bucket rotor, 4 °C, 38 000 r.p.m., for 48 h. After centrifugation, the gradient was divided into 1-ml fractions that were analyzed by SDS–PAGE. Fractions containing equimolar amounts of core histone proteins were dialyzed against buffer (10 mM HEPES, 2 M NaCl, and 0.1 mM EDTA) for ~36 h to remove sucrose and stored at −10 °C.

**Linker histone isolation**

The supernatant from the 0.65-M centrifugation was used for linker histone isolation through ion exchange chromatography. The procedure used a 1-ml High Trap Sp Hp ion exchange column (GE Healthcare) equilibrated with buffer A (20 mM HEPES, 50 mM NaCl, 0.02% sodium azide, 0.5 mM PMSF, pH=7.5) and eluted using a two-step gradient with buffer B (20 mM Hepes, 1.5 M NaCl, 0.02% sodium azide, pH=7.5), with 0–33% buffer B over 5 column volumes, and 33–100% buffer B over 35 column volumes. H5 fractions were dialyzed ~12 h against buffer (10 mM HEPES, 50 mM NaCl, 0.1 mM EDTA, pH=7.5) and stored at 4 °C.

**Oligonucleosome array reconstitution**

Reconstituted 12-mer (and 24-mer) arrays were assembled using DNA templates (described above) and chicken erythrocyte histones. Reconstitutions were performed essentially as described in (Springhetti et al, 2003), but in the presence of carrier DNA (Huynh et al, 2005) to ensure proper loading. The carrier DNA was created through digestion of pUC19 vector backbone by DraI, and used at a template:carrier ratio of 2:1. Core histones and purified DNA were combined in a final mixture containing 2 M NaCl and 1 mM PMSF. Samples were dialyzed in 3500 MWCO membranes against buffer (2 M NaCl, 10 mM HEPES, 0.2 mM EDTA, 0.1%
NP-40, 5 mM β-mercaptoethanol, pH=7.5) for 2 h. The salt was then lowered to 1.5 M NaCl and
dialyzed for 2 h, followed by dialysis against 1 M NaCl for 3 h, and a 0.75 M NaCl dialysis for 3
h (all other buffer component concentrations were kept constant). The salt concentration was
then decreased to 0.5 M NaCl and dialyzed ~18 h. Reconstitutes were next dialyzed against a
buffer containing 5 mM NaCl, 10 mM HEPES, 0.2 mM EDTA, 0.1% NP-40, 5 mM β-
mercaptoethanol, pH=7.5, for 3 h. Then the buffer was changed to one lacking NP-40 (all other
buffer components remained the same) and dialyzed for an additional 3 h to remove NP-40. The
reconstituted core arrays were then quantified by spectrophotometry (A260 nm) measurements,
run on 1% Type IV agarose (DNP) gels to ensure proper histone loading, and assayed through
SDS–PAGE to ensure integrity of core histones.

In instances where linker histone is added to the chromatin arrays, additional
reconstitutions were performed by mixing reconstituted core arrays with linker histone H5 at a
molar ratio of one molecule H5 per nucleosome and dialyzing against 500 mM NaCl, 0.1 mM
PMSF, and 0.025% NP-40 using 3500 MWCO membranes for 2 h at 4 °C, followed by ~18-h
dialysis against 5 mM NaCl, 10 mM HEPES, 0.1 mM EDTA, 0.025% NP-40, pH=7.5 buffer,
followed by a 4-h dialysis against 5 mM NaCl, 10 mM HEPES, 0.1 mM EDTA, pH=7.5, to
remove NP-40. The linker arrays were then removed from dialysis, 0.1 mM PMSF was added,
A260 nm measurements were taken, and samples were stored at 4 °C.

Carrier DNA was removed from reconstituted arrays through purification on a 5–25%
sucrose gradient in the following buffer: 10 mM Tris, 1 mM EDTA, 0.5 mM PMSF, and
centrifuged at 35 000 r.p.m., 4 °C, for 8 h on the ultracentrifuge in the Beckman SW41Ti
swinging bucket rotor. After centrifugation, samples were divided into 1-ml fractions and
analyzed on a 1% agarose gel, and fractions containing arrays lacking carrier DNA were
dialyzed against 10 mM HEPES, 5 mM NaCl, 0.1 mM EDTA, pH=7.5, for ~36 h to remove sucrose. Samples were analyzed by SDS–PAGE to confirm histone integrity.

**Analytical ultracentrifugation**

The extent of nucleosome array folding was assayed through sedimentation velocity experiments on the analytical ultracentrifuge (Beckman Optima XL-A) (Carruthers et al, 2000). Chromatin arrays with A260 nm=0.5 OD were run in buffer containing 10 mM HEPES, 0.1 mM EDTA, pH 7.5 under varying ionic strength conditions (5 mM, 60 mM, 150 mM NaCl and 0.6 mM, 1 mM, 2 mM MgCl2). Scan data were collected using ProteomeLab XL-A, at a wavelength of 260 nm while samples were spun at 20 000 r.p.m., at 20 °C, for ~2.5 h. Data analysis from the analytical ultracentrifuge runs was carried out using the continuous c(S) distribution model of the SEDFIT software (Schuck, 2000). (http://www.analyticalultracentrifugation.com). Boundary analysis was also conducted by the method of van Holde and Weischet (van Holde and Weischet, 1975) using the program Ultrascan II (http://www.ultrascan2.uthscsa.edu/).

**Electrophoretic techniques**

Low electro-endosmosis agarose gels of 1% (Lonza SeaKem LE agarose, #50002) in TAE buffer (BioRad #161-0743) were used to visualize histone-free DNA. Gels were run for 40 min at 80 V. Medium electro-endosmosis Type IV agarose (Sigma-Aldrich #A3643) was used to analyze oligonucleosome reconstitutes by DNP (deoxynucleoprotein) electrophoresis. For these gels, Type IV agarose was dissolved in HE buffer (20 mM HEPES, 0.1 mM EDTA) at a concentration of 1%. These gels were run at 80 V for 70–80 min and stained with ethidium bromide. Type IV agarose (Sigma-Aldrich #A3643) gels were also used to analyze mononucleosomes and free DNA resulting from test restriction digestions of oligonucleosome
reconstitutes. For these gels, Type IV agarose was dissolved in TAE buffer at 1% concentration. The gels were run at 80 V for 70–80 min. All agarose gels were stained with ethidium bromide (~1 μg/ml) for 30 min before imaging.

Polyacrylamide gels were run throughout the experiments with nucleosome arrays to assay the stability and/or concentration of histones. These gels were typically 15% acrylamide with 1 mm thickness. For gels where band quantification was required, the NIH program Image J was used (http://rsbweb.nih.gov/ij/). Gel bands were marked using the rectangular selection tool and, for each lane, a background reading was taken from the area above the band and subtracted from the measured band intensity. In the case of magnesium precipitation assays, relative band intensities were averaged from a minimum of three gels.

Self-association assay

The extent of self-associated material was assayed through selective precipitation in magnesium similar to (Schwarz et al, 1996). Arrays with A260 nm=0.5 were incubated with increasing concentrations of MgCl2 (Sigma-Aldrich, #M1028-1 ml) for 20 min on ice. The reactions were then centrifuged at 12 000 r.p.m., 4 °C, for 10 min. Supernatants were removed and loaded on a 1% agarose gel, while pellets were resuspended in 25% glycerol, 50 mM EDTA, 0.5% SDS, and loaded on a 1% agarose gel. The percentage of DNA in the supernatant was determined by DNA band quantification using Image J software. The concentration of magnesium at which 50% of the reconstituted chromatin was precipitated was reported.

Electron microscopy

EM samples were made with ~25 μl of reconstituted arrays with A260 nm between 0.5–2.5 OD. Arrays were incubated at various salt conditions (5 mM NaCl, 150 mM NaCl, or 1 mM
MgCl2), and 0.1% glutaraldehyde at 4 °C for ~5 h, followed by ~12-h dialysis against buffer (10 mM HEPES, 5 mM NaCl, 0.1 mM EDTA, pH=7.5) at 4 °C using 10 000 MWCO membranes. Specimens were diluted 10–20 times with 50 mM NaCl and applied to carbon-coated and glow-discharged EM grids (T1000-Cu, Electron Microscopy Sciences). Grids were stained with 0.04% uranyl acetate. For EM images (indicated in the figures), the uranyl acetate-stained grids were dried under high vacuum in a turbo-pumped evaporator (BA080, Baltec Inc., Hudson, NH) and shadowed with platinum at an angle of 10° as we described (Springhetti et al, 2003). Dark-field images were obtained and digitally recorded as before (Grigoryev et al, 2009), and also using a JEM-1400 electron microscope (JEOL USA, Peabody, MA) at 120 kV with SC1000 ORIUS 11 megapixel CCD camera (Gatan Inc. Warrendale, PA).

Conflict of Interest

The authors declare that they have no conflict of interest.

Supplementary Information

Supplementary information is available at:

http://emboj.embopress.org/content/embojnl/31/10/2416/DC1/embed/inline-supplementary-material-1.pdf

Acknowledgements
We thank Sangita Chakraborty and undergraduate interns at Penn State College of Medicine, Hershey: Jason Richards, Neihl Williamson, Adrienne Baksi, Michael Freiberg, and Monica Kriete for experimental assistance. We are thankful to J Widom for clone 601 DNA and discussing the internucleosome rotational experiments, and to C Woodcock for advice on electron microscopy. We thank R Myers for technical assistance with electron microscopy at Penn State Hershey Imaging Facility. This work was supported by NSF grant MCB-1021681 to SA Grigoryev and AHA pre-doctoral graduate fellowship 0815402D to SJ Correll.

Author contributions: SA Grigoryev designed the experiments, SJ Correll, MH Schubert, and SA Grigoryev performed the experiments and analyzed the data, SJ Correll and SA Grigoryev wrote the manuscript, and MH Schubert edited the manuscript.

References


Birnboim HC, Doly J (1979) A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res 7: 1513–1523


111


Copyright © 2012 European Molecular Biology Organization