Investigating higher-order chromatin structure and SAGA cooperativity using existing and modified in vitro chromatin systems

Divya Sinha
Iowa State University
Investigating higher-order chromatin structure and SAGA cooperativity using existing and modified in vitro chromatin systems

by

Divya Sinha

A dissertation submitted to the graduate faculty
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Major: Molecular, Cellular and Developmental Biology

Program of Study Committee:
Michael Shogren-Knaak, Major Professor
Guru Rao
Kristen Johansen
Mark Hargrove
Yanhai Yin

Iowa State University
Ames, Iowa
2013

Copyright © Divya Sinha, 2013. All rights reserved
To Maa, Papa, Niti Di, Jeejaji, Neha and Anvesh for your unconditional love, continuous faith and unwavering support
# TABLE OF CONTENTS

DEDICATION ii

ACKNOWLEDGMENTS v

ABSTRACT ix

## CHAPTER 1. GENERAL INTRODUCTION

- Chromatin 1
- Chromatin Structure 2
  - Nucleosome: The basic structural unit of chromatin 3
  - Hierarchical levels of chromatin organization 4
- Chromatin modifiers 5
  - Histone modifications: Writers, Readers and Erasers 5
  - Histone acetyltransferases 6
  - Nucleosome positioning and chromatin remodeling 8
  - Chromatin remodelers 9
- In vitro study of chromatin 10
  - Nucleosomal arrays as in vitro chromatin model systems 10
  - Histone assembly 10
  - DNA templates for nucleosome assembly 11
- Chromatin compaction studies 12
  - Requirement of histone tails in formation of secondary and higher-order chromatin structures 13
- Dissertation organization 15
- References 16
- Figures 25

## CHAPTER 2. ROLE OF DIRECT INTERACTIONS BETWEEN THE HISTONE H4 TAIL AND THE H2A CORE IN LONG RANGE NUCLEOSOME CONTACTS

- Abstract 28
- Introduction 29
- Materials and Methods 30
- Results 33
- Discussion 38
- Footnotes 48
- References 57
- Figures 58

- 25
CHAPTER 3. DEVELOPMENT OF NUCLEOSOMAL SYSTEMS TO STUDY CHROMATIN REMODELERS AND THE EFFECTS OF NUCLEOSOMAL POSITIONING ON CHROMATIN STRUCTURE

Abstract 72
Introduction 73
Materials and Methods 74
Results 77
Discussion 82
References 85
Figures 89

CHAPTER 4. INVESTIGATING THE ROLE OF ADA3 ACETYLATION IN SAGA COOPERATIVITY

Abstract 94
Introduction 95
Materials and Methods 97
Results 99
Discussion 101
Acknowledgments 102
References 103
Figures 106

CHAPTER 5. CONCLUSIONS AND FUTURE WORK 108
ACKNOWLEDGEMENTS

Behind the success of every individual, there is a lot of hard work and sweat, not just of that person but also of an entire team of people who shape up the path to that success in myriad ways. My experience at Iowa State University has been an enriching one and it was possible because of the support of several individuals. I would like to extend my deepest gratitude to them.

My major professor, Dr Michael Shogren-Knaak, has been a great mentor and has been pivotal in shaping my critical and scientific thinking. He is extremely patient with training students and uses questions as great learning tools. One of the several invaluable lessons I have learned from him is to stay focused and calm irrespective of the results obtained.

I am grateful to my committee members Dr Guru Rao, Dr Kristen Johansen, Dr Mark Hargrove and Dr Yanhai Yin for their support and for being so easily approachable. Dr Guru Rao has always given me suggestions and feedback that have strongly encouraged me to perform better. I would also like to thank Dr Kristen Johansen for her insightful suggestions and candid guidance.

This work would not have been possible without the help and support of the members (present and past) of the Shogren-Knaak lab: Dr Melissa Blacketer, Dr Hamid Azzaz, Chitvan Mittal, Sannie Olson, Isaac Young, Maggie Gannon, Haitham Alabsi, Sarah Wendler and Junko Yamamoto. They all deserve special thanks for making the work environment very supportive and
friendly, we have shared good laughs and innumerable discussions. Hamid taught me some of the initial techniques when I joined the lab. Chitvan, graciously and very patiently, taught me the techniques that she had been using on one of the projects. Sannie has been a great addition to the lab and has beautifully taken the project forward. I express my sincere gratitude to Melissa for being an excellent scientist, colleague and a very precious friend. Not only have the countless insightful scientific discussions with her enriched my knowledge, but she has also always found out a way to help me during my difficult times and provided constant encouragement and support.

I am thankful to the wonderful staff members at Iowa State University. Kathryn deserves a very special thank you for helping me settle down in Ames when I first came to the United States. She has been an amazing friend. I am grateful to Katie, Kelly, Connie and Stacey for their timely help in completing paperwork and other requests. Special thanks go to Gary Polking at the DNA facility for his willingness to help and for useful suggestions during problems with PCR and cloning. I would like to thank Pete Lelonek and Herman Huang who have readily helped me at multiple occasions. I would also like to thank Trish Stauble who is a great person and brings a lot of positivity to all the students she knows.

My time in Ames has been so much fun because of my friends Soniya, Smita, Vijay, Pooja, Deepak, Neha, Kuldeep and Dustin. I am very grateful to my friend Rupai for always believing in me and for her constant support.
I would like to acknowledge Dr Ichha Purak, Dr Sadhana Nighojkar, Dr Anand Nighojkar and Grand Master Yong Chin Pak who have been great teachers. They have taught me the value of hard work, to never set limitations for myself and to always strive for the best.

Most important of all, I want to thank my family for everything. My parents, Nalini Sinha and Birendra Kumar Sinha are the best! They have loved me unconditionally and always believed in me. They have taught me, by setting examples, to keep smiling under all conditions, to work hard, persevere, give importance to family, be considerate and to value people. They inspire me every single day. My sisters Niti Bariar, Kirti Gopal and Neha Verma are the best sisters and friends. We have shared the best laughs and worst fights but among all that they have always loved, encouraged and helped me. I would like to thank my brothers-in-law Ratnesh Bariar, Akshay Gopal and Awanish Verma for their constant love and support. I specially thank my brother-in-law Ratnesh Bariar, without whose tireless help and support I may not have been able to apply for graduate programs on time. My sisters’ kids Shubhangi, Suhani and Aditya fill my life with love, joy and invaluable friendship. I am also thankful to my parents-in-law for their support. I would like to thank my sister-in-law Deepali, brother-in-law Akhilesh and their kids Armaan and Annika for their love and support.

I thank my husband Yash for his love and support. He has encouraged me to pursue my dreams and motivated me when I doubted myself. He is not only a great husband and dad but also an excellent scientist and inspires me to
be my best. My son Anvesh is the best thing to have happened to me. He fills my life with unexpected moments of fun and laughter. He has taught me how to love unconditionally and keeps reminding me what’s important in life. Thank you sweetheart!
ABSTRACT

The organization of genomic DNA with histones and other proteins ensures the proper storage, utilization and segregation of DNA’s encoded information in a healthy eukaryotic cell. These components comprise chromatin, and knowledge of its structure, mechanisms of its formation, and the dynamic changes it undergoes is essential to gain insights into chromatin’s role in controlling gene regulation and the defects associated with misregulation. Here we describe our efforts to understand the different aspects of chromatin by using existing and modified in vitro chromatin model systems.

To better understand histone-histone interactions involved in the formation of higher-order chromatin structure, we employed a disulfide cross-linking strategy previously used to study short-range nucleosomal interactions. Using in vitro assembled nucleosomal arrays, we show that histones H4 and H2A, belonging to nucleosomes on different arrays, directly interact with each other under conditions that promote array-array associations. Additionally, prior intra-array cross-linking of nucleosomal arrays has an antagonistic effect on inter-array self-association. Together, our data show the role of H4-H2A contacts in the interplay between short-range nucleosomal compaction and higher-order chromatin structure.

Nucleosomal arrays used above, generated by assembling histone octamers on DNA templates, provide an excellent model system due to their
homogeneity and reproducibility in assembly. However, there is still a need for improved and novel systems to further expand the scope of *in vitro* chromatin studies. We have generated new DNA templates for nucleosome assembly and have improved our native chemical ligation technique for generating post-translationally modified histones, reducing protein racemization. Acetylated histones generated by the modified histone ligation method were successfully used for the study of ATP-dependent chromatin remodeling by SWI/SNF and RSC complexes.

Finally, previous studies from our group using nucleosomal arrays showed that SAGA, a histone acetyltransferase complex, acetylated nucleosomes cooperatively. Preliminary results indicated that this cooperativity requires functional Gcn5 bromodomain and acetylation of Ada3 subunit. Our follow-up experiments to dissect the role of individual lysine acetylation on Ada3 show that Ada3 lysine 8 pre-acetylated peptide binds tighter to the Gcn5 bromodomain and is a better substrate for acetylation by the Gcn5/Ada2/Ada3 subcomplex as compared to the Ada3 lysine 14 pre-acetylated peptide.
CHAPTER 1

GENERAL INTRODUCTION

Chromatin

Since the first experiments of Gregor Johann Mendel shed light on the existence of hereditary elements in living organisms, our understanding of these elements and their regulation has grown exponentially. Milestones include the identification of the hereditary units as genes and their chemical nature that of DNA (1). DNA is the blueprint of life in all free-living organisms and expresses itself by guiding the formation of RNA and proteins, which in turn regulate the expression of DNA along with performing other cellular functions.

Genomic DNA in each cell of eukaryotic organisms can be meters in length when fully extended. To be able to fit these long strands of DNA into a cell nucleus of a few micrometers in size, while maintaining its accessibility and regulation, DNA in every eukaryotic cell nucleus exists in the form of a nucleoprotein complex called chromatin.

Chromatin is a complex assembly of DNA, histone proteins and other chromatin associated proteins that packages millions of base pairs of DNA in an extremely small nuclear volume with a packaging ratio up to ~1:400,000 (2). Apart from functioning in DNA packaging and storage, chromatin organizes the genome both structurally and spatially for the differential genomic expression at various developmental stages, and in diverse cells and tissues (3). Additionally,
this structural organization prevents entanglement of DNA molecules and is a prerequisite for faithful segregation of the genome during cell division (4).

Through the organization and utilization of genomic DNA, chromatin provides epigenetic control of gene expression during processes like development and aging. Misregulation of chromatin functions can lead to developmental disorders and cancers (5-7). Thus, knowledge of chromatin structures, mechanisms of their formation and the dynamic changes it undergoes is required for complete molecular understanding of DNA-related processes like transcription, replication and repair, which take place in a chromatin-environment. This knowledge is also essential to gain insights into chromatin related diseases and identify possible treatment options (8,9).

**Chromatin Structure**

Since chromatin packages DNA, it seems reasonable that chromatin structure would be a way to regulate these biological processes. Indeed, some of the earliest staining of interphase nuclei showed the presence of two states of native chromatin: more open and lightly stained euchromatin that corresponds to most of the expressed genes; and, darker-stained and compact heterochromatin corresponding to mostly silenced genes (10). In addition, the densest form of chromatin was visualized as chromosomes during eukaryotic cell division, where the formation of chromosomes helps in equal distribution of the replicated DNA to the daughter cells. With the advancement and refinement
of biochemical, biophysical and molecular imaging techniques it was soon discovered that chromatin has a hierarchical organization that corresponds with the earlier visualized compaction states and that it is highly dynamic in nature (11,12).

**Nucleosome: The basic structural unit of chromatin**

The basic structural unit of chromatin, the nucleosome (Figure 1), is composed of 147 bp of DNA wrapped ~1.65 times around histone octamer as a left-handed superhelix. The histone octamer that forms the nucleosome core, is a disc-shaped structure composed of two copies each of the four core histones H2A, H2B, H3 and H4 (13,14). Histones are very basic proteins, which neutralize highly negatively charged DNA and increase its ability to be bent upon association.

One of the most striking features of histones is that each type of histone has been highly conserved through evolution (up to ~95% sequence conservation for histone H4) (15). Such high levels of sequence conservation for individual histones point toward functional similarity and conservation of structures they form. Moreover, even though the different types of core histones do not have any sequence similarity with each other they still share structural homology (16). For example, although histones H3 and H4 differ in their amino acid sequence, they form similar structural motifs (Discussed later).
In addition to the regular histones, all the eukaryotic organisms also express evolutionarily conserved histone variants (17,18). These variants are expressed from genes different from the ones that encode major core histones and are specialized to replace specific core histones at specific genomic locations (17,18). Examples include histone H3 variant CENP-A localized at centromeres (4,19) and Histone H2A variant H2A.Z flanks transcriptional start sites (20,21). Some histone variants, like H2A.Bbd, even carry out tissue-specific transcription (22).

Hierarchical Levels of Chromatin Organization

Stretches of histone-free DNA, called linker DNA, link the nucleosomes together to form the primary chromatin structure called the nucleosomal arrays. The linker DNA typically varies between ~10-50 bp in length (23) and allows binding of linker histones at sites where DNA enters and exits the nucleosome. Linker histones, categorized into two major classes - H1 and H5, are histones that are not a part of the histone octamer that forms the nucleosome core. Instead, by binding at the ‘entry-exit’ site they further stabilize the nucleosome structure and form a chromatosome (12,24,25). The term ‘chromatin array’ is often used to describe nucleosomal arrays bound with linker histones (12,26). These arrays, with or without linker histones, are fully extended ~10 nm diameter fibers and have a “beads-on-string” appearance.
The nucleosomal arrays further compact to form secondary (the traditional “30 nm” fiber), tertiary and additional levels of chromatin structure and, thus, can eventually achieve chromosomal levels of compaction (27,28). The structures beyond the 30 nm fiber are commonly referred to as higher-order chromatin structures and most of the chromatin, even in the interphase nuclei, is believed to exist in this form (29). Most studies to date have concentrated on the 30 nm fiber structure but the structure and existence of this fiber in vivo is still controversial (30-33). It is hypothesized that 30 nm fibers further compact to form higher orders of chromatin structure, however, it is still unclear how (Figure 2).

**Chromatin modifiers**

Transcriptional regulation, as well as all other chromatin functions, are highly influenced and regulated by local and global chromatin structure. Post-translational modifications of histone protein side chains and alterations in nucleosomal positioning in the genome are two of the major ways to achieve this regulation (34).

**Histone modifications: Writers, Readers and Erasers**

Post-translational modifications (PTMs) are a recurring theme in a multitude of cellular functions, ranging from cell signaling to transcription to targeting proteins for degradation. Histones undergo a wide variety of histone
modifications, including acetylation, methylation, phosphorylation, and ubiquitination (35,36). These modifications predominantly take place on specific amino acid residues on the unstructured N-terminal regions; however, the central globular domains of histones can also be modified (37).

Multiple enzymes target histones to put modification marks on them. These enzymes can be broadly categorized as ‘writers’ (38). ‘Readers’, a second group of enzymes/proteins, can read these modifications through the presence of specific domains and carry out the intended function (38). Examples include chromodomain to recognize methylated lysines (39,40) and bromodomain to recognize acetyl groups on lysines (41,42). Finally, a third group of enzymes, classified as ‘erasers’, can remove these modifications once their intended purpose is fulfilled.

In the following sections, only certain aspects related to histone writers and readers, specifically with respect to histone acetylation, will be briefly discussed. ‘Erasers’ are beyond the scope of this dissertation and will not be discussed further.

_Histone Acetyltransferases_

Of all the known histone modifications, histone acetylation was among the first histone post-translational modifications discovered (43). It is also the modification most associated with opening of the chromatin structure (44,45) and gene activation (46,47). The enzymes that bring about histone acetylation
are referred to as Histone acetyltransferases (HATs) and usually exist as multisubunit complexes within the cell (48). HATs can modify histones with varying specificity towards different histones and towards different lysine residues within each histone (48). However, multiple non-histones substrates have been identified for the HATs as well and due to this reason many HATs are now referred to as Lysine acetyltransferases (KATs).

The first multisubunit HAT complex discovered was SAGA (Spt-Ada-Gcn5-acetyltransferase) from *Saccharomyces cerevisiae* (49) (Figure 3). It is also the most extensively studied and best understood of all HATs. More aspects of SAGA will be reviewed in Chapter 3.

Based on the substrate recognition motifs, HATs are classified into two major classes, GNAT family and MYST family (50,51). There is an additional class of HATs consisting of p300/CBP and Taf1 and referred to as an “orphan-class” since these do not contain the same consensus motif as the major HATs (50). GNAT (Gcn5 N-acetyltransferases) family is named after Gcn5 and includes Gcn5, PCAF, Elp3, Hat1, Hpa2 and Nut1 as the catalytic subunit of this family of HAT complexes. MYST family of HATs is named after its members Myst, Ybf2, Sas2 and Tip60 (48).

The existence of all the HATs in the form of multisubunit complexes in cells contributes to their specialized functions and regulation. In addition to the histone acetyltransferase activity several of these complexes possess additional functions including acetylation of non-histone substrates like transcription factors
and chromatin remodelers (53), histone deubiquitination, DNA repair (54), recruitment to specific genomic locations, and regulation of these functions. In view of the diverse roles played by HAT, it is not surprising that misregulation of their activity has been associated with congenital disorders and tumorigenesis (55-57).

Nucleosome positioning and chromatin remodeling

DNA sequences govern, at least in part, the organization of nucleosomes both in vitro and in vivo, thereby determining nucleosome positioning within the cells to a major extent (58).

Nucleosome distribution in the eukaryotic genome, as observed from global nucleosome positioning maps, is mostly non-uniform, with certain regions enriched with nucleosomes and others being nucleosome deficient (59,60). Even within the nucleosome-enriched regions, the nucleosomal spacing is frequently non-regular, and this variability plays a role in chromatin fiber compaction and function.

There exist regions in the eukaryotic genome containing long ranges of regularly spaced nucleosomes and these regions have been associated with heterochromatin and transcription-silencing (61). Additionally, some chromatin-remodeling complexes (discussed ahead), like those belonging to the ISWI family, have the ability to reposition nucleosomes at regular intervals (62). Such requirements suggest a role of nucleosome position and spacing in regulation of
chromatin structure by affecting compaction.

**Chromatin Remodelers**

Although nucleosome organization exhibits DNA sequence preferences, chromatin landscape is not completely dictated by genomic DNA sequence and varies among different tissues of the same organism. Even within the same cell, chromatin is highly dynamic and this property is required for transcription, gene silencing, replication and repair processes.

To be able to modify nucleosome positions and provide access to underlying DNA, cells employ a specialized set of enzymes called ‘chromatin remodelers’. These remodelers are enzyme complexes that have been highly conserved throughout eukaryotic evolution and utilize the energy of ATP to move, evict, disassemble or restructure a nucleosome (63).

There are four families of chromatin remodeling complexes: SWI/SNF, ISWI, CHD and INO80 (63). Although these families perform specialized remodeling functions, they have certain features in common. Some of these common features are: presence of ATPase subunit, domains to recognize specific histone modifications (readers), and subunits to recruit additional proteins (63). Among these remodeler families, members of the ISWI family have the ability to regularly position the nucleosomes.
**In vitro study of chromatin**

Much of our current understanding about chromatin structure and the contribution of various chromatin-associated components in regulation of genomic DNA comes from *in vitro* studies. Initial studies depended on chromatin or histones purified from eukaryotic nuclei. Although this chromatin represents “native” state, there are multiple disadvantages/limitations of this approach including heterogeneity in nucleosome composition, unknown types and levels of histone modifications, variability in DNA sequences, and mixtures of sources – euchromatic vs heterochromatic. These limitations are solved by the use of *in vitro* assembled nucleosomal systems.

**Nucleosomal arrays as *in vitro* chromatin model systems**

Recombinant histones that are expressed and purified using standard techniques can be used for the generation of histone octamers (64). These octamers can be assembled on DNA templates containing multiple repeats of nucleosome positioning sequences creating nucleosomal arrays (65).

**Histone assembly**

All of the four core histones consist of a central ‘histone-fold’ domain comprised of three α helices connected by two less structured loops (14,66). Histone H2A-H2B and H3-H4 heterodimers form via interaction between these histone-fold domains in a ‘handshake-motif’ (14,67). *In vitro*, under high salt
concentrations, two H2A-H2B dimers flank one H3-H4 tetramer to form the histone octamer. Under physiological conditions, however, the H2A-H2B dimer and H3-H4 tetramer interact only in the presence of DNA (67,68).

In vitro, for histone octamer assembly on the DNA to form nucleosomes, H3-H4 tetramer deposits first, followed by subsequent and cooperative addition of the H2A-H2B dimers (68,69). In the nucleosomal unit, the first ~20-35 N-terminal amino acid residues of each histone extend past the nucleosomal DNA forming largely unstructured histone ‘tails’. In addition, H2A also has a C-terminal tail (14).

Multiple histone-histone interactions as well as multiple Histone-DNA contacts stabilize the nucleosome structure (14).

DNA templates for Nucleosome assembly

Two DNA templates most commonly used for nucleosomal array assembly are: (1) the 208-12 DNA sequence containing 12 repeats of 208 bp 5S rDNA nucleosome positioning sequence from *Lytechinus variegates* (70), and the 601-177-12 DNA template with 12 repeats of a 177 bp sequence artificially generated through multiple rounds of selection for strong nucleosome positioning preference (31,71). While the ‘601’ sequence is used to form arrays with highly regularly spaced nucleosomes, the 208-12 DNA is also used often as it represents a repeat of a natural DNA sequence.
The ability to generate nucleosomal arrays using pure components provides the opportunity to prepare arrays containing unmodified or specifically modified versions of histones individually or in various combinations and study their effects on array compaction (65,72). Additionally, the effects of varying linker lengths and nucleosome positions can be studied by generating desired DNA templates for array assembly (32).

Lately, novel approaches are being used to develop additional chromatin model systems. These include modified histone tail peptides, modified ligated histones (73,74), ligatable tetranucleosomal array system (75), and the generation of genetically encoded acetylated histones (76). These advancements have continued to further our ability to tease apart the role of individual chromatin components.

**Chromatin Compaction studies**

In solution, nucleosomal arrays undergo compaction in the presence of cations. At low concentrations of divalent cations (1 to 2 mM Mg$^{2+}$) the arrays fold via intra-array nucleosome-nucleosome interactions (31). When the cation concentrations are increased, a reversible and highly cooperative oligomerization of arrays is observed (26). While array folding mimics the formation of the 30 nm fiber, array oligomerization is believed to reflect long-range chromatin fiber-fiber interactions that occur in eukaryotic cells. Correspondingly, chromatin folding is associated with local changes in chromatin
structure, perhaps made possible by local regulation of divalent cation concentrations, while oligomerization is related to globally condensed ‘beyond 30 nm fiber’ chromatin structures (12,77,78).

The intra-array folded state which resembles formation of 30 nm fiber sediments as 55S species while the oligomerized arrays sediment as >300S species when analyzed using analytical ultracentrifugation techniques (45). Although inter-array species can assemble from the 30 nm fiber-like structures, they can also form independently. This is also confirmed by sub-saturated arrays, which can form array oligomers but not 30 nm fiber like structures (79). These differences suggest that different types of interactions might be involved in formation of 30 nm fiber structures and higher compacted states.

Inter-array compaction can be distinguished from intra-array compaction by differential centrifugation assay where nucleosomal arrays are mixed with Mg$^{2+}$ and centrifuged. The fraction of arrays (non-oligomerized) remaining in solution after centrifugation, as measured through change in absorbance, is plotted as a function of Mg$^{2+}$ concentration (45). Thus, more the self-association, more the sedimentation of arrays and, hence, lower their levels in solution.

Requirement of histone-tails in formation of secondary and higher-order chromatin structures

In the 2.8 Angstrom crystal structure of the nucleosome core particle amino acid residues 16-25 of the H4 tail were observed to interact with the acidic
patch, a stretch of seven amino acids containing mostly glutamate, of the histone H2A-H2B dimer of an adjacent nucleosome (14). This observation suggested a role of this specific interaction in the formation of secondary and tertiary chromatin structures.

Indeed, array compaction studies done prior to, and after, this revelation clearly indicate that histone tails play structural roles. Loss of tails individually or in combination affects both intra-array folding and inter-array oligomerization (77,80,81). Moreover, without core histone tails, other chromatin associated proteins like the linker histones cannot induce the formation of higher-order chromatin structures (82). Further, via disulfide cross-linking, the Richmond lab (31) demonstrated that the interaction between H2A and H4 histones observed in the crystal structure also occurs in formation of the 30 nm fiber and involves contact between residues H4V21 and H2AE64.

The roles of histone tails at distinct levels of chromatin organization might be different from each other. While in nucleosomes these domains bind and neutralize charges on the DNA, they play different roles in formation of higher order structures since the electrostatic neutralization by Mg2+ cannot substitute for tail loss (83). These interactions might vary for short-range and long-range interactions, as they would require different orientations of nucleosomes (parallel and anti-parallel, respectively) with respect to each other.

Overall, while the current state of chromatin-related information demonstrates that we have significantly advanced our understanding of the
hereditary elements and their regulation, it is also a humble reminder that we still have a long way to go and new milestones to uncover.

**Dissertation Organization**

This dissertation is organized into five chapters. Chapter 1 provides a general introduction to the field of chromatin and various aspects associated with it. Chapter 2, which is also a research article published in the Journal of Biological Chemistry, describes the role of interactions between histones H2A and H4 in mediating long-range nucleosomal interactions to form higher-order chromatin structure. Chapter 3 addresses the need for development of novel *in vitro* systems to study chromatin. It also describes our attempts to improve existing systems and develop new ones to address specific questions. Chapter 4 covers the experiments performed to understand the possible source of cooperativity shown by yeast histone acetyltransferase SAGA during *in vitro* nucleosome acetylation. Chapter 5 concludes the studies and discusses possible future work.
References


architecture in Schizosaccharomyces pombe. *Nature structural & molecular biology* 16, 1286-1293


melanogaster pericentric heterochromatin. *Molecular and cellular biology* 21, 2867-2879


**Figure 1.** Nucleosome core particle, top and side views (from (14)). Color codes for histones: Yellow - H2A; red - H2B; blue - H3; green - H4
Figure 2. Packaging scheme for the eukaryotic genome. A DNA molecule, 2 nm in diameter, wraps around the core histone octamer to form a nucleosomal array which is further stabilized by the binding of linker histones. The 30 nm fibers are hypothesized to form 100-400 nm chromatin fibers which exist in the interphase nucleus and are then further compacted by the help of other proteins and divalent cations to form a maximally condensed state of chromatin, the chromosome (From (11)).
Figure 3. Molecular architecture of the yeast SAGA complex (Wu & Winston, 2004)
CHAPTER 2

The role of direct interactions between the histone H4 tail and the H2A core in long-range nucleosome contacts

Divya Sinha¹ and Michael A. Shogren-Knaak¹

¹Department of Biochemistry, Biophysics, and Molecular Biology, Iowa State University, Ames, Iowa 50011

Address correspondence to: Michael Shogren-Knaak, 4214 Molecular Biology Building, Ames, IA 50011. Tel: 515-294-9015 Fax: 515-294-0453; E-mail: knaak@iastate.edu
Abstract

In eukaryotic nuclei the majority of genomic DNA is believed to exist in higher-order chromatin structures. Nonetheless, the nature of direct, long-range nucleosome interactions that contribute to these structures is poorly understood. To determine whether these interactions are directly mediated by contacts between the histone H4 amino-terminal tail and the acidic patch of the H2A/H2B interface, as previously demonstrated for short-range nucleosomal interactions, we have characterized the extent and effect of disulfide crosslinking between residues in histones contained in different strands of nucleosomal arrays. We show that in 208-12 5S rDNA and 601-177-12 nucleosomal array systems, direct interactions between histones H4-V21C and H2A-E64C can be captured. This interaction depends on the extent of initial cross-strand association, but does not require these specific residues, as interactions with residues flanking H4-V21C can also be captured. Additionally, we find that trapping H2A-H4 intra-array interactions antagonizes the ability of these arrays to undergo intermolecular self-association.
Introduction

In eukaryotic nuclei, DNA is packaged into chromatin in order to facilitate and regulate the storage, segregation, organization and utilization of the genome. Chromatin is a complex DNA-protein assembly that exhibits multiple levels of structures. Its most basic structural unit, the nucleosome, is composed of an octamer (two copies each of histone proteins H2A, H2B, H3, and H4), wrapped by 147 base pairs of DNA (1,2). In general, the majority of genomic DNA is sequestered in nucleosomes (3), and in their most extended form these nucleosomal arrays form a 10 nm fiber. However, even for cells in interphase, it is widely believed that most of the chromatin adopts higher-order structures. In these cells a variety of chromatin fibers have been observed, including fibers that are greater than 100 nm thick (4).

The understanding of higher-order chromatin structures has been significantly aided by in vitro studies of isolated and reconstituted nucleosomal array systems, where reversible short-range intra-array and long-range inter-array nucleosome associations can be induced even in the absence of additional chromatin-associated proteins (5,6). From these studies a number of factors important for higher-order chromatin structure have been identified. Within the nucleosome, the amino-terminal portions of the histones that extend past the nucleosomal DNA, the histone tails, have been shown to affect both intra- and inter-array associations (7-9). Amongst these tails, the histone H4 tail has the largest effect on both types of association (9-11), in a manner dependent on the
charge (12), modification state (13,14), and position of the H4 tail (12). Another important region within the nucleosome is the acidic patch interface of histones H2A and H2B (1), where mutations to this region can change both intra- and inter-array interactions (10,15,16).

How these nucleosome components contribute to inter-nucleosomal interactions is not completely clear. However, based on crystal contacts observed in the first high-resolution structure of a mononucleosome, i.e. a nucleosome not linked to other nucleosomes through intervening linker DNA, it was proposed that one way in which nucleosome interactions occur in nucleosomal arrays is through direct contact between the H4 tail of one nucleosome, and the H2A/H2B interface of another nucleosome (1,17). Indeed, for intra-array nucleosomal interactions, this contact has been captured by disulfide and photo-affinity crosslinking (18,19). Whether such interactions also occur for inter-array associations is not as clear, since support for such a model has been indirect and potentially conflicting. In some cases, changes to the H4 tail and the H2A surface result in similar changes in intra- and inter-array associations (9-11,13,15), suggesting either that both types of interactions share a common mechanism or that intra-array associations facilitate inter-array associations. In contrast, there are a number of examples where changes to the H4 tail-H2A/H2B interface interaction result in different intra- and inter-array association effects (10,16). This is seen, perhaps most dramatically, in the case of studies with tetramer arrays that lack the H2A and
H2B subunits entirely. Studies of these systems have shown that they are highly defective in intra-array association (20). However, despite the absence of any H2A or H2B histone, these arrays are just as capable of forming inter-array associations as arrays with a complete complement of histones (8). These results may indicate that the major mechanism of inter-array association is different from that for intra-array associations, or that these two types of interactions are interrelated in a way such that changes in one type of nucleosome to nucleosome interaction can affect the nature of the other.

To provide a foundation for interpreting these observations, we sought to determine the extent and nature of direct cross-strand interactions between the H4 tail and the H2A/H2B surface. Due to the complexity and large size of the self-associated array species, standard structural techniques are not readily applicable. Moreover, because the H4 interaction with the H2A/H2B surface is already known to be directly involved in intra-array interactions, any technique used must be able to separate this contribution from those involved in inter-array interactions. By adapting a nucleosomal array system previously developed to trap intra-array nucleosome interactions (18), we have been able to isolate and better understand the interactions of the H4 tail and H2A/H2B surface in inter-array self-association.
Materials and Methods

Template and Carrier DNA Preparation

208-12 5S rDNA and 601-177-12 DNA templates were excised from plasmids and purified by gel filtration chromatography according to standard methods (9,21). The 174 bp carrier DNA was prepared by PCR amplification of the purified 196 bp fragment that results from complete EcoRI-digestion of the 208-12 template. Carrier DNA was purified by phenol-chloroform extraction and gel electroelution. The purity and quantity of the template and carrier DNA were determined by gel electrophoresis and absorbance spectroscopy.

Histone Octamer Preparation

Xenopus histones were recombinantly expressed, purified, and characterized using standard methods (22). Cysteine-containing histones H2A-E64C, H4-K20C, H4-V21C, H4-L22C and H4-R23C were generated by QuikChange mutagenesis (Stratagene) of the bacterial expression vectors containing histone genes (22). Octamers were assembled and purified in the presence of 0.1 mM tris (2-carboxyethyl) phosphine (TCEP) reductant using standard methods (22). All octamers were prepared using H3-C110A. Octamers were characterized by SDS-PAGE gel electrophoresis and absorbance spectroscopy (22).
**Nucleosomal Array Assembly**

Nucleosomal arrays were assembled by mixing histone octamers and DNA components followed by step-wise salt dialysis method as previously described (13,21). 0.1mM TCEP reductant was added to the DNA and octamer mixture and to the dialysis solutions. For 208-12 arrays, molar ratios of octamer to template varied from 0.9 to 1.1. For 601-177-12 arrays, octamer to template to carrier DNA molar ratios were 1.15-1.3:1.0:0.3. Carrier DNA and mononucleosomes were removed by differential centrifugation as previously described (9). The final composition of the array solutions included array buffer (2.5 mM NaCl, 10 mM Tris-HCl pH 8.0, 0.25 mM EDTA for 208-12; 2.5 mM NaCl, 10 mM HEPES pH 8.0 for 601-177-12) and 0.1 mM TCEP, and arrays were stored at 0 °C. The arrays were quantified based on amount of DNA by measuring absorbance at 260 nm. Array saturation was analyzed by restriction endonuclease analysis, sedimentation velocity analysis and differential centrifugation assay as described below. Arrays with similar nucleosome saturation were used for studies.

**Disulfide Crosslinking of Nucleosomal Arrays**

2X crosslinking solutions were prepared by adding 1M Tris-HCl pH 9.0 (2X final concentration of 100 mM), 100 mM glutathione (2X final concentration of 2 mM, molar ratios of oxidized (Sigma) to reduced (Acros Organics) glutathione of 1:15, 1:7, 1:3, 1:1, 3:1, 7:1, 15:1), and 500 mM MgCl₂ (2X final
concentrations from 0-12.0 mM for inter-array crosslinking, 2.0 mM for intra-array crosslinking) to the appropriate array buffer. To initiate crosslinking, this solution was mixed in equal volume with nucleosomal arrays (concentration of 50ng/µl of DNA template). For inter-molecular crosslinking, the arrays were incubated at RT for 16h followed by addition of an equal volume of EDTA solution (20 mM final) and further incubation at RT for 2h. For intra-molecularly crosslinked arrays, the samples were incubated at 37°C for 16 h. These samples were then dialyzed three times at 4°C in array buffer. After dialysis, the absorbance values of samples at 260 nm were measured before and after centrifugation at 14,000g for 10 min at RT to quantify and remove any highly crosslinked species. Array concentrations were determined by subtracting the 260 nm absorbance of a mock reaction consisting of all of the components except the array.

*Preparation of Crosslinked Histone Standards*

Ideal solution conditions for generating crosslinked histone standards are those in which the histone is fully reduced and denatured, and where the solution is at a high ionic strength to allow close approach of the highly basic histones. 7M Guanidine hydrochloride fulfills these criteria and was used for the reduction and crosslinking steps. Urea was used in the dialysis steps used to remove the reductant because of its lower cost. Specifically, lyophilized histones H2A-E64C and H4-V21C were resuspended in unfolding buffer (7M Guanidine-
HCl, 20 mM Tris-HCl pH 7.4, 10 mM DTT), dialyzed two times in 7M urea and 20 mM Tris-HCl pH 7.4, and finally in 7M guanidine-HCl and 20 mM Tris-HCl pH 7.4. Dialyzed histones were then quantified by absorbance. H2A-E64C alone, H4-V21C alone, or a one-to-one mixture of H2A-E64C and H4-V21C, were mixed with 4 mM 5,5’-dithiobis-(2-nitrobenzoic acid) (DTNB) prepared in 0.1 M phosphate buffer pH 7.6 to achieve a final concentration of 100:50 µM histone to DTNB for single histones, and 200:100 µM histones to DTNB for mixed histones. This mix was incubated at RT for 12-16 h and then dialyzed in 0.1% trifluoroacetic acid at 4°C. The histones were then dried, resuspended in protein loading dye with no reducing agents, and stored at -20°C.

Restriction Endonuclease Analysis

As previously described, EcoRI and Scal digestion were used to characterize the extent of nucleosome saturation of the 208-12 and 601-177-12 arrays, respectively (9,21). Resulting mononucleosomes and free DNA were analyzed on 4% native PAGE gel followed by ethidium bromide staining.

Sedimentation Velocity

Sedimentation velocity experiments were performed with a Beckman XLA ultracentrifuge. Nucleosomal arrays were analyzed at concentrations ranging from 18ng/µl to 25ng/µl of DNA template and at speeds from 12,000 to 16,000 RPM. TCEP to a final concentration of 0.1 mM and 0.1 mM additional EDTA
were added to non-crosslinked nucleosomal arrays prior to analysis. For intramolecularly crosslinked arrays the mock sample was used as reference and no TCEP or EDTA was added to the samples. The data was analyzed using the method of van Holde and Weischet on Ultrascan data analysis software (Dr. B. Demeler, University of Texas Health Science Center, San Antonio, TX) as described previously (23).

**Differential Centrifugation**

Differential centrifugation analysis of nucleosomal arrays prior to crosslinking was performed largely as previously described (8). In short, nucleosomal arrays (~30 ng/µl of DNA template) were mixed with an equal volume of array buffer containing both 0.1 mM TCEP reductant and MgCl$_2$ at twice the desired final concentration. Following 15 min incubation at room temperature (RT), the arrays were centrifuged at 14,000g for 10 min at RT. The absorbance of array in the supernatant was then determined at 260 nm. To calculate the fraction of nucleosomal array remaining in solution, this absorbance was divided by the absorbance of array that remains in solution when treated similarly, but with a 0 mM final MgCl$_2$ concentration. The differential centrifugation analysis of crosslinked arrays was performed in a similar manner, but with the following differences: No TCEP reductant was present and the absorbance of array in the supernatant was determined relative to a mock crosslinked sample, which lacked nucleosomal arrays, but was
otherwise treated identically. For the inter-molecularly crosslinked arrays, no additional array buffer or MgCl₂ was added prior to measuring the absorbance. For arrays containing no cysteine residues, around 75% recovery of initial signal was observed, potentially due to incomplete magnesium ion sequestration or a different array subpopulation, and the absolute recovered absorbance has been reported. For intra-molecularly crosslinked array, 10X MgCl₂ in array buffer was added in a ratio of 1:9 instead of the 2X MgCl₂ solution prior to measuring the absorbance. All trials were repeated at least three times mostly with the same array preparation, and presented either as representative data or mean values, with error bars representing the standard deviation.

**Non-reducing SDS-PAGE**

Histones from crosslinked arrays were TCA precipitated (20% trichloroacetate, final) and resuspended in protein loading dye with no reducing agent. Histones were separated on an 18% SDS-PAGE gel and visualized using Coomassie blue.

**Results**

*Inter-array Crosslinking in 208-12 Arrays via H4-V21C and H2A-E64C*

Direct interactions between the histone H4 tail of one nucleosome and the H2A/H2B acidic patch of another have been previously demonstrated within the same nucleosomal array by Richmond and coworkers using oxidative
crosslinking (18). In this system, substitution of cysteines for histone residues H4-V21 and H2A-E64 allowed them to trap interactions that occurred under conditions where nucleosomal arrays exhibit intramolecular, but not intermolecular compaction. To adapt this technique to study the extent of inter-array contacts between the H4 tail and H2A/H2B acidic patch, our strategy was to generate two different sets of arrays, where each array consisted of nucleosomes that included either H4-V21C histones or H2A-E64C histones. With a mixture of these arrays, if this interaction is involved in inter-array association, then disulfide crosslinking would be induced under conditions which cause inter-array self-association. The products of this crosslinking can then be characterized to assess how much of the array association persists under conditions which do not normally promote array self-association. Additionally, the nature of the crosslinked histone species can be determined, where any observed H2A-H4 crosslinking is only possible via inter-array contacts.

To generate well-defined nucleosomal arrays, wild-type and cysteine-containing *Xenopus laevis* histones were recombinantly expressed, purified, and assembled into histone octamers. These octamers were then deposited onto recombinantly expressed and purified DNA templates by step-wise salt dialysis. In our initial experiments 208-12 DNA templates containing 12 head to tail repeats of the naturally occurring *Lytechinus variegatus* 5S rDNA sequence were used (24). Arrays were assembled to be significantly saturated, but not oversaturated, and to be closely matched in saturation. To confirm this, arrays
were characterized by EcoRI digestion (Figure 1A), sedimentation velocity analysis (Figure 1B), and cation-dependent differential centrifugation assays (Figure 1C). Comparison of the results for the three arrays indicate that the arrays are well matched in saturation, that the array are nearly saturated (21,25), and that the presence of the cysteine residues does not change their properties under non-oxidizing conditions. In the differential centrifugation analysis (Figure 1C), in contrast to arrays assembled from isolated, endogenous histones (26), no plateau is observed for lower magnesium ion concentration. Nonetheless, this behavior is consistent with previous studies of arrays reconstituted from recombinantly expressed core histones (10,25), and occurs regardless of whether or not the histones in the arrays contain cysteine residues (Figure 1C).

To capture inter-array interactions, equal amounts of H2A-E64C and H4-V21C arrays were mixed with a solution containing magnesium chloride and glutathione. With a final concentration of divalent magnesium ion of either 4.0 or 6.0 mM, the arrays were expected to be fully self-associated and form species, which exhibit a large sedimentation coefficient (Figure 1C). The glutathione in solution contained equal molar amounts of its oxidized and reduced forms, creating a redox buffer in which histone disulfide formation could reach equilibrium with respect to the redox potential established by the ratio of glutathione species. Prior to analysis of the crosslinked arrays, EDTA was added to chelate the divalent magnesium and thereby limit inter-array association due to noncovalent interactions.
Differential centrifugation analysis of these arrays, in which the amount of non-sedimented arrays is assessed after centrifugation, revealed that the glutathione-treated array mixture generated one or multiple species with very large sedimentation coefficients relative to wild-type arrays, consistent with extensive inter-array crosslinking (Figure 1D). Significant sedimentation was not observed with similar treatment of arrays containing H2A-E64C histones (Figure 1D), indicating that H2A-H2A crosslinking was not responsible for the inter-array association. However, arrays containing only H4-V21C did show significant sedimentation, suggesting that H4-H4 crosslinking could be responsible for the inter-array association in the mixed array experiment. To directly assess the nature of the histone crosslinking, the histone composition of these reactions was analyzed by non-reducing denaturing gel electrophoresis (Figure 1E). Consistent with the differential centrifugation, this analysis shows that H2A-E64C arrays do not generate crosslinked histones, while H4-V21C arrays, as well as mixtures of H2A-E64C and H4-V21C arrays, do. The H4-V21C arrays generate H4-H4 crosslinked histones, while the H2A-E64C and H4-V21C array mixture predominantly captures direct inter-array interactions between histones H4-V21C and H2A-E64C, while also producing some of the H4-H4 crosslinked species.

To determine the relative stability and role of the H2A-H4 crosslink in inter-array association, we investigated the effects of decreasing the oxidizing potential in the crosslinking reaction. For the H2A-E64C and H4-V21C array
mixture the formation of inter-array crosslinked species that sediments remains constant over decreasing ratios of oxidized to reduced glutathione (Figure 2A). In contrast, formation of such species for the H4-V21C arrays alone decreases with decreasing oxidizing potentials. At an oxidized to reduced glutathione ratio of 1:15, the array mixture shows complete sedimentation, while all individual array species show no discernable sedimentation (Figure 2B). Analysis of the histones under these conditions shows that for the mixed array experiments, crosslinked H2A-H4 histones are the predominant crosslinked species, although some H4-H4 crosslinking is still observed (Figure 2C). For the H4 arrays alone, H4-H4 crosslinking is still present, although slightly reduced relative to the higher oxidation potentials (data not shown). This indicates that the nature of the remaining H4-H4 crosslinked histones is insufficient to facilitate inter-array sedimentation. In the mixed arrays, the amount of H4-H4 crosslinking is even less and suggests that, since the larger amounts of H4-H4 crosslinking observed in the H4-V21C arrays alone were not sufficient to promote inter-array sedimentation, the smaller amounts present in the mixed arrays are not likely to be the species responsible for the inter-array sedimentation, i.e., the H2A-H4 crosslinked species are the predominant species responsible for the inter-array sedimentation.

**Dependence of Inter-array Crosslinking on Array Self-association**

In self-associated arrays, inter-array crosslinking between H4-V21C and H2A-E64C is favored over the other potential modes of crosslinking. We
expected that this crosslinking was because these sites are brought into close spatial proximity with one another in self-associated arrays. To directly test this idea we determined the relationship between array self-association and inter-array crosslinking.

As has been previously shown (8), and is apparent in Figure 1C, self-association of arrays is facilitated by divalent cations. Thus, if the observed inter-array crosslinking requires array self-association, the degree of inter-array crosslinking should decrease with decreasing amounts of divalent cation. Indeed, formation of species sufficiently crosslinked to sediment in the absence of divalent cations occurs only when initial crosslinking is performed at higher concentrations of divalent cation (Figure 3A). Formation of this stably associated species appears to require H2A-H4 interactions, as arrays with only one of these components do not show significant differential centrifugation. Further, analysis of the histones from the H2A-E64C and H4V21C array mixture after crosslinking treatment shows the H2A-H4 crosslinked pair to be the predominant species, with its presence increasing with increasing concentrations of divalent cation (Figure 3B). Thus, these data suggest that the observed inter-array crosslinking requires array self-association.

**Specificity of Inter-array H4 to H2A Interactions**

As inter-array interactions between H4-V21C and H2A-E64C are captured in self-associated arrays, we wondered how specific this interaction
was. To test the specificity, we investigated crosslinking of other sites in the H4 tail to histone H2A-E64C. Like H4-V21, substitution of residues directly adjacent to this site with cysteine, i.e. H4-K20C and H4-L22C, resulted in arrays with an ability to crosslink to themselves in the absence of H2A-E64C, but preferentially crosslink to H2A-E64C when it is present (Figure 4A). Also like H4-V21C, the products generated by H4-H4 crosslinks resulted in less differential centrifugation than products generated by the H4-H2A crosslinking (Figure 4B). The similarity in behavior suggests that the inter-array interaction between the H4 tail and the H2A/H2B acidic patch does not have to involve specific residue contacts, since multiple H4 tail residue contacts with H2A-E64C can be captured. However, the preference for H2A-E64C contacts appears to be constrained to a specific region of the H4 tail, since substitution of H4-R23 with cysteine does not favor crosslinking to H2A-E64C, but rather predominantly crosslinks to itself (Figure 4A).

**Intra-array Crosslinking Affects Inter-array Associations**

Prior studies indicate that the H4 histone tail contributes to intra-array interaction through direct contacts with the surface of histone H2A, while our results suggest the same interactions contribute to inter-array nucleosome contacts. The dual role for this interaction raises the question as to how intra- and inter-array nucleosome associations are related. To address this we wanted to see to what extent inter-array associations would change when intra-array H4-
H2A associations were favored. Because these intra-array associations might not persist to a significant extent under the usual conditions necessary to observe inter-array associations, we sought to trap the intra-array association irreversibly through a covalent interaction.

To generate arrays with intramolecular nucleosomal crosslinks, we turned to the strategy devised by Richmond and coworkers, where nucleosomal arrays containing both H4-V21C and H2A-E64C are oxidized under conditions that form only intra-array disulfide crosslinks (18). Although 208-12 arrays containing both cysteine-containing histones were readily generated and were similar to our wild-type arrays (Figure S1), attempts to apply this strategy to 208-12 arrays proved problematic, as sedimentation velocity analysis of the oxidation products showed them to be highly heterogeneous (data not shown). Since a different nucleosomal array DNA template, 601-177-12, was utilized for the majority of the previous intra-array crosslinking studies (18), this template was utilized for our subsequent experiments.

The 601-177-12 DNA template is composed of twelve head-to-tail repeats of a SELEX-selected 177-bp octamer-binding sequence (9,27). This template has a shorter linker length and stronger positioning sequence than the 208-12 template, and arrays assembled on this template show greater homogeneity of octamer positioning and saturation than 208-12 arrays (9,21). Additionally, non-H4 histone tails appear to have a less significant role in inter-array associations than in 208-12 arrays (9,11). Nonetheless, in both types of arrays, the H4 tail is
the most important mediator of inter-array association, suggesting that the H4 tail in 601-177-12 arrays might also mediate direct inter-array contacts with histone H2A.

To confirm that direct H4-H2A contacts mediate 601-177-12 array self-association, inter-array crosslinking studies were performed with these arrays. Arrays that were well matched and nearly saturated (Figure S2) (9), exhibited crosslinking-properties very similar to the analogous 208-12 arrays (Figure 5 and S3). In particular, under 1:15 ox/red glutathione oxidation conditions, solutions with a mix of H4-V21C and H2A-E64C arrays show a preference for heterotypic crosslinking and generate a crosslinked species that is readily sedimented (Figure 5). Thus, the 601-177-12 array system also appears to exhibit direct inter-array H4-H2A interactions, making it a suitable system for investigating the relationship between intra- and inter-array interactions.

To assess the effect of intramolecular crosslinking on inter-array self-association, 601-177-12 arrays were assembled with either octamers containing wild-type histones or with octamers containing both H4-V21C and H2A-E64C histones. Hydrodynamic and self-association characterization of these arrays confirmed that the arrays were nearly saturated and similar in both composition and gross structure (Figure S4) (9), making them appropriate for assessing the effects of intra-array cysteine crosslinking.

The arrays were subjected to oxidation in 1.0 mM MgCl$_2$, a divalent cation concentration in which intramolecular nucleosome interactions predominate.
Following removal of the divalent magnesium cation, sedimentation velocity analysis showed that the cysteine-containing array had a sedimentation coefficient of about 10S greater than the wild-type array (Figure 6A, left). This crosslinked species was the predominant form of the array, as relatively little crosslinked species with a large sedimentation coefficient was observed to pellet during initial sedimentation (Figure S5). Coupled with the relative uniformity of the distribution of the crosslinked species, these results suggest that array oxidation under these conditions results in reasonably homogeneous arrays with a similar extent of crosslinking. The magnitude of the sedimentation coefficient for the internally crosslinked array is consistent with previous reports, where intra-array crosslinking traps individual arrays in a more compacted state that has an increased propensity to undergo intramolecular compaction to the fully compacted 30 nm fiber (18). Further, this change in array compaction appears to be due to H2A-H4 crosslinking, as it is the predominant crosslinking species (Figure 6A, right).

The arrays subjected to crosslinking conditions were then assessed for their ability to undergo inter-array self-association. Comparison of the wild-type and intramolecularly crosslinked arrays shows the crosslinked species requires a greater amount of divalent cation to induce comparable differential centrifugation (Figure 6B). This difference is observable throughout the range of magnesium ion concentrations in which the greatest differential sedimentation occurs (Figure 6B, left) and is highly statistically significant, as the p-value for
the observed differential centrifugation at 1.5 mM MgCl$_2$ is 0.00021 (Figure 6B right). Thus, even though the 601-177-12 intra-array crosslinked species exhibits greater compaction that the analogous untreated array, this crosslinking has an opposing effect on inter-array association.

**Discussion**

*Role of Direct H4-H2A Inter-array Interactions*

In chromatin, the acid patch of the histone H2A/H2B dimer serves as a key protein interaction site. Chromatin-associated proteins, such as the Kaposi’s Sarcoma Herpesvirus latency-associated nuclear antigen protein directly interacts with this surface (15,28). Additionally, direct interaction of the histone H4 tail with this site occurs in forming short-range intra-array nucleosome contacts involved in 30 nm chromatin fiber formation (18). Our inter-array crosslinking results demonstrate that this site is also directly contacted in long-range nucleosome interactions that can mediate higher-order chromatin structures beyond the 30 nm fiber. Importantly, this crosslinking increases with increasing inter-array interaction, suggesting that the two components are brought into closer proximity during the association process. Such an increase in direct contact is not a general effect of increasing the proximity of cysteine-containing residues during array associations, as crosslinking of H2A-E64C with itself is not observed, and crosslinking of H4-V21C with itself does not increase with increasing array self-association.
Interestingly, the initial divalent magnesium concentration required for half-sedimentation of the cross-linked arrays occurs at 3.0-3.5 mM (Figure 3A), a midpoint concentration greater than that observed for inter-array association of non-crosslinked arrays (Figure 1C). This suggests that the crosslinked species generated at a given divalent magnesium concentration is not equivalent to the inter-array associated species induced by divalent magnesium at that same concentration. This observation may be due to several non-mutually exclusive reasons. One possibility is that divalent magnesium can induce multiple types of species that cannot be distinguished by differential centrifugation. Some of these species may not generate H2A-H4 crosslinking, but the amount of the species that can undergo H2A-H4 crosslinking increases with increasing divalent cation concentration. Another possibility is that because the glutathione system is only weakly oxidizing, at lower magnesium ion concentrations there might not be enough of a driving force to achieve extensive crosslinking. However, at higher magnesium concentration, an increase in the stability of inter-array association could drive crosslinking under the relatively weak oxidizing conditions.

The presence of inter-array contact between the H4 tail and the H2A core is further supported by prior observations. In a recent photo-crosslinking study, the amount of crosslinking between these histones was shown to increase with the formation of more extensive higher-order chromatin structure, although the extent to which these contacts occurred within versus between arrays was not distinguishable (19). Additionally, the importance of this direct contact is
consistent with the wealth of experimental data in which changes in either the H4 tail or H2A core result in changes in array self-association (9-11,15,16,20).

_Flexibility in H4-H2A Inter-array Interactions_

The pattern of inter-array crosslinking as a function of the sites of cysteine substitutions indicates some conformational flexibility in the mode of H4-H2A interactions. In the crystal contacts of the first high resolution mononucleosome structure the H4 tail residue H4 valine 21 is oriented toward histone residue H2A glutamate 64, while the H4 tail residues directly flanking this valine residue, lysine 20 and leucine 22, do not (1). Nonetheless, both of these flanking residues show preferential H2A-H4 crosslinking with a magnitude similar to crosslinking between H4-V21C and H2A-E64C. This shows that some slippage in the mode of interaction between the H4 tail and the H2A/H2B acidic patch can be accommodated without significantly disrupting the extent of nucleosomal array self-association. This flexibility in interaction may be a feature specific to inter-array associations, as only a minor degree of crosslinking was observed between H4-K20C and H2A-E64C when intra-array crosslinking was studied (18).

When crosslinking of residue H2A-E64 with residues flanking H4-V21 occurs, the changes in the overall H4 tail-H2A interaction may be confined to residues near this site of interaction. Alternatively, this change might result in movement of the entire H4 tail relative to the H2A/H2B acidic patch. While the
latter possibility presents a more dramatic change in structure, recent experimental data suggests that such a structural change is possible. Specifically, Hansen and coworkers have shown that in nucleosomal arrays, the H4 tail can be replaced with a number of other histone tail sequences with similar charge densities with maintenance or improvement in the degrees of array self-association (12). This suggests that the H4 tail interactions with the H2A/H2B acidic patch may not require a single, defined mode of interaction. Nonetheless, there does appear to be limits as to how much reorientation of the tail can occur. Replacement of H4-R23 with a cysteine does not result in crosslinking to H2A-E64C. This suggests that the proximity of these two sites is reduced under the physical constraints of the orientation of the residues imposed by array self-association.

Role of the H4 Tail and H2A Core Outside of Their Direct Contacts

The less than full extent of inter-array crosslinking suggests that the H4 tail and H2A-H2B acid patch have other potential roles in facilitating array self-association. In our experiments, H2A-H4 crosslinking is not quantitative, as significant non-crosslinked H4 histone is observed. This is presumably true of H2A, as well. However, the recombinant H2A histones are not readily resolved from H2B histones under our experimental conditions. This lack of total H2A-H4 crosslinking could be accounted for in a variety of ways. One, up to half of the actual H2A-H4 interaction may not result in disulfide crosslinks because not
every interaction in the H2A-E64C/H4-V21C self-associated arrays would necessarily have two cysteine residues present. For example, a nucleosome in an H2A-E64C array could be equally likely to be next to an array containing H2A-E64C nucleosomes (and wild-type H4) as to be next to one containing H4-V21C nucleosomes (and a wild-type H2A). Two, not every H2A-H4 interaction would result in a disulfide linkage if the reaction did not go to completion. This could be due to glutathione oxidation not reaching a steady state. However, in our system we do not believe that this is the case, as longer reaction times do not appear to significantly change the extent of crosslinking (data not shown). Alternatively, such incomplete reaction might be due to cysteines being inaccessible within densely packed self-associated arrays. However, it is likely that such inaccessibility does not explain incomplete crosslinking in its entirety.

Similar non-quantitative disulfide crosslinking is observed for intra-array crosslinking, where accessibility is not expected to be an issue (Figure 6 and (18)). As a third possibility, the lack of quantitative H2A-H4 crosslinking may reflect that not all H4 tails and H2A histones are involved in direct contacts, and are available to play alternative roles in facilitating array self-association. Prior data supports several alternative roles. The H4 histone alone can interact with DNA, H2A/H2B histones and H3/H4 histones (15), and in the context of nucleosomal arrays, the H4 tail has been shown to interact with array DNA in trans, especially at residues near the amino-terminus (19). It is also important to note, that the H4 tail and H2A/H2B patch are not the only mediators of inter-
array association, as neither the H4 tail nor the H2A/H2B histones are absolutely required (9,20).

In addition to H2A-H4 crosslinking, H4-H4 crosslinking can also be observed. In mixed arrays with the H4 cysteine at position 21, some of this H4-H4 crosslinking occurs between arrays, as these arrays can be differentially sedimented (Figure 1D). However, it appears that stable H4 interactions are not predominantly inter-array in nature. H4-H4 crosslinking persists under less forcing oxidation conditions in which differential sedimentation does not occur (Figure 2), but does not increase at the higher MgCl$_2$ concentrations that increase inter-array association (Figures 3 and S3). The extent, and potentially nature of these H4-H4 interactions also appear to vary with H4 cysteine position (Figure 4), suggesting that certain locations on the H4 tail might be better positioned to engage in H4-H4 interactions.

**Interplay between Intra- and Inter-array Associations**

In nucleosomal array systems the interplay between intra- and inter-array interactions is poorly understood. When arrays containing wild-type histones are subjected to lower concentrations of divalent magnesium cation, intra-array interactions occur preferentially over inter-array interactions (Figures 1C and S2) (9,21). This suggests that at the higher divalent magnesium concentrations required for stable inter-array interactions (Figure 1C and S2) and H2A-H4 crosslinking (Figures 3 and S3), intra-array interactions precede inter-array
interactions. However, how existing intra-array interactions influence inter-array interactions is not clear. Formation of the intra-array compacted species could present interaction sites in a way that facilitates or hinders nucleosome interactions between strands (or potentially not affect them at all). Indeed, existing data with arrays containing mutated or truncated wild-type histones and histone variants can be used to support each of these models (8-11,13,15,16,20). Further, the necessity that intra-array interactions precede inter-array ones in wild-type array is not certain, as the relative stability of intra-versus inter-array associations may be different at higher divalent magnesium concentration.

Our results indicate that stabilizing inter-nucleosomal associations through intra-array H2A-H4 crosslinking antagonizes inter-array self-association. The magnitude of the shift in divalent magnesium sensitivity is statistically significant and is greater than one-third the effect of the complete loss of the H4 tail (13), which of all of the histone tails is the only one whose loss affects inter-array association of 601-117-12 arrays (9). Moreover, we expect that what we observe is less than the full magnitude of the antagonism, because complete antagonism is difficult to reproduce experimentally. In our experiments the intra-array crosslinked species does not have a sedimentation coefficient of 50-60S, characteristic of fully compacted 30 nm fiber (9). This result is similar to those obtained by other groups (18) and might be due to incomplete intra-array H2A-H4 crosslinking (Figure 6B). With some of these interaction sites still available to
facilitate inter-array association, the full degree of potential antagonism between intra- and inter-array associations is likely to be masked. The antagonism of inter-array association through intra-array crosslinking can be interpreted in at least two different ways. The observed effect may be a specific result of using 601-177-12 arrays. These arrays were originally designed to facilitate intra-array association by virtue of their well-positioned nucleosomes and relatively short linker lengths. This results in quantitative differences between these arrays and less homogeneous array systems, such as 208-12 arrays. Specifically, 601-177-12 arrays form inter- and intra-array interactions, and inter-array crosslinks at lower concentrations of divalent magnesium than the analogous 208-12 arrays ([8,9,21] and Figures 1C, 3A, S2 and S3).

However, in the inter-array association experiments with 601-177-12 arrays that have been crosslinked within the array (Figure 6B), this difference is minimized, as the midpoint association MgCl$_2$ concentration (1.8 mM) is nearly the same as that for non-crosslinked 208-12 arrays (Figure 1C). This could mean that intra-array crosslinking of the 601-177-12 array makes this array function more like the non-crosslinked 208-12 array. However, this resemblance seems limited to inter-array association. Prior studies with the 601-177-12 intra-array crosslinked species showed that this crosslinking increased its propensity to undergo intra-array compaction (18), making its behavior resemble a 208-12 array even less so than when the 601-177-12 array is not crosslinked.
An alternative interpretation of the antagonism of inter-array association by intra-array crosslinking is that these forms of interactions are antagonistic for all chromatin. Qualitatively, 601-177-12 and 208-12 arrays behave similarly with respect to inter-array and intra-array interactions ((8,9,21) and Figures 1C and S2). Similarly, the H4-H4 inter-array cross-linking in the 601-array is largely independent on the MgCl2 concentration while the H4-H2A inter-array cross-linking is highly dependent (Figures 3A and S3). The quantitative differences seen in 601-177-12 arrays could be attributed to the propensity of the well-positioned nucleosomes and relatively short linker lengths to facilitate interactions that already occur in 208-12 systems. With respect to linker length differences, experimental data supports this structural similarity. While 601 arrays with a repeat length less than 177 bp, such as those used for the first tetranucleosome structure (29), do show structural differences from 601 arrays with a longer repeat length (30), arrays with nucleosome repeat lengths from 177 to 207 bp can form similar structures (31).

In a general model of intra- and inter-array antagonism, the direct H4-H2A interaction between nucleosomes in a strand would prevent such interactions between strands. This could occur by direct competition, where tying-up H2A-H4 contacts in one type of interaction would prevent them from being used for other interactions. Additionally, the structure formed by one type of interaction, such as the intra-array compacted species, could be less compatible with inter-array associations. For example, if inter-array associations required interdigitation
between nucleosomes within an individual strand, intra-array compaction would disfavor this interaction. For such models, some histone mutations or variants would stabilize one type of interaction at the cost of the other (10, 16), while others, such as the loss of the H4 tail, as well as other tails, could detrimentally affect the stability of both interactions (9-11).

Altogether, our data sheds new light on the nature and role of the interaction between the histone H4 tail and the H2A/H2B acidic patch in intra-array compaction and inter-array self-association. Further studies to place this interaction in the context of other factors in intra- and inter-array association will provide a fuller understanding of higher-order chromatin structure.

**Footnotes**

* This work was supported in part by a grant from NIH to MSK (GM79663). We thank Melissa Blacketer for helpful discussion.

The abbreviations used are: DTNB, 5,5'-dithiobis-(2-nitrobenzoic acid); DTT, dithiothreitol; EDTA, ethylenediaminetetraacetic acid; HCl, hydrochloric acid; HEPES, 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid; MgCl$_2$, magnesium chloride; NaCl, sodium chloride; PAGE, polyacrylamide gel electrophoresis; RT, room temperature; SDS, sodium dodecyl sulfate; TCEP, tris(2-carboxyethyl)phosphine hydrochloride; TRIS, tris(hydroxymethyl)aminomethane.
References


Figure 1. For self-associated nucleosomal arrays, disulfide crosslinking captures direct inter-array interactions between histones H4-V21C and H2A-E64C. A, EcoRI digestion of arrays prior to crosslinking treatment. Mononucleosomes and free DNA liberated from the 208-12 arrays via cleavage of the linker DNA between nucleosomes were separated by electrophoresis on a native 4% PAGE gel and stained with ethidium bromide. B, Sedimentation velocity characterization of arrays prior to crosslinking treatment. Shown are representative integrated sedimentation coefficient distribution plots for 208-12 5S rDNA nucleosomal arrays in the absence of divalent cation. Each array contains either wild-type, H4-V21C, or H2A-E64C histones. S_{20,W} is the sedimentation coefficient corrected for water at 20^\circ C. C, Cation-dependent self-association of arrays prior to crosslinking treatment. Shown is a representative differential centrifugation plot for non-crosslinked arrays, where the fraction of array remaining in the supernatant is plotted as a function of MgCl₂ concentration. D, Differential centrifugation of individual arrays and array mixtures after treatment of the self-associated arrays with 1:1 ox/red glutathione and MgCl₂ removal. Shown is the amount of array absorbance remaining in the supernatant, where averages and standard deviations were derived from three independent trials. E, Histone composition of the arrays in section 1D. Histones were separated by SDS-PAGE analysis under non-reducing conditions and
stained with Coomassie Blue. The identities of the crosslinked histones were assigned by comparison to H4-V21C and H2A-E64C histones crosslinked individually and as a mixture.
Figure 2. Stable inter-array interactions via H4-V21C to H2A-E64C crosslinking persist at decreased oxidizing potentials. A, Differential centrifugation characterization of individual arrays and array mixtures following glutathione treatment of the self-associated arrays at varying ratios of oxidized to reduced glutathione, where total glutathione remained constant at 1.0 mM. Shown is the amount of array absorbance remaining in the supernatant for a representative trial. B, Differential centrifugation characterization of individual and mixed arrays following 1:15 ox/red glutathione treatment. Array remaining in the supernatant was analyzed as described in 1D. C, Histone composition of individual and mixed arrays following 1:15 ox/red glutathione treatment. Histone components were analyzed as described in 1E.
Figure 3. Array self-association is required for H4-V21C to H2A-E64C crosslinking. A, Differential centrifugation characterization of individual and mixed arrays following 1:15 ox/red glutathione treatment at varying MgCl₂ concentrations. Shown is a representative plot of the amount of array absorbance remaining in the supernatant after the removal of MgCl₂. B, Histone composition of mixed arrays containing both H2A-E64C and H4-V21C arrays following 1:15 ox/red glutathione treatment at varying MgCl₂ concentrations. Histone components were analyzed as described in 1E.
Figure 4. Multiple, but not all, positions on the H4 tail preferentially crosslink to histone H2A-E64C. A, Histone composition of crosslinked arrays (1:15 ox/red glutathione treatment, 6.0 mM MgCl₂) containing a cysteine residue at various histone H4 tail locations. Histone components were analyzed as described in 1E. B, Differential centrifugation characterization of arrays analyzed in section A. Shown is the amount of array absorbance remaining in the supernatant in the absence of MgCl₂ for various H4 arrays alone (left) or combined with arrays containing H2A-E64C histone (right). Analysis was performed as described in 1D.
Figure 5. 601-177-12 arrays also undergo direct inter-array crosslinking. A, Differential centrifugation characterization of individual and mixed arrays following 1:15 ox/red glutathione treatment in 4.0 mM MgCl₂. Arrays remaining in the supernatant were analyzed as described in 1D. B, Histone composition of individual and mixed arrays following 1:15 ox/red glutathione treatment. Histone components were analyzed as described in 1E.
Figure 6. Intra-array crosslinking disrupts inter-array self-association. A, Glutathione treatment (1:1 ox/red glutathione with 1.0 mM MgCl$_2$) results in intra-array crosslinking. Shown is the characterization of the arrays described in A following crosslinking treatment. On the left are representative integrated sedimentation coefficient distribution plots obtained in the absence of divalent magnesium ion. On the right are the histone components of these arrays. B, Intra-array crosslinking decreases cation-dependent self-association. Shown on the left are differential centrifugation plots for the arrays described in A following crosslinking treatment. The average and standard deviations of the fraction in solution are calculated from three independent crosslinking experiments. The differential centrifugation of these arrays at 1.5 mM MgCl$_2$ is highlighted on the right. The p-value is 0.00021 and is calculated from the single tail Student’s t-test for unpaired data with equal variance.
Supplementary Information

**Figure S1.** Analysis of non-crosslinked 208-12 5S rDNA nucleosomal arrays containing H2A-E64C and H4-V21C histones. **A,** Arrays were digested with EcoRI and resulting mononucleosomes and free DNA were analyzed on a native 4% PAGE gel followed by staining with ethidium bromide. **B,** A representative integrated sedimentation coefficient distribution plot resulting from sedimentation velocity characterization of arrays in absence of divalent cation. **C,** Differential centrifugation plot for non-crosslinked arrays, where the fraction of array remaining in the supernatant is plotted as a function of MgCl₂ concentration.
Figure S2. Analysis of non-crosslinked 601-177-12 nucleosomal arrays containing H2A-E64C or H4-V21C histones. 

A, Restriction endonuclease ScaI cleaves at the linker DNA sites of 601-177-12 nucleosomal arrays resulting in mononucleosomes and free DNA. These products were electrophoresed on native 4% PAGE gel followed by ethidium bromide staining. B, Sedimentation velocity analysis and; C, Cation-dependent self-association of arrays as described in S1.
Figure S3. Like 208-12 5S rDNA arrays, 601-177-12 arrays require self-association for H4-V21C to H2A-E64C crosslinking. A, Differential centrifugation characterization of individual and mixed arrays and; B, Histone composition of mixed arrays containing both H2A-E64C and H4-V21C arrays following 1:15 ox/red glutathione treatment at varying MgCl₂ concentrations. Histone components were analyzed as described in 1E.
Figure S4. Analysis of non-crosslinked 601-177-12 nucleosomal arrays containing both H2A-E64C and H4-V21C histones. A, Scal digestion; B, Sedimentation velocity characterization, and; C, Cation-dependent self-association analysis of arrays as described in S2.
**Figure S5.** Intra-molecularly crosslinked arrays form the majority of crosslinked species following crosslinking at 1 mM MgCl$_2$. Shown are the amount of array absorbance before and after centrifugation at 14000 g for 10 min following crosslinking in presence of 1 mM MgCl$_2$. 

208-12 nucleosomal arrays
CHAPTER 3

Development of nucleosomal systems to study chromatin remodelers and the effects of nucleosomal positioning on chromatin structure

Techniques described in the first portion of this chapter have been incorporated into the manuscript “Histone H3 tail acetylation modulates ATP-dependent remodeling through multiple mechanisms” published in the Journal Nucleic Acids Research [39(19):8378-8391]

Abstract

Regulation of genomic DNA associated processes take place at the level of chromatin, and in vitro chromatin model systems have been instrumental in understanding structural and mechanistic aspects of chromatin function. To expand the scope of in vitro chromatin studies, we have developed new nucleosomal systems. Here we describe a refinement of the native chemical ligation technique used to incorporate post-translational modifications into histones. This refinement effectively eliminates the problem of racemization at the peptide thioesterification step. Acetylated histones, generated using this modified histone ligation method, were successfully used for studying ATP-dependent remodeling by SWI/SNF and RSC complexes. We also describe the development of new DNA templates for nucleosomal assembly.
Introduction

Ubiquity of histones within the cell nucleus in addition to the complexity of chromatin often poses challenges in teasing apart the role of individual chromatin components \textit{in vivo}, even in a comparatively simple model system like yeast. In this context, \textit{in vitro} nucleosomal array systems provide powerful tools for understanding chromatin structure and functions (1-3). However, these array systems often do not fully reflect the state of \textit{in vivo} chromatin. For example, the nucleosomes of specific genomic regions often possess specific patterns of post-translational histone modifications (4). However, most nucleosomal array systems used are either fully unmodified or possess a mixture of modification states depending on the use of recombinantly expressed histones (5) or histones purified from eukaryotic cell nuclei (6,7), respectively.

Several methods to overcome this limitation have been devised, including disulfide-directed histone ubiquitylation (8,9) and synthesis of histones with site-specific modifications \textit{in vivo} by expanding the genetic code of host organism (10). One technique for incorporating H3 and H4 histone tail modifications is native chemical ligation (11-13). In this technique, histone N-terminal tail peptides, with specific modifications and protection on amino acid side chains, are synthesized via solid phase peptide synthesis. The remaining C-terminal region of the histone is recombinantly expressed and purified. Following cleavage from the resin, activation of the C-terminal residue and removal of the protecting groups, the peptide is ligated with its C-terminal histone counterpart.
With this strategy, racemization of the C-terminal amino acid can occur during the peptide thioesterification step. In some cases, the peptide epimers can be isolated. However, for some sequences, purification cannot be readily performed. In order to improve the histone native chemical ligation strategy, we sought to find thioesterification conditions that would limit racemization.

Another limitation of standard nucleosomal array systems is that the linker DNA length within an array is uniform (1,14,15). This restricts our understanding of the effect of non-uniform nucleosome positioning on short- and long-range nucleosomal interactions and, therefore, on chromatin compaction. Even for the mononucleosomes, the DNA templates are usually derived from the nucleosomal array DNA template and thus have fixed linker lengths. Sometimes this aspect may pose limitations in the study of factors/complexes that bind to both the nucleosome and the linker DNA. Here we describe our efforts to address this issue by modifying the linker DNA lengths for both, the mononucleosomes and array DNA templates.

**Materials and Methods**

*Native chemical ligation*

Histone peptides from amino acid residues 1-22 for histone H3 and 1-25 for histone H4 were generated using standard Fmoc-based solid phase peptide synthesis (Baylor College of medicine, Houston, Texas). The remaining C-terminal fragment of histones were recombinantly expressed and purified using
standard methods (5). Acetylated histones were generated by native chemical ligation strategy (12) with modification at the peptide thioesterification step. Here, the weak acid cleaved peptide was resuspended in Dimethylsulfoxide (DMSO) and then 95 mg Dicyclohexylcarbodiimide (DCC) per 100 mg of peptide was added. After mixing DCC, ~55 mg of benzyl mercaptan was added and the mixture was incubated at room temperature for 3 hours with constant stirring. The DMSO was then removed by rotovap and the remaining steps were performed as described previously (12). The ligated histones were quantified on 18% SDS-PAGE gel by comparison with a known histone standard.

**Mononucleosomal DNA sequence generation**

601-177-12 DNA containing plasmid was digested with ScaI and purified by gel purification, butanol extraction and ethanol precipitation to generate 177 bp ‘601’ fragments. PCR was used to generate mononucleosome DNA templates with these ‘601’ nucleosome positioning sequences and extended linker DNA at one end. The same forward primer was used for each of the three templates while three different reverse primers were employed to accommodate the extended linker DNA (30 bp, 40 bp, or 60 bp) at one end.

**Histone octamer preparation**

Xenopus histones were recombinantly expressed, purified, and characterized using standard methods (5). Acetylated histones were generated
using Native chemical ligation strategy as discussed above. Octamers were assembled and purified in the presence of 0.1 mM tris (2-carboxyethyl) phosphine (TCEP) reductant as described previously (5). Octamers were characterized by SDS-PAGE gel electrophoresis and quantified using absorbance spectroscopy (5).

**Mononucleosome Assembly**

Mononucleosomes were assembled by mixing histone octamers and DNA template followed by step-wise salt dialysis method as previously described (13,16). 0.1mM TCEP reductant was added to the DNA and octamer mixture and to the dialysis solutions. The final composition of the mononucleosome solutions included 2.5 mM NaCl, 10 mM HEPES pH 8.0 and 0.1 mM TCEP. The mononucleosomes were quantified based on amount of DNA by measuring absorbance at 260 nm. They were stored and shipped at 0°C.

**Unphased array generation**

601-177-12 DNA containing plasmid was digested with ScaI and purified by gel purification, butanol extraction and ethanol precipitation to generate 177 bp fragments. These fragments were then used as templates for PCR to generate 10-147-20 or 12-147-18 fragments, where 147 bp of nucleosome positioning sequence from ‘601’ was flanked by 10 and 20 base pairs or 12 and 18 base pairs, respectively. Fragments generated by PCR were digested with
EcoRI and either cloned directly or post-self-igation into plasmid vectors. The ligation products were used for transformation of DH5α or XL1-Blue competent cells. Plasmids from the resulting colonies were purified using IBI Scientific plasmid miniprep kit. The plasmids were then digested using BamHI and Scal and resulting fragments were analyzed on 1% agarose gel.

**Results**

*Peptide thioesterification for native chemical ligation of histones*

The native chemical ligation strategy to obtain histones with desired modifications has proved useful for studying the role of histone modifications *in vitro* (13,17,18). A key step in this strategy is the generation of the histone tail peptide activated at the C-terminus as a thioester. However, in our previous method, the intermediate in generating the thioester is susceptible to racemization of the C-terminal residue, making the peptide thioester a mix of stereoisomers. The reverse phase HPLC used to separate this thioester from the other components of the reaction mixture often does not resolve the stereoisomers well, leading to a mixture of thioesterified peptides or a loss of significant amounts of the peptide thioester.

To solve this problem, we tested different carboxylate activation strategies previously shown to limit racemization of individual amino acids during C-terminal activation. We found that the use of Dicyclohexylcarbodiimide (DCC) in dimethyl sulfoxide (DMSO) efficiently activates the carboxyl group of the C-
terminal amino acid of the side-chain protected histone tail peptide. The addition of DCC was immediately followed by addition of benzyl mercaptan to yield thioester. This thioester was then treated under acidic conditions to remove the thioester peptide side-chain protection group, as performed previously.

Purification of the peptide thioester by reverse phase HPLC showed that racemization under these conditions was significantly limited, with no detectable amounts of stereoisomer present (Figure 1). Moreover, the change in the activation chemistry does not affect peptide’s efficiency to ligate to the globular domain of histone as confirmed during generation of tetra-acetylated histones H3 (H3K9,14,18,23Ac) and H4 (H4K5,8,12,16Ac). With their recombinantly expressed counterparts, these histones were assembled into octamers (either acetylated H3 or acetylated H4 octamers) and then mononucleosomes (19). The mononucleosomes were successfully used for studying recruitment and ATP-dependent remodeling activity of SWI/SNF and RSC complexes (19).

*Generation of mononucleosomes with variable linker DNA lengths for RSC-nucleosome structure study using reconstructive EM*

RSC complex is an abundant chromatin remodeler in *S. cerevisiae* and is involved in transcription as well as DNA damage repair (20). Previous research has shown that RSC can bind to nucleosomal DNA and prevent digestion by nucleases (21). RSC complex contains multiple bromodomains (22), domains that can recognize and bind to acetylated lysines (23), that are believed to be involved in recruitment of RSC to specific sites in the genome (24,25). Earlier
work to understand RSC complex structure had used Electron Microscopy (EM) reconstruction method. In addition to the surface topography of RSC at 25-30 Å, a central cavity within the structure, that could fit a nucleosome, was observed (21). Based on these observations, we devised to assemble acetylated mononucleosomes with varying linker DNA lengths on one end. These mononucleosomes were assembled for a collaborative project to study the RSC complex-nucleosome structure using reconstructive EM.

To this end, we used sequences from the linker DNA region of the 208-12 DNA template to extend one end of the 177 bp sequence obtained from the 601-177-12 DNA template (Figure 2A). This was done through PCR, using primers encompassing sequences from one end of the 177 bp fragment and from the 208-12 DNA linker region. The 208 nucleosome positioning sequence of the 208-12 DNA template, obtained from the 5S rDNA sequence of Lytechinus variegatus (26), is weaker than the strong positioning ‘601’ DNA sequence (27), and the 208 linker DNA has even weaker nucleosome binding ability.

The DNA templates, purified from multiple rounds of PCR, were then used to assemble mononucleosomes using histone octamers containing tetra-acetylated histone H3 (Figure 2). This was done by step dialysis method where, through stepwise decrease of salt concentration, the octamer deposits on the DNA (5,28). The mononucleosomes were then analyzed on a 4% native gel to confirm proper assembly (Figure 2B) and shipped to our collaborator for RSC structural studies.
Although the initial screening results looked promising, before further work could be done another group had published the RSC-nucleosome structure using rat liver histone octamer assembled on the Xenopus 5S rDNA nucleosome positioning sequence (29).

**DNA templates with variable linker DNA lengths**

The use of nucleosomal arrays in numerous *in vitro* chromatin studies has significantly advanced our understanding of chromatin, mostly in the context of chromatin structure and remodeling (1,13,15). However, all the studies done to date have used nucleosomal arrays with uniform linker DNA length within the array (1,14,15). Within the eukaryotic cell nucleus, the nucleosome spacing can be variable. While most of the nucleosomes within heterochromatic regions have uniform linker DNA length (30,31), the nucleosome spacing in euchromatin is not as uniform. This variability in linker DNA length may be an important factor in differential compaction levels of actively transcribed and silenced regions of the genome (30). To understand how nucleosome positioning affects short- and long-range nucleosomal interactions and to study the role of chromatin remodelers in nucleosome spacing, we set out to develop nucleosomal DNA template with non-uniform linker lengths.

Among the DNA templates used for nucleosomal array assembly, 208-12 and 601-177-12 DNA are the two most commonly used ones (3,32). Between these two, ‘601’ DNA contains a very strong nucleosome positioning sequence
(27). We decided to use the ‘601’ positioning sequence to generate arrays with defined positioning sequences separated by non-uniform linker lengths. For ease of understanding, hereafter, the nucleosomal arrays with uniformly spaced nucleosomes will be referred to as ‘phased’ arrays while ones with non-uniform yet defined nucleosome spacing will be referred to as ‘unphased’ arrays.

Traditionally, the linker DNA length used in nucleosomal arrays has been a multiple of 10 (1,3,15,32). However, through personal communication with Dr Jonathan Widom, we learned that the first base-pair resolution map generated by his lab revealed that the linker DNA length in *S. cerevisiae* follows a pattern of $10n +5$ (this work was later published in *Nature*. See (33)). Based on this information we decided to generate two types of unphased DNA templates: one with a combination of 20, 30 and 40 base pairs of linker DNA (a pattern of 10n), and second with a mix of 24, 30 and 36 base pairs of linker DNA (a pattern of $\sim10n+5$).

For this purpose, 601-177-1 fragments generated by Scal digestion were used. Using PCR, new EcoRI sites were introduced either at 10 and 20 base pairs flanking the ‘601’ positioning sequence or at 12 and 18 base pairs flanking the ‘601’ positioning sequence. Since EcoR1 produces recleavable sticky ends, ligation of these inserts would produce templates with randomly positioned linker lengths of 20, 30, 40 base pairs or 24, 30, 36 base pairs.

There were two approaches used for generating 12-mer unphased arrays (Figure 3A) both of which were to generate plasmids with six inserts first, for...
ease of sequencing to know the linker DNA lengths, and then ligate two 6-mer fragments and clone into a plasmid. The first strategy was to use high insert to vector ratio and screen for plasmids containing 4-8 repeats of insert. The second approach was to self-ligate the inserts first, then gel purify the 6-mers and clone in plasmids. The latter strategy seemed more promising as most of the transformants post-ligation were expected to have plasmids with six inserts and fewer transformant colonies would have to be screened.

Both the strategies, however, yielded plasmids with single or no inserts (Figure 3B), even when linearized vector with dephosphorylated ends were used. Several attempts at optimizing the ligation conditions, transformations and multiple screenings did not yield plasmids containing multiple inserts. Hence, this work was not pursued further.

Discussion

Peptide thioesterification and racemization

In the native chemical ligation strategy to generate modified ligated histones, modification and optimization of the step to activate peptide by using DCC in dimethyl sulfoxide instead of 2-(1H-benzotriazole-1-yl)-1,1,3, tetramethyluronium hexafluorophosphate (HBTU) in dimethyl formamide significantly limits racemization, similar to published reports of limiting racemization during amino acid activation (34). How this reagent reduces racemization is not clear, as it was not discussed in the original published
method, nor directly tested by us. However, because activation-mediated racemization of the \( \alpha \)-proton on amino acids is thought to proceed via an intermolecular cyclization intermediate, we speculate that formation of such an intermediate is limited under our conditions, either due to the increased steric bulk of the DCC substituents, and/or by tuning of the reactivity of the activated species by the DCC substituents and the solvent. Because these conditions eliminated the stereoisomer side-product and the need to remove it the yield of the peptide thioester was increased by 1.5 to 2-fold, thus also making the process more cost-effective. The ligated histones, with different sites of acetylation and methylation, generated using this method have been successfully used in chromatin remodeling (19) and HP1 binding studies (Azzaz et al, manuscript submitted).

**Mononucleosomes for RSC studies**

Successful assembly of mononucleosomes on DNA templates of varying length shows that the use of Widom ‘601’ strong positioning DNA in combination with weak binding linker DNA of ‘208-12’ prevents the problem of multiple nucleosome positioning on a DNA template that is longer than 147 bp in length. Here, we saw two positions of nucleosomes, end-positioned and centrally positioned nucleosomes, which is a common observation for mononucleosome assembly on the ‘601’ DNA using step dialysis method. For studies that require single-positioned mononucleosomes, rapid dilution method can be used to
obtain uniformly positioned nucleosomes (35). Although another group published the RSC-nucleosome structure beforehand, our work with the generation of new DNA templates shows the possibility of modifying mononucleosomes for specific applications. Since this work, similar PCR-based approach has been used for other studies (36,37).

*Varying linker length in nucleosomal array DNA*

Both of the two different strategies we used here to develop DNA templates containing varying linker lengths did not work. This failure possibly occurred at the vector ligation step or during cell growth post-transformation with ligated insert-containing plasmid or most likely both. For the first strategy where high ratio of insert to vector was used, high number of transformants was seen post-ligation. Additionally, the inserts were able to self-ligate in the absence of vectors. Both these results rule out the possibility that inserts had unligatable ends.

*In vitro*, DNA sequence is a major determinant of octamer deposition (33,38) indicating that certain sequences of DNA are more bendable than the others due to the stacking properties of adjacent bases. Ligation of linear fragments of DNA in a plasmid vector leads to circularization of DNA. It is possible that this circularization, in addition to further supercoiling required in the bacterial cell (39), puts constraints on the plasmid. Under selection pressure, this may lead to the loss of inserts without the loss of plasmid. Additionally,
during propagation and purification of the plasmid containing the 601-177-12 DNA sequence, we have often experienced changes in the number of repeats resulting from recombination within the 177 bp repeat containing region of the plasmid (unpublished work). Thus, even in the case of unphased arrays, the repetitive nature of the inserts may have been a reason for loss of these inserts within the plasmid. In future, multiple DNA sequences may have to be tested along with use of strains like Stbl2 and JM109 to reduce the problems with repetitive DNA sequences to address the stability issues of repetitive DNA sequence in a plasmid.

Overall, through this work we have explored the possibility and challenges of expanding the scope of *in vitro* chromatin studies by describing the approaches to add to the repertoire of current chromatin systems.

**References**


**Figure 1.** Modified native chemical ligation strategy. Native chemical ligation scheme to generate histones with site-specific modification is shown here. At the peptide activation step, the use of HBTU gives racemic mixture of thioesterified peptide (left), while the use of DCC prevents the formation of stereoisomers, as seen from the peptide peak from the reverse phase HPLC purification (right).
A

- **Fwd**
  - 177 bp
  - Rev30
  - Rev40
  - Rev60

- **PCR**
  - +30 bp
  - +40 bp
  - +60 bp

Mononucleosomal DNA templates

Histone Octamers (Ac or UnAc)

- **Salt step dialysis**

- Mononucleosomes
Figure 2. Mononucleosome template generation. (A) Strategy to develop mononucleosome templates with 30, 40 and 60 base pairs of linker DNA at one end is shown here. The templates generated using PCR were then used for assembling histone octamers to obtain mononucleosomes with varying linker DNA lengths at one end. (B) Analysis of the PCR amplified DNA templates and mononucleosomes on 1% agarose gel and 4% native PAGE gel, respectively.
A

- PCR

Inserts

Vector

B

Self-ligated inserts

Vector

Plasmids containing multiple copies of inserts

Transformation into bacterial amplification strain
Figure 3. (A) Strategy to generate nucleosomal array unphased arrays. (B) Analysis of self-ligated 12-147-18 fragment (left), cloned 10-147-20 fragment (top, right), and cloned 12-147-18 fragment (bottom, right) on 1% agarose gel.
CHAPTER 4

Investigating the role of Ada3 acetylation in SAGA cooperativity

Results from this chapter regarding Gcn5 bromodomain recognition of Ada3 acetylation will be incorporated into a future manuscript to be submitted to Protein Science.

Abstract

The SAGA family of transcriptional coactivators is associated with inducible gene expression in eukaryotes. One of the ways in which SAGA carries out this function is through histone H3 acetylation by its Gcn5 subunit. Previously, in vitro studies done to understand the mechanism of SAGA acetylation had shown that yeast SAGA acetylates nucleosomal histones cooperatively. Initial studies performed to elucidate the source of this cooperativity had concluded that SAGA self-acetylates two of the lysines on its Ada3 subunit, which then bind to Gcn5 bromodomain and lead to the dimerization of SAGA. Here we discuss our efforts to dissect the role of acetylation on individual lysines of Ada3. We show that pre-acetylated Ada3 peptides bind more strongly to the Gcn5 bromodomain compared to H3K14Ac peptide. This observation may provide further insights into the mode of binding of acetylated lysine by bromodomains.
Introduction

Chromatin plays a key role in differential gene expression in response to developmental and environmental cues. One of the ways to achieve this regulation is through the post-translational modifications of histones (1), of which histone acetylation is one of the most prevalent. Histone acetyltransferases (HATs) are the enzymes that put these acetyl marks, predominantly on lysines of histone tails (2).

HATs mostly exist as multisubunit complexes and in many cases also target non-histone substrates (3,4). The first acetyltransferase complex purified was the SAGA (Spt-Ada-Gcn5 acetyltransferase) complex from budding yeast, *Saccharomyces cerevisiae* (5). Yeast SAGA is the founding member of SAGA family of transcriptional coactivators (5,6). The members of this family are highly conserved both in terms of structure and composition (2), and are involved with expression of stress-induced and developmental genes in response to environmental signals (7-9). Improper functioning of SAGA attributes to certain neurodegenerative diseases (10) and tumorigenesis (11). Thus, it is essential to investigate the mechanism of SAGA’s functions not only to understand inducible gene expression but also to gain insights into the disease states and possible treatments.

Yeast SAGA, a 1.8 MDa complex, carries out its HAT function through its Gcn5 subunit (5). Gcn5 consists of a catalytic domain that performs direct transfer of acetyl group primarily to lysine 14 of histone H3 (12). With respect to
histone acetylation, SAGA also has acetyl lysine recognizing bromodomains on its Gcn5 and Spt7 subunits (13-15). These, as well as additional subunits, aid in recruitment of SAGA to specific genomic locations (16).

Structurally, Gcn5 is localized with two other subunits, Ada2 and Ada3, within SAGA (17). Recombinantly expressed full-length Gcn5 has very weak acetyltransferase activity by itself. However, when over-expressed alongside Ada2 and Ada3 it forms a subcomplex that has similar HAT activities as the full SAGA complex in vitro (18). Indeed, several in vitro studies, to gain insights into SAGA HAT function, have been done using the recombinantly expressed Gcn5/Ada2/Ada3 subcomplex.

Previous studies done in our lab to characterize in vitro nucleosomal acetylation by yeast SAGA had showed that SAGA acetylates nucleosomes cooperatively (19). Similarity in lysine composition and positions between the histone H3 tail and the Ada3 N-terminal region, as well as requirement of Ada3 subunit for full catalytic activity of the SAGA HAT domain, led to the hypothesis that Ada3 acetylation may be involved in SAGA cooperativity. Some of the initial experiments had showed that SAGA does acetylate on its Ada3 subunit and this leads to subsequent dimerization. Furthermore, this ability was lost when two of the N-terminal lysines on Ada3 were mutated to Arginine (20). To better understand the role of each of these two lysines in SAGA cooperativity, we have studied how the acetylation on these individual sites is recognized by the Gcn5
bromodomain and if this recognition affects the rate of acetylation of the second site by the Gcn5/Ada2/Ada3 subcomplex.

Materials and Methods

Gcn5/Ada2/Ada3 subcomplex expression and purification

Wild type Gcn5/Ada2/Ada3 subcomplex was expressed and purified as described earlier (18). Protein concentrations for the subcomplex were determined by comparison to a known BSA standard using Bradford total protein quantification method. The purified subcomplex was analyzed on a 12% SDS-PAGE gel, stained with Coomassie blue and its activity was tested for known amounts of H3 peptide (Anaspec).

Gcn5 bromodomain cloning, expression and purification

Gcn5 bromodomain (121 C-terminal amino acid residues of Gcn5) DNA sequence was amplified from plasmid containing Gc5/Ada2/Ada3 DNA sequence using PCR primers to introduce EcoRV and HindIII cut sites. The PCR-amplified fragment was then digested using EcoRV and HindIII, purified and cloned into pET30a vector. The resulting expressed protein had 6-His tag at the N-terminal of bromodomain. With inclusion of other tags and an enterokinase cleavage site the total molecular weight of the expressed bromodomain was 19.16 KDa. The tagged bromodomain was then affinity purified using Ni-NTA
resin and quantified using Bradford protein quantitation method using known amounts of BSA as the standard.

*Peptide synthesis and purification*

Standard F-moc based solid-phase peptide synthesis was used to generate Ada3 and H3 peptides (University of Wisconsin Biotechnology Center). For fluorescent labeling, fluorescein-5-EX succinimidyl ester was coupled to the amino-terminal of the resin-bound peptides. Both unlabeled and fluorescein-labeled peptides cleavage from the resin, removal of the side chain protecting groups, and purification were performed by standard methods.

*Ada3 peptide acetylation*

Known amounts of Ada3, Ada3K8Ac and Ada3K14Ac peptides were tested for acetylation by Gcn5/Ada2/Ada3 subcomplex in 25 mM HEPES, pH 7.5, 50 mM KCl, 5% glycerol buffer containing 1:11 ratio of $^3$H-AcCoA: AcCoA and protease inhibitors. The acetylation signal was measured using standard filter binding assay.

*Fluorescence anisotropy*

Purified Gcn5 bromodomain was dialyzed in 10 mM NaH$_2$PO$_4$, pH 7.5, 0.1 mM EDTA, 5 mM DTT. The peptide stocks were diluted in the same buffer as well. 2 µM of each of the peptides were titrated against increasing
concentrations of Gcn5 bromodomain, ranging from 20 to 977.3 µM. The mix was incubated for 30 minutes at room temperature and the fluorescence anisotropy was measured using Varian Fluorescence spectrophotometer fitted with a polarizer.

Results

*Gcn5 bromodomain preferentially binds to Ada3K8Ac peptide*

Previous work to define the source of SAGA cooperativity in nucleosomal acetylation showed that SAGA autoacetylates on lysines 8 and 14 of the N-terminal of its Ada3 subunit. This acetylation led to dimerization of SAGA seemingly mediated by binding of acetylated lysines to the Gcn5 bromodomain (20). This conclusion was based on the observation that the simultaneous mutations of lysines 8 and 14 to arginines in Ada3 or mutation in the Gcn5 bromodomain resulted in loss of SAGA dimerization and cooperativity (20,21). To pursue a deeper mechanistic understanding of the role of Ada3K8K14 acetylation on SAGA dimerization and cooperativity, we sought to characterize the effect of acetylation on individual lysines on Ada3.

We started by quantifying the binding affinity of Gcn5 towards Ada3 peptides (residues 1-21) using fluorescence anisotropy. The peptides were pre-acetylated at either lysine 8 or lysine 14, or at both the lysines, and contained a fluorescent label. The use of recombinantly expressed and purified full-length Gcn5 for these studies proved problematic due to serious issues with the
stability of purified Gcn5 in solution. We then shifted to the use of Gcn5 bromodomain for binding studies with Ada3 peptides. This approach turned out to be beneficial for specifically looking at the bromodomain binding of the peptides by ruling out any concerns of binding contributions from the catalytic domain of Gcn5.

The fluorescence anisotropy binding experiments show that the Gcn5 bromodomain can bind all of the acetylated Ada3 peptides (Figure 1). Interestingly, the Gcn5 bromodomain binds the dual acetylated Ada3 peptide with greater affinity than the H3K14Ac peptide ($K_d = 260.1\pm42.9$), its presumed in vitro target. Also, Gcn5 bromodomain has higher affinity for the Ada3K8Ac peptide ($K_d = 59.6\pm7.6$) than the Ada3K14Ac peptide ($K_d = 421.5\pm29.4$). Additionally, its binding affinity for the dual acetylated peptide ($K_d = 78.8\pm7.6$) is comparable to that for Ada3K8Ac peptide. Thus, much of the binding interaction of the Gcn5 bromodomain to acetylation of the Ada3 N-terminus appears to be mediated by recognition of lysine 8 acetylation.

*Acetylation of Ada3 peptides by Gcn5/Ada2/Ada3 subcomplex*

We also tested the ability of the Gcn5/Ada2/Ada3 subcomplex to acetylate the N-terminal region of Ada3. For this purpose, Ada3K8Ac, Ada3K14Ac and unacetylated Ada3 peptides were assayed for their initial rates of acetylation using a filter binding assay to quantify the amount of radiolabeled acetate transferred to the peptide (Figure 2). We found that all of the Ada3
peptides could be acetylated. We also found that preacetylating Ada3 peptide affected the rate of acetylation. The rate of acetylation of Ada3K8Ac peptide is 2-fold more than that of the Ada3 peptide pre-acetylated at lysine 14, but 3-fold less than the unacetylated Ada3 peptide. These results suggest that both lysine 8 and 14 can serve as sites of Gcn5/Ada2/Ada3-mediated acetylation, but that it is easier to acetylate lysine 14, once lysine 8 is acetylated compared to acetylating lysine 8, once lysine 14 is acetylated.

Discussion

Factors and complexes associated with altering the state of chromatin are being continuously discovered to undergo post-translational modifications to regulate their activity. To date, examples of factors/complexes undergoing lysine acetylation include Gcn5 acetylation of RSC complex (22), and autoacetylation of histone acetyltransferase Rtt109 (23). Additionally, lysine acetylation has been linked with protein associations, e.g., Rtt109 (24); as well as complex dissociations, e.g., Lysine acetyltransferases (KATs) such as p300 (25) and TIP60 (26). Thus, initial findings in our lab that SAGA autoacetylation plays a role in dimerization and cooperativity observed during nucleosomal acetylation seemed feasible (20).

Our in vitro studies of Ada3 peptide acetylation and bromodomain binding present some interesting findings with respect to questions of molecular recognition. Published studies of the determinants of Gcn5 bromodomain
recognition have indicated that the sequence specificity of acetyl-lysine peptides is largely confined to the second and third residues C-terminal to the acetylated lysine (13,15). In Ada3, the second and third residues following lysine 8 and 14 conform to the consensus characteristics of known binding partners of the Gcn5 bromodomain. They are identical at these sites, suggesting they should have identical affinities. However, our fluorescence anisotropy studies show that the binding affinities to lysine 8 acetylated site is significantly greater than to the lysine 14 acetylated site. This suggests that the mode of Gcn5 bromodomain recognition may be more complicated than previously thought, and additional studies will help clarify the factors involved in this binding interaction.

Additionally, our results with Gcn5/Ada2/Ada3 acetylation of the preacetylated Ada3 peptides indicate that preacetylation of lysine 8 allows lysine 14 acetylation much more than the reversed situation. This result could mean that lysine 14 is simply a better substrate site. However, another possibility, in light of our bromodomain binding studies, is that preacetylation of lysine 8 and Gcn5 bromodomain binding facilitates subsequent lysine 14 acetylation but not vice versa. Future studies will be devoted to addressing these possibilities.

**Acknowledgements**

I would like to thank Dr Melissa Blacketer for insightful discussions and work on the SAGA mutant yeast strains. I would also like to thank Chitvan Mittal
and Sannie Olson for their help with some of the SAGA and Gcn5/Ada2/Ada3 subcomplex experiments, respectively.

References


*Methods* 41, 271-277


Figure 1. Gcn5 bromodomain binding studies with Ada3 peptides using fluorescence anisotropy. (A) Schematic representation of expressed Gcn5 bromodomain, and Coomassie stained 15% SDS-PAGE gel showing increasing amounts of the purified Gcn5 bromodomain protein. (B) Representative plots of measured anisotropy (r) for Ada3K8Ac, Ada3K14Ac, Ada3K8AcK14Ac and H3K14Ac peptides against increasing concentrations of Gcn5 bromodomain protein. (C) Comparison of the measured binding constants (K_d) for Ada3K8Ac, Ada3K14Ac, Ada3K8AcK14Ac and H3K14Ac peptides.
Figure 2. Acetylation of Ada3 peptides by Gcn5/Ada2/Ada3 subcomplex. (A) Ada3 peptides used for acetylation assays. (B) The apparent initial rates for each peptide at different peptide concentrations are plotted here. Apparent initial rates for the acetylation of each peptide were calculated using the signal for radioactive acetyl group transfer. (C) Comparison of slopes obtained from the apparent initial rates for each peptide.
CHAPTER 5
Conclusions and Future Directions

Overall, through the work presented in this dissertation I have probed into the formation of higher-order chromatin structure and mechanistic aspects of nucleosomal acetylation by SAGA. Using nucleosomal arrays containing specific cysteine containing histones, we show that histones H4 and H2A interact under conditions that promote array-array-associations. These inter-array associations involve inter-nucleosomal contacts and are believed to resemble long-range nucleosomal interactions in vivo. We also show that prior involvement of histones H4-H2A contacts in short-range nucleosomal interactions, resembled by intra-array folding, antagonizes the formation of long-range nucleosomal associations. This observation may imply, to a limited extent, mutually exclusive nature of chromatin structure in the genome.

This structural study was done using two commonly used DNA templates for in vitro array assembly, 208-12 and 601-177-12. Nucleosome positioning sequence of each of these templates is also often used for mononucleosome preparation, sometimes including modified histones generated using native chemical ligation technique. Here we have further improved the native chemical ligation method by solving the problem of peptide thioesterification. Furthermore, although the use ‘208’ and ‘601’ DNA provide excellent model systems, they can be further used to develop new DNA templates for specialized applications as discussed in this dissertation. To expand the study on long-range nucleosomal interactions, we also attempted to assemble a DNA template containing varying
linker lengths using the ‘601’ positioning sequence. The nucleosomal arrays assembled on such DNA template have the potential to be used for studying the effects of nucleosome spacing and distribution on chromatin compaction. Although the strategy of cloning to synthesize this novel DNA template did not work here, a different strategy is currently being developed in the lab to achieve this goal. The plan is to sequentially ligate the inserts on a resin with the use of non-palindromic ligatable ends.

Finally, here we have shown that SAGA does not autoacetylate or dimerize as shown previously. However, while pursuing the role of Ada3 acetylation, we observed, previously unseen, significantly tighter binding of pre-acetylated Ada3 peptide with Gcn5 bromodomain as compared to the binding of H3K14Ac peptide. This work seems promising in terms of providing additional insights into the nature of bromodomain binding and, thus, will be pursued further.