2015

Investigating the role of chromatin structure in modulating SAGA mediated nucleosomal acetylation

Chitvan Mittal

Iowa State University

Follow this and additional works at: https://lib.dr.iastate.edu/etd

Part of the Biochemistry Commons

Recommended Citation

Mittal, Chitvan, "Investigating the role of chromatin structure in modulating SAGA mediated nucleosomal acetylation" (2015). Graduate Theses and Dissertations. 14939.

https://lib.dr.iastate.edu/etd/14939

This Dissertation is brought to you for free and open access by the Iowa State University Capstones, Theses and Dissertations at Iowa State University Digital Repository. It has been accepted for inclusion in Graduate Theses and Dissertations by an authorized administrator of Iowa State University Digital Repository. For more information, please contact digirep@iastate.edu.
Investigating the role of chromatin structure in modulating
SAGA mediated nucleosomal acetylation

by

Chitvan Mittal

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

DOCTOR OF PHILOSOPHY

Major: Biochemistry

Program of Study Committee:

Michael Shogren-Knaak, Major Professor
Amy Andreotti
Reuben Peters
Scott Nelson
Yanhai Yin

Iowa State University
Ames, Iowa
2015

Copyright © Chitvan Mittal, 2015. All rights reserved.
To Mummy, Papa, Manu, Dadi, Dadu, Naniji and Nanaji for your unconditional love, faith and rock solid support. You are my pillars of strength.
# TABLE OF CONTENTS

DEDICATION .................................................................................................................... ii
NOMENCLATURE ............................................................................................................. v
ACKNOWLEDGEMENTS .................................................................................................. vi
ABSTRACT ......................................................................................................................... ix

## CHAPTER 1  GENERAL INTRODUCTION ................................................................. 1

- Chromatin .................................................................................................................... 1
- Histone code hypothesis – the tails tell a tale ............................................................. 3
- SAGA – The HAT on the chromatin ........................................................................... 6
- Kinetic mechanism of yeast Gcn5 acetyltransferase ............................................... 11
- Dissertation organization ......................................................................................... 13
- References .................................................................................................................. 14
- Figures ......................................................................................................................... 20

## CHAPTER 2  NUCLEOSOMAL ACETYLATION SEQUENCING TO STUDY THE ESTABLISHMENT OF CHROMATIN ACETYLATION .................................. 25

- Abstract ...................................................................................................................... 26
- Introduction ................................................................................................................ 27
- Materials and Methods ............................................................................................ 28
- Results ........................................................................................................................ 33
- Discussion .................................................................................................................. 38
- Acknowledgements .................................................................................................... 42
- Figures ......................................................................................................................... 43
- References .................................................................................................................. 51

## CHAPTER 3  ACTIVATOR AND LINKER DNA MODULATE THE HAT ACTIVITY OF YEAST SAGA .................................................................................. 54

- Abstract ...................................................................................................................... 54
- Introduction ................................................................................................................ 55
- Materials and Methods ............................................................................................ 57
- Results ........................................................................................................................ 62
- Discussion .................................................................................................................. 68
- Acknowledgements .................................................................................................... 72
- References .................................................................................................................. 73
- Figures ......................................................................................................................... 76
<table>
<thead>
<tr>
<th>CHAPTER 4</th>
<th>ADDITIONAL DATA AND FUTURE STUDIES</th>
<th>84</th>
</tr>
</thead>
<tbody>
<tr>
<td>Studies....</td>
<td>..................................................................................................</td>
<td>84</td>
</tr>
<tr>
<td>Figures....</td>
<td>..................................................................................................</td>
<td>90</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CHAPTER 5</th>
<th>CONCLUSIONS</th>
<th>93</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>APPENDIX</th>
<th>EXPRESSION AND PURIFICATION OF HISTONE H3 PROTEINS CONTAINING MULTIPLE SITES OF LYSINE ACETYLATION USING NONSENSE SUPPRESSION</th>
<th>100</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstract...</td>
<td>.................................................................................................................................</td>
<td>101</td>
</tr>
<tr>
<td>Introduction</td>
<td>............................................................................................................................</td>
<td>102</td>
</tr>
<tr>
<td>Materials and Methods</td>
<td>..................................................................................................................</td>
<td>104</td>
</tr>
<tr>
<td>Results...</td>
<td>........................................................................................................................</td>
<td>108</td>
</tr>
<tr>
<td>Discussion</td>
<td>.........................................................................................................................</td>
<td>114</td>
</tr>
<tr>
<td>Conclusion</td>
<td>..........................................................................................................................</td>
<td>117</td>
</tr>
<tr>
<td>References</td>
<td>..........................................................................................................................</td>
<td>118</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
<td></td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
<td></td>
</tr>
<tr>
<td>DNA</td>
<td>Deoxyribonucleic Acid</td>
<td></td>
</tr>
<tr>
<td>ATP</td>
<td>Adenosine Triphosphate</td>
<td></td>
</tr>
<tr>
<td>CENP-A</td>
<td>Centromere Protein-A</td>
<td></td>
</tr>
<tr>
<td>SWI-SNF</td>
<td>Switch-Sucrose Non Fermentable</td>
<td></td>
</tr>
<tr>
<td>ISWI</td>
<td>Imitation SWI</td>
<td></td>
</tr>
<tr>
<td>SAGA</td>
<td>Spt-Ada-Gcn5-Acetyltransferase</td>
<td></td>
</tr>
<tr>
<td>PTM</td>
<td>Post Translational Modifications</td>
<td></td>
</tr>
<tr>
<td>HAT</td>
<td>Histone Acetyltransferases</td>
<td></td>
</tr>
<tr>
<td>SLIK</td>
<td>SAGA Like</td>
<td></td>
</tr>
<tr>
<td>UV</td>
<td>Ultraviolet</td>
<td></td>
</tr>
<tr>
<td>VP16</td>
<td>Viral Protein 16</td>
<td></td>
</tr>
<tr>
<td>TBP</td>
<td>TATA Binding Protein</td>
<td></td>
</tr>
<tr>
<td>EMSA</td>
<td>Electrophoretic Mobility Shift Assay</td>
<td></td>
</tr>
<tr>
<td>CPM</td>
<td>Counts Per Minute</td>
<td></td>
</tr>
<tr>
<td>UAS</td>
<td>Upstream Activating Sequence</td>
<td></td>
</tr>
<tr>
<td>RNA</td>
<td>Ribonucleic Acid</td>
<td></td>
</tr>
<tr>
<td>ADP</td>
<td>Adenosine Diphosphate</td>
<td></td>
</tr>
<tr>
<td>EM</td>
<td>Electron Microscopy</td>
<td></td>
</tr>
</tbody>
</table>

ACKNOWLEDGEMENTS
My journey at Iowa State University began six years ago, and I would like to thank the people who made it possible and worthwhile, every step of the way.

First and foremost, I would like to thank my advisor, Dr. Michael Shogren-Knaak, for his guidance and patience throughout my graduate career. He has been an incredible mentor and taught me to be a better scientist and a better person. His laser sharp focus and perseverance through tough times have been extremely inspirational, and I hope I can live up to the high standards that he has instilled in me.

I would also like to thank my committee members, Dr. Amy Andreotti, Dr. Reuben Peters, Dr. Scott Nelson and Dr. Yanhai Yin for their valuable support, helpful suggestions, and insightful discussions that helped me steer the research in the right direction.

My experience with the Preparing Future Faculty (PFF) program at Iowa State University has been incredible and extremely valuable. This was made possible due the vision of two extremely talented and generous people, Karen Bovenmeyer and Dr. Holly Bender. They tirelessly organized excellent workshops week after week to give us endless exposure to the life in academia. They have been extremely helpful and were instrumental in my continuing journey. My PFF mentors Dr. Guru Rao and Dr. Amy Andreotti were fantastic and truly were incredible mentors. As a naïve graduate student, they shaped up my thoughts and helped me grow and mature during this journey.

I cannot thank the members of Shogren-Knaak laboratory enough, and I would not have been able to finish this journey without their support and friendship. I would especially like to thank Dr. Hamid Azzaz, Dr. Divya Sinha and Dr. Melissa Blacketer.
They were the first ones I had the pleasure to work with. Not only did they teach me different techniques and different aspects of science with patience, but also maintained an enthusiastic and fun environment in the lab. I am thankful for their friendship and without their constant support and faith in me, I would not have been able to graduate successfully. I am especially thankful to Divya and Melissa for believing in me during my worst moments and encouraging me every step of the way. Words cannot express enough gratitude, thank you! I am also thankful to Sannie Olson, who is wonderfully carrying the torch ahead. I have watched her grow and learn phenomenally, and I am confident that she will achieve great success. Her enthusiasm and go getter attitude is contagious, and our friendship has been extremely valuable to me. I would also like to thank other present and past members of the lab, Isaac Young, who is driven and curious, Maggie, Sarah and Junko.

Dr. Raji Joseph has been incredibly helpful during my entire journey as a graduate student. She is one of the most hardworking people I have known, and does so with a smile on her face. She has taught me to be perseverant and strong, and has also helped me in my research career with valuable discussions and encouragement.

I would like to thank the staff members at the department of BBMB, for helping with all the paperwork and making my journey smooth. Thank you Stacey, Connie, Pete and Herman. You are valuable to our department.

My family has been instrumental at every stage in my life and without their constant support, encouragement and unconditional love, I would not have had the courage to take a leap of faith to follow my dreams. I am extremely grateful and lucky to have the best parents in the world, and I owe my success to them. My Mother, Anju
Mittal is my hero. Her strength, perseverance and loving nature nurtured me and I hope I can be half as good as she is. Thank you mummy for believing in me when I didn’t see any light at the end of the tunnel. My father, Sanjay Mittal has inspired me to be a better person and to keep following my dream with passion and vengeance. Papa, I have always admired your courage, strength and give it all attitude to achieve your goals. I will forever be grateful and indebted to the sacrifices that both of you made in order to make my life better. I can only hope that I can fulfill all your wishes and can make your lives better too. I have the best brother in the world, Manu Kumar Mittal. Manu, I love you very much and am thankful for your love, friendship and faith in me. Your constant encouragement helped me survive some of the worst days. You make us very proud, and I am lucky to be your elder sister. I would also like to thank my grandparents, who believed in me more than anyone else did. Their love filled my childhood with endless joy. I miss you and wish you were here with me on my special day.

Last and certainly not the least, I would like to thank my friends who made Ames my home away from home. I am thankful to my best friend, Santhosh Vammi, for making this journey one of the best experiences of my life. I hope to have many more with you. I started my life at ISU with Abhijeet Kapoor and Navjot Singh, and the four of us have had the best times since then. I would also like to thank Kinit, Swati, Pratiksha, Hemant, Shambhavi, Gunjan, Jenifer, Divya, CJ, Nidhi, Nalini and Jagan.

In the end, I would like to thank my undergraduate advisors, Dr. Meenakshi Kulhar and Dr. Nandita Narayanasamy for their guidance and support. Also, my research mentors, Dr, Suneel Kateriya and Dr. Suman Kundu, who taught me the first lessons in research and getting me curious and inquisitive. It is because of all of you that this journey began. Thank you very much!
ABSTRACT

The SAGA family of co-activators play a major role in regulating eukaryotic gene expression. These multiprotein complexes are highly conserved across all eukaryotic species, from budding yeast to humans. Transcription of inducible genes is facilitated by SAGA, a multi-protein complex, which mediates acetylation of ε-amino tails of the histone H3. The acetyltransferase activity of SAGA is harbored in the Gcn5 subunit of the complex. Previous studies characterized the enzymatic properties of the recombinantly expressed HAT domain of Gcn5 from *Saccharomyces cerevisiae* on synthetic H3 peptides. We were interested in studying the properties of endogenous Gcn5, in the context of full SAGA complex. We also utilized nucleosomes as substrates, since the basic structural unit of chromatin is a nucleosome.

Towards this end, we first turned our efforts into developing a reliable, reproducible and sensitive assay for detecting acetylation on nucleosomes. Previous studies utilized filter binding assays, which work great for short peptides. Since these peptides are positively charged, they bind to the filter in an efficient and reproducible manner. Nucleosomes, being net negatively charged and structurally fragile, required an alternative approach. We developed bead based strategies which immobilize the nucleosome substrates, and thus circumvented the above-mentioned issues.

We utilized a trinucleosome system to study the pattern of acetylation established across the nucleosomal array under exhaustive and initial-rate conditions. We observed a 2 fold preference for nucleosomes harboring longer linker lengths on one side. By further optimizing the bead based assay for performing initial rate, steady state kinetics, we
designed several nucleosomes harboring different lengths of linkers on both sides. Our data suggest that linker DNA longer than 80 bp on one end stimulate the HAT activity of SAGA. We also show that a long linker in combination with linker on the other side of the nucleosome stimulates SAGA to the maximum potential.

Steady state data on different nucleosomes revealed that SAGA binds to all substrates pretty tightly, but the turnover rates were lower than expected. Our bead based assay also allowed us to perform single turnover experiments, which give important information on the nature of chemical step. Preliminary data suggests that indeed, the first turnover is pretty fast, but subsequent cycles are much slower. This indicates a step post catalysis, such as product release, is limiting. We are currently working towards understanding this burst phase in more detail.

*In vivo*, different factors work in concert to regulate gene transcription. For example, DNA bound activators recruit SAGA to target genes to initiate transcription. Previous studies have followed a combination of genetic and biochemical approaches to qualitatively assess this interaction. In order to investigate the effect of this interaction from a quantitative perspective, we incorporated an activator binding site in the linker DNA of one of the nucleosomes. Our preliminary data suggests that activator modulates SAGA complex in potentially two different ways, even beyond the stimulation mediated by linker DNA. This might explain how the cell is able to quickly respond to stress environments and turn on genes to achieve homeostasis.

Overall, our results help us in gaining better insights into the mechanisms employed by the SAGA complex in maintaining basal levels of acetylation under normal
conditions, and establishing hyperacetylated domains under inducing conditions.
CHAPTER 1

GENERAL INTRODUCTION

Chromatin

The discovery of DNA as the genetic material all living organisms remains one of the most crucial discoveries in understanding the fundamentals of all living things. Realizing genes are short segments of DNA which contribute to inheritable transmittance of characteristics led to the golden decade in the history of science. However, contrary to a long-standing belief, our genetic material is more than just naked DNA floating freely in the nucleus. Long stretches of DNA, often meters in length, need to be efficiently packaged within microns of a cell nucleus. This dilemma is addressed by packaging the DNA into a highly condensed form, by wrapping it around proteins, to form a nucleo-protein complex known as chromatin, which in its most extended state resembles beads on a string.

Nucleosomes: The basic structural unit of chromatin

The characteristics of the basic structural unit of chromatin, the nucleosome, have been investigated in detail. Pivotal studies conducted by Luger and co-workers shed light on the bead-like structure of nucleosomes. A nucleosome consists of 147 bp of DNA wrapped roughly 1.65 times around a spool of proteins known as histones, in a left
handed superhelical fashion [1; 2]. Two copies each of histone H2A, H2B, H3 and H4 constitute an octamer, and sequester genomic DNA. Histones are highly positively charged and thus interact quite strongly with the negatively charged DNA. The N-terminal tails of histones are highly unstructured, and protrude out from the core structure and beyond the DNA.

All histone proteins are highly conserved across all eukaryotic species and deletion mutants are lethal. Even though DNA sequences govern the stability of resultant nucleosomes, variations in DNA and histone sequences are generally readily accommodated [3]. This is necessary to be able to assemble nucleosomes across a wide variety of types of sequences within genomes. However, some DNA sequences are better suited to be wrapped around nucleosomes, and this sequence preference likely contributes at least partially to the positioning of nucleosomes throughout the genome.

In the genome, nucleosomes are linked to each other by short stretches of naked DNA, known as linker DNA. The lengths of these stretches vary largely across the genome. Transcriptionally silent loci tend to have regularly spaced nucleosomes, whereas more active loci have irregularly phased nucleosomes packing density, which results in variable lengths of accessible linker DNA. The linker DNA often harbors transcription factor and other modifying complex binding sites, and is often utilized as a means to regulate transcription and other cellular processes.

All eukaryotic genomes are studded with nucleosomes, the frequency of which varies from organism to organism. For example, in budding yeast, a nucleosome can be found every 200 bp ± 40 bp. Thus, a significant proportion of the eukaryotic genome is wrapped around histone proteins, and this has important ramifications for cellular
processes that utilize DNA as a substrate, such as DNA replication, repair and transcription. Because DNA wrapped around nucleosomes is less accessible to other proteins, chromatin is largely repressive in nature. Yet, the cell is able to faithfully replicate and segregate the genetic information to the daughter cells, and is able to respond and adapt quickly to changes in the environment to access genomic DNA for transcription. The cell uses a number of different mechanisms to change the structural properties of nucleosomes. Histone variants that are either more or less stable than the canonical histones can be incorporated [4; 5]. For example, CENP-A which is a variant of H3, is found exclusively in the centromeric regions of the genome [6]. H2A.Z, on the other hand, is often present in the promoter regions of actively transcribing genes [7; 8]. Additionally, the position and stability of nucleosomes can be directly changed by ATP-dependent remodeling complexes, such as SWI-SNF and ISWI family of remodelers [9; 10; 11]. The properties of nucleosomes are regulated by the cells by addition of post-translational modifications to the histone proteins [12]. Thus, there are mechanisms existing in the cell to overcome this repressive nature when required, and mechanisms to return chromatin back to the original state afterwards.

**Histone code hypothesis – the tails tell a tale**

Studies performed in the 1960s by Allfrey’s group correlating nucleosome acetylation with gene activation brought a paradigm shift in the field of chromatin biology [13]. Subsequent studies showed that histones are heavily modified with a multitude of post translational modifications. A majority of these modifications occur on
the tails, with some in the globular region. As mentioned earlier, the tails protrude out of the nucleosome, and thus provide a surface for potential interactions. The tails are also involved in mediating histone-DNA and histone-histone contacts and modifications of the residues have important consequences for the structure and stability of the nucleosome [12; 14].

A multitude of modifications have been detected on the histone tails. Initial attention was brought to this aspect when it was discovered that a lot of complexes that are usually tethered to the chromatin had modifying activities. Proteomic approaches, such as mass spectrometry analysis of histones, rapidly led to accumulation of a staggering amount of data, all converging to the fact that histone tails serve as a crucial target of various modification enzymes [12; 15]. Figure 2 lists the possible modifications that have been detected till date. As the figure shows, the same residue can be marked with different modifications, and also a small stretch can harbor multiple marks [16].

Chromatin is a highly dynamic structure with respect to these marks. Not all of them are present at the same time, and the landscape rapidly and continually evolves with different upstream signals. This led to the genesis of the histone code hypothesis, which states that the downstream cellular effect is a result of combined effects of the various modifications present at and surrounding the target loci. Since all of these modifications are reversible, and are deposited or erased rapidly, there is no one code, but a language, where each mark is a letter that makes sense in the context of all the other marks constituting the word [12; 15; 17].
Acetylation

Of the several modifications mentioned above, lysine acetylation has been most extensively studied, and it’s effects well established. The $\varepsilon$-amino group of the side chain of lysine is a target of various modifying enzymes. Transfer of the acetyl group to the positively charged lysines neutralizes the charge, and thus weakens the interactions between the modified histone and DNA. This leads to a number of structural changes. For example, the stability of the nucleosome is changed in some cases, whereas in others, it imparts different mobility to the octamer. More extensively, it is a mechanism to open the chromatin structure to a more de-condensed state. Indeed, Shogren-Knaak and coworkers have shown that acetylation of lysine 4 on H4 results in a less compacted nucleosomal array [18]. This has important consequences for mediating transcriptional control. One can imagine that unraveling of the chromatin into a more open structure will help in facilitating transcription. This has been supported over the years, since histone acetylation is largely correlated with transcriptionally active chromatin. Lysine acetylation can also affect transcription through a mechanism separate from directly changing chromatin structure and stability. Histone acetylation creates new platforms for tethering of chromatin modifying complexes. Effector proteins that harbor acetylation recognition domains get recruited on specific sites, and thus mediate downstream effects. It can also inhibit recruitment of other classes of enzymes, by destroying the un-acetylated sites. This prevents deposition of repressive marks, thereby inhibiting transcriptional repression and/or chromatin compaction.
Spt-Ada-Gcn5-Acetyltransferase (SAGA) – The HAT on the chromatin

The enzymes that mediate acetylation on chromatin are referred to as histone acetyltransferases (HATs). The first link between histone acetylation and gene regulation came from the discovery that budding yeast protein Gcn5, which previously was known to be involved in gene transcription, was also shown to possess histone acetyltransferase activity [19; 20]. Since then, a number of other HATs have been discovered, and a combination of genetic, biochemical and structural studies unequivocally established the role of histone acetylation in regulation of chromatin structure and associated processes, such as transcription, DNA replication, DNA repair etc. [21; 22].

Acetyltransferases have been divided into five different families, GNATs (Gcn5-related-N-acetyltransferases); MYST (MOZ, Ybf2, Sas2 and Tip60) related HATs; p300/CREB-binding protein (CBP) HATs; general transcription factor HATs including TAF (TBP-associated-factor) and finally, the nuclear hormone related HATs [23; 24; 25]. Some of the now known HATs were previously known to assist in transcription as adapters or co-activators, but later shown to harbor histone specific acetyl transferase activity. It is now known, that these HATs also have the ability to acetylate non-histone substrates, such as transcription factors and other chromatin modifying complexes, and are thus referred to as general ATs, or Lysine ATs (KATs).

Biochemical characterization of the yeast Gcn5 protein showed that it acts specifically on the lysine 14 of H3, *in vitro*. However, researchers were left puzzled for a while when it failed to act on core histones and nucleosomes, as observed *in vivo*. The purification and characterization of several Gcn5 containing endogenous complexes
solved this mystery. Gcn5 by itself can only target H3K14, but has an expanded specificity in the context of the entire complex [26; 27]. Song Tan and co-workers showed that additional subunits Ada2 and Ada3 are necessary and sufficient to expand this specificity and to also endow the ability to act on full histones, as well as chromatin substrates [28].

Further, Gcn5 is present in three different complexes in yeast. They are the 1.8 MDa complex known as the SAGA complex, the 800 kDa complex known as the ADA complex, and a complex very similar to SAGA, known as SALSA/SLIK (SAGA altered. Spt8 absent/SAGA like). Chromatography and other biochemical approaches established that these three are distinct complexes, despite a large overlap of subunits [29]. As expected, they have redundant roles in the cell, and are able to compensate for each other. However, the presence of unique subunits also makes them indispensable for certain pathways [25].

The SAGA complex is highly conserved across all eukaryotic species and overall has a very similar composition of subunits. In yeast, it is a 1.8 MDa complex, comprising of at least 19 different polypeptides [21; 25]. Multiprotein complexes offer several advantages in cellular context. By tethering different activities together, the cell is poised to function much more efficiently and respond quickly to any changes in the extracellular environment. Low resolution Electron Microscopy (EM) studies showed that the complex is organized into five different modules, each of which harbor a different function [30]. The recruitment module is involved in interactions with transcription activators and other factors involved in transcription. Tra1 is the main subunit of this module, and is 400 kDa in size. This offers several interaction opportunities and is often seen as a scaffold [31;
The acetylation module comprises of Gcn5, Ada2, Ada3 and Sgf29. Gcn5 is the HAT, but needs proper interactions with Ada2 and Ada3 to be able to act on in vivo chromatin substrates, i.e., nucleosomes either one or multiple in tandem. The TBP interaction unit, which comprises mainly of Spt3 and Spt8, is directly involved in interacting with the TATA binding protein, and thus possess the classic co-activator properties. The de-ubiquitination (Dub) module comprises of Ubp8, Sus1, Sgf11 and Sgf73, and accounts for the second enzymatic activity of the complex [30; 33]. Ubp8 is a deubiquitinase and removes the ubiquitin moiety from the lysine side chain of residue 123 on H2B, in yeast. Similar to Gcn5, Ubp8 works in the context of the Dub module and is a poor enzyme by itself. More and more of such examples are being discovered, underscoring the importance of context in the world of chromatin. The fifth module is the architectural module and is crucial in maintaining the overall structural integrity of the complex. Spt7, Spt20, Ada1 and some general TAFs such as TAF5, 6, 9 and 12 constitute this module. The overall subunit composition is strikingly similar across different species, such as yeast, Drosophila and humans, suggesting similar functions and utilization of similar principles to modulate chromatin [21; 22; 34; 35].

As shown, different subunits harbor different domains which allow them to mediate different activities. For instance, Gcn5 has a HAT domain which catalyzes the acetyltransferase reaction and a bromodomain, which recognizes acetylated lysines on the histone, thereby facilitating recruitment on the acetylated regions of chromatin.

Given the fact that SAGA is conserved in all eukaryotes and harbors a multitude of functions, it’s not surprising to know that it is involved in a variety of cellular processes, not just transcription. Also, misregulations in these activities leads to
developmental issues and diseases. For instance, gcn5 null mice do not survive post 10.5 days, and gcn5 null fruit flies die due to failure of metamorphosis [36; 37]. In humans, the Sca7 subunit can undergo polyglutamine expansion and this results in neurodegeneration. The mutant subunit gets incorporated into the complex and results in mistargeting and aberrations in the activity. Studies have shown that it directly inhibits the HAT activity of Gcn5 [38; 39]. Mutations in Ubp8 enzyme leads to a variety of cancers, and is in fact a part of 11 signature mutations that result in a wide variety of cancers [40; 41]. A lot of current drug targets being developed are aimed against the domains present in the SAGA complex, thereby giving impetus to a need for a greater understanding of how this complex functions to ensure the success of an organism.

**Promoter specific targeted acetylation**

Genome-wide transcriptional profiling of SAGA deletions in budding yeast indicates that SAGA regulates the transcription of about 5-10% of the total genes [42; 43]. Most of these genes are up-regulated during stress conditions, such as glucose deprivation, amino-acid starvation, phosphate shortage, heat shock and ER specific stresses [44; 45; 46; 47]. Recent studies have also shown that Gcn5 promotes resistance to UV-mediated damage and helps in maintaining the overall integrity of the genome [48; 49]. Thus, SAGA is largely involved in the regulation of stress specific pathways. This specificity is ensured so that the promoters of only a subset of genes undergo Gcn5 mediated hyper-acetylation, to initiate transcription.

GCN5 was identified during one of the genetic screens that looked for factors which directly interacted with Gal4VP16, a hybrid transcriptional activator that has the
DNA binding domain of endogenous GAL4 activator, and the activation domain of VP16, a viral activator protein from herpes simplex virus [50; 51]. Since then, Gcn5 and SAGA have been shown to physically interact with a variety of other transcription activators. These DNA bound proteins are highly specific for their target loci, which is usually the UAS region in the promoter. Once the signal arrives, these activators then recruit SAGA to the promoters to kick start transcription. At these target loci, SAGA can facilitate transcription in several different ways. For instance, Gcn5 acetylates the neighboring nucleosomes which establishes hyper-acetylated domains. This might lead to nucleosome eviction. Another way is to recruit the TATA binding protein (TBP), which then proceeds to establish the pre-initiation complex, thereby sending the message to RNA Polymerase II to begin transcribing the downstream coding region [52; 53].

Several studies have been done to look at the profile of acetylated histones generated, following the targeted recruitment of SAGA. It was found, that at several loci, SAGA acted on 600 bp, proximal to the target site, thereby spanning 2-3 nucleosomes. Thus, a sharp spike in acetylation occurs, but on a very limited region. Other HATs, such as NuaA4 on the other hand, create broader domains of these hyper-acetylated regions [54]. These profiles are dependent on activator-SAGA interactions, as relocation of the activator target site shifts the recruitment of SAGA, in accordance to the new site, and the neighboring nucleosomes are now acetylated, more than the basal levels. And the nucleosomes that were hyper-acetylated, now return to basal levels [54]. Also, HATs that do not physically interact with activators, such as NuA3, function in a more untargeted manner.
Global, untargeted acetylation

Given the fact that SAGA is involved in regulating only a subset of genes, it was surprising and perhaps confusing to find a large part of the yeast genome occupied by SAGA. The genome was found to be largely acetylated, at levels more than expected background levels. Gcn5 deletion resulted in drastic decrease in widespread acetylation, suggesting that SAGA is also involved in establishing basal levels of acetylation throughout the genome, which are different from the promoter targeted hyperacetylation domains on specific loci [55; 56]. It is speculated that non targeted acetylation maintains a poised state for transcription, so that the cell is able to quickly respond to any extracellular changes. The role of basal levels of acetylation and widespread genomic occupancy of SAGA is still unclear and remains an active area of research.

Kinetic mechanism of yeast Gcn5 acetyltransferase

Denu and coworkers have used enzymology to gain insights into the mechanism employed by Gcn5 acetyltransferase to acetylate histone H3. Initial studies were performed using calf thymus histones, and later with synthetic H3 peptides. Both were in good agreement, allowing researchers to use synthetic peptides as probes to study the role of specific modifications. Synthetic peptides offer several advantages over biologically isolated histones. First, synthetic peptides are homogeneous in nature, and also allow
researchers to modify them at desired locations, thus studying the roles and effects of that specific modification.

Enzymes that utilize acetyl-CoA as a cofactor can employ two different mechanisms to mediate acetylation of the desired substrates. One of them involves a transfer of the acetyl group on the enzyme, which is in turn transferred to the substrate. This mechanism involves the formation of a covalent complex between the enzyme and acetyl CoA. The second mechanism involves the formation of a ternary complex between the enzyme, acetyl CoA and the substrate. A direct transfer of acetyl group occurs from acetyl CoA to the second substrate. Using the first 20 residues of H3, and the HAT domain of Gcn5, Denu and coworkers were able to show that the Gcn5 class of acetyltransferases employ the latter, by forming a ternary complex with the two substrates. A series of elegant and detailed approaches proved that the Glutamic Acid 173 on Gcn5 is essential for the proper functioning of the enzyme under physiological conditions, and acts by deprotonating the ε-amino group of the lysine side chain of the histone. This allows direct nucleophilic attack by acetyl CoA, to ultimately transfer the acetyl moiety to the histone, within the active site of the enzyme. The reaction follows an ordered Bi-Bi mechanism, where acetyl CoA binds first, followed by histone to the active site of Gcn5. A ternary complex is formed, and the acetyl group is directly transferred to the lysine of histone. The acetylated peptide is then released, followed by CoA [57; 58].

The abovementioned studies were done on a histone peptide, using the HAT domain of Gcn5. The enzyme acts on lysine 14 of H3 in vitro, but has an expanded specificity towards lysine 9, 18, 23 of H3, 8 and 16 of H4 and some residues on H2B, when incorporated into larger, multi-subunit complexes. We have been investigating the
substrate specificity and kinetic mechanisms of the entire complex on more biologically relevant substrates, such as nucleosomes and longer arrays of nucleosomes. These will be discussed in Chapter 3.

**Dissertation Organization**

The dissertation has been organized into five different chapters and one Appendix section. Chapter 1 encompasses a general introduction about chromatin, and how different aspects of it are utilized by co-activators, such as SAGA, to facilitate eukaryotic transcription. Chapter 2 describes a published method that was developed in order to gain quantitative insights into chromatin acetylation by SAGA. The work has been published in the journal Analytical Biochemistry. Chapter 3 builds on an observation made in Chapter 2, that DNA flanking nucleosomes appears to regulate SAGA activity. Additionally, experiments to determine how activator proteins affect SAGA activity are described. This work is nearly completed and will be submitted for publication shortly after this defense. Chapter 4 discusses ongoing work and future experiments, in collaboration with another graduate student in the lab. Results obtained will be a part of a future manuscript. Chapter 5 summarizes the conclusions reached in our studies. The Appendix section describes another published report that was developed as a collaborative effort with a graduate student in the lab, during my graduate career. The method described in Chapter 2 was utilized in this study.
REFERENCES


A. Durand, J. Bonnet, M. Fournier, V. Chavant, and P. Schultz, Mapping the deubiquitination module within the SAGA complex. Structure 22 1553-9.

Y. Han, J. Luo, J. Ranish, and S. Hahn, Architecture of the Saccharomyces cerevisiae SAGA transcription coactivator complex. EMBO J 33 2534-46.


Figure 1. The structure of a mononucleosome, as determined by X-ray crystallography. (A) The left panel shows the view down the DNA superhelix axis,
whereas it’s perpendicular to it, for the right panel. (B) Half of the nucleosome core particle, as viewed down the DNA axis [14].
Figure 2. Histones are extensively enriched with post-translational modifications.

The tails are modified more than the central core. Shown are the different residues that get modified in the 4 core histones in human cells. In addition to the depicted modifications, other kinds of modifications also occur, such as sumoylation, ADP-ribosylation and proline isomerization. Ac, acetylation; Me, methylation; P, phosphorylation; Ub, ubiquitination [16].
Figure 3. Molecular architecture of the SAGA complex. (A) Low resolution cryo-EM studies established the modular nature of SAGA [30]. (B) More recent studies using cross-linking approaches gave insights into a more detailed orientation of the subunits [35].
Figure 4. Pleiotropic effects of SAGA on eukaryotic gene transcription. (A) The recruitment module of SAGA interacts with activator protein to mediate localization on target genes. This allows the HAT module to acetylate neighboring nucleosomes. (B) The TBP interaction module furthers the interactions between SAGA and the transcriptional machinery, to establish the PIC. (C) Upon initiation, SAGA interacts with Pol II and helps in elongation. The HAT module further acetylates and often evicts the nucleosomes from the coding region. (D) The Dub module de-ubiquitinates H2B, ultimately causing phosphorylation of Pol II, to extend elongation [21].
CHAPTER 2

NUCLEOSOME ACETYLATION SEQUENCING TO STUDY THE
ESTABLISHMENT OF CHROMATIN ACETYLATION

Chitvan Mittal, Melissa J. Blacketer, and Michael A. Shogren-Knaak

1. Roy J. Carver Department of Biochemistry, Biophysics and Molecular Biology,
Iowa State University, Ames, Iowa 50011

E-mail address: knaak@iastate.edu

Short title: Nucleosome acetylation sequencing
Abstract

The establishment of post-translational chromatin modifications is a major mechanism for regulating how genomic DNA is utilized. However, current \textit{in vitro} chromatin assays do not monitor histone modifications at individual nucleosomes. Here we describe a strategy, nucleosome acetylation sequencing, that allows us to read the amount of modification at each nucleosome. In this approach, a bead-bound trinucleosome substrate is enzymatically acetylated with radiolabeled acetyl CoA by the SAGA complex from \textit{S. cerevisiae}. The product is digested by restriction enzymes that cut at unique sites between the nucleosomes and then counted to quantify the extent of acetylation at each nucleosomal site. We find that we can sensitively, specifically, and reproducibly follow enzyme-mediated nucleosome acetylation. Applying this strategy, when acetylation proceeds extensively, its distribution across nucleosomes is relatively uniform. However, when substrates are used that contain nucleosomes mutated at the major sites of SAGA-mediated acetylation, or that are studied under initial rate conditions, changes in the acetylation distribution can be observed. Nucleosome acetylation sequencing should be applicable to analyzing a wide range of modifications. Additionally, because our trinucleosomes synthesis strategy is highly modular and efficient, it can be used to generate nucleosomal systems in which nucleosome composition differs across the array.

Keywords: Chromatin; Acetylation; Enzymology
Introduction

In eukaryotes, the utilization of genomic DNA for processes such as transcription, replication, or repair, is usually preceded by post-translational modification of its chromatin [1]. Typically, these modifications occur over multiple nucleosomes and include different types of modifications on distinct histone residues. For example, during the transcription of inducible genes, nucleosomes of promoter regions can exhibit histone H3K4 methylation, H3 and H4 histone tail acetylation at multiple lysines, and histone H3S10 phosphorylation [2-8]. These modifications can be highly dynamic and can affect the establishment of additional modifications. For example, H2BK123 ubiquitinination precedes and is required for H3K4 methylation [9, 10].

*In vitro* assays offer a potentially powerful tool for elucidating how chromatin modifications are established by controlling the composition of the reaction components and by providing quantitative information on reaction rates and distributions. However, current *in vitro* assays do not provide information on how histone modifications are established across a span of nucleosomes. On one hand, assays have been used to carefully define the sites and kinetics of enzymes toward histone tail peptides [11-14], but provide a relatively limited approximation of the physiological substrates, chromatin. On the other hand, biochemical methods can generate nucleosomal arrays, but discerning the extent of modification across different nucleosomes is difficult.

To remedy these problems, we have established a method, nucleosome array acetylation sequencing, which allows us to read-out the extent of nucleosome modification across a well-defined oligonucleosome substrate. The enzyme we have chosen to focus on for our initial studies is the Spt-Ada-Gcn5-Acetyltransferase (SAGA)
complex from *Saccharomyces cerevisae*. This multi-subunit complex, first isolated in budding yeast and subsequently found to be present in a wide range of eukaryotes [15], facilitates the transcriptional initiation of inducible genes, such as developmental genes in humans [16-18]. One way it accomplishes this is by acetylating histones lysines in chromatin, where four lysine residues in histone H3 are the major sites of acetylation [19, 20]. Further, SAGA contains domains that recognize histone acetylation, which potentially helps mediate the spread of SAGA-mediated nucleosome acetylation across multiple nucleosomes in inducible gene promoters [21]. Studying how SAGA acetylates a multiple nucleosome system offers the potential to generate new insights into the mechanism by which histone modifications are established across a chromatin region.

**Materials and Methods**

**Chemicals**

Buffer components were purchased from either Fisher Scientific or Sigma Aldrich. [H3]-Acetyl CoA with a specific activity of 16.9 Ci/mmol was purchased from Moravek. Restriction endonucleases, MNase, T4 DNA ligase, and Streptavidin coated magnetic beads were purchased from NEB.

*Histone octamer preparation and mononucleosome assembly*
Recombinant wild type and tetra-alanine (K8A, K14A, K18A, K23A) mutant H3 *X. laevis* histones were expressed in *E. coli* and purified as described previously [22]. The purity of the histones was assessed by SDS PAGE and quantitated using UV spectroscopy. Histone octamer was prepared using the four canonical histones H2A, H2B, H3 and H4, as described previously [23]. The 601-177-1 strong positioning DNA sequence was prepared, as described previously [24, 25]. Mononucleosomes were assembled by rapid dilution as described previously [26, 27] and analyzed by 4% native PAGE to assess the degree of homogeneity and saturation. We observed that the innermost nucleosome, nucleosome 1, could adopt two different positions, with respect to the DNA.

*SAGA expression and purification*

Endogenous *S. cerevisiae* SAGA complex containing Spt7-TAP was affinity purified as described previously [28]. The purified enzyme was quantitated by Western Blotting, using recombinant Gcn5 as a standard. The activity of the purified complex was tested on H3 peptide using a filter binding assay [29].

*Biotinylated adapter*

Oligonucleotides 5’/-5PCBio/ACGACG GCCAGTGAACCACGATT-3’ and 5’/-5Phos/GTGGTTCACTGGCCGTCGT - 3’ were purchased from IDT. To form a double stranded adapter, the oligos were mixed, each at a final concentration of 30 µM, in 1X
NEB T4 DNA Ligation buffer (50 mM Tris-HCl at pH 7.5, 10 mM MgCl₂, 1 mM ATP, and 10 mM DTT), in a final volume of 50 µl. The mixture was heat denatured at 96-98°C in a heating block for 10 minutes. Subsequently, the heating block was turned off and the temperature was allowed to cool to 33°C. The oligos were then placed at RT for 5 minutes before transfer to ice and subsequent storage at -20°C. The annealed product, a 19 bp double stranded adapter with a 4 nucleotide overhang on the non-biotinylated end, was analyzed by 20% SDS-PAGE to assess the purity.

**Synthesis of the trinucleosome on beads**

For a typical synthesis, 4.66 pmoles of biotinylated adapter were attached to 50 µl of Streptavidin coated Dynabeads slurry (4mg/ml suspension) by incubation in 1X binding buffer (10 mM Tris-HCl pH 7.4, 100 mM NaCl) at room temperature for 10 minutes with occasional agitation. The beads were washed 3 times with 50 µl 1X binding buffer prior to adding the adapter. The adapter-bound beads were pulled to the side of the tube using a magnet and the supernatant was removed. 6.99 pmoles of 601-177-1 mononucleosomes were added one at a time in a set of sequential ligations in the following order: mononucleosome 1, mononucleosome 2, mononucleosome 3. Each ligation was done in a final volume of 40 µl in 1X ligation buffer (50 mM Tris, 10 mM MgCl₂, 1 mM ATP, 10 mM DTT; pH 7.5) and 800U of T4 DNA ligase. The ligation reactions were allowed to proceed at room temperature for 12-16 hours after which, excess unbound nucleosome was washed off.
This first nucleosome, referred to as nucleosome 1, has 80 bp of additional sequence upstream in the linker DNA as compared to nucleosomes 2 and 3, and the upstream end forms a BstXI restriction site upon ligating with the adapter. The linker DNA between ligated nucleosomes 1 and 2 contains a BglII restriction site, whereas the linker DNA between ligated nucleosomes 2 and 3 contains a PflMI site.

The resultant trinucleosome was digested with BstXI, BglII or PflMI at 37°C for 5 hours and analyzed by 4% native PAGE to assess the quality of the ligation product. The trinucleosome was stored, on beads, at 4°C at a final concentration of 150 nM.

**Acetylation sequencing of trinucleosome**

WT or mutant trinucleosome, at a final concentration of 75 nM of trinucleosome, was acetylated under saturating or sub-saturating conditions by the SAGA complex under standard HAT assay conditions [30]. For both saturation and sub-saturating conditions, the substrate was acetylated at 30°C in 1X HAT buffer (25 mM Tris-Cl pH 7.5, 50 mM KCl, 5% glycerol, 1 mM DTT, 1 mM PMSF, 10 mM sodium butyrate, 2.66 µM unlabeled, 5.33 2.66 µM radiolabeled tritiated acetyl CoA with a specific activity of 2.66 µM/ml) at a final volume of 50 µl. For saturation conditions, acetylation was allowed to occur for 4 hours by 60 nM final SAGA, whereas the reaction took place for 10 minutes for sub-saturation conditions by 10 nM final SAGA. The beads were washed 5 times at RT with 50 µl 1X washing buffer (100 mM NaCl, 50 mM Tris-HCl, 10 mM MgCl₂, 1 mM DTT, 0.1% Triton X-100). They were then washed an additional 11 times, with a 7 minute incubation at 37°C between washes, for saturation conditions and an additional 7
times, with 37°C incubation as above, for sub-saturating conditions. After washing, the beads were resuspended in 50 µl 1X digestion buffer (100 mM NaCl, 50 mM Tris-HCl, 10 mM MgCl₂, 1 mM DTT, 0.1% Triton X-100). 10 µl aliquots were transferred to 4 fresh tubes. In order to minimize handling and reduce the total time for the assay, we chose to count nucleosomes generated from a single round of digestion. The samples were either digested with PflMI (4U), BglII (4U), BstXI (4U), or mock digested at 37°C for 5 hours in a final volume of 20 µl. Because 10-20% of nucleosomes were not liberated by BstXI digestion alone, these samples were further digested with MNase (1U) under conditions similar to the other digestions, for 10 minutes. The supernatants were added directly to vials containing 9 ml of Scintiverse BD LSC cocktail (Fisher Scientific). All samples were counted using a Packard tri-carb 1600 TR liquid scintillation analyzer. Results were analyzed from 4 independent trials of each experiment.

Testing of nucleosome stability under assay conditions

Wild type trinucleosome was subjected to the assay conditions described above with one exception: only cold acetyl CoA was used. The trinucleosomes were then digested with PflMI or BglII. Following removal of the supernatant, additional digestions were performed on the remaining bead-bound product. The beads were resuspended in 1X digestion buffer and digested with BstXI or mock digested in a final volume of 20 µl, at 37°C for 5 hours. The supernatants were then analyzed by 4% native PAGE.
Comparison of nucleosome acetylation on bead and in solution.

DNA for mononucleosome 1 and the biotinylated adapter were ligated and purified. Mononucleosome was then prepared with this DNA and wild-type octamer. For acetylation assays, 7.5 nM SAGA was used with 55 nM mononucleosome under the reaction conditions described above. For the solution assay, all components were mixed and the reaction was allowed to proceed for eight minutes before streptavidin beads were added. After two minutes, the beads were washed extensively and then counted as above. For the bead assay, mononucleosome were first incubated with the streptavidin beads for two minutes. The remaining reaction components were then added and the reaction was allowed to proceed for 10 minutes. Beads were then washed and counted. It was found that two-minutes of incubation was sufficient for complete nucleosome binding. To determine background signal, mock reactions were performed without SAGA.

Results

To better understand how chromatin-modifying enzymes establish histone acetylation across a chromatin substrate, we developed an on-bead acetylation sequencing assay (Fig. 1). In this strategy, a standard histone acetyltransferase assay is performed on a trinucleosome substrate that is attached to streptavidin coated magnetic beads, where the SAGA complex is the enzyme and radiolabeled acetyl CoA is the acetyl
After removing unincorporated acetyl CoA, the acetylated trinucleosome is digested with different restriction endonucleases to liberate specific nucleosomes. The amount of radioactivity is quantified by liquid scintillation counting, allowing us to ultimately measure the extent of acetylation for each nucleosome.

*Synthesis of cleavable trinucleosome substrate*

As a first step, we needed to generate a bead-bound trinucleosome substrate with unique restriction sites between each nucleosome (Fig. 1). In previous work [26], we had generated short nucleosomal arrays by ligating mononucleosomes with different non-palindromic overhangs together in solution. While this strategy was effective, it was difficult to separate ligated products that were incomplete from the desired array. As an alternative, we investigated a solid-phase nucleosomal array synthesis approach (Fig. 2A). The first step in this strategy was to immobilize a biotinylated, double stranded DNA adapter to streptavidin coated magnetic beads. The binding between the biotin and streptavidin occurred very rapidly and completely (data not shown). Following binding of the adapter, mononucleosomes were then ligated sequentially, first to the adapter, then to each other, with extensive washing between ligation steps. An excess of mononucleosomes was used in each ligation step to drive ligations to completion and could be observed in the ligation supernatant (Fig. 2B, lanes “Sup 1-3”).

To characterize the fully ligated product, the trinucleosome was liberated from the bead by BstXI digestion. Native gel analysis of this product (Fig. 2B) suggests that it is a single product with no incomplete ligation intermediates present. Further, it appears that
using mononucleosome with different non-palindromic overhangs prevents incorporation of more than one nucleosome per round of ligation.

Validation of acetylation sequencing assay

For nucleosome acetylation sequencing to be useful, a number of criteria must be fulfilled. One requirement is that acetylation signal should be due to enzyme-mediated acetylation of the nucleosomes, not because of non-specific acetylation or acetyl CoA retention. Additionally, this acetylation should be detectable with high sensitivity and reproducibility. To this end, we looked at acetylation levels of bead-based substrates with different components of the acetylation reaction present. In the absence of acetyl CoA, SAGA, or nucleosome, we saw that the level of acetylation was relatively low on the bead prior to digestion (Fig. 3A left panel) and in the supernatant after digestion (Fig. 3A right panel). This indicates that there was not significant signal from something besides the acetyl group, non-enzymatic substrate acetylation, bead acetylation, or non-specific acetyl CoA retention. On the other hand, when all components were present, a robust signal was seen on the bead prior to digestion (Fig. 3A left panel) and in solution after digestion (Fig. 3A right panel). This signal was significantly greater than those observed in our control reactions, indicating that enzyme-mediated nucleosome acetylation could be detected with good specificity and sensitivity. The amounts of signal on the bead prior to digestion and in solution after digestion were relatively similar, suggesting that counting on bead and in solution are comparable. However, because initial results with the liberated nucleosomes were more reproducible, as judged from the relatively small
size of the error bars relative to the total signal (Fig. 3A), all subsequent experiments were done in solution.

Another requirement of the acetylation sequencing assay is that each nucleosome be removable from the bead quantitatively, specifically, and without disturbing the integrity of the nucleosomes. After wild-type trinucleosomes were acetylated by SAGA, they were subjected to restriction enzyme digestion and analyzed by native gel electrophoresis (Fig. 3B). Digestion of the acetylated trinucleosome by PflMI liberated nucleosome 3 specifically, as no other digestion products were observed. The liberated mononucleosome was also stable to the assay conditions, as no free DNA was seen. Moreover, the digestion was complete. Following PflMI digestion and removal of the liberated product, further digestion of the bead by BstXI generated dinucleosome product. If the PflMI digestion had been incomplete, some trinucleosome product would have been expected in the subsequent BstXI digestion. Similarly, digestion of the acetylated trinucleosome product between the first and second nucleosomes with BglII resulted exclusively in dinucleosome product (nucleosomes 2 + 3) with only mononucleosome (nucleosome 1) left on the bead. BstXI digestion of the acetylated trinucleosome product generated only trinucleosome product. Thus, our restriction digestion analysis indicates that digestion at all three restriction sites is highly efficient and non-disruptive. Additionally, they further support the integrity of the synthesized trinucleosome substrate. While these results suggest that performing multiple digestions is possible, subsequent experiments were performed with single digestion for consistency and to minimize the handling of the substrates.
To determine if acetylation is affected by attachment of the trinucleosome to the bead, assays were performed on mononucleosomes in solution and immobilized on a bead (Fig. 3C). Comparison of the total acetylation between the two substrates show that they are identical, suggesting that attachment of the nucleosomes to a bead does not significantly change how SAGA acetylates its substrate.

Characterization of model trinucleosome acetylation

Because our initial experiments demonstrated the feasibility of our acetylation sequencing strategy, we applied this strategy to studying SAGA-mediated nucleosome acetylation with different nucleosomal substrates. Our initial studies were performed under saturating acetylation conditions, where acetylation reactions were run with concentrations of SAGA near those of the nucleosome substrate (60 nM and 75 nM, respectively), and over long periods of time (4 hours). Under these conditions, we expected that nucleosome acetylation should be largely complete, and thus not differ significantly between the three nucleosomes. Indeed, all three nucleosomes were labeled to nearly the same extent under these conditions (Fig. 4A). However, acetylation sequencing could also report on differences in the pattern of nucleosome acetylation. When assays were performed under saturating acetylation conditions on a substrate in which the central nucleosome lacked the major sites of SAGA acetylation (Fig. 4B), nucleosome 2 was acetylated significantly less, while the other two nucleosomes were acetylated to comparable levels.
These experiments were also performed under sub-saturating conditions. For the wild-type trinucleosome (Fig. 4C), the total amount of acetylation is significantly diminished, as is expected because the acetylation assay was performed with significantly less enzyme (6 fold decrease) and time (24 fold decrease) than for the saturating conditions. The pattern of nucleosome acetylation is also altered. There is a preference for nucleosome 1, the innermost nucleosome (closest to the bead) with the other two nucleosomes, nucleosomes 2 and 3, showing roughly half as much acetylation. For experiments under similar conditions with the trinucleosome with the mutated central nucleosome, the central nucleosome (nucleosome 2) is poorly acetylated, while nucleosome 3 is again acetylated to a level of about half that of nucleosome 1.

Discussion

Interpretation of acetylation sequencing patterns

While insightful from a qualitative perspective, the acetylation distribution can also provide quantitative kinetic information. During the course of an acetylation reaction, each nucleosome within an array competes as a substrate, where the extent of acetylation of a particular nucleosome depends on the strength with which it binds to the enzyme, its acetylation turnover, and the concentration of the unacetylated nucleosome:

\[
\frac{V_{\text{nuc}_x}}{V_{\text{nuc}_y}} = \left(\frac{K_{\text{cat}}}{K_M}\right)_{\text{nuc}_x}[\text{nuc}_x] \left(\frac{K_{\text{cat}}}{K_M}\right)_{\text{nuc}_y}[\text{nuc}_y]
\]
Where, \( v_{nuc} \) is the instantaneous rate of nucleosome acetylation under steady state conditions \( (d[][]/dt) \), \( k_{cat}/K_M \) is the specificity constant, and \([nuc]\) is the equilibrium concentration of free nucleosome.

Under certain circumstances, it is possible to directly compare the relative amounts of nucleosome acetylation at two nucleosomes to determine their relative ratio of specificity constants, thereby quantifying relative substrate preference. However, under saturating acetylation conditions (Fig. 4A & B), this is not possible. In these experiments, large amounts of enzyme are used (1:1.25 enzyme to each nucleosome) and reactions are allowed to proceed for long times (4 hours total). Under these conditions, the total amount of radioactive signal corresponds to nearly 8 acetylations/nucleosome, suggesting complete acetylation of the major nucleosomal sites (four sites in each of the two H3 tails). Thus, the total amount of acetylation at each nucleosome does not reflect its acetylate rate \( (v_{nuc}) \), and it would be expected that little difference in acetylation distribution between nucleosomes would be observed even with differences in specificity constants. Interestingly, under exhaustive acetylation conditions, some acetylation is observed for the nucleosome in which the major sites of H3 acetylation have been mutated to alanine (Fig. 4B). Others have observed that additional lysines on H2B can be acetylated by SAGA, although the extent of this acetylation tends to be relatively modest [19]. However, under exhaustive acetylation conditions the modification of these sites could account for acetylation signal above background.

In contrast to the assays performed under exhaustive acetylation conditions, it should be possible to determine approximate specificity constant information for the
assays performed under sub-saturating acetylation conditions (Fig. 4C & D). In these experiments significantly less enzyme is used (1:7.5 enzyme to each nucleosome) and the reaction is run for a much shorter time (10 minutes). Under these conditions, the amount of acetylation is roughly 5% of that observed under the saturating conditions. Thus, we can approximate the equilibrium concentration of unbound, unacetylated nucleosome as the initial nucleosome concentration (i.e. $[\text{nuc}_x] = [\text{nuc}_x]_o$). Further, because the initial concentrations of each unacetylated nucleosome in the trinucleosomal array are the same ($[\text{nuc}_x]_o = [\text{nuc}_y]_o$), the ratio of their concentrations is unity (i.e. $(\text{nuc}_x)/\text{nuc}_y = 1$). Since only 5% of the major acetylation sites appear to be acetylated, the ratio of the instantaneous rates for each nucleosome can also be approximated as the ratio of total acetylation per nucleosome ($v_{\text{nuc}_x}/v_{\text{nuc}_y} = (d[\text{acet}]/dt)_{\text{nuc}_x}/(d[\text{acet}]/dt)_{\text{nuc}_y} = (\Delta[\text{acet}]/\Delta t)_{\text{nuc}_x}/(\Delta[\text{acet}]/\Delta t)_{\text{nuc}_y}$). Using this information, our data suggests that nucleosomes 2 and 3 have identical specificity constants, while the specificity constant for nucleosome 1 is roughly 2-fold greater than the other nucleosomes. It should be also noted that, because of the approximations made and because not all acetylation sites within a nucleosome have identical acetylation kinetics, the specificity constants should more appropriately be referred to as apparent specificity constants.

Because the specificity constant is composed of both $k_{\text{cat}}$ and $K_M$, we cannot tell which kinetic parameter contributes to the apparent difference in specificity constant between innermost nucleosome 1 and the others. It is expected that the chemical step in acetylation does not differ between nucleosomes; making it likely that $k_{\text{cat}}$ is unaltered. This would suggest that the difference is due to altered $K_{MS}$ between nucleosomes and is consistent with an increased binding affinity for nucleosome 1. Unpublished data from
our lab suggests that SAGA has a strong affinity for free DNA, and nucleosome 1 does contain a longer stretch of DNA relative to the other nucleosomes. Additionally, a number of chromatin-interacting factors are known to show preferential binding of nucleosomes with large regions of free DNA [32, 33]. Future studies will be necessary to conclusively establish the basis for this difference in specificity.

Potential extensions of acetylation sequencing

Our acetylation sequencing assay can be extended a number of different ways. By using a more extensive range of time points, it should be possible to view the progression of nucleosome acetylation from initial establishment to completions to better understand how intermediate nucleosome acetylation states affect the course of the acetylation reaction. With more initial time points and variation in substrate concentrations, it may also be possible to perform more traditional initial rate steady state kinetics to obtain separate $k_{cat}$ and $K_M$ values for each nucleosome. To do so requires rapid quenching of acetylation, which should be possible because magnetic beads can be removed from solution relatively quickly (less than 10 seconds). Currently, our assay does not report on the distribution of acetylation within a given nucleosome, but our ability to isolate individual nucleosomes should make this feasible. Specifically, resolving the histones electrophoretically and detecting which bands are radiolabeled can determine which histones are acetylated [30]. Microsequencing of the isolated histones can determine which amino acids are acetylated and their extent [24]. Finally, we note that acetylation sequencing reactions do not require that either substrate or enzyme concentration be
below the $K_M$ of the reactions. This is because substrate nucleosomes are in direct competition with one another, and acetylation activity will simply partition between substrates according to the specificity factor equation described above.

The nucleosome acetylation sequencing strategy should also be applicable to other substrates and enzymes. Because the on-bead solid-phase ligation strategy is highly modular and efficient, it should be suitable for a wide range of oligonucleosome products and lengths, including those incorporating mononucleosomes with different histone variants, post-translational modifications, DNA sequences, and DNA lengths. Studies with different histone modifying enzymes should be feasible, assuming that the high degree of specificity and sensitivity exhibited by SAGA-mediated acetylation is demonstrated. Finally, it should be possible to study how other chromatin-associated factors, such as transcriptional activators and ATP-dependent chromatin remodeling complexes, affect the establishment of chromatin modifications.

**Acknowledgements**

This work was supported by a grant from the American Cancer Society (RSG-12-065-01-DMC).
**Fig. 1. Nucleosome acetylation sequencing strategy.** The immobilized oligonucleosomal substrate is enzymatically acetylated with radiolabeled acetyl CoA and subsequently washed to remove unincorporated acetyl CoA. Desired nucleosomes are then cleaved off of the bead using specific restriction endonucleases to yield mono-, di-, and trinucleosomes. Released fragments are counted by scintillation to quantify the extent
of nucleosome acetylation. The acetylation pattern shown is meant as a possible example of differential nucleosome acetylation.
Fig. 2. Solid-phase synthesis of bead-bound trinucleosome substrates. (A) Schematic representation of substrate synthesis. A double-stranded 19 bp biotinylated adapter with a 4 nt single-stranded overhang opposite the biotinylated end is first bound to Streptavidin-coated magnetic beads. Mononucleosomes are then sequentially ligated, allowing different kinds of nucleosomes to be linked together. Unique non-palindromic restriction endonuclease sites are present in the linker DNA between each nucleosome. (B) Characterization of reactants and products in the solid-phase synthesis of bead-bound
trinucleosome substrate. 4% native PAGE was used to resolve mononucleosome in the supernatant from the ligation steps (Lanes 2-4) and to characterize the fully ligated product liberated from the bead by digestion with BstXI (first lane). DNA-containing components were visualized by SYBR Gold staining. The mononucleosome attached to the bead first contains 80 bp more DNA than the other two mononucleosomes.
Fig. 3. Demonstration of the specificity and robustness of the acetylation sequencing assay. (A) Investigation of the specificity of radiolabeled acetyl CoA incorporation. Shown are the amounts of radioactive signal after acetylation assays run with different components. The left panel shows signal on the beads, while the right panel shows the signal in the supernatant fractions. Error bars represent standard deviations from three independent trials. (B) Characterization of the extent of nucleosome liberation, as well as
nucleosome stability. Following restriction enzyme digestion, products were analyzed by 4% native PAGE. Products of a single digestion are shown in lanes labeled with a single restriction enzyme. Lanes labeled with two restriction enzymes indicate that after the beads were treated with the first enzyme and washed, the beads were then digested with the second enzyme, and this liberated product was analyzed. The two bands in lane PflMI/BstXI are consistent with variable nucleosome positions on the same dimer DNA. (C) Comparison of nucleosome acetylation reaction efficiency on beads versus in solution. Shown is the total amount of nucleosome acetylation on the bead, where nucleosomes were either first acetylated and then bound to beads, or first bound to beads and then acetylated. Acetylation experiments were performed in triplicate. Background reactions (no SAGA) were performed once and subtracted from the acetylation bead counts.
Fig. 4. Acetylation sequencing of wild type and mutant trinucleosomes under 
saturating and sub-saturating acetylation conditions. Trinucleosomes were acetylated 
by the yeast SAGA complex and then subjected to acetylation sequencing. The 
trinucleosome substrate contained either three wild type nucleosomes (A and C) or two 
wild type nucleosomes flanking a central nucleosome in which the four major sites of 
SAGA acetylation (K9, K14, K18, K23) were mutated to alanine in the histone H3 tail. 
Assays were performed under saturating acetylation conditions (A and B) or sub-
saturating acetylation conditions (C and D). Error bars represent standard deviations from four independent trials.
REFERENCES


CHAPTER 3

ACTIVATOR AND LINKER DNA MODULATE THE HAT ACTIVITY OF YEAST SAGA

The work presented in this chapter will be submitted to a peer reviewed journal.

ABSTRACT

The Spt-Ada-Gcn5-Acetyltransferase (SAGA) family of co-activators acetylates the histone H3 tails of the nucleosomes. SAGA both maintains genome wide basal levels of acetylation and also establishes hyper-acetylated domains on targeted regions of genome, upon induction of gene transcription. To better understand how these activities are regulated, we were interested in how the interaction of SAGA with nucleosomes and chromatin-associated factors affect the activity of SAGA. Recent studies from our lab hinted at a preference of SAGA for nucleosomes harboring longer linker DNA. Expanding on those studies, we show that linker DNA both decreases the $K_M$ for nucleosomes by 2 fold and increases the $k_{cat}$ by 4 fold. Linker DNA on both sides of nucleosome is required for maximum stimulation, with differing amounts of DNA required on each side. Additionally, we show that activator proteins not only decrease the $K_M$ for nucleosomes, as expected, but also appear to increase $k_{cat}$, uncovering a novel role of transcriptional activators. To better understand how SAGA activity is stimulated, we
probed the mechanism of turnover. We show that the first turnover is very rapid, whereas
the subsequent turnovers are much slower, indicating that a post-chemical step, such as
product release, is rate limiting. Overall, our results give us a better understanding of how
SAGA works with other factors to distinguish favorable regions on chromatin from other
regions.

INTRODUCTION

The eukaryotic genome is largely wrapped into nucleosomes, the basic structural
unit of chromatin. Nucleosomes contain 147 bp DNA wrapped 1.65 times around a
histone octamer containing 2 copies each of H2A, H2B, H3 and H4 [1; 2]. The frequency
of nucleosomes is roughly every 200 ± 40 bp on an average. However, different regions
of chromatin differ in nucleosome density and regularity of spacing. Heterochromatin
regions, for example, are relatively uniformly spaced with a high density of nucleosomes
to repress transcription. Euchromatin regions, on the other hand, generally have fewer
nucleosomes on average that are more irregularly spaced, to facilitate transcription and
other cellular processes, such as DNA replication and repair.

The SAGA (Spt-Ada-Gcn5-acetyltransferase) family of co-activators is one such
factor, which works in different ways to facilitate cellular processes. In yeast, it is a 1.8
MDa multisubunit complex, which is highly conserved across all eukaryotic species [3].
It consists of at least 19 different polypeptides. The Gcn5 subunit, along with Ada2 and
Ada3 were discovered during genetic screens to identify proteins that functionally
interact with transcriptional activators such as Gal4VP16 [4; 5; 6]. Subsequent studies showed that SAGA can also directly interact with the TATA binding protein (TBP) and recruit it to initiate the formation of PIC. In that sense, SAGA acts as a traditional co-adaptor, by bridging the interactions between activator and the transcriptional machinery. Another way it can facilitate transcription is by directly acetylating histones H3 and H2B on the chromatin. The Gcn5 subunit of SAGA harbors the HAT activity, and histone acetylation can lead to de-compaction of the chromatin, destabilization of the nucleosomes, and increased nucleosome mobility, thereby facilitating transcription factor binding [7; 8].

Different subunits of SAGA harbor different domains, which allow it to interact with different aspect of chromatin. For instance, the SWIRM domain of Ada2 has been shown to bind to the DNA on the chromatin [9; 10]. Similarly, the bromodomain of Gcn5 recognizes acetylated histones and helps tether SAGA on the acetylated domains [11; 12]. This interaction seems to be specific, since the bromodomain of Spt7, another subunit, does not seem to participate in such an interaction.

We have been interested in understanding the role of different features of chromatin, and how they impart specificity towards targeting of acetylation by SAGA. In order to better understand how SAGA interacts with a nucleosome in the context of neighboring nucleosomes, we previously developed a method, called nucleosome acetylation sequencing assay, to assess the spread of acetylation on each nucleosome, in a quantitative manner [13]. Using a trinucleosome as a substrate, we found a preference for nucleosomes that contain longer linker lengths than others. In order to fully understand this effect, we further investigated the nature of interactions between SAGA and
nucleosomal linker DNA. By incorporating an activator binding site in the linker DNA, we were also able to take a quantitative approach to differentiate between activator bound targeted acetylation from untargeted global acetylation. Finally, preliminary pre-steady state kinetics shed light on a previously unknown mechanistic aspect of the HAT activity of GCN5, in context of SAGA.

**MATERIALS AND METHODS**

**Nucleosomal template preparation**

The strong positioning 601-177-1 sequence was prepared as described previously [14; 15]. The 147 bp linker-less template was generated by PCR amplifying the region using the following set of primers (include primer sequences) on the template X. The forward primer was purchased from IDT, which was modified on the 5’ end to incorporate a biotin moiety with a 15 atom spacer. Other nucleosomes harboring different lengths of linker DNA either on one or both sides were also generated using the above strategy, with different sets of primers and templates. The PCR amplified templates were cleaned up using the PCR purification strategy from Qiagen. The purified DNAs were quantitated and assessed for homogeneity on a 1% agarose gel.

**Recombinant histone purification and octamer assembly**

Wild type histones H2A, H2B, H3 and H4 from *Xenopus laevis* were expressed recombinantly and purified to homogeneity, as previously [16]. The core histones were
used to assemble the octamer, as described before [17]. The homogeneity and integrity of purified histones and assembled octamer were assessed on 18% SDS PAGE.

**Mononucleosome assembly**

Nucleosomes harboring different linker DNAs were assembled by depositing the histone octamer on different lengths of template DNAs by the process of rapid dilution, as described previously [18; 19]. The homogeneity and degree of saturation of the assembled nucleosomes were assessed on a 4% Native PAGE.

**Immobilization of nucleosomes to streptavidin coated magnetic beads**

The nucleosome containing an extra 80 bp linker harboring Gal4VP16 target motif is hereby referred to as GB1XM. The preparation of GB1XM was done as described previously [13]. Other nucleosomes containing biotinylated DNA were directly attached to the streptavidin bound magnetic beads and resuspended in 1X NEB T4 DNA ligation buffer (50 mM Tris-HCl at pH 7.5, 10 mM MgCl₂, 1 mM ATP, and 10 mM DTT) to get a final nucleosome concentration of 150 nM. The binding efficiency of each nucleosome was assessed by comparing the supernatants with known amounts of nucleosomal DNA on a 4% Native PAGE.

**ySAGA expression and purification**

Endogenous SAGA complex from *Saccharomyces cerevisiae* was purified using the TAP purification strategy, as previously [20]. The purified enzyme was quantitated by
using known amounts of recombinantly purified Gcn5, on a Western Blot. The activity of the purified enzyme was tested on synthetic H3 peptide, as before [13].

**Recombinant Gal4VP16 expression and purification**

Recombinant Gal4VP16 was purified from transformed *E.coli* Xa90 cells according to a previously published protocol, with a few modifications. Transformed colonies were grown in LB amp media to an O.D\textsubscript{600} of 0.7, and were induced with 1 mM IPTG at 37°C for 4 hours. The cells were harvested at 4000 rpm for 20 minutes. The cells were resuspended in 1X lysis buffer (20 mM HEPES, 200 mM NaCl, 20 mM βME, 10 µM Zn(Ac)\textsubscript{2}, 0.2 mM PMSF; pH 7.5) and lysed by sonication. The sample was centrifuged at 16000 rpm for 30 minutes. The soluble fraction was loaded on SP IEC column, pre-equilibrated with HEMG-100 (20 mM HEPES, 100 mM NaCl, 1 mM EDTA, 10 µM Zn(Ac)\textsubscript{2}, 10% glycerol, 5 mM MgCl\textsubscript{2}; pH 7.5) and purified over a linear gradient from HEMG-100 to HEMG-1000 (20 mM HEPES, 1 mM EDTA, 1000 mM NaCl, 10 µM Zn(Ac)\textsubscript{2}, 10% glycerol, 5 mM MgCl\textsubscript{2}; pH 7.5). The eluted fractions were analyzed on 15% SDS-PAGE, and appropriate samples were pooled and loaded on QP IEC column, pre-equilibrated with HEMG-0 (20 mM HEPES, 1 mM EDTA, 10 µM Zn(Ac)\textsubscript{2}, 10% glycerol, 5 mM MgCl\textsubscript{2}; pH 8.5) and purified over a linear gradient from HEMG-100 to HEMG-1000 (20 mM HEPES, 1 mM EDTA, 1000 mM NaCl, 10 µM Zn(Ac)\textsubscript{2}, 10% glycerol, 5 mM MgCl\textsubscript{2}; pH 8.5). Appropriate fractions were pooled and purified on a S-300, XK 16/40 column, pre-equilibrated with HEMG-100, pH 7.5. Pure fractions, as assessed on SDS PAGE, were pooled, concentrated and aliquoted in
different tubes. All samples were flash frozen in liquid nitrogen and stored in -80°C, until further use.

**EMSA using Gal4VP16**

Electrophoretic Mobility Shift Assays were performed on 2 nM GB1X DNA, with increasing concentrations of recombinant Gal4VP16 (0, 2, 4, 8, 16 and 32 nM) on a 4% Native PAGE, at 60V for 45 minutes in 0.5X TBE (composition). The gel was further stained in 150 mL 1X TBE buffer (composition) and SYBR Gold stain at RT, for 30 minutes. The dissociation constant was estimated by quantitating the activator bound and free DNA on Typhoon. To demonstrate specificity towards target loci, similar assays were also performed on a template DNA which is identical in length to GB1X DNA, but does not harbor any activator binding site.

**Activity assays to test saturation of activity in presence of Gal4VP16**

Initial rate steady state experiments were performed using 10 nM GB1XM nucleosomes and increasing concentrations of Gal4VP16, to determine the activator concentration that would result in saturation of ySAGA activity. X µL of GB1XM (resuspended in LDB) were mixed with different volumes of Gal4VP16. X µL of 2X HAT buffer were added and the samples were equilibrated at 30°C for 5 minutes. Reactions were initiated by adding X µL of ySAGA, and acetylation was allowed to proceed. 10 µL aliquots were taken out at 0’, 3’, 6’ and 9’ and added directly to eppendorf tubes containing 10 µL of WB 1. The samples were washed for a total of 4 times, at RT with WB1. They were then washed with WB 2 at 37°C for five times, with
25 minutes of incubation between each wash. The beads were finally resuspended in X µL of WB1 and added to 6/9 mL of scintillation cocktail, which were then counted in LSC.

**Steady state kinetics**

Steady state kinetics were performed on ySAGA by titrating increasing concentrations of nucleosomes. Reactions were performed under initial rate conditions, by using limiting concentrations of enzyme. Unless specified otherwise, 25 µL of bead bound substrate (in LDB) was mixed with 25 µL of 2X HAT buffer (composition) and incubated at 30°C for 5 minutes. Reactions were initiated by adding 2 nM final enzyme. 10 µL aliquots were taken at 0’, 3’, 6’ and 9’ and were directly added to eppendorf tubes containing 10 µL of wash buffer 1. Each sample was washed 4 times with WB1 at RT and five times with wash buffer 2 at 37°C, with 25 minutes of incubation between each wash. The beads were resuspended in X µL of WB1 and added to scintillation vials containing 9 mL of cocktail. All samples were counted with LSC. The CPM obtained were plotted as a function of time, for each [S]. The resultant slopes were then converted to [P] to get $k_{cat}$ and $K_M$ values.

**Single turnover experiments**

Single turnover experiments were performed on 200 nM nucleosomes by using 50 nM ySAGA in a final 1X reaction buffer (composition). Nucleosomes (in 2X HAT buffer) were equilibrated at 30°C for 5 minutes. ySAGA was then added in a final reaction volume of 50 µL, and incubated at 30°C for different time intervals. The beads
were quickly pulled to the side of the eppendorf tubes using a magnet and washed 4 times with 50 µL wash buffer 1 at RT followed by 50 µL wash buffer 2 at 37°C, five times with 25 minutes between each wash. The beads were then resuspended in 20 µL WB1 and added to a scintillation vial containing 9 mL of scintillation cocktail. All the vials were counted in the LSC. The obtained counts were converted into turnovers and plotted as a function of time.

**Data fitting**

For obtaining saturation curves, the CPM values were plotted as a function of time, and fitted to straight lines to generate slopes. These slopes were then converted to rates and plotted as a function of substrate concentration to obtain $K_M$ and $k_{cat}$ values. The data obtained by performing radioactive HAT assays were converted into either amount of product generated, number of turnovers, or rates depending on the type of experiment. All data fitting was done using the software Kaleidagraph.

**RESULTS**

**DNA flanking the nucleosome positioning sequence stimulates ySAGA.**

Acetylation sequencing from our lab suggested that extra DNA surrounding nucleosomes increases the specificity constant of SAGA toward such nucleosomes. This specificity constant could reflect a change in apparent binding affinity and/or a change in the turnover rate. However, the acetylation sequencing assay was not suited to determine the relative contributions of these factors. To better understand the source of specificity,
we developed a bead-based, initial rate, steady state acetylation assay (Fig. 1A). In this assay, a mononucleosome is immobilized on a bead, and the acetylation reaction is initiated by addition of SAGA and radiolabeled acetyl CoA. Beads were removed at specific timepoints, and the reaction was stopped and background reduced by washing. Beads were then counted to generate initial rates. A key issue that had complicated the analysis was the high level of non-specific background signal relative to the small amounts of nucleosome acetylation generated. This was overcome using multiple washes, including some with hydroxylamine. Hydroxylamine can cleave the thioester bond of acetyl CoA, and this cleaves the tritiated acetyl group from the rest of the molecule.

Using this assay with the nucleosome that showed the greatest specificity in our acetylation assay, hereby referred to as GB1XM, we observed that a reproducible saturation curve of acetylation activity could be generated (Fig. 1B, black circles). Fitting these data to the Michaelis Menten equation, we found that this substrate had a $K_M$ of 21 nM and demonstrates a $k_{cat}$ of 1 min$^{-1}$.

To determine how much the flanking DNA influenced SAGA activity, a second substrate containing only the 147 bp positioning sequence from the 601 template, with no linker DNA was generated (Fig. 1A). The kinetic behavior of this substrate is noticeably different (Fig. 1B, white circles). Its $K_M$ was 50 nM and its $k_{cat}$ was 0.25 min$^{-1}$. These kinetic constants show that addition of the flanking DNA changes both the $K_M$ and $k_{cat}$ for the substrates. Combining the effects of both $K_M$ and $k_{cat}$, allows us to calculate an overall catalytic efficiency/specificity constant ($k_{cat}/K_M$) for our substrates, and indicates an 8-fold stimulation is observed when nucleosomes containing a long linker DNA is used. To investigate the possibility that immobilizing the nucleosomes with no flanking
DNA to the magnetic beads might be contributing to this difference, we also performed a control reaction where the extent of acetylation for the on-bead substrates was compared with a substrate free in solution. As can be seen in Fig. 1C, the reaction rate was largely unchanged, ruling out the possibility that the observed kinetic differences were from bead attachment and more generally, confirming the usefulness of the bead bound approaches.

**ySAGA engages linker DNA on both sides of nucleosome.**

To map how much of the flanking DNA was necessary to stimulate SAGA activity, we investigated substrates containing different lengths of linker DNA. Initially, we looked at the effect of increasing amounts of flanking DNA on just one side (Fig. 2A). Flanking DNA as short as 40 bp results in an increase in acetylation, as seen in Figure 2A, although the increase was not equivalent to that of the GB1XM. Increasing the amount of flanking DNA to 80 bp did not increase the rate of acetylation, and the pattern remains similar across this regime.

GB1XM also includes a short 15 bp linker on the other side of the positioning sequence (Fig. 1C). We postulated that ySAGA might engage both the linker tails to achieve maximum stimulation. To test this hypothesis, we investigated a nucleosome that contained 80 bp linker on both sides, and one with 15 bp on both sides. We found that 15 bp on both sides result in a very modest increase in acetylation rate relative to the nucleosome without any flanking DNA, 147M, but much less of an increase compared to GB1XM. However, the nucleosome with 80 bp on both the sides showed a striking increase in acetylation rate relative to 147M. These data indicate that flanking DNA on both sides of the nucleosome are necessary for stimulation. Since the GB1XM only has
15 bp on one side of the nucleosome, it suggests that one side needs no more than 15 base pairs. On the other hand, having 15 bp on both does not provide much stimulation. Thus, the other side requires flanking DNA longer than 15 base pairs.

Because flanking DNA for the GB1XM nucleosome changes both the KM and kcat parameters relative to nucleosome without flanking DNA, and flanking DNA on both sides of the nucleosome stimulate acetylation activity, we were curious as to how flanking DNA individually affected the kinetic parameters. A preliminary substrate acetylation saturation curve was performed for nucleosomes with 80 bp on either one or both sides of the nucleosome (Fig. 3A).

**Transcriptional activator modulates ySAGA in potentially two different ways to ensure targeted chromatin acetylation**

SAGA is involved in the regulation of transcription of about 10% genes of a eukaryotic genome [21; 22]. A key question then is, how is it targeted to specific promoter loci to aid in inducible gene transcription? Studies have shown that one way this is achieved is by physically interacting with transcriptional activators [23; 24]. We took a quantitative approach to understand this scenario better. Towards this end, we incorporated a target binding site for the chimeric transcriptional activator Gal4VP16 in the linker DNA of GB1XM. Gal4VP16 has the DNA binding domain of the yeast Gal4 activator, and the activation domain of the Herpes Simplex Virus. To demonstrate binding activity of the protein to this target locus, we performed electrophoretic mobility gel shift assays to resolve the activator bound DNA from free DNA. The DNA template used for this purpose was the GB1X DNA harboring the activator binding site. As can be
seen in Figure 4B, increasing amounts of activator protein results in conversion to a slower migrating species. As little as 3 nM results in roughly 50% loss of free DNA, and very little free DNA is observed with 8 nM of activator. This suggests that the activator is capable of DNA binding and readily saturate binding at low concentrations of activator.

Next, we decided to perform preliminary tests to see if Gal4VP16 modulated the activity of SAGA in the context of nucleosome acetylation. We chose saturating concentrations of Gal4VP16 on GB1XM and performed nucleosomal HAT assays, as described. In the absence of activator, basal levels of acetylation were observed (Fig. 4C). However, the extent of acetylation increased to almost 10 times upon addition of activator. Based on the concentrations of substrate and activator used in this assay, we predicted an increase of about 3 fold in the acetylation if activator modulated $K_M$ alone. An increase of 10 fold suggests that activator not only results in tighter binding of SAGA on target loci containing nucleosomes, but also affects the turnover rate. In order to test that more exhaustively, we first performed activity saturation experiments on 10 nM GB1XM, by using increasing concentrations of activator, until we no longer observed any further increase in product formation (Fig. 4D). Now that we have results from these studies, we are set/ready to tease out the effect of activator on GB1XM acetylation. Steady state experiments will be performed on GB1XM nucleosomes, similar to what we did in Figure 1C, in the presence of saturating concentrations of Gal4VP16. The resultant $K_M$ and $k_{cat}$ values that will be obtained will then be compared with those obtained in Figure 1C, in order to get a mechanistic view of the interactions taking place between activator and SAGA.
Pre-steady state kinetics of ySAGA experiments suggest that a mechanistic step following acetyl transfer is rate limiting

To better understand how acetylation turnover can be stimulated, we wanted to determine the rate limiting step in enzymatic turnover. In steady state kinetics, $k_{cat}$ represents the rate limiting unimolecular step and could be the chemical step, acetyl transfer, or could be a subsequent step, such as product release. For our assay, since acetylated histone is detectable on the bead even without product release, it gave us an opportunity to potentially directly measure the rate of the chemical step for the enzymatic turnover. Using saturating concentration of enzyme and substrate, we followed the time course of acetylated histone production (Fig. 5A). These preliminary results show that there is a rapid burst phase in GB1XM acetylation, where, atleast 3 turnovers are observed within 10 seconds. This is then followed by a linear slower phase, that has a rate constant comparable to what was seen in our steady state experiments. These data suggest that the first enzyme-mediated acetyl transfer is very fast, to the point that we cannot currently resolve its rate constant. This is then followed by another unimolecular mechanistic step that is much slower, and becomes rate limiting for subsequent turnovers. Thus our data suggests that the acetyl transfer step is not rate limiting, and that the rate limiting step in SAGA-mediated nucleosome acetylation occurs after the chemical step.

Model of ySAGA towards inducible eukaryotic gene transcription

Based on our results, we can begin to understand how the SAGA complex might function *in vivo*. Under non inducing conditions (Fig. 6A), it can be predicted that regions harboring longer stretches of DNA between nucleosomes would be acetylation hot spots
by virtue of providing a better platform for tethering of SAGA, and in turn stimulating the generation of acetylated nucleosomes. On the other hand, regions containing shorter stretches would be less optimal substrates, and would also result in SAGA being trapped at such locations, since product release becomes limiting without linker DNA. Thus, one can imagine high occupancy of SAGA on such spots, without a corresponding increase in acetylation levels. As stress conditions arrive, the transcriptional activators bound to the UAS of target genes would physically recruit SAGA to the promoters of the stress induced genes (Fig. 6B). This interaction will result in rapid acetylation of nearby nucleosomes, thereby generating hyperacetylated domains. It is possible that other factors in the cell, such as remodeling complexes might act synergistically or antagonistically in order to achieve the desired outcome. It would be interesting to see how these factors work together.

**DISCUSSION**

We developed and utilized quantitative approaches in this study to understand the mechanism of acetylation mediated by the full SAGA complex, on nucleosomal substrates. By generating nucleosomes that harbored different lengths of flanking linker DNA, we show that the length of this DNA differentially regulates the activity of ySAGA. Long stretches of DNA stimulate this activity by not only creating a platform that results in tighter binding, but also modulating the catalytic activity towards more efficient turnover. We also show that transcription activators work in more than just
potentiating tighter binding of SAGA on target loci. We show for the first time, that these activators also stimulate SAGA for faster turnover, by affecting the catalytic efficiency. By optimizing \textit{in vitro} acetyltransferase assays, we also demonstrate for the first time that SAGA undergoes burst phase kinetics, indicating that the first acetyltransferase step is very rapid. Subsequent step, such as product release, is rate limiting which results in a slower, linear second phase.

\textit{Potential acetylation hotspots on genome}

The eukaryotic genome has a highly dynamic chromatin landscape that evolves rapidly in response to changes in extracellular and intracellular conditions [25; 26]. One of the features of this landscape is the spacing between tandem nucleosomes. For example, nucleosomes in the heterochromatin regions consist of regularly phased arrays, which results in dense packing of nucleosomes to create transcriptionally repressive domains [27]. Studies have shown that members of ISWI family of remodelers, such as ACF, are actively involved in creating such domains [28]. ACF has been shown to be a DNA ruler, and is modulated by the lengths of free DNA flanking the core positioning nucleosomal DNA. By utilizing the energy obtained from hydrolysis of ATP, ACF slides the nucleosomes to centrally positioned nucleosomes [29].

Recently published studies from our lab hinted that ySAGA might have preferences for longer stretches of linker DNA [13]. To explore further into this scenario required a sensitive and reproducible assay, which is conducive to steady state kinetics using nucleosomes as substrates. Towards this end, we developed a bead based assay
which allowed us to immobilize biotinylated nucleosomes to streptavidin conjugated magnetic beads. Optimizations were done to achieve high signals in a reproducible fashion, and more importantly, low background levels. Using this assay, we were able to test the effect of linker DNA on the activity of ySAGA, and found that indeed, longer linker DNA containing nucleosomes are better targets over linker-less nucleosomes. Our results suggest that nucleosomes that have long flanking linker would be hot spots for acetylation by SAGA, in the eukaryotic genome. Note that SAGA binds linker-less and sub-optimal nucleosomes with nanomolar affinities too, thereby explaining how it is able to establish untargeted basal levels of acetylation genome wide. We argue that nucleosomes with longer linkers contribute to acetylation hotspots and result in hyper-acetylation domains, above from basal levels of acetylation.

Targeting of SAGA at specific loci by transcription activators uncovers a novel effect of this interaction

Gcn5 was discovered along with Ada2 and Ada3 during genetic screens to identify factors which interacted with transcriptional activators. Subsequent studies confirmed that the full SAGA complex is recruited at different promoters by the respective activator proteins. Additional subunits, such as Tra1 have been shown to be involved in this interaction [30; 31]. These responses are often quite rapid in order to achieve homeostasis, and so far, qualitative studies have been performed to understand this interaction. We were interested in taking a quantitative approach to gain further insights into the mechanism of this interaction. Our preliminary results suggest that
activators mediate this by physically recruiting SAGA to target loci, in excellent agreement with previous studies. We also discovered a previously unknown mechanism. The physical interaction between activator and SAGA also modulates the catalytic activity to result in much faster turnover, almost as much as 10 times. This explains how the cell is able to quickly respond to the upstream stimuli and turn inducible genes on. Past studies have shown that upon targeted recruitment, SAGA establishes a very restricted region of hyperacetylation in vivo, spanning roughly 2-3 nucleosomes, in contrast to other HAT complexes such as NuA4 [32]. It would be interesting to watch the extent of spread of this acetylation in vitro, quantitatively. It would also be interesting to uncover the source of this stimulation. Mutant SAGA complexes and mutant activator proteins will shed a light on this aspect, to further our understanding.

_Burst phase kinetics of SAGA hint at rate limiting product release_

Our bead based assay allowed us to perform single turnover kinetics which led us to discover an unprecedented feature of the acetyltransferase reaction mechanism. The first chemical step of acetyl group transfer to the histone tail is quite rapid and results in a burst phase. We observed as many as 3 turnovers in under 10 seconds under these conditions. In contrast, subsequent cycles were much slower, and linear, suggesting a step after the chemical step, such as product dissociation, or conformational change, might be the rate limiting step. This might also suggest that in vivo, SAGA might get stuck on some loci because of this. The functional ramifications of this would be interesting. The source of this burst phase and the nature of the rate limiting step are currently under
investigation. In our current set up of these assays, it is very difficult for us to go less than 10 seconds in a reproducible fashion. We are trying to optimize our reaction conditions that would allow us to chase the burst phase in a more quantitative manner. We are also in the process of repeating above experiments to get a reproducible rate constant. The rate constant for the slower, linear phase obtained from our single turnover experiments was found to be very similar to the one obtained from our steady state experiments, further validating our results. The next set of experiments will be performed in similar manner as above, with different concentrations of SAGA, to show a linear dependence of amplitude as a function of enzyme concentration. We will also be employing 147M nucleosomes to investigate if the source of this burst phase is due to longer flanking linker.

Overall, through our quantitative approaches, we now have a better picture of how SAGA maintains some substrate specificity and also how it is able to work along with other factors in the cell to ensure the ultimate success of the organism.

**AKNOWLEDGEMENTS**

We would like to thank Dr. Steve Triezenberg for a generous gift of the pJL2A and pJL2S plasmids, and Dr. Blaine Bartholomew for *E. coli* Xa90 cells. This work was supported by the funding from American Cancer Society.
REFERENCES


Figure 1. Linker DNA stimulates ySAGA mediated nucleosomal acetylation. (A)
Schematic representation of the nucleosomes (left panel) and assay (right panel) used.
The nucleosome containing the longer linker DNA and activator binding site is referred
to as GB1XM and contains 95 base pairs on one side of the core nucleosome, and 15 on
the other. The nucleosome containing only the 147 bp positioning sequence is referred to
as 147M and contains no base pairs flanking the core nucleosome. (B) Steady state
kinetics analysis to determine the mechanistic factors contributing to the difference in
observed acetylation. (C) Comparison of acetylation for immobilized nucleosomes and nucleosomes free in solution.
Figure 2. Determination of the features of DNA flanking the nucleosome that stimulate ySAGA. (A) Comparison of the acetylation effects for substrates containing different lengths of flanking DNA on a single side of the nucleosome. Shown are a schematic representation of the substrates tested (above panel) and a bar graph showing the associated extents of acetylation (lower panel). (B) The effects of proximity of the linker DNA to the bead were tested by constructing templates harboring the linker on the
opposite side (upper panel). The extents of acetylation are shown below (lower panel).

(C) Comparison of the acetylation effects of substrates containing different lengths flanking DNA on both sides of the nucleosome. The upper panel shows a schematic of the substrates tested and the lower panel compares the associated extents of acetylation.
Figure 3. (A) Steady state kinetics performed on nucleosomes that harbor 80 bp linker on one or both sides of the positioning sequence. (B) A table summarizing the kinetics constants obtained using different nucleosomes.
Figure 4. Effect of activator protein on nucleosomal acetylation. (A) Assessment of the homogeneity of purified Gal4VP16 on 15% SDS PAGE. (B) Analysis of activator binding to DNA. Electrophoretic mobility gel shift assays were performed using GB1X DNA and increasing amounts of recombinantly purified Gal4VP16 were titrated with 2 nM DNA. Activator-DNA complex was resolved from the unbound DNA by 4% Native PAGE and stained with SYBR-Gold (C) Extent of GB1XM nucleosome acetylation with or without saturating amounts of Gal4VP16 activator. (D) Pilot nucleosome activation saturation study.
Figure 5. Mechanistic analysis of acetylation turnover, $k_{\text{cat}}$. (A) Shown is a plot of product turnovers as a function of time under conditions where enzyme and substrate concentrations significantly above $K_M$ (200 nM GB1XM and 50 nM ySAGA, respectively). Product turnover is calculated from the concentration of acetyl-lysine produced divided by the concentration of SAGA. Pre-steady state acetylation occurs before the first time point, and the post-steady state behavior is fit to the linear equation shown.
Figure 6. A model of potential interactions between SAGA, chromatin and transcription activators to regulate eukaryotic gene expression.
CHAPTER 4

ADDITIONAL DATA AND FUTURE STUDIES

We have been working towards gaining a more comprehensive picture of the mechanisms employed by the SAGA complex, either by itself, or in conjunction with additional factors, to regulate eukaryotic gene expression. This section describes our ongoing studies, which will be optimized further to accomplish our overall goal.

Exhaustive mapping of linker requirements

We are currently mapping the features of linker DNA that modulate the HAT activity of the SAGA complex. We have already designed an exhaustive library of nucleosomes differing in lengths of linker on both sides of the core positioning sequence. The next thing to do is to perform standard HAT assays under conditions mentioned before, and compare the extents of acetylation across the entire panel. We aim to obtain information such as length, orientation and sequence preferences of SAGA.

Using minimal enzyme and minimal substrate systems

Previous studies showed that a subcomplex containing Gcn5, Ada2 and Ada3 can recapitulate the HAT activity of the full complex. We started by comparing the kinetic
properties of the full SAGA complex with the minimal SAGA complex, which we call as the subcomplex, on synthetic H3 peptides. We were interested in seeing if modules or domains outside the HAT module potentially regulates the acetyltransferase activity of SAGA. Figure 1 shows steady state saturation curves obtained by using WT SAGA and WT subcomplex on increasing concentrations of the H3 peptide.

Using minimal enzyme on nucleosomal substrates

With the beads based assay optimized to perform steady state kinetics, a wealth of information can be generated by using different forms of SAGA (WT and mutants), and by changing the characteristics of the nucleosome.

The significant differences observed between SAGA and subcomplex on peptides prompted us to test the system in the other way too, by using subcomplex on mononucleosomes. We were curious to see if similar differences might be seen using chromatin substrates. Fig 2 shows our preliminary results obtained.

Future experiments

More applications of the bead-based assay

The bead-based assay can also be utilized to gain more insights to explain biologically relevant scenarios. For instance, we can incorporate histone variants, such as H2AZ, which is usually found in the promoter regions of actively transcribing genes. It would be interesting to see if the kinetic parameters obtained by using such a substrate
can explain locus specific presence. Similarly, we can incorporate H3 histone which is acetylated on all the four primary sites targeted by SAGA (14, 18, 9, 23), or mutated into alanine to so it can’t be further acetylated. It would be interesting to see how these substrates would change the kinetics of turnover.

**More applications of the nucleosome acetylation sequencing strategy**

Using the nucleosome acetylation sequencing strategy, we can incorporate different kinds of nucleosomes in tandem and observe how the spread of acetylation across these systems differ. For instance, we can incorporate histone variants on one of the nucleosomes, or build a system that has different post-translational modifications on the neighboring nucleosomes. We can also vary the spacing between nucleosomes to see how things change. The promoter region of active genes, for instance, has a nucleosome free region flanked by nucleosomes harboring H2AZ, instead of the core H2A. It would give us clues to how the modification landscape on chromatin evolves with respect to neighboring environment.

**Discovering domains/modules involved in regulating HAT activity of SAGA**

We would ultimately like to map the modules/domains within full SAGA complex that are responsible for above activities. For instance, the SANT domain of Ada2 has been shown previously to have the ability to bind naked DNA. An obvious
direction would be to create a SANT Δ yeast strain and purify SAGA from it. It would be interesting to see if that abolishes the preference towards longer linker DNA, that we observed for the WT strain. Results from this study would shed light on the role of SANT domain and other potential domains involved. Similarly, we would also like to perform steady state experiments using Gcn5Δ and bromodomain Δ strains to understand the source of burst phase and also preferences towards linker DNA.

The effect of different modules on the activity of the HAT module is currently being investigated by another graduate student in the lab, Sannie Olson. She generated some knockout strains and her preliminary results. She is currently in the process of generating more knockouts, to get a more comprehensive picture. Current studies involve histone peptides as substrates, and filter binding assays. I will be performing bead-based assays on nucleosomal substrates to study differences observed. The work from this collaboration will be published as a paper in a peer reviewed journal, and I will be an author on it.

**Investigating the source of burst kinetics**

We are interested in investigating the source of the burst kinetics that were observed by using GB1X nucleosomes. Our first experiment will be to perform similar experiments using the linkerless 147 bp nucleosome, to test if the linker DNA contributes to the rapid first turnover. In addition, we will also perform above experiments by adding Gal4VP16 in our reactions, to see if the presence of activator changes the kinetic properties under single turnover conditions.
Ultimately, we would also like to optimize our assay to be able to perform true pre-steady state experiments. Towards this end, we will have to try a quenching mechanism that does not disrupt the structural integrity of the nucleosome, and would allow us to perform these experiments in the millisecond phase, to extract the true burst constant.

**Gaining better understanding of activator effect**

The current studies probing into the mechanism of transcriptional activators will include performing steady state experiments under initial rate conditions, by using increasing concentrations of GB1X nucleosomes and saturating concentrations of Gal4VP16, similar to Figure 1C in chapter 1. It would be interesting to see the effect of activator on the kinetic properties of SAGA. We will also perform similar experiments on 147 bp nucleosomes, under conditions that allows specific binding of activator to target sequence. Thus, it would be interesting to see if activator can mediate any effect independent of target locus binding. Any difference observed will suggest new mechanisms, and thus will further our understanding. These effects can also be mapped by utilizing mutant activator proteins, to tease out the role of different domains in the activator protein. Towards that end, it would also be useful to employ different forms of SAGA complex, such as subcomplex and other deletion mutants that result in loss of a subunit or a module. Previous studies have shown that Tra1 is an activator interacting
subunit of SAGA, but it might not be the only one. Above results will uncover additional activator interacting domains within the full complex.

We also generated different activator knockout yeast strains. As shown in Figure 3, the knockout strains grow poorly under respective stress conditions. These activators require Gcn5 \textit{in vivo}, to initiate gene transcription. These knockout strains are valuable tools for performing in vivo studies, such as ChIP to monitor changes in SAGA occupancy on target loci.
Figure 1. Comparison of kinetic properties between endogenous SAGA and recombinant subcomplex. (A) The recombinantly expressed subcomplex was purified to near homogeneity. (B) Steady state kinetics were performed using full SAGA and the subcomplex, on synthetic H3 peptides. (C) Silver stain analysis of purified SAGA complexes from different knockout yeast strains. (D) Summary of kinetic parameters obtained from different Gcn5 containing complexes.
Figure 2. Reaction progress of recombinant subcomplex on GB1X nucleosomes.

Shown is a plot of turnovers as a function of time. 2.5 nM subcomplex was used to initiate the HAT reaction on 40 nM nucleosomes.
Figure 3. Plate growth assays of different yeast strains. Different knockout and mutant strains were generated and grown under different stress conditions. (A) Serial dilutions of WT, Gcn5Δ, and Gcn5 BD mutants were grown under the absence of tryptophan (left panel). When the corresponding activator, Gcn4, was deleted, the growth phenotype is worse (right panel), validating activator KO strains as useful tools to perform in vivo studies. (B) Serial dilutions of WT, Gcn5Δ, and Gcn5 BD mutants were grown under the absence of galactose (left panel). Deletion of Gal4 results in worse phenotype (right panel).
CHAPTER 5

CONCLUSIONS

The low resolution crystal structure of a nucleosome, and identification of p55 from Tetrahymena as a histone acetylating enzyme remain the two most crucial discoveries in the field of chromatin, in my opinion. These studies provided a great impetus to understand chromatin, beyond acting as a store of genetic material. Soon it was appreciated that it indeed is a complex and dynamic structure that is essentially involved in a plethora of cellular processes such as DNA replication, repair, transcription, stress survival and development. More and more processes are now being linked with chromatin, such as metabolism and circadian rhythm, which makes sense since gene expression regulates the overall success of an organism.

Soon after demonstrating the HAT activity of p55, Gcn5 it’s yeast ortholog was also shown to possess this activity, suggesting lysine acetylation of histones is a highly conserved mechanism to regulate chromatin, across species. Higher eukaryotes, from Drosophila to humans, were subsequently shown to harbor Gcn5, establishing this fact. Elegant in vitro studies were performed on different histone peptides by using recombinant HAT domain of Gcn5 and it was shown that Gcn5 targets lysine 14 on histone H3. However, unexpected results showed that it failed to acetylate core histones and mononucleosomes, the basic structural unit of chromatin. This led to purification of
Gcn5 from biological sources, such as *Saccharomyces cerevisiae*, and shown that Gcn5 is a part of large multisubunit complexes, that were chromatographically different.

Subsequent characterization using various biochemical and structural approaches proved that despite having large overlaps in subunit compositions, these complexes were indeed distinct from each other by virtue of regulating distinct/unique cellular processes. This also explained why recombinant Gcn5 alone could not acetylate histones or nucleosomes. Gcn5’s specificity is expanded in the context of these complexes, and it can now target all forms of chromatin. It can also acetylate H3K14, 18, 9 and 23, and also H2B, to a lesser extent. Other enzymes were also shown to work differently in the context of their respective complexes. Context is a very important theme in chromatin biology. While it is certainly important to understand the role of any protein or histone modification in isolation, and quite often is a necessary step towards understanding the bigger picture, it should always be kept in mind that *in vivo*, this is often not the case and thus, the next step is to study the roles of these factors in combination with other factors.

Towards this end, our lab has been investigating the role of different aspects of chromatin, i.e., the role of acetylation, methylation and the nature of higher order chromatin structures to gain a more comprehensive view of how different factors work in a concerted fashion to regulate biology. My attempt during these past 6 years has been to further this knowledge and advance the field. Specifically, I have been working towards understanding the function of the SAGA complex, one of the three Gcn5 containing complexes, from *Saccharomyces cerevisiae*. As mentioned before, SAGA is highly conserved across all eukaryotes, with large overlaps in subunit composition and
functions. Thus, the knowledge gained out of this project is predicted to help us extrapolate the information to higher eukaryotes such as humans, in an educated manner. SAGA harbors two different enzymatic functions, the acetyltransferase activity contained in Gcn5, and a ubiquitin protease activity harbored by Ubp8. These activities are present in distinct modules within the complex, and are often independent.

My work has been aimed towards characterizing the HAT activity of Gcn5 in the context of full complex towards chromatin substrates, such as nucleosomes using *in vitro* approaches. Denu and coworkers previously characterized the HAT domain of recombinantly purified Gcn5 on peptide substrates, also using *in vitro* approaches. Elegant experiments established that Gcn5 forms a ternary complex with acetyl CoA and histone and catalyzes a direct transfer of the acetyl group to the histone. This reaction follows an ordered Bi-Bi mechanism whereby acetyl CoA binds the catalytic domain of the enzyme before histone. Following catalysis, acetylated histone leaves first followed by CoA.

These experiments were pioneering towards a better understanding of chromatin acetylation by histone acetyltransferases. However, the studies utilized highly simplified substrates. For instance, only the HAT domain of Gcn5 was used as the enzyme. The substrates were simple too, being short peptides from the amino terminal tails from histones. *In vivo* scenario is much more complex. As mentioned before, Gcn5 does not act alone, but rather in context of large multisubunit complexes. Also, it targets chromatin, which constitutes of nucleosomes present in an array fashion. Therefore, we started by purifying endogenous SAGA complex from *Saccharomyces cerevisiae*. For our substrates, we employed the well characterized 601 template, which is a strong
nucleosome positioning sequence. We realize that this is not a naturally occurring sequence and in vivo, nucleosomes of all sorts of stabilities are observed. However, for our in vitro studies, it is important that the substrate used is homogeneous and stable so that we get reliable and reproducible results.

The first challenge that we tackled was to develop a reliable, reproducible and a sensitive assay to be able to perform kinetics using nucleosomes as a substrate. Filter binding assays offer great potential for peptide substrates, which is routinely used in our lab. However, these peptides are usually positively charged due to the basic nature of histones. Also, they are quite stable and thus spotting them on filters doesn’t compromise their integrity. Nucleosomes, on the other hand, are overall negatively charged and quite fragile. With these caveats in mind, we tried using filter binding in our studies, but found that such an approach is not well suited for substrates such as nucleosomes. This necessitated development of an alternative method.

The work described in chapter 2 was my contribution towards it, in collaboration with Dr. Melissa Blacketer in the lab. We turned our efforts towards developing a bead based assay, wherein the nucleosomal substrates were designed to incorporate a biotin moiety, which would allow us to immobilize them to streptavidin conjugated magnetic beads. An added advantage of developing such a strategy was the fact that multiple nucleosomes can be ligated in tandem, thus allowing us to study a specific nucleosome, in the context of neighboring nucleosomes. We showed that attaching nucleosomes to the beads does not compromise the ability of ySAGA to acetylate the target residues. The efficiency of the reaction was similar to one where substrates are present in solution. We also showed that substrates ligated with 100% efficiency and were stable throughout the
assay period. By incorporating different endonuclease sites in the linker DNA, we were also able to liberate the nucleosomes into solution in a reproducible and efficient manner, thereby allowing us to do quantitative studies. We call this strategy as nucleosome acetylation sequencing strategy. We were interested in studying the patterns of acetylation established across multiple nucleosomes. Using a trinucleosome substrate, we started comparing the extent of acetylation across individual nucleosomes, and found that under saturating acetylation conditions, all nucleosomes were targeted to similar extents, as long as they harbor wild type, unacetylated targets. Differences started to unravel under limiting conditions, and interesting patterns were observed. We observed that when one of the nucleosomes was flanked by a much longer linker DNA than others, it was a preferred target over nucleosomes flanked with shorter linkers. This got us intrigued and we decided to pursue it further, which is described in more detail in chapter 3. In summary, our nucleosome sequencing strategy is highly modular, reproducible and sensitive. It can also be easily adapted to other systems to understand the working of other chromatin modifying complexes.

The results obtained from our work in chapter 2 got us intrigued to pursue the role of linker DNA on SAGA mediated nucleosome acetylation. We decided to use mononucleosomes for these studies. A new challenge now was to optimize the above bead based strategy to make it suitable for steady state kinetics. Once that was accomplished, we designed nucleosomes that were flanked with different lengths of linker DNA and attached all of them to the streptavidin coated magnetic beads. Steady state and initial rate reactions were performed on these substrates and we found that the length of the linker DNA has a profound effect on the activity of ySAGA. Nucleosomes
with longer linkers, and linkers on both sides of positioning sequence are better targets. This stimulation is not only due to better binding, as one would expect, but linker DNA modulates the $k_{cat}$ too, thereby increasing the overall turnover efficiency. Further, it appears that ySAGA engages not only a long linker, but also linker DNA on both sides of the core positioning sequence to achieve maximum potential. We are currently in the process of investigating this in an exhaustive fashion to gain a better understanding of this preference.

Another aspect that has been of interest to us is to understand the mechanism of targeted histone acetylation by SAGA. Only about 10% of yeast genes are under the regulation of SAGA. How is this specificity maintained. Several models have been suggested, and physical recruitment by a transcriptional activator has been most studied. It is thought that the activation domain of an activator directly interacts with SAGA, to recruit it to specific loci on the genome. To understand this effect in a more quantitative manner, we incorporated binding site of a well studied activator, Gal4VP16, in the linker DNA of one of the substrates. Preliminary studies suggest that activator stimulates nucleosomal acetylation beyond linker DNA, in potentially two ways, by affecting the binding affinity and by modulating the turnover rate. This might explain how the cell is able to quickly respond to environmental and cellular stresses. These studies are very preliminary at this stage, and we are currently working towards gaining better insights into the mechanistic aspects.

Finally, we are also characterizing the enzyme using pre-steady state kinetic approaches. Using GB1XM nucleosomes as substrates, we found that ySAGA undergoes a burst phase, which is suggestive of a rapid chemical step. It is then followed by a
slower, linear phase which we captured earlier during our steady state experiments. So far, we believe that a step subsequent to catalysis, such as product release, is rate limiting. We are also working towards performing more exhaustive kinetic experiments to understand the source of this burst phase.
APPENDIX

EXPRESSION AND PURIFICATION OF HISTONE H3 PROTEINS CONTAINING MULTIPLE SITES OF LYSINE ACETYLATION USING NONSENSE SUPPRESSION

Isaac A. Young\textsuperscript{a}, Chitvan Mittal\textsuperscript{a}, and Michael A. Shogren-Knaak\textsuperscript{a,*}

\textsuperscript{a} Roy J. Carver Department of Biochemistry, Biophysics and Molecular Biology, Iowa State University, 1210 Molecular Biology Building, Ames, Iowa, USA 50011

* Corresponding author: knaak@iastate.edu
Abstract

Lysine acetylation is a common post-translational modification, which is especially prevalent in histone proteins in chromatin. A number of strategies exist for generating histone proteins containing lysine acetylation, but an especially attractive approach is to genetically encode acetyl-lysine residues using nonsense suppression. This strategy has been successfully applied to single sites of histone acetylation. However, because histone acetylation can often occur at multiple sites simultaneously, we were interested in determining whether this approach could be extended. Here we show that we can express histone H3 proteins that incorporate up to four sites of lysine acetylation on the histone tail. Because the amount of expressed multi-acetylated histone is reduced relative to the wild type, a purification strategy involving affinity purification and ion exchange chromatography was optimized. This expression and purification strategy ultimately generates H3 histone uniformly acetylated at the desired position at levels and purity sufficient to assemble histone octamers. Histone octamers containing four sites of lysine acetylation were assembled into mononucleosomes and enzymatic assays confirmed that this acetylation largely blocks further acetylation by the yeast SAGA acetyltransferase complex.
**Introduction**

Lysine acetylation is a pervasive post-translational protein modification, where approximately 3,600 sites of lysine acetylation in over 1,700 different proteins have been identified through proteomic studies of human cells [1]. This modification helps to regulate proteins with a wide range of functions, including those involved in metabolism [2], cell structure [3], and cell signaling [4].

Histone proteins that comprise chromatin constitute an especially large class of acetylation targets. Histone proteins H2A, H2B, H3 and H4, are highly basic proteins that contain a high density of lysine residues. Two copies of each of these proteins can form a histone octamer, around which DNA can wrap to form a nucleosome [5], the fundamental structural unit of chromatin. In nucleosomes, the bulk of the histone sequence is contained within the wraps of the DNA in the globular histone domain. However, less structured regions of each histone extend past the DNA to form histone “tails.” At least thirty different sites of lysine acetylation have been identified in histones, with multiple sites identified in each [6]. Modifications occur both in the globular and tail regions of the histone, although a significantly greater number of sites are present in the histone tails.

One strategy, to better understand how histone acetylation affects chromatin structure and function, is to reconstitute chromatin *in vitro* using acetylated histones. In this strategy, the main challenge is to obtain histones that contain lysine acetylation at the desired sites. Direct purification of uniformly acetylated histones is generally impractical, because the heterogeneity of acetylation sites coupled with their similar physical properties makes isolation difficult. Enzymatic acetylation of histones has had some success. However, ensuring that only desired lysines are acetylated, and are acetylated fully, is often not easy
As an alternative, we have previously adapted the strategy of native chemical ligation to incorporate a range of histone modifications in the H3 and H4 tails [8, 9]. In this strategy, a histone tail peptide containing desired modifications is generated by standard solid-phase peptide synthesis strategies, and then ligated to a recombinantly expressed histone fragment corresponding to the remaining sequence. This strategy yields uniformly modified, full-length histone in reasonable yields. However, because peptide synthesis is relatively costly, and the steps involved are laborious, we were interested in exploring alternative strategies for incorporating different patterns of lysine acetylation into the histone tails.

A promising strategy that has recently been applied to generating acetylated histone is nonsense suppression expression. In nonsense suppression expression [10], the idea is to genetically encode for a non-standard amino acid that becomes incorporated during translation. In its most common form, the non-standard amino acid is introduced into a specific site in a protein by first mutating its sequence to an amber stop codon. By including a modified tRNA that contains an amber anticodon and a site for attaching a desired nonstandard amino acid, as well as a tRNA synthetase that can charge the tRNA with the desired amino acid, the nonstandard amino acid becomes incorporated into the protein during translation. Chin and coworkers showed that a modified pyrrolysyl-tRNA synthetase derived from M. barkeri can efficiently charge a tRNA with acetyl-lysine, and that this residue can be incorporated into histone H3 with good yields [11]. With this strategy acetyl-lysine was incorporated at multiple sites within the H3 histone. However, in all cases, each histone contained only a single acetyl-lysine.

In vivo, acetylation of multiple sites within the same histone is common [12], where the pattern of this acetylation can change the binding of chromatin-associated proteins or
directly change chromatin structure and stability. For examples, dual bromodomains found in a number of nuclear proteins recognize and bind to specific pairs of lysine acetylation, thereby targeting these proteins to specific chromatin regions [13]. Similarly, because lysine acetylation reduces the total change on histone tails, multiple acetylations can work synergistically to reduce the folding of chromatin into 30 nm fibers and reduce its stability [14, 15]. Because of our interest in how different combinations of lysine acetylation can modulate the structure and function of chromatin, we were interested in exploring to what extent this strategy could be extended to incorporating multiple sites of acetylation within a single H3 histone.

**Materials and Methods**

*Plasmid Production*

*Xenopus laevis* H3 histone in a pET3c expression plasmid [16] was mutated to incorporate 1-4 amber codons by sequential application of Quikchange mutagenesis (Strategene). The mutated H3 histone ORFs were then cloned into a pCDF-PylT-H3K14amb plasmid to replace the single amber codon containing H3 histone [11].

*Expression of acetylated histone H3*

Expression methods were adapted from those previously published [11]. BL21 DE3 cells were co-transformed with the pAcKRS-3 plasmid containing the tRNA synthetase and the pCDF-PylT plasmid containing both the H3 histone with the desired number of amber codons and the amber suppressor tRNA. Transformed cells were grown overnight at 37°C in 50 mL standard Luria Broth media under selection of 50 µg/ml kanamycin and 50 µg/ml
spectinomycin. 250 mL of prewarmed non-standard 2xYT broth (2xYT-KS: 0.5% w/v sodium chloride, 1% w/v yeast extract, 2% w/v tryptone), containing 50 µg/ml kanamycin and 50 µg/ml spectinomycin, was inoculated with overnight culture to a final OD₆₀₀ of 0.08 at 37°C. The dilute 2xYT-KS culture was grown to 0.7 OD₆₀₀. N-ε-Acetyl-L-lysine (≥98% pure, Novabiochem) was then added to a final concentration of 10, 20, 40, or 80 mM. Nicotinamide (NAM) was added to a final concentration of 20 mM. After 30 minutes, protein expression was induced with the addition of 0.5 mM IPTG. After 4 hours, 250 mL cultures of cells were pelleted and resuspended in 30-ml of Wash Buffer (WB: pH 7.5, 50 mM Tris-HCl, 100 mM NaCl, 1 mM Na-EDTA, 1 mM benzamidine, 1 mM DTT, 20 mM NAM) then frozen at -80°C overnight.

Purification of acetylated histone H3

Inclusion bodies from 250 ml of culture were isolated as previously described [17]. The insoluble final pellet containing histone was macerated with 0.25 mL DMSO and then thoroughly resuspended in 15 mL unfolding buffer (UB: 6M guanidine hydrochloride, 20 mM Tris hydrochloride, 5 mM β mercaptoethanol, pH 8.0) and stirred at room temp for 1 hour. The mixture was clarified via centrifugation (12,000 RCF, 10’). The supernatant was added to 2.5 mL of pre-equilibrated and drained Ni²⁺-NTA Bead Resin (Qiagen) in a 30-ml disposable column and stirred at room temperature for 1 hour. The column was drained and the resin washed two times with 30-ml Guanidine Wash Buffer (GWB; 6M guanidine hydrochloride, 100 mM monosodium phosphate, 5 mM β-mercaptoethanol, pH 6.3). The resin was then washed with 100 mL 1X TEV Cleavage buffer (50 mM Tris hydrochloride, 50 mM imidazole, 0.5 mM EDTA, 5 mM β mercaptoethanol, pH 8.0). The resin was
resuspended in 23 mL 1X TEV cleavage buffer and 10,000 units of TURBOTEV protease (Nacalai USA) were added. Digestion was carried out with nutation for 16 hours at room temperature. The column was then drained and resuspended in 15 mL UB and stirred at RT for 1 hour. UB eluent containing histones was dialyzed into SAU buffer (7M urea, 20 mM sodium acetate, 1 mM DTT, 1 mM sodium EDTA, pH 5.2). The histone sample was then purified via cation exchange on a salt gradient from 0 mM to 600 mM NaCl over 225 mL using a HiTrap SP HP 5 mL column (GE Life Sciences). Histones eluted in a highly pure state and were concentrated and desalted using Sep-Pak C-8 reverse phase purification columns (Waters) before lyophylization. Purified acetylated histone H3 was checked via MALDI-TOF analysis to verify the correct number of acetylations were present (John Leszyk at University of Massachusetts Medical School). Histone was quantified using densitometry against standard curves of H3-tetra-alanine containing histone.

Assembly of Mononucleosomes

Wild type histones and the tetra-alanine H3 histone were expressed, purified, and quantified according to literature protocols [17]. 177 bp DNA templates containing the 601 strong positioning sequence [18] and BglII non-palindromic sticky ends were prepared as previously described [19]. Octamers were assembled from the appropriate recombinant histones and purified via size exclusion chromatography as published previously [17]. Mononucleosomes were reconstituted from the DNA template and histone octamer via rapid dilution methods [20], and then were characterized via 4% native PAGE Gel with staining for the DNA. The assembled nucleosomes were dialyzed into native buffer (2.5 mM NaCl, 10 mM Tris, 0.25 mM EDTA, pH 7.4) and concentrated 7-10 fold by volume using 30 kDa MWCO
concentrators (EMD Milipore). Mononucleosomes were immobilized onto beads largely as previously described [19]. Briefly, 4.66 pmol of the biotinylated adapter were bound to 200 mg of paramagnetic streptavidin beads (NEB) at RT for 20 minutes in 1X binding buffer (10 mM Tris-HCl, pH 7.4, 100 mM NaCl) to get a final reaction volume of 100 ml. The supernatant was removed and beads were washed three times with 50 ml of 1X ligation buffer (50 mM Tris, 10 mM MgCl₂, 1 mM ATP, 10 mM DTT; pH 7.5). 6.99 pmol of WT, TA or TAc nucleosomes were added to the beads in 1X ligation buffer, in a final reaction volume of 100 ml. 800 U of T4 DNA ligase (NEB) were used to ligate the nucleosomes to the adapter. The reactions were carried out at RT for 5-6 hours. Excess unbound nucleosomes were then washed off. The beads were resuspended in 31 ml of 1X ligation buffer to get a final nucleosome concentration of 150 nM. The integrity of the assembled substrates was analyzed by digesting the beads with PstI at 37°C, for 5 hours. The liberated nucleosomes were analyzed on 4% native PAGE.

**Histone acetylation assay**

To compare the extent of acetylation of WT, TA and TAc nucleosomes, substrates were subjected to standard acetylation assays by ySAGA (Spt-Ada-Gcn5-acetyltransferase complex from *Saccharomyces cerevisiae*), under initial rate, sub-saturating nucleosome concentrations, as previously described [19]. Briefly, 1.5 nM ySAGA was used to acetylate 10 nM of each kind of nucleosome in a 50 ml reaction volume of 1X HAT buffer (25 mM Tris-HCl, pH 7.5, 150 mM NaCl, 5% glycerol, 1 mM DTT, 1 mM PMSF, 10 mM sodium butyrate, 4.0 μM acetyl CoA with a specific activity of 5.78 Ci/mmol (Moravek Biochemicals)). The acetylation reaction was carried out at 30°C for 10 min. The beads were
washed 4 times at RT with 50 ml 1X wash buffer 1 (WB1- 100 mM NaCl, 50 mM Tris-HCl, 10 mM MgCl₂, 1 mM DTT, 0.1% triton X-100) and for additional 5 times at 37°C with 15 ml 1X WB2 (50 mM NH₂OH , pH 7.5, 0.1% triton X-100), with 25 minutes incubation between each wash. The beads were then resuspended in 15 ml of 1X WB1 and added directly to 6 ml of scintillation cocktail. The samples were counted in Tri carb 4910 TR Liquid Scintillation Analyzer (Perkin Elmer). To account for non-H3 nucleosomal acetylation, the counts obtained from the WT and TAc nucleosomes were subtracted from those obtained from TA nucleosomes. The resultant extent of acetylation obtained from TAc substrate was compared to that of WT, which was normalized to 100%.

**Results**

We have been interested in how histone H3 tail acetylation is established by the SAGA family of transcriptional co-activitors [9, 19, 21], and how this acetylation affects chromatin structure and function, such as subsequent histone acetylation [21]. The SAGA complex from budding yeast has been shown to target four major residues on the H3 tail, with lysine 14 being the most prominent acetylation site, followed by lysine 9, lysine 18, and then lysine 23 (Figure 1) [22]. Nonsense suppression has been successfully employed to generate H3 histones containing lysine 14 acetylation [11], and we sought to determine if such a strategy could be applied to generating histones containing as many as all four primary acetylations. Histone proteins can be recombinantly expressed in *E. coli* [16], and H3 K14 acetylated histone was previously generated from an overexpression plasmid containing an amber codon at amino acid residue 14 (Figure 1) [11]. Building off this sequence, we introduced additional putative acetyl-lysine residues by using site directed mutagenesis to convert lysine
codons to amber stop codons. Plasmids were constructed to contain up to four amber codons at the major sites of SAGA acetylation, where histones with intermediate levels of acetylation were made to match the known preference of lysine acetylation - i.e. H3 histone containing one acetylation contained H3 K14 acetylation; K14 and K9 acetylation for two acetylations; and K14, K9, and K18 acetylation for three acetylation sites (Figure 1).

![Figure 1. Sequences of acetylated histone H3 proteins.](image)

To test the efficiency of acetyl-lysine incorporation, we first over-expressed singly acetylated H3 K14 acetylated histone according to literature protocols [11]. In this protocol, H3 protein expression is induced by IPTG addition in *E. coli* BL21 DE3 cells. However, for suppression of the amber stop codons, the cells also must contain a plasmid that constitutively expresses a tRNA containing an amber anticodon and a tRNA synthetase that can charge the tRNA with acetyl-lysine. Acetyl-lysine is added to the media, as well as NAD, which inhibits acetyl-lysine deacetylation. As expected, we observed visible amounts of H3
histone in whole cell extracts (data not shown). In the inclusion bodies, where overexpressed histones localize, the full-length histone was especially prevalent (Figure 2A). However, when the same conditions were used for overexpression of the tetra-acetylated H3 histone, no H3 histone was observed in the inclusion body (Figure 2A). Because full length expression of the tetra-acetylated histone requires efficient suppression of all four non-sense codon, and suppression involves a number of components, lack of expression could be due to a number of factors. During optimization, we found that increasing the concentration of acetyl-lysine in the cell media dramatically increased the amount of full-length histone expressed (Figure 2B). Further, expression of histones with fewer than four acetyl-lysines was also feasible, with greater expression occurring when fewer amber codons are suppressed (Figure 2C).

**Figure 2. Expression of acetylated histone H3 proteins.** Proteins were resolved on an 18% SDS-PAGE gel and stained with Coomassie Blue. A) Comparison of amounts of full-length mono- and tetra-acetylated histone protein in inclusion bodies under published nonsense suppression conditions. B) Comparison of amounts of full-length tetra-acetylated histone protein in inclusion bodies with increasing concentrations of acetyl-lysine in the media. 1X represents previously utilized acetyl-lysine concentrations of 10 mM. C) Comparison of expression of full-length histone containing one to four acetyl-lysines in whole cell extracts using 40 mM acetyl-lysine concentrations.
With improved tetra-acetylated H3 histone expression, we set out to purify the histone. Standard purification protocols for histones overexpressed in E. coli involve isolating inclusion bodies and then further purifying denatured histone by size exclusion and cation exchange chromatography. With these steps, we were able to increase the purity of the H3 histone, but not enough to obtain pure protein (data not shown). In the literature protocol for the singly acetylated histone [11], a TEV-cleavable, six-His affinity tag was fused to the N-terminus to facilitate purification, and a similar tag was investigated for the multiply acetylated histones. For the tagged, tetra-acetylated histone, induction and inclusion body purification proceeded to enrich the amount of full-length H3 histone (Figure 3A, lanes 1-3). Histone could then be unfolded and captured on nickel-NTA beads under denaturing conditions to further enrich purity (Figure 3A, lane 4). After solvent exchange into a non-denaturing buffer compatible with TEV protease, H3 histone was liberated from the nickel beads (Figure 3A, lane 5). We found that to get complete cleavage required optimization of the both the binding and cleavage steps. For histone binding, we utilized just enough nickel-NTA beads to get complete histone binding, because additional beads decreased cleavage efficiency. Because commercially available TEV protease is also His-tagged, we believed that the nickel-NTA beads could be sequestering away TEV protease. Consistent with this idea, we found that addition of up to 50 mM imidazole improved the cleavage efficiency of the immobilized histone (higher concentrations of imidazole resulted in elution of uncleaved histone). The amount of TEV protease was also optimized to minimize the amount of protease required to give complete cleavage of the tetra-acetylated histone. Despite full cleavage, the histone could only be eluted from the beads under denaturing conditions, where a significant amount of contaminating proteins coelute. A cation exchange chromatography
step under denaturing conditions resulted in histone protein sufficiently pure for subsequent applications. (Figure 3A, lane 6). Because we were working with relatively small amounts of protein, we found that reversed phase C8 Sep-Pak purification was an efficient way to concentrate and desalt the purified tetra-acetylated histone after cation exchange chromatography.

**Figure 3.** A) Analysis of tetra-acetylated H3 histone purity at various steps of the modified purification protocol. Proteins were resolved on an 18% SDS-PAGE gel and stained with Coomassie Blue. Pre- and post-induction of whole cell extracts are shown in lanes 1 and 2, respectively. Lane 3 shows protein composition following inclusion body purification. Lanes 4 and 5 show proteins bound to nickel-NTA-bead and then released by TEV protease, respectively. Lane 6 shows the protein composition following cation exchange chromatography. C) MALDI-TOF mass spectrometry analysis of purified tetra-acetylated H3 protein.

While the purification step generated full-length H3 histone, additional characterization was necessary to confirm the acetylation state of the histone. Full-length H3 protein lacking tetra-acetylation could result from amber suppression by non-acetyl-lysine codons, or by enzymatic deacetylation of the acetyl-lysine residues. To rule out these possibilities, the mass of the purified histone was determined by MALDI mass spectrometry (Figure 3B), and this analysis confirmed that the isolated protein was tetra-acetylated (15,425
Da expected, 15,428 Da observed, 42 Da per acetyl group). The ultimate yield of the tetra-acetylated H3 histone was 0.06 mg/g of cells. While this is significantly less than the 0.82 – 2.45 mg/g of cells that is typically obtainable for unacetylated H3, it is not significantly worse than the 0.17 mg/g of cells that was obtained from original single-site nonsense suppression studies [11]. A similar purification protocol was applied to each of the other acetylated histones, and yielded 2.45 mg/g of cells, 0.65 mg/g of cells, and 0.24 mg/g of cells for H3 histones containing one, two and three acetylated lysines, respectively.

The purity and amount of tetra-acetylated H3 was sufficient to incorporate into a histone octamer and then a mononucleosome. To generate histone octamer, the tetra-acetylated H3 was combined with recombinant H2A, H2B, and H4 histones under denaturing conditions and dialyzed into a high salt (2M NaCl) solution [16]. Histone octamer was resolved from incomplete assembly products by size exclusion chromatography, and eluted identically to wild-type octamers. Denaturing protein gel electrophoresis of the wild type, tetra-acetylated H3, and tetra-alanine H3 octamers look similar (Figure 4A), with some changes in electrophoretic mobility of acetylated and alanine-containing H3 histone relative to the wild-type H3. Equal amounts of octamers were then incorporated into mononucleosomes by deposition onto a 177 base pair 601 double stranded DNA template by rapid dilution [20], and all three mononucleosomes appear largely similar in their extent of assembly and their electrophoretic mobility (Figure 4B). Each mononucleosome was immobilized onto a paramagnetic bead, and then used as a substrate for a SAGA-mediated acetyltransferase assay (Figure 4C). As expected, the nucleosome containing acetyl-lysine at the major SAGA acetylation sites showed a significant reduction of new acetylation relative to a nucleosome that was not pre-acetylated. The reduction in the amount of acetylation was
similar to that observed for a nucleosome in which the four major H3 tail acetylation sites were mutated to alanine, indicating that the tetra-acetylated histone prevents SAGA-mediated histone acetylation in a manner similar to loss of the lysine residue.

**Figure 4. Utilization of tetra-acetylated histone H3 protein.** A) Denaturing protein gel analysis of histone octamers containing unacetylated, tetra-alanine, or tetra-acetylated H3 histones. Histones were resolved on an 18% SDS PAGE gel with Coomassie Blue staining. B) Native gel analysis of mononucleosomes containing tetra-alanine, tetra-acetylated, or unacetylated H3 histone. Species were resolved on a 4% native PAGE gel with ethidium bromide staining. C) Comparative extent of SAGA-mediated mononucleosome acetylation relative to mononucleosomes in which the major H3 tail acetylation sites were mutated to alanine. To determine the extent of acetylation, the amount of radioactive acetyl incorporation was subtracted from that of the tetra-alanine mononucleosome and then normalized to the wild-type mononucleosome. Data represents four independent trials.

**Discussion**

We have shown that nonsense suppression can be used to incorporate multiple acetyl-lysine residues into the tail of H3 histone protein. One major hurdle we encountered was expressing full-length protein. This difficulty presumably arises because full-length expression requires efficient nonsense suppression at every amber codon. Consistent with this idea, we observed that the amount of full-length protein decreases with increasing numbers of amber mutations
to suppress (Figure 2C). Thus, conditions to optimize nonsense suppression were necessary. We found that increasing the amount of acetyl-lysine present in the media increases full-length protein expression (Figure 2B), suggesting that a key difficulty in nonsense suppression in this system is charging the nonsense suppression tRNA with acetyl-lysine. It is likely that acetyl-lysine is limiting in the cell and by increasing its concentration in the media, intracellular concentrations increase to drive the action of the synthetase. For our purposes, a four-fold increase in acetyl-lysine concentrations was sufficient for generating amounts of tetra-acetylated histone we needed. However, for the increasing acetyl-lysine concentrations used (Figure 2B), we did not observe saturation of histone expression, suggesting that even greater levels of expression should be possible. In this case, the only trade-off may ultimately be the cost of acetyl-lysine. Further, because nonsense suppression is a complicated process, other steps in the reaction might also benefit from optimization. This could include improving the expression levels and properties of the tRNA and synthetase, as well as optimizing the length and timing of histone induction relative to acetyl-lysine addition.

The other major hurdle encountered was in the purification of the multiply acetylated histones. Inclusion body preparations, nickel-NTA beads binding, and cation exchange chromatography were highly efficient, while proteolytic cleavage of the histone from the nickel-NTA required significant optimization in the bead, imidazole, and TEV protease amounts. Even under these optimized conditions, a significant amount of TEV protease was required. The inefficiency of TEV cleavage could be due in part to sequestration of the protease to the nickel-NTA beads, and alternative forms of TEV-protease that do not contain a His-tag might improve cleavage efficiency. Another potential contributor could be the
inability of H3 histone to form a well-behaved globular structure under native conditions. In such a case, it is likely that affinity captured histones form aggregates on the surface of the nickel-NTA beads when solvent is exchanged for proteolysis, making the cleavage site less accessible. A cleavage strategy that can be performed under denaturing conditions would likely avoid this issue.

In our studies, we focus on incorporating multiple acetyl-lysine residues into the H3 histone at known sites of SAGA-mediated acetylation. However, we feel our strategies should be applicable to other sites within the H3 histone, to other acetylation sites in other histones, and even to acetylation of other proteins. Within the H3 histone, Chin and coworkers have previously shown that many different sites within the H3 histone can be individually acetylated, suggesting that, in general, nonsense suppression is not highly sequence dependent in the H3 histone. This idea is further supported by the fact that we observe a similar degree in decrease in histone expression with every amber codon added. Thus, we expect that other combination of multiple lysine acetylations in H3 histone should be possible. With respect to other histones, multiple acetylations may also be possible. For example, while the H4 histone had proven recalcitrant to single site acetyl-lysine nonsense suppression, recent studies have shown that with codon optimization, H4 K16 acetylated histones can be generated at level comparable to those found for single sites acetylation H3 histones in the original Chin study. Our improvements in H3 expression and purification may be directly applicable to H4 histones, as well as others. Finally, for non-histone proteins, especially those that can be purified more effectively than histone, increasing acetyl-lysine concentrations might prove sufficient to allow for improved incorporation of single or multiple acetyl-lysine residues.
Conclusion

We have shown that we can extend a nonsense suppression strategy for incorporating acetyl-lysine into H3 histones to generate histones containing up to four acetyl-lysine residues. Key to achieving this is improved histone expression via an increase in the amount of acetyl-lysine added to the cell media, and a combined affinity capture and ion exchange chromatography purification strategy. The tetra-acetylated H3 histone is generated in sufficient yield and purity to be incorporated into histone octamers and nucleosomes, and we expect that the insight gained from our study could aid in utilizing nonsense suppression to incorporate acetyl-lysine residues into other histones and non-histone proteins.

Acknowledgements and Funding: This work was supported by American Cancer Society Research Scholar Grant 1206501 to M. A. S.-K.
REFERENCES


