Université de Montréal

Quantitative Proteomics Methods for the Analysis of Histone Post-translational Modifications

par

Nebiyu Ali Abshiru

Département de Chimie
Faculté des arts et des sciences

Thèse présentée à la Faculté des études supérieures
et postdoctorales en vue de l'obtention du grade de
philosophiæ doctor (Ph.D.) en chimie

Septembre 2015

Résumé

Les histones sont des protéines nucléaires hautement conservées chez les cellules des eucaryotes. Elles permettent d'organiser et de compacter l'ADN sous la forme de nucléosomes, ceux-ci representant les sous unités de base de la chromatine. Les histones peuvent être modifiées par de nombreuses modifications post-traductionnelles (PTMs) telles que l'acétylation, la méthylation et la phosphorylation. Ces modifications jouent un rôle essentiel dans la réplication de l'ADN, la transcription et l'assemblage de la chromatine. L'abondance de ces modifications peut varier de facon significative lors du developpement des maladies incluant plusieurs types de cancer. Par exemple, la perte totale de la triméthylation sur H4K20 ainsi que l'acétylation sur H4K16 sont des marqueurs tumoraux spécifiques a certains types de cancer chez l'humain. Par conséquent, l'étude de ces modifications et des événements determinant la dynamique des leurs changements d'abondance sont des atouts importants pour mieux comprendre les fonctions cellulaires et moléculaires lors du développement de la maladie.

De manière générale, les modifications des histones sont étudiées par des approches biochimiques telles que les immuno-buvardage de type Western ou les méthodes d'immunoprécipitation de la chromatine (ChIP). Cependant, ces approches présentent plusieurs inconvénients telles que le manque de spécificité ou la disponibilité des anticorps, leur coût ou encore la difficulté de les produire et de les valider. Au cours des dernières décennies, la spectrométrie de masse (MS) s'est avérée être une méthode performante pour la caractérisation et la quantification des modifications d'histones. La MS offre de nombreux avantages par rapport aux techniques traditionnelles. Entre autre, elle permet d'effectuer des analyses reproductibles, spécifiques et facilite l'etude d'un large spectre de PTMs en une seule analyse. Dans cette thèse, nous présenterons le développement et l'application de nouveaux outils analytiques pour l'identification et à la quantification des PTMs modifiant les histones.

Dans un premier temps, une méthode a été développée pour mesurer les changements d'acétylation spécifiques à certains sites des histones. Cette méthode combine l'analyse des histones intactes et les méthodes de séquençage peptidique afin de déterminer les changements d'acétylation suite à la réaction *in vitro* par l'histone acétyltransférase (HAT) de levure Rtt109 en présence de ses chaperonnes (Asf1 ou Vps75).

Dans un second temps, nous avons développé une méthode d'analyse des peptides isomériques des histones. Cette méthode combine la LC-MS/MS à haute résolution et un nouvel outil informatique appelé Iso-PeptidAce qui permet de déconvoluer les spectres mixtes de peptides isomériques. Nous avons évalué Iso-PeptidAce avec un mélange de peptides synthétiques isomériques. Nous avons également validé les performances de cette approche avec des histones isolées de cellules humaines érythroleucémiques (K562) traitées avec des inhibiteurs d'histones désacétylases (HDACi) utilisés en clinique, et des histones de *Saccharomyces cerevisiae* liées au facteur d'assemblage de la chromatine (CAF-1) purifiées par chromatographie d'affinité. Enfin, en utilisant la méthode présentée précédemment, nous avons fait une analyse approfondie de la spécificité de plusieurs HATs et HDACs chez *Schizosaccharomyces pombe*. Nous avons donc déterminé les niveaux d'acétylation d'histones purifiées à partir de cellules contrôles ou de souches mutantes auxquelles il manque une HAT ou HDAC. Notre analyse nous a permis de valider plusieurs cibles connues des HATs et HDACs et d'en identifier de nouvelles. Nos données ont également permis de définir le rôle des différentes HATs et HDACs dans le maintien de l'équilibre d'acétylation des histones. Dans l'ensemble, nous anticipons que les méthodes décrites dans cette thèse permettront de résoudre certains défis rencontrés dans l'étude de la chromatine. De plus, ces données apportent de nouvelles connaissances pour l'élaboration d'études génétiques et biochimiques utilisant *S. pombe*.

Mots clés : Spectrométrie de masse, histones, modifications post-traductionnelles, peptides isomérique, histones acétyltranférases, histones déacétylase.

Summary

Histones are highly conserved, basic proteins found in eukaryotic cell nuclei. They organize and package DNA strands into nucleosome core particles (NCPs), the fundamental repeating units of eukaryotic chromatin. The histones are subject to a wide variety of posttranslational modifications (PTMs) including acetylation, methylation and phosphorylation. These PTMs play an essential role in DNA-replication, transcription, and chromatin assembly. Alterations in histone PTM abundances have been implicated in several types of cancer. For example, the global loss of trimethylation at H4K20 and acetylation at H4K16 is a hallmark of human cancers. Thus, characterization of histone PTMs and their dynamics is extremely useful for elucidating normal cellular functions and molecular pathways that lead to diseases.

Traditionally, histone PTMs are analyzed using antibody-based approaches such as western blot and chromatin immunoprecipitation (ChIP) assays. These methods, however, suffer from several limitations including antibody cross-reactivity, epitope occlusion, and the cost and difficulty in producing and validating antibodies. Over the last decade, mass spectrometry (MS) has emerged as a powerful technique for the characterization and quantification of histone PTMs. MS offers several advantages over the traditional approaches including reproducibility, specificity, and ability to rapidly analyze numerous PTMs in a single experiment. In this thesis, the development and applications of novel analytical tools for the identification and quantification of histone PTMs are presented.

First, a method useful for measuring the global and site specific changes in histone acetylation is described. This method combines intact mass analysis and peptide sequencing approaches to study the global and site specific changes in histone acetylation during *in vitro* assays with yeast Rtt109 and its chaperone (Asf1 or Vps75). Second, a method for analysis of isomeric histone peptides is presented. This method combines a high resolution LC-MS/MS with a novel bioinformatics tool called Iso-PeptidAce to deconvolute mixed spectra of co-eluting isomeric peptides. We benchmarked Iso-PeptidAce using mixtures of synthetic isomeric peptides. We demonstrated its capability in histones isolated from human erythroleukemic (K562) cells treated with clinically relevant histone deacetylase inhibitors (HDACi) and in affinity-purified *S. cerevisiae* histones bound to chromatin assembly factor-1 (CAF-1). Third, by employing the above methods, an in-depth quantitative analysis of the substrate specificities of several fission yeast HATs and HDACs was assessed. We

determined the acetylation site occupancy of multiple lysines in histones isolated from a control or mutant strains lacking specific HAT or HDAC activities. Our analysis identified several known and novel HAT and HDAC target sites. Our data also defined the division of labor between the different HATs and HDACs in maintaining the steady-state level of histone acetylation. Overall, we anticipate that the methods described in this thesis will address some of the existing challenges facing the chromatin field. Moreover, the data presented will provide valuable insights for future genetic and biochemical studies involving the fission yeast.

Key Words: Mass spectrometry, histones, posttranslational modifications, isomeric peptides, histone acetyltransferases, histone deacetylases

Table of Contents

List of Tables

List of Figures

List of abbreviations

3-D	Three dimensional
ac	Acetylation
CoA	Co-enzyme A
Asf1	Anti-silencing factor 1
BCA	Bicinchoninic Acid
BET	Bromodomain and extra terminal domain
bp	Base pair
BRD	Bromodomain
CAA	Chloroacetamide
CAF1	Chromatin assembly factor 1
ACN	Acetonitrile
CE	Collision energy
ChIP	Chromatin immunoprecipitation
CID	Collision induced dissociation
csv	Comma separated value
Da	Dalton
DC	Direct current
DNA	Deoxyribonucleic acid
DSB	Double strand break
DTT	Dithiothreitol
EI	Electron ionization
EIC	Extracted ion chromatogram
ELISA	Enzyme-linked immunosorbent assay
ESI	Electrospray ionization
ETD	Electron transfer dissociation
FA	Formic acid
FC	Fold change
fmol	Femtomole
FTICR	Fourier transform ion cyclotron resonance
FWHM	Full width at half maximum
µg	Micro-gram
GC	Gas chromatography

GST	Glutathione S-transferase
HAT	Histone acetyltransferase
HCD	High-energy collision dissociation
HDAC	Histone deacetylase
HDACi	Histone deacetylase inhibitor
HDMT	Histone demethylase
HMT	Histone methylase
HP1	Heterochromatin protein 1
HPLC	High performance liquid chromatography
i.d.	Internal diameter
IAA	Iodoacetamide
IDs	Identifications
K_d	Dissociation constant
µL	Micro-liter
LC-MS/MS	Liquid chromatography tandem mass spectrometry
LTQ	Linear trap quadrupole
µ/mM	Micro-/milli-molar
m/z	Mass to charge ratio
µm	Micro-meter
MALDI	Matrix-assisted laser desorption ionization
me	Methylation
me1	Mono-methylation
me2	Di-methylation
me3	Tri-methylation
m-H_2O	Milli-Q water
MS	Mass spectrometer
MW	Molecular weight
NAD^+	Nicotinamide adenine dinucleotide
NCP	Nucleosome core particle
ND	Not determined/detected
nLC	Nano-liquid chromatography
PDB	Protein data bank
ppm	Part-per-million
PRM	Parallel reaction monitoring

pr	Propionylation
PTMs	Post-translational modifications
QIT	Quadrupole ion trap
QQQ	Triple quadrupole mass spectrometer
QTOF	Hybrid Quadrupole-Time of flight mass spectrometer
RF	Radio frequency
RP	Reverse phase
Rtt109	Ty1 transposition gene product 109
SAHA	Suberoylamide hydroxamic acid
SDC	Sodium deoxycholate
SD	Standard deviation
SILAC	Stable isotope labeling by amino acid in cell culture
SRM	Single reaction monitoring
TAP	Tandem affinity purification
TCA	Trichloroacetic acid
TFA	Trifluoroacetic acid
TIC	Total ion chromatogram
UV	Ultraviolet
v/v	Volume per volume
Vps75	Vacuolar Protein Sorting 75
WT	Wild type
XIC	Extracted ion chromatogram

Acknowledgements

‘A single stick may smoke, but it will not burn’ Ethiopian proverb

It is with great pleasure that I thank all the friends, families and everyone I have worked with over the past five years who have made this work possible.

First, I would like to express my deepest gratitude to my supervisor Dr. Pierre Thibault for giving me the opportunity to pursue this graduate program and for his continuous guidance, encouragement, and support. Before coming to Montreal I did not know much about the city, which is not surprising for someone from Ethiopia (well, I understand some of you might be trying to figure out where ‘Ethiopia’ is located). Yet, I had read several outstanding scientific contributions by Pierre and his team to the field of mass spectrometry. At the time I was on the verge of completing my Master’s degree in Norway and Pierre’s work caught my attention. I immediately contacted him by email expressing my interest to join his team and luckily he accepted my application. It wasn’t easy starting life in Montreal; it was a huge cultural shift for me. Pierre’s support and encouragement along the way was very helpful. Apart from this, his capacity as one of the leading scientists in Canada- the depth of knowledge he has, his dedication and his attention to details makes him very unique. And it was an honor and a great privilege working with him over the past five years.

I would also like to thank my co-supervisor Dr. Alain Verreault who has supported me in so many ways during my study. Much of the research activities I have been involved in required a deeper understanding of biological processes though I was trained as analytical chemist. Alain helped me understand most of these processes, which made my study progress a lot more manageable. He has always been ready to share his knowledge and expertise. I also owe a big thank you to Alain’s current and former team members, a very lively and wonderful team. Especially, Roshan Elizabeth, who has been at the center of all the research works I have conducted. She has been tirelessly providing me with all the biological samples that I needed to do my work. Also, together with Alain, she has immensely improved my understanding of the role of various epigenetic processes. I’m very thankful for all this.

I would like thank Prof. Joelle Pelletier, the Chair of my thesis committee, and Prof. Martine Raymond, member of my thesis committee, for the great advice and encouragement over the

course of my study. Their guidance has been instrumental in continuously developing my knowledge and skills.

I have been fortunate enough to work with a number of wonderful labmates over the years. I would like to first thank Eric Bonneil for his great enthusiasm to share his expertise in mass spectrometry. He has always been there to train and guide me on how to operate the various instruments in the lab. Thank you, Eric! I also want to thank all the other former and current members of the lab including Christina Bell, Frederic Lamoliatte, Sibylle Pfammatter, and Chongyang Li, who are more friends than labmates. A special thank you to Frederic Lamoliatte (he likes to be called Fred) who generously accepted to translate my 'Summary' section into French.

I would like to thank all my friends and families in Montreal, Chicago, and Ethiopia. My parents Senait and Ali may not read this thesis but their support, encouragement and prayer have never ceased till this day. In my home country not many people go to school till they turn thirty and I guess that is why my parents have always wondered when I would stop going to classes.

This thesis would not have been possible if not for the support and encouragement of the love of my life Feben Woldemariam. She has been always close to my heart even when she is thousands of miles away. Since we married three years ago she has constantly stood by me in times of difficulties and has been a wonderful friend and wife to me. It is with great pleasure that I dedicate the Ph.D. Thesis to her. I also want to thank all my family in Chicago Bresh, Abby, Woudu, Teklu, Nate, Nuhamin, Sapphier and others for their encouragement.

Finally, I praise God almighty for the wonderful life and people He has blessed me with!

“It is a miracle that curiosity survives formal education.”

Albert Einstein

CHAPTER 1: Introduction

1.1. How is the DNA stuffed into a tiny microscopic cell nucleus?

The DNA contains all the genetic instructions required for an organism to grow and reproduce [1]. It is composed of a phosphate group, a sugar molecule (called deoxyribose) and a nitrogen-containing base namely adenine (A), guanine (G), cytosine (C) and thymine (T). These components are linked in the order phosphate-sugar-base to form the structural unit of DNA known as the nucleotide (Fig. 1.1a). In turn, the nucleotides are joined to one another in a chain by a phosphodiester bond between the phosphate of one nucleotide and the deoxyribose of the next to form a DNA strand (Fig. 1.1b). The DNA exists in a double-stranded form whereby two adjacent strands running in opposite directions are joined together by hydrogen bonds between complementary bases: A with T, and G with C (Fig. 1.1c). In the cell, the DNA double-strand mainly exists in right-handed helix structure (Fig. 1.1d).

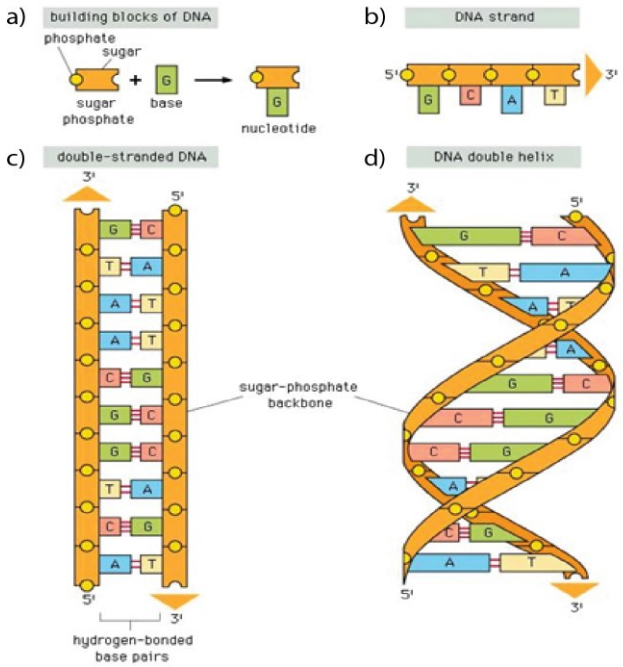


Figure 1.1 The building blocks of DNA.

From reference [1] with permission from Garland Science - Books)

The nucleotide is formed from a phosphate group, a deoxyribose sugar and a nitrogen-containing base (a). Successive nucleotides are linked by phosphodiester bonds to form a DNA strand (b). Adjacent DNA-strands are joined by hydrogen bonds between complementary bases to form a DNA double-strand (c). In the cell, the DNA double-strand is present as a helix structure (d).

The discovery of the DNA double-strand by James Watson and Francis Crick in 1953 [2] was indeed a major advance in the field of science. This discovery gave scientists a vital clue on how the DNA carries genetic information and how this information is transmitted between generations. But, the idea that a very long DNA double-strand can fit into a tiny microscopic cell nucleus and still be accessible for gene-regulatory processes is extremely fascinating. Just to illustrate the extent of packaging- the total length of the DNA double-strand in a single human cell is approximately 2 meters; and it is believed that the human body is comprised of tens of trillions of cells that contain this length of DNA. So, if the entire DNA sequence in all the cells were to be stretched end-to-end, it would extend to tens of billions of kilometers. Instead, the DNA sequence in each cell is condensed by about 10,000-fold and marvelously fit into a nucleus that is only about 10μm wide. The DNA undergoes various layers of organizations to achieve such high level of packaging.

The first level of DNA packaging involves the wrapping of the DNA double-strand around a complex of proteins known as histones [3, 4]. These proteins are highly basic and contain multiple positively charged amino acid residues that form electrostatic interactions with the negatively charged phosphate group of the DNA backbone. Approximately 147 base-pairs (bps) of DNA sequences are wrapped almost twice around the histone complex to form a bead-like structure called the nucleosome core particle (NCP) [5] (see Fig. 1.2). This structure was first elucidated by electron microcopy in the mid-1970s [6]. The packaging of DNA into nucleosomes shortens the long, linear double-strand sequence by about seven-fold. Adjacent nucleosomes are then joined together by linker DNA sequences (approximately 10 to 80 bps) to form an 11 nm beads-on-a-string like structure, where the beads represent nucleosomes and the string represents the linker DNA. In the next level of DNA packaging, arrays of nucleosome cores condense into a 30 nm higher-order chromatin fiber [7]. This is achieved through a complex DNA folding processes that involve histone H1 and several non-histone proteins [8, 9].. Histone H1, also known as linker histone, binds to the entry and exit sites of the linker DNA sequence of the NCP.

The chromatin fiber is further compacted into different layers of higher order chromosomal domains. The two major chromosome structural domains are euchromatin and heterochromatin [9]. The euchromatin is structurally less condensed and contains a higher proportion of genes that are transcriptionally active. In contrast, the heterochromatin region is highly condensed and genes located in this region are generally less accessible for transcription. The heterochromatin is further classified into two broad structural domains.

The first type, which is called constitutive heterochromatin, is found in all cell types and comprises genes that are generally permanently inactivated [10]. This type of heterochromatin is mainly found near chromosomal regions that contain highly repetitive DNA sequences, such as centromeres and telomeres. The second type, facultative heterochromatin, refers to regions of the chromosome that are temporarily condensed [11]. Most inactive euchromatic genes exist in this form of heterochromatin.

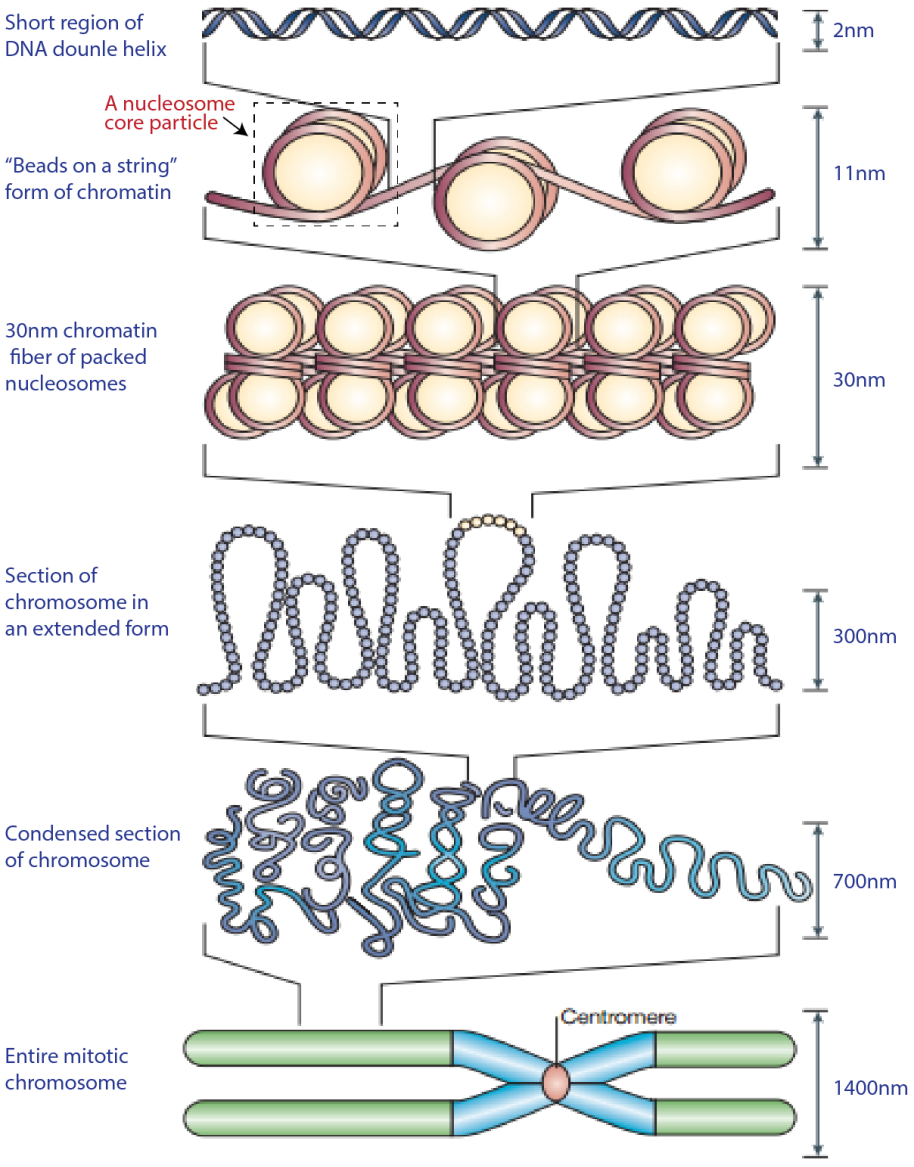


Figure 1.2 Overview of eukaryotic DNA packaging.
From reference [12] with permission from Nature Publishing Group.

1.2. Histones

Extensive research in the composition of the cell nuclei resumed in the late 1860s when Friedrich Miescher first isolated the "nuclein" (now called DNA) from human white blood cells [13]. He later extended his research to other cell types such as salmon spermatocytes, and found the DNA bound to a basic molecule, which he named "protamin". Together, the DNA-protamin complex amounted to almost the entire mass of the sperm head. Meanwhile, Miescher was deeply interested in the DNA component and paid little attention to the protamin. Later in 1884, Albrecht Kossel isolated DNA-bound proteins from the nuclei of avian red blood cells [14, 15]. He found major similarities between the proteins and Miescher's protamine. He named the proteins 'histones'. Kossel later attempted to examine the physical and chemical features that distinguish histones from protamins and other nuclear proteins. In this respect he found that histones exist in complex with nucleic acids and, are highly basic in nature. Moreover, by exploiting the basic property of histones Kossel described the first protocol for acid-extraction of histones from the cell nuclei. Kossel received the Nobel Prize in Physiology or Medicine in 1910 for his major contribution in the discovery of the chemical compositions of the cell nucleus.

However, it wasn't until the late 1950s that researchers began to investigate the different classes of histone found in eukaryotic cells. Extensive purification and characterization of histones in calf thymus and pea seedlings revealed five major classes: Histones H1, H2A, H2B, H3 and H4 [16, 17]. Histones H2A, H2B, H3 and H4 (collectively known as core histones) form the nucleosomal core around which the DNA is wrapped, whereas H1 acts as a scaffold protein for the linker DNA [18]. Early chemical cross-linking experiments showed that the nucleosome consists of two copies of each core histone assembled into an octameric structure [19]. Moreover, primary sequence analyses showed that histones contain highly basic N-terminal sequences and C-terminal fold domains that include amino acid distributions characteristic of globular proteins [20].

1.2.1. Histone variants

Each class of histone described in the previous section, except histone H4, has a number of primary sequence variants. A total of 55 histone variants have been reported in humans (Appendix A, Table S1.1) [21]. Histone H3, for instance, is present in six different variants: H3.1, H3.1t, H3.2, H3.3, H3.3C, and CenH3 (see Fig. 1.3). The first five variants share about 90% amino acid sequence similarity. The CenH3, also known as CENP-A, has longer N-

terminal domain and less than 50% amino acid sequence similarity with the other five variant.

The canonical histone H3.1 differs from H3.2 and H3.3 by one and five amino acid residues, respectively. In H3.1, the residue at position 96 is cysteine, where as in H3.2 and H3.3 the cysteine is replaced by serine (highlighted by a blue box in Fig.1.3). The other four residues of H3.1 that differ from H3.3 are alanine, serine, valine and methionine located at positions 31, 87, 89 and 90 (highlighted by a red boxes in Fig.1.3), respectively. Histones H2A and H2B have the largest number of variants, each with at least nineteen variants.

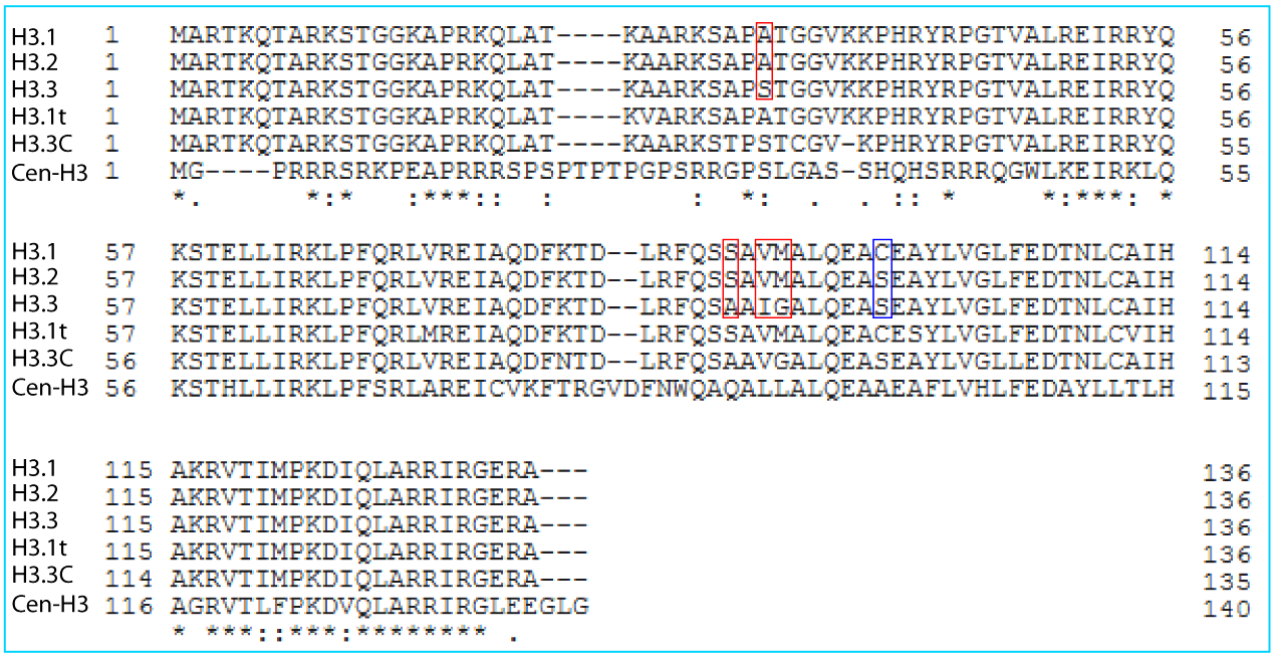


Figure 1.3 Sequence alignment of human histone H3 variants.
(Sequence alignment prepared using the 'align' program at http://www.uniprot.org/)

Histone variants play an important role in marking and organizing the nucleosome in specialized regions of the chromatin. For instance, CenH3 (also known as CENP-A) is incorporated specifically into nucleosomes at the centromere [22], a specialized region on a chromosome where two sister chromatids are linked. On the other hand, histones H3.1/H3.2 and H3.3 mediate distinct nucleosome assembly pathways. *In vitro* nucleosome assembly assays show that deposition of H3.1/H3.2 into the nucleosomes is coupled with DNA-replication processes, whereas H3.3 deposition takes place independent of DNA replication [23]. In accordance with this, studies show that the synthesis of H3.1/H3.2 is highly elevated during DNA replication in the S phase of the cell cycle, whereas the H3.3 variant is expressed at all stages of the cell cycle [24-26].

Histone H2A variants are also implicated in several DNA-associated process including DNA damage response, X chromosome inactivation and transcriptional regulation [27-29]. Histone H2A.Bbd has less than 50% amino acid sequence identity with the canonical H2A. It is believed to promote destabilization of the nucleosome core in mammals [30]. Histone H2A.X is highly linked to DNA damage response (DDR) pathways [31]. It is rapidly phosphorylated at Ser 139 (in humans) or Ser 129 (in budding yeast) following DNA damage response. The phosphorylated H2A.X is commonly referred to as ɣ-H2A.X. Accumulation of ɣ-H2A.X at DNA damage sites serves as a mark for the recruitment of several DNA repair protein complexes [32]. Recently, histone H2A.Z was also shown to have a role in DNA strand break (DSB) repair [33]. It facilitates this process by creating an open chromatin structure at the DSB site. On the other hand, despite the crucial role of histone H2B in chromatin formation, the functions of its variants remain highly understudied.

1.2.2. The structure of core histones

1.2.2.1. Primary and secondary structures

Although histones were discovered more than a century ago, analysis of their primary and higher order structure did not begin until the late 1950s. Meanwhile, significant advances were made in developing analytical strategies for the isolation and fractionation of histones from different types of cells [16]. These strategies had enabled investigators to carry out fundamental structural analysis of the core histones. Initial studies were focused on determining the primary sequences of histones. In this regard, histone H4 was the first histone for which the complete amino acid sequence was determined [34, 35]. Sequencing of tryptic peptides generated from calf thymus or pea seedling histone H4 gave rise to a total of 102 amino acid residues. The N-terminal region of H4 was found to be highly rich in positively charged amino acids such as arginine and lysine, suggesting a potential binding site for the DNA phosphate group. Histone H3 is the second histone for which a complete sequence was reported [36]. Sequencing of calf thymus H3 peptides generated by trypsin, chymotrypsin and cyanogen bromide digestion identified a total of 135 amino acid residues. Similar to histone H4, the N-terminal region of H3 contained numerous arginine and lysine residues. Moreover, histone H3 had two cysteine residues which are absent in histone H4. The complete primary sequences of histones H2A and H2B were also determined around the same period. Both histones comprised highly basic N-terminal region and a hydrophobic C-terminal domain. The primary sequences of the four core histones are highly conserved across eukaryotes,

suggesting a similar role for all. The sequence alignment for canonical histone H3 and histone H4 in five different species is illustrated in Appendix A, Figure S1.1.

Soon after sequence analyses of the core histones, researchers began a thorough investigation into how the individual histones are assembled into an octameric structure. A number of biochemical and structural analyses revealed a pairwise association between core histones [37, 38]. Equimolar mixtures of calf thymus H3 and H4 form a very stable $(H3\text{-}H4)_2$ tetramer at high salt concentrations (~ 2M NaCl). Moreover, CD (circular dichroism) spectra analyses showed increased α-helical properties of the tetrameric structure compared to the individual histones. Similarly, fluorescence anisotropy and CD measurements indicated that histones H2A and H2B form a stable H2A-H2B dimer in high salt solutions [39]. Formation of the dimer increases the α-helical content of the H2A-H2B by about 15 residues compared to the individual histones. Moreover, CD spectra analyses of trypsin-digested chromatin showed increased secondary structure properties at the less basic (or C-terminal) regions of histones H2A and H2B [40].

The discovery of the tetrameric $(H3\text{-}H4)_2$ and dimeric H2A-H2B structures led to the formulation of the first model for chromatin structure. As such, the chromatin repeating unit (or the nucleosome) is comprised of one $(H3\text{-}H4)_2$ tetramer and two H2A-H2B dimers assembled into an octamer and approximately 200 base-pairs of DNA. These structural features were later tested by several research groups and were found to agree substantially with the proposed model [18, 19, 41, 42]. At high salt concentrations, the histone octamer complex is highly stable even in the absence of the nucleosomal DNA or chemical cross-linking reagents. Meanwhile at physiological salt concentrations the octamer dissociates into one $(H3\text{-}H4)_2$ tetramer and two H2A-H2B dimers. These key physical properties were fundamental to the successful crystallization of the histone octamer in subsequent structural studies.

1.2.2.2. Higher order structures

By the late 1970s it was generally accepted that the core of the nucleosome structure is comprised of a histone octamer assembled from two copies of each core histone. Meanwhile, the physical association between the histones in the octameric complex was not fully understood. An early study by Eickbush et al. [42] suggested that the octamer is held together by two types of protein-protein interactions. The first type involves hydrophobic interactions which hold together the individual subunits of the tetrameric $(H3\text{-}H4)_2$ and dimeric H2A-

H2B structures. The second type involves weak interactions that promote the association between one H3-H4 tetramer and two H2A-H2B dimers to form an octamer. The later type mainly involves hydrogen bond interactions between specific regions of the tetrameric and dimeric subunits. Nonetheless, these interactions were first proposed based solely on results obtained from studying the effects of temperature, urea, and pH titration on the dissociation of the octamer and hence, further structure analysis of the octamer was needed.

Initially, researchers were investigating low resolution structures of the histone octamers using techniques such as neutron and x-ray diffraction, and electron microscopy [43-47]. These studies revealed the disc-like shape of the octamer and the left handed DNA superhelix wrapped around the octamer (Fig.s 1.4a-b). Moreover, it was shown that the arrangement of histones in the octamer involves two H2A-H2B dimers flanked by one centrally located $(H3\text{-}H4)_2$ tetramer. Meanwhile, due to the low resolution of these models accurate assessment of the secondary and tertiary structures of the individual histones could not be achieved. A more defined structure of the octamer was later determined by X-ray crystallography at 3.1 Å resolution [48]. In addition to the previously described features, the new model revealed three dimensional structural motifs (which are now called the histone fold) in each of the core histones. The motifs were later confirmed by high resolution X-ray crystallography analysis (Fig. 1.4c) of *Xenopus laevis* and human nucleosome core particles solved at 2.8 [49] and 2.5 Å [50], respectively.

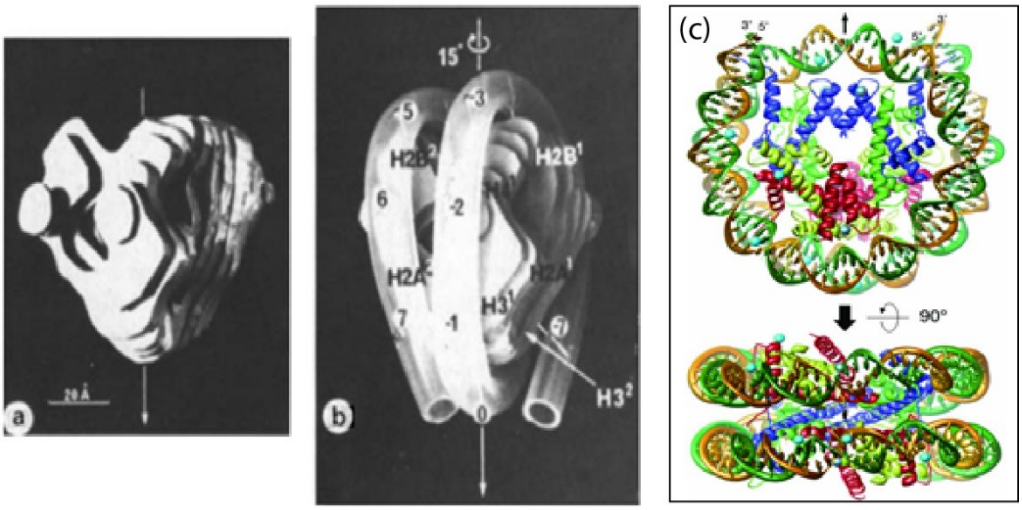


Figure 1.4 Three dimensional structural models of the histone octamer/NCP
From references [43, 50] with permission from Nature Publishing Group and Oxford University Press, respectively. (models a & b published in 1980 and c in 2005).
(a) A 3D model of the octamer reconstructed from an electron micrograph. (b) A model showing the left handed supercoiling of the DNA around the octamer. (c) A ribbon model (top panel) and an axial view (bottom panel) of a 2.5Å crystal structure of the human NCP.

Although the four core histones have very low sequence homology, they all share the histone fold motif. This motif is characterized by a long central helix flanked on either side by a loop and a short helix (Fig. 1.5a). The histones H3 and H4 form two pairs of heterodimers (H3-H4 and H3'-H4') mainly through the interaction of their central helices (Fig. 1.5b, left panel). The heterodimers are further assembled into a $(H3\text{-}H4)_2$ tetramer through hydrogen bonding and weak interactions of the central helices, C-terminal helices, and the loops (Fig. 1.5c). The two H4 molecules (H4-H4') comprising the histone octamer are aligned side by side in opposite directions (Fig. 1.5c). In contrast, the two H3 molecules in the homodimer H3-H3' are arranged in head-to-tail or 'handshake' motif (Fig. 1.5c). In the same way, the H2A and H2B histone folds are joined together through several interactions to form the two H2A-H2B dimers (Fig. 1.5b, right panel). These dimers, however, do not form tetrameric structures. Instead, they each are linked to the $(H3\text{-}H4)_2$ tetramer through the interaction of multiple residues of H2B and H4. Histone H2B residues E73 and R90 form hydrogen bonds with histone H4 residues R92 and E90, respectively. These hydrogen bonds and several other hydrophobic interactions promote the formation of the histone octamer (Fig. 1.5d).

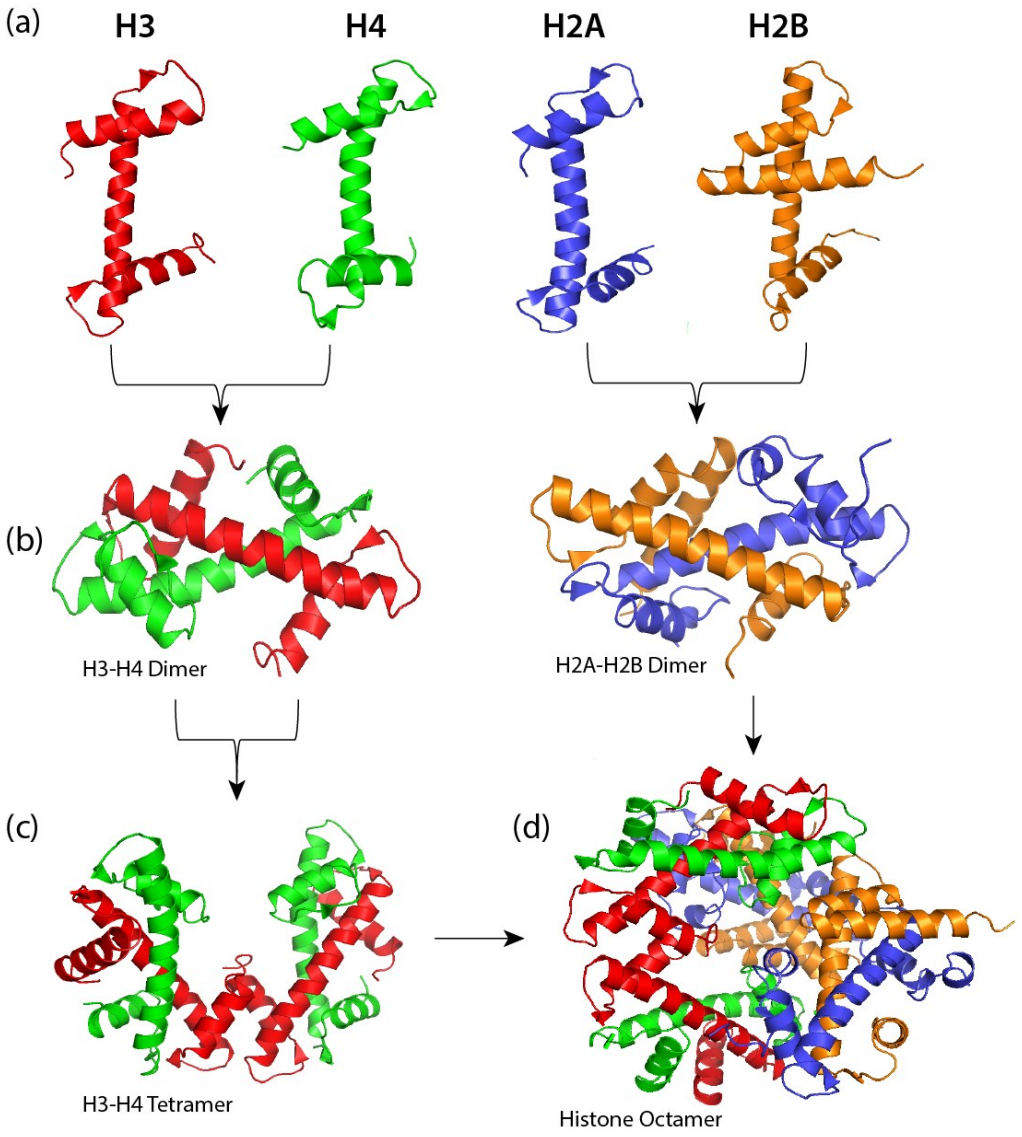


Figure 1.5 The individual core histones assemble into a histone octamer.
A representation of (a) the helix-loop-helix-loop structural motif of histones H3, H4, H2A and H2B; (b) the structure of H3-H4 (left panel) and H2A-H2B (right panel) heterodimers; (c) $(H3\text{-}H4)_2$ tetrameric structure; (d) the histone octamer assembled from the $(H3\text{-}H4)_2$ tetramer and the H2A-H2B dimer. (The figures were adapted from Wikimedia Commons-https://commons.wikimedia.org, and prepared using The Pymol Molecular Graphics System, Version 1.2r 1, based on the crystal structure of Xenopus laevis nucleosome core particle (PDB entry 1AO1).

The histone fold domains contact the DNA superhelix chain at a number of sites. One of the primary contact sites involves the N-terminal tail region of the core histones [49, 51]. This region is highly flexible and as a result it produces weak electron density in X-ray diffraction experiments. A segment of about eight and five amino acids lengths of histones H2B and H3 tails, respectively, pass through the DNA superhelix minor groove channel. The side-chain of

these amino acids is involved in stabilizing the histone-DNA interactions by forming a hydrogen bond with the DNA phosphate group or by creating favorable positively charged density inside the superhelix channel. Another major histone-DNA contact site involves the amino acid side-chain of the α-helices. In this regard, a number of arginine side-chains in the histone fold were shown to interact with the DNA superhelix. Some of these arginines are H2A-R42, R77; H2B-R30, H3-R49, R63, R83; and H4-R48. Other residues are also involved in stabilizing the histone-DNA interactions. For example, the side-chains of two proline residues in the histone fold (H3-P66 and H4-P32) interact with the DNA deoxyribose moiety.

1.2.3. Histone post-translational modifications

Post-translational modifications (PTMs) are covalent modifications that occur at distinct amino acid residues in a protein [52]. Most PTMs are mediated by enzymatic activities. They are present in cellular proteins of various prokaryotic and eukaryotic organisms, ranging from bacteria to humans. PTMs play versatile roles in several cellular processes including the cell cycle, gene regulation, signal transduction, and protein-protein interactions [53-56]. Some of the most common protein modifications include acetylation, methylation, ubiquitination of lysine residues; and phosphorylation of serine, threonine or tyrosine residues. A protein can be post-translationally modified at one or more sites. When a PTM occurs on different sites of the same protein it creates several versions of that protein. Such proteins are referred to as PTM-isoforms.

While it is known that there are only four types of core histones in humans, extensive biochemical and mass spectrometry analyses have identified a number of PTM-isoforms of each type. The core histones possess multiple lysine residues on their N-terminal tail. These lysines are subject to a wide variety of modifications including acetylation, methylation, ubiquitination, butyrylation, crotonylation, and many other modifications [57]. In addition to lysines, the histone tails possess multiple arginine residues that are subject to modifications such as mono- and di-methylation, and citrullination. Phosphorylation of serine, threonine and tyrosine residues was also reported. Among the four core histones, H3 is by far the most modified. Almost all of the above PTMs have been identified in various combinations on H3.

Most histone PTMs carry out their function in a combinatorial fashion although there are cases where individual histone PTMs have been shown to have important roles. One of the mechanisms by which combinatorial functions are achieved involves histone PTM ‘crosstalk’, whereby the addition or removal of a PTM at a specific site affects the

maintenance of other neighboring PTMs. In the following subsections I will briefly introduce two of the most characterized histone PTMs - acetylation and methylation of lysine residues.

1.2.3.1. Acetylation

Acetylation is one of the most widely studied histone PTMs. This modification is established by the addition of an acetyl group to the ε-amino group of lysine residues, a process catalyzed by a group of enzymes known as histone acetyltransferases (HATs) [58]. In this reaction, a conserved glutamate residue (Glu) of the HAT activates the ε-amino group of lysine for nucleophilic attack on the carbonyl moiety of acetyl-coenzyme A (CoA-SH) which acts as an acetyl donor (Fig. 1.6) [59]. The reverse reaction is catalyzed by histone deacetylases (HDAC) enzymes. The two groups of enzymes (HATs and HDACs) will be discussed in detail in sections 1.2.5.1 and 1.2.5.2.

Histone acetylation has been extensively studied in many eukaryotic organisms. Most of the identified acetylation sites are conserved from yeast to humans, suggesting a crucial role of the modifications in each species. Several studies have shown the role of histone acetylation in modulating important DNA-associated processes including replication, transcription, gene regulation and the DNA double-strand break (DSB) repair [60, 61].

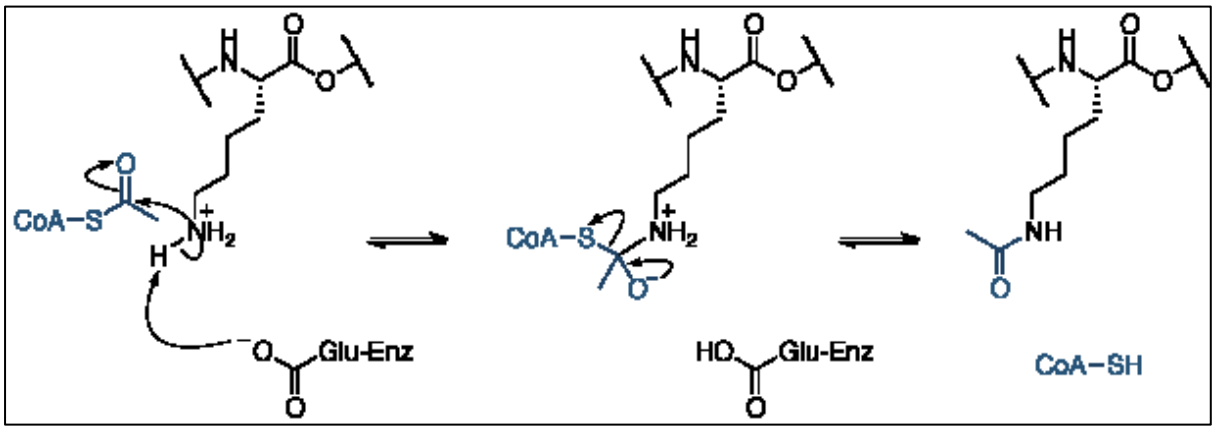
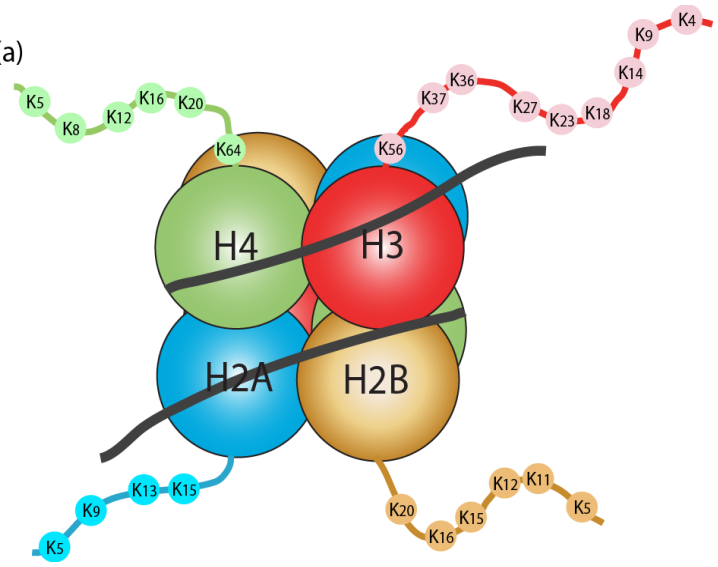


Figure 1.6 Mechanism of lysine acetylation by HATs.
Adapted from reference [59] with permission from the American Society for Biochemistry and Molecular Biology.

Previous comprehensive mass spectrometry-based analyses have identified over twenty acetylation sites on the four core histones in human cells. Some of the known acetylation sites are shown in Figure 1.7a-b. The majority of these acetylations are found on the N-terminal tails, yet a small number of acetylations are also located in the histone fold domain, for example H3K64ac and H4K91ac.

(a)



(b)

H3	H4	H2B	H2A
K4me, ac	K5ac	K5ac	K5ac
K9me, ac	K8ac	K11ac	K9ac
K14ac	K12ac	K12ac	K13ac
K18me, ac	K16ac	K15ac	
K23me, ac	K20me	K16ac	
K27me, ac	K91ac	K20ac	
K36me, ac		K46ac	
K56ac		K120ac	
K64ac			
K79me			
K122ac			

Figure 1.7 Known histone acetylation sites.
(a) An illustration of the N-terminal tails of core histones and sites of acetylation; (b) lists of known histone acetylation and methylation sites, labeled 'ac', and 'me', respectively. Figures prepared using Adobe® Illustrator® CS6, version 16.0.0.

Histone acetylation is involved in the modulation of diverse chromatin-associated processes. For instance, acetylation of lysine 16 of H4 was shown to inhibit chromatin condensation and formation of higher order chromatin fiber [62]. Histone acetylation is also associated with transcriptional activity [63], although the mechanism underlying this association is still not fully elucidated. One of the proposed mechanisms by which histone acetylation could facilitate transcription is via disruption of the histone-DNA interactions [64]. According to this model, acetylation neutralizes the electrostatic interactions between the positively charged Ɛ-amino group of lysines and the negatively charged phosphate group of DNA, thereby increasing the accessibility of DNA for transcriptional machineries.

In an alternative model, increasing evidences suggest that histone acetylation may mediate transcription by acting as docking platforms for recruitment of chromatin-modifying complexes [65, 66]. This model is supported by the findings that chromatin regulators often recruit specific HATs or HDACs to promoter and transcription start sites of actively transcribing genes [67]. This has been observed in budding yeast where CHIP-ChIP assays showed elevated levels of acetylated histones in the vicinity of active genes; sites where HATs Gcn5 and Esa1 were also found to be recruited [68]. This study found a genome wide increase in acetylation of Gcn5 target sites H3K9 and H3K14, and Esa1 target sites H4K5ac,

H4K8ac, H4K12ac and H4K16ac. These modifications were found to also correlate with transcriptional rates genome-wide, suggesting the association between histone acetylation and transcriptional activity. Similar observations were made in humans [69]. ChIP-seq analysis of histone modifications in human $CD4^+$ T cells showed elevated levels of H2BK12ac, H2BK20ac, H2BK120ac, H3K4ac, H4K5ac, H4K8ac, H4K12ac and H4K16ac in the promoter and transcription start sites of active genes, whereas the PTMs H2AK9ac, H2BK5ac, H3K9ac, H3K18ac, H3K27ac, H3K36ac, H4K91ac were mainly identified at regions surrounding the transcription start sites. It is not known whether these combinatorial acetylation patterns have distinct functions, or whether they simply reflect differences in the specificity of HATs recruited at various stages of transcription. It should be noted that transcriptional activation does not always result in increased histone acetylation. Previous studies in budding yeast have shown a dramatic decrease in histone H4 acetylation at target promoter regions upon activation by specific gene regulators [70].

1.2.3.2. Methylation

Histone methylation is catalyzed by a group of enzymes known as histone methyltransferases (HMTs) [71, 72]. The general mechanism of lysine methylation whereby an HMT catalyzes the transfer of a methyl group from the cofactor *S*-adenosylmethionine (SAM) to the ε-amino group of a specific lysine residue is illustrated in Figure 1.8. Lysine residues can be mono-, di-, or tri-methylated by accepting one, two or three methyl groups on their side chain. Unlike acetylation, however, methylation of lysines does not alter the overall charge state of histones. Some of the known methylation sites on the four core histones are shown in Figure 1.7b (labeled ‘me’). Majority of the methylation sites reside in the N-terminal tail. Histone H3 is by far the most methylated core histone, with abundant methylation detected at lysines 4, 9, 27, 36, and 79. The biological role of histone methylation has been implicated in several DNA-associated processes including transcription, gene silencing, and chromatin condensation [73, 74]. Histone arginine residues are also known to be mono- or di-methylated, which serve important roles in epigenetic regulation of various genes [75, 76].

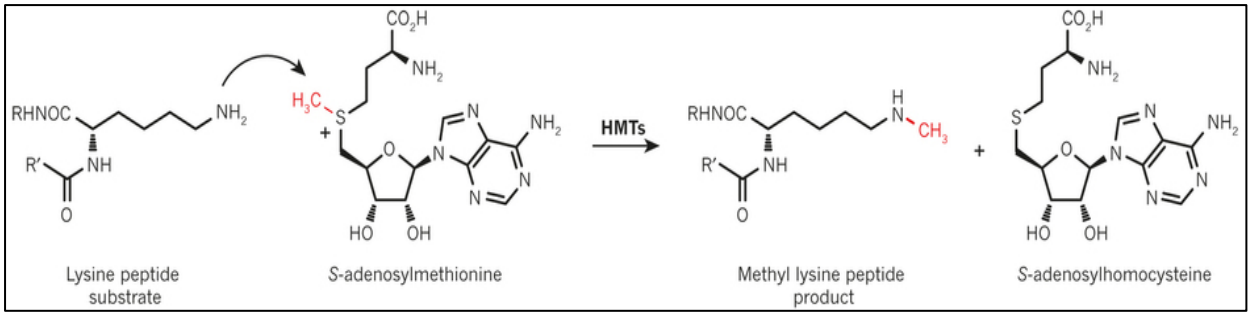


Figure 1.8 Mechanism of lysine methylation by HMTs.
From reference [77] with permission from Nature Publishing Group.

Histone methylation can be associated with transcriptional activation or repression depending on the site and level of modification. Increased level of H3K4me3 was found at highly active genes, whereas H3K4me1 and H3K4m2 were found at intermediately active group of genes. [78-80]. In contrast, the levels of H3K9me2, H3K9me3, H3K27me2, H3K27me3, and H4K20me3 are highly associated with repressed or silent genes [80-83]. In contrast, however, H3K9me1, H3K27me1 and H4K20me1 are enriched at promoter regions of active genes. In a recent report it was shown that non-nucleosomal histone H3 in budding yeast contains high levels of H3K9me1, but is devoid of H3K9me2 or H3K9me3 [84], suggesting that this modification may not have a major role in transcriptional regulation. On the other hand, H3K36me3 and H3K79me3 are found abundantly at active regions, whereas H3K36me1 and H3K79me1 did not show any major preference toward either active or silent regions.

1.2.4. Combinatorial histone PTMs

Although histone PTMs correlate with various transcriptional states of genes, their precise mechanism of action is still not fully understood. As illustrated in Figure 1.9, a diverse array of closely-spaced PTMs co-exists on a short segment of the N-terminal tail of histones. While there are few instances where a specific PTM is associated with distinct cellular function [85], the majority of studies suggest that histone PTMs exert their function through combinatorial patterns [86-88]. For example, synergic interplay between H3S10ph and H3K14ac was shown to serve a major role in marking mammalian epidermal growth factor stimulation [89]. In another study, phosphorylation at H3S10 was shown to induce increased acetylation of H3K14 by the HAT Gcn5 both *in vitro* and *in vivo* [90], suggesting a strong functional relationship between the two modifications.

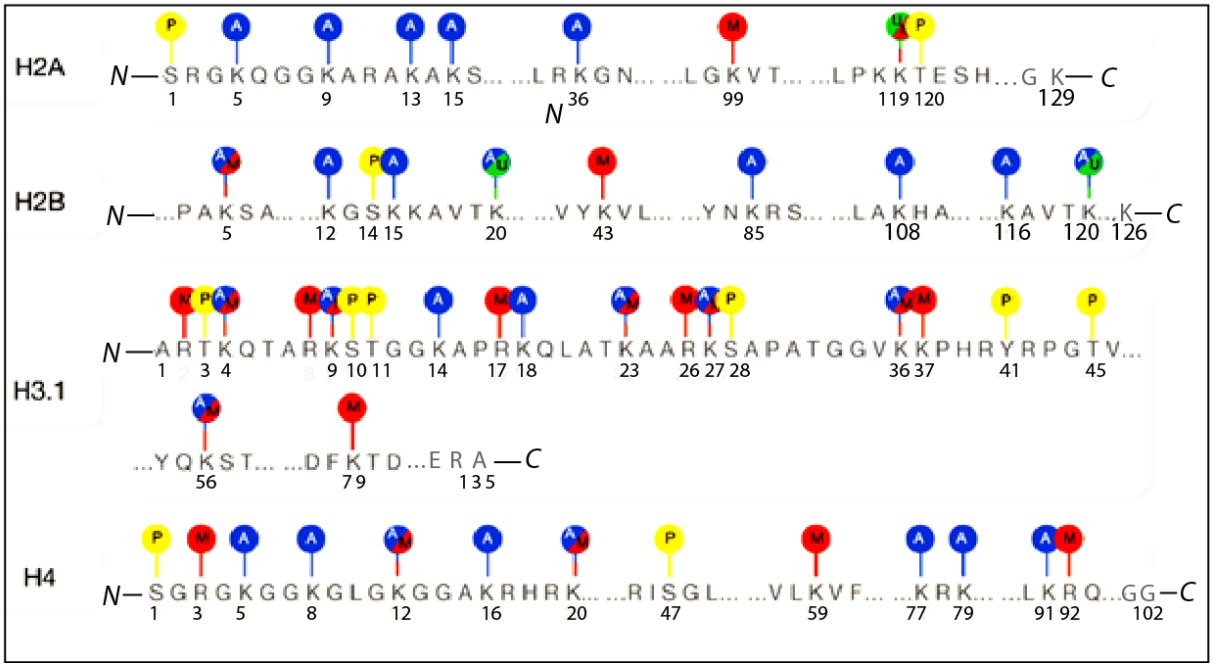


Figure 1.9 Histone tail modifications.
Adapted from reference [91] with permission from Nature Publishing Group).
Legends: acetylation (A, blue), methylation (M, red), phosphorylation (P, yellow) and ubiquitination (U, green).

Functional interactions between PTMs of different core histones have also been identified in budding yeast. It was shown that mono-ubiquitination of H2BK123 by Rad6 acts as a signal for activation of methylation of H3K4 by HMT complex known as COMPASS, a process termed as 'trans-tail' regulation of histone modifications [90]. In agreement with this observation, *S. cerevisiae* cells with H2B-K123R mutation lacked H3K4 methylation, whereas H2BK123 mono-ubiquitination was not affected in H3-K4R mutants [92].

Moreover, these studies showed that activation of H3K4 methylation via H2BK123 mono-ubiquitination is functionally linked with transcriptional silencing at telomere regions. The same type of PTM interactions has been recently identified in human cells [93].

Crosstalk between two different PTMs of the same residue has also been reported. Recently, it was shown that H3K4 methylation regulates the global level of H3K4 acetylation in budding yeast [94]. Deletion of specific subunits of COMPASS, the complex responsible for methylating H3K4, resulted in elevated level of H3K4ac. However, strains lacking Gcn5 or Gcn5/Rtt109, the HAT responsible for H3K4 acetylation, did not affect the level of H3K4 mono-, di- or tri-methylation. According to these results, the global abundances of H3K4me, particularly H3K4me2 and H3K4me3, affected the level of H3K4ac, whereas the converse was not true. These modifications are mutually exclusive and are enriched at promoter regions of actively transcribing genes with H3K4ac located upstream of H3K4me.

The different types of crosstalk described above represent few examples of how histone PTMs could act synergistically or in combination to regulate distinct cellular events. It was previously postulated that these type of PTM interactions form a ‘histone code’ [95]. According to this hypothesis, first, a histone code is established through a series of addition and removal of PTMs by histone modifying enzymes such as HATs, HMTs and HDACs. Second, the code is interpreted into meaningful biological outcome. The later process is facilitated by protein modules that can specifically recognize or ‘read’ a particular type of PTM or combination of PTMs on histones [65]. Some of the most common ‘reader’ modules are listed in Table 1.1.

Table 1.1 Histone PTM readers.

From reference [65] with permission from Nature Publishing Group

Reader module		PTM mark
	Bromodomain	Many histone Kac, (Kac)
Royal	Chromodomain	H3K9me2/3, H3K27me2/3
	Double chromodomain	H3K4me1/2/3
	Chromo barrel	H3K36me2/3
	Tudor	(Rme2s)
	Double/tandem tudor	H3K4me3, H4K20me3 H4K20me1/2, (Kme2)
	MBT	H4K20me1/2, H1K26me1/2 H3K4me1, H3K9me1/2
	PHD finger	H3K4me3, H3K4me0 H3K9me3, H3K36me3
	WD40 repeat	H3R2/K4me2, (R, Sph, Tph)
	14-3-3	H3S10ph, H3S28ph, (Sph, Tph)
	BRCT	H2AX-S139ph, (Sph, Tph)

Bromodomains (BRDs) are protein domains that recognize acetylated lysines in histones and non-histone proteins [96]. In humans, a total of 46 BRD-containing proteins have been identified, representing various types of transcriptional co-regulators and chromatin modifying enzymes [97]. Some of these proteins possess multiple BRDs; for example, the transcription co-activator TAF1 contains double BRDs; also the yeast chromatin remodeling protein Rsc4 contains tandem BRDs. Bromodomains can recognize a single or combinations of acetylated lysines in histones. The Rsc4 tandem BRDs, for example, bind to acetylated H3K14 [98]. TAF1 (also denoted as TAFII250) have two adjacent bromodomains that efficiently recognize di-acetylated histone H4 [99]. *In vitro* peptide binding assays revealed that TAF1 bromodomains have stronger affinity for the di-acetylation mark H4K5_12ac (K_d = 1.4 μM) compared to H4K8_16ac (K_d = 5.6 μM), H4K5_8_12_16ac (K_d = 5.3 μM) and H4K16ac (K_d = 39 μM), and no binding affinity for the non-acetylated counterpart. These results are, however, inconclusive as the binding assays were conducted with only limited combinations of acetylated lysines derived from the N-terminal tail of H4.

The Bromodomain and ExtraTerminal (BET) family members Brd2, Brd3 and Brd4 and Brdt are additional examples of proteins that recognize multiply acetylated lysines [100, 101]. Recently, isothermal calorimetric analysis of the interactions between Brdt and acetylated histone peptides showed that Brdt recognizes tetra-acetylated H4 peptide H4K5_8_12_16ac with a K_d value of 11.4 μM [102]. This study also showed that Brdt interacts with multiply acetylated peptides through the binding pocket of only one of its two bromodomains, suggesting a combinatorial binding property of bromodomains.

On the other hand, unlike acetylated lysines, methylated lysines can be 'read' by more than one type of protein module including chromodomains, PHD fingers, and MBT proteins (Table 1.1). One of the well-characterized chromodomain-containing proteins is HP1 (heterochromatin-associated protein 1). Analysis of HP1-peptide binding affinity using ITC (isothermal titration calorimetry) showed that it interacts with the histone marks H3K9me3 (K_d = 2.5 μM) and H3K9me2 (K_d = 7 μM), but had no detectable affinity for the non-methylated counterpart [103]. Meanwhile, it has been shown that HP1 binding to H3K9me is diminished in the presence of phosphorylated H3S10 [104]. In addition, methylation of H3K9 was inhibited in peptides containing phosphorylated H3S10. These results suggest a negative correlation between HP1 binding or H3K9 methylation and H3S10 phosphorylation.

1.2.5. Histone modifiers - 'Writers' and 'Erasers'

Histone modifications are generally mediated by two opposing groups of enzymes, namely 'writers' and 'erasers'. The 'writers' are enzymes such as HATs, HMTs and kinases that modify histones by inserting respectively acetyl, methyl and phospho groups on specific residues. The 'erasers' are enzymes such as HDACs, histone demethylases (HDMTs) and phosphatases that reverse the activities of the 'writers' by removing the corresponding PTMs. Shown in Figure 1.10a is a schematic representation of the opposing activity of writers and erasers.

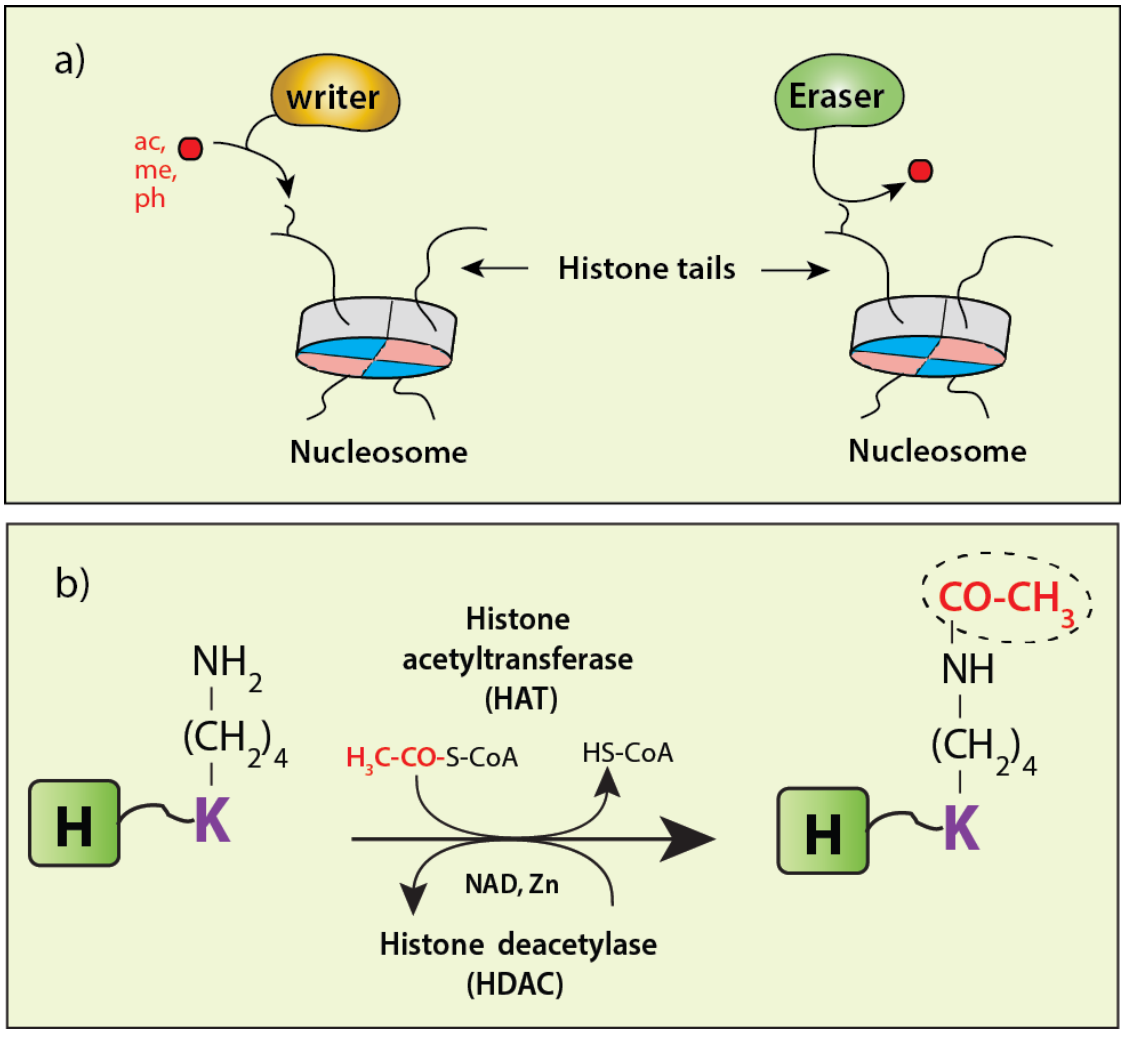


Figure 1.10 Schematic showing the reciprocal activities of 'writers' and 'erasers'.
(a) The 'writing' and 'erasing' of histone PTMs. (c) The opposing activity of HATs and HDACs toward histone lysine residues. (ac-acetyl, me-methyl, ph-phospho). Figures prepared using Adobe® Illustrator® CS6, version 16.0.0.

1.2.5.1. HATs

HATs catalyze the transfer of an acetyl moiety from acetyl-CoA to the amino group of histone lysine residues (Fig. 1.10b). Their activity toward histones neutralizes the electrostatic interaction between the positively charged amino-group of lysines and the negatively charged phosphate group of the DNA. This process relaxes the nucleosomal DNA and increases its accessibility for transcription. HATs function in a variety of DNA-associated processes [105]. Several genome-wide analyses have shown that a large proportion of acetylation marks and the HATs that installs them are localized at promoter and enhancer regions of highly expressed genes [106-108]. They are recruited to these regions via interactions with specific activator proteins [67, 109, 110]. For example, in budding yeast, the HATs Gcn5 and Esa1 were found to be recruited to the promoter regions of active genes, leading to elevated levels of acetylation at Gcn5 target sites- H3K9 and K14, and Esa1 target sites- H4K5, H4K8, H4K12 and K16 [111]. These observations suggested a strong correlation between HATs activity and transcriptional activation. In addition to histones, HATs also catalyze the acetylation of several non-histone proteins [112, 113].

Based on their catalytic domains and sequence homology, eukaryotic HATs are grouped into three major families [114, 115]: i) Gcn5-related N-acetyltransferases (GNATs), named after the founding member Gcn5; ii) the MYST family HATs, named for the founding members of this family: MOZ, Ybf2, Sas2 and Tip60; and iii) CBP/P300. Members of each family share high amino acid sequence similarity. For instance, CBP vs. P300, GCN5 vs. PCAF, MOZ vs. MORF, and TIP60 vs. MOF share respectively 87%, 89%, 87% and 65% sequence similarity [116]. The HATs Rtt109 (found in yeast) and HAT1 (found in most eukaryotes) also catalyze histone acetylation in specialized DNA-associated processes, but they share very low sequence homology with the three HAT families or with each other. TAF1, one of the components of the transcription factor TFIID, also exhibits acetyltransferase activity toward histones H3 and H4, and is required in transcriptional activation of genes. Table 1.2 shows the conservation of the different HATs between species.

Table 1.2 Conservation of HATs between species and phyla.
Adapted from reference [116], Caister Academic Press., 2011

Name used in text	Common synonymous abbreviations	*Homo sapiens*	*Mus musculus*	*Danio rerio*	*Drosophila melanogaster*	*Caenorhabditis elegans*	*Saccharomyces cerevisiae*
TIP60	HTATIP/KAT5	✓	✓	✓	✓	✓	✓
MOF	MYST1, KAT8	✓	✓	✓	✓	✓	✓
HBO1	MYST2, KAT7	✓	✓	✓	✓	–	–
MOZ	MYST3, KAT6A	✓	✓	✓	–	–	(✓)
QKF/MYST4	MORF, KAT6B	✓	✓	✓	✓	–	
GCN5	KAT2A, PCAF-B, Gcn5l2	✓	✓	✓	✓	✓	–
PCAF	KAT2B	✓	✓	✓	✓	✓	✓
CBP	CREBBP, KAT3A	✓	✓	✓	✓	–	–
P300	Ep300, KAT3B	✓	✓	✓	✓	–	–
NCOA1	SRC1, SRC-a KAT13A	✓	✓	✓	–	–	–
NCOA3	SRC3, pCIP, ACTR, AIB1, RAC3, TRAM1, KAT13B	✓	✓	✓	–	–	–
CLOCK	KAT13D	✓	✓	✓	✓	–	–
HAT1	KAT1	✓	✓	✓	✓	✓	✓
ELP3	KAT9	✓	✓	✓	✓	✓	✓

1.2.5.2. HDACs

HDACs catalyze the deacetylation of histone lysine residues (Fig. 1.10b). Their activity is historically associated with transcriptional repression and heterochromatin formation because the removal of acetylation marks from histones often leads to condensed chromatin structure. A number of HDACs are present in association with multiple protein transcription corepressor complexes [117]. One example is the Sin3 corepressor which was found in complex with the HDAC Rpd3 in *S. cerevisiae* [118] and HDAC1/2 in humans [119]. Most HDAC complexes do not directly bind to nucleosomal DNA; instead they are targeted to specific genomic loci via interaction with other DNA-binding proteins. For example, in *S. cerevisiae*, the Sin3/Rpd3 complex interacts with the DNA-binding protein called Ume6 to target the deacetylation of histone H4K5ac and repress transcription of the adjacent genes [120, 121]. The function of HDACs is also tightly coupled with dynamic changes in the levels of histone methylation. In addition to hypoacetylation, silenced genes and heterochromatin domains are characterized by high levels of H3K9 di- and tri-methylation. It could be that the deacetylation of lysine residues by HDACs serves as a prerequisite for establishing repressive histone methylation marks.

Mammalian HDACs are classified into four major classes based on sequence similarity of their catalytic domain with the corresponding yeast HDACs [121-123]. These are: Class I: HDAC-1, -2, -3 and -8; Class-II: HDAC-4, -5, -6, -7, -9 and -10; Class III: sirtuins or SIRT 1–7; and Class IV: HDAC11. The Class I HDACs are mostly localized in the nucleus and are known to catalyze the deacetylation of histones and several non-histone proteins. In

comparison, the Class II HDACs are located in both the cytoplasm and nucleus. Their catalytic domain is located close to the C-terminus. Depending on the number of deacetylase domains they contain the Class II HDACs are divided in two subgroups, Class IIa and IIb; the latter has two deacetylase domains. HDACs Class I, II, and IV require the Zn^{2+} cofactor for their deacetylase activity. Class III HDACs are commonly referred to as sirtuins. Their activity requires the cofactor known as nicotinamide adenine dinucleotide (NAD^{+}). The different classes of mammalian HDACs and their yeast counterpart are summarized in Table 1.3 below.

Table 1.3 Classification of mammalian HDAC isotypes.
From reference [122] with permission from Elsevier.
(N-nucleus, C-cytoplasm, M-mitochondia)

Class	Isoform	Yeast counterpart	Size	Co-factor	Location	Expression	Catalytic domain
(I)	HDAC1	spHos2	58	Zn^{++}	N	Ubiquitous	482 aa
	HDAC2	spClr6 scHos1	59		N		488 aa
	HDAC3	scHos2	50		N, C		428 aa
	HDAC8	scRPD3	44		N		377 aa
(IIa)	HDAC4	scHos3 scHDA1 spClr3	120	Zn^{++}	N, C	Specific	1084 aa
	HDAC5		130		N, C		1122 aa
	HDAC7		110		N, C		952 aa
	HDAC9		160		N, C		1011 aa
(IIb)	HDAC6	scHDA2	160	Zn^{++}	N, C	Specific	1215 aa
	HDAC10		70		N, C		(Inactive) 669 aa
(III)	SIRT1	scHst1 sc/spHst2 scHst3 sc/spHst4 sc/spSir2	120	NAD^{+}	N	Variable	NAD 747 aa
	SIRT2		45		C		NAD 389 aa
	SIRT3		28		M		NAD 399 aa
	SIRT4		35		M		NAD 314 aa
	SIRT5		36		M		NAD 310 aa
	SIRT6		39		N		NAD 355 aa
	SIRT7		48		N		NAD 400 aa
(IV)	HDAC11	RPD3/HDA1	39	Zn^{++}	N	Ubiquitous	347 aa

TRENDS in Microbiology

1.2.5.3. HDAC inhibitors – Emerging drugs for cancer therapy

The balance between histone acetylation and deacetylation is crucial for the normal development of cells. Maintaining this balance requires continuous regulation of the interactions of HATs and HDACs with their target sites [124, 125]. Several gene deletion studies have shown that dysregulation of the acetylation equilibrium adversely affects numerous biological processes including DNA replication, transcription, and chromatin assembly [126-128]. While normal cells have dedicated control mechanisms that regulate the steady-state dynamic levels of acetylation of histones, as well as non-histone proteins, most cancer cells have been shown to possess altered global histone modifications [129-132]. For example, a global loss in H4K16ac was found to be a hallmark of human tumor cells [133]. The H4K16ac mark was previously shown to have a pivotal role in the regulation of the structure of chromatin and its functional interactions with chromatin assembly and remodeling factors [62]. The HAT known as MOF is responsible for acetylating H4K16 [134], whereas the HDAC named SirT1, the orthologue of budding yeast Sir2, is responsible for deacetylating H4K16ac [135].

Numerous studies have reported aberrations in HDAC activities in cancer cells. Increased expression of HDAC2 was found in human gastric cancer [136]. HDAC1 was shown to be upregulated in hormone refractory (HR) cancer and targets numerous transcription factors including the tumor suppressor protein p53 [137]. Increased recruitment of HDACs has also been associated with the pathogenesis of acute promyelocytic leukaemia [138]. These abnormalities in the recruitment and activities of HDACs in different types of cancers have made them strong therapeutic targets.

As such, a great deal of effort has been made in the discovery and development of new drugs that have the potential to reverse the detrimental effects of aberrant HDAC activities. These drugs are collectively known as HDAC inhibitors (HDACis). So far, numerous HDACis have been discovered and tested for antitumor activity in either cultured cells or model organisms. However, only a few were approved or are being investigated at various levels of clinical trials. The HDACis are divided into five groups depending on their chemical structure, which are, hydroxamates, cyclic tetrapeptides, benzamides, aliphatic acids, and electrophilic ketones (see Table 1.4) [139, 140]. The hydroxamates exhibit strong inhibitory activity toward HDAC I, and II, and also IV in some cases. SAHA (also known as Vorinostat) was the first to be FDA-approved for the treatment of cutaneous T-cell lymphoma [141]. It induces global hyperacetylation of histones [142], represses telomerase activity in human lung

adenocarcinoma cells [143], and arrests the growth of several types of cancer cells [144-146]. Romidepsin (FK228), inhibitor of HDAC I and II, was also FDA-approved for treatment of cutaneous T-cell lymphoma [147]. Most recently, the FDA approved a new broad-spectrum HDACi called Panobinostat (Farydak) for the treatment of patients with multiple myeloma, a type of blood cancer [148]. Panobinostat belongs to the hydroxamate HDACi group and has been shown to have activity against HDACs Class I, II, and IV [149]. Despite these successes, the three FDA-approved HDACs described above are all pan-HDAC inhibitors, meaning that they act indiscriminately on several HDAC classes [150, 151]. In addition, these drugs have been shown to have major side effects both in the clinical and experimental settings due to their inability to discriminate between normal and cancerous cells [152-155]. Thus, target-specific HDAC inhibitors are much needed for a more efficient cancer treatment.

Table 1.4 HDACis currently under clinical trial or FDA-approved.
Adapted from from reference [139]

Group	Compound	HDAC target	Current state
Hydroxamic acid	Vorinostat (SAHA, Zolinza)	class I, II, IV	**FDA approved**
	Panobinostat (LBH589)	class I, II, IV	**FDA approved**
	Belinostat (PXD101)	class I, II, IV	phase II CT
	Abexinostat (PCI24781)	class I, II	phase II CT
	Resminostat (RAS2410)	class I, II	phase II CT
	Givinostat (ITF2357)	class I, II	phase II CT
	Dacinostat (LAQ824, NVP-LAQ824)	class I, II	phase I CT
	Pracinostat (SB939)	class I, II	phase II CT
Cyclic tetrapeptide	Romidepsin (Depsipeptide, FK228)	HDAC1, 2	**FDA approved**
	Apicidin	HDAC2, 3	Phase II CT
	Trapoxin A	HDAC1, 4, 11	ND
Benzamide	Mocetinostat (MGCD0103)	HDAC1, 2, 11	phase II CT
	Entinostat (MS-275, SNDX-275)	HDAC1, 9, 11	phase II CT
	Rocilinostat (ACY-1215)	HDAC6	phase II CT
Aliphatic acid	Valproic acid (VPA)	class I	phase III CT
	Pivanex (AN-9)	ND	phase II CT
	Butyrate	class I, IIa	Phase II CT
Electrophilic ketone	Trifluorometchylketone	ND	ND

ND – not determined, CT-clinical trial

1.3. Mass spectrometry-based proteomics in histone analysis

Early studies involving histone modifications were conducted mainly using chemical approaches. For example, Phillips used a method known as hydrazinolysis to identify acetylation on the N-terminal amino group of histone proteins [156, 157]. In this method, first N-acetyl groups were derivatized through successive reactions with anhydrous hydrazine and 1-fluoro-2,4-dinitrobenzene at pH 3. The final product from this reaction was acetic acid 2, 4-dinitrophenylhydrazide, which changed its color from yellow to brown upon treatment with an alkaline solution. Confirmation of the presence of the N-acetyl group was made by a two-dimensional chromatographic technique and by measuring the absorbance of the product and comparing it with the absorbance of a reference hydrazide derivative of acetic acid. Although this method enabled the detection of N-acetylated histones, it was not suitable for analysis of acetylations that were localized on the ε-amino group of lysine side chains (ε-N-acetyl-lysine).

Evidence showing the latter type of acetylation came from the work of Allfrey and colleagues [158, 159]. In order to identify acetylated lysines in histones, Allfrey and colleagues relied on radiolabeling techniques, where calf thymus nuclei were incubated *in vitro* with a ^{14}C-labelled sodium acetate. The labelled histones were fractionated and subjected to amino acid analysis to characterize and quantify their amino acid composition. The presence of acetylated lysines was confirmed based on the radioactivity of standard amino acid mixtures containing the reference compound, ε-N-acetyl-lysine. Alternatively, following digestion of the histones with trypsin (cleaves c-terminal to Arg or Lys amino acids) and Pronase (digests proteins to free amino acids), the peptides and amino acids were fractionated by exclusion chromatography and ion exchange chromatography. All radioactivity of the ^{14}C-acetate-labelled histones was recovered in one chromatographic peak identified as ε-N-acetyl-lysine. This method, however, was only able to detect abundant acetylations in histones H3 and H4.

Following the discovery of histone lysine acetylation by Allfrey and colleagues, a number of subsequent studies demonstrated the mechanisms through which this PTM can modify chromatin structure and transcriptional activity [63, 160-163]. Among the major advances achieved during this period is the development of antibodies that can distinguish between acetylated and non-acetylated forms of histones [164, 165]. These antibodies mainly contributed to our current understanding of the role of histone PTMs in various biological processes. Pfeffer and colleagues developed an antibody specific for a tetra-acetylated H4

peptide (a.a 1-37) [165]. Their procedure involves two major steps: Preparation of acetylated histone peptides and assessment of site-specificity of the antibody. In the first step, Pfeffer and colleagues prepared the N-terminal H4 peptide (amino acids 1-37) by dissociating the intact histone H4 isolated from calf thymus nuclei. This peptide was subjected to hyperacetylation by *in vitro* chemical reaction with excess acetic anhydride. The acetylated peptides were then separated from the reaction mixture by exclusion chromatography. The extent of acetylation of the peptide was estimated from its reaction with a fluorophore reagent called fluorescamine under conditions which limit the reaction of the reagent with free amino groups. This preparative step yielded a large proportion of the tetra-acetylated form of the H4 peptide. In the second step, Pfeffer and colleagues immunized a particular breed of rabbits with the tetra-acetylated peptides, which acted as antigens. Antiserum activity against the antigens was assessed using ELISA and further validated by peptide competition assays. Ideally, in addition to the ELISA and competition assays, the specificity of an antibody is also validated by assessing its activity toward mutant histones (e.g. histone H4 with Lys5 mutated to Arg or Ala) expressed in model organisms such as yeast. The later approach provides a more precise verification of the specificity of the antibodies.

The antibody-based approach is still a powerful tool in the chromatin field. Its applications are demonstrated in several traditional biochemical techniques including Western blot, ChIP, ChIP-DNA microarray (ChIP-on-chip), and immunofluorescence assays. Despite these achievements, however, the antibodies used in histone PTM analysis still suffer from several limitations. Site-specific antibody detection requires one specific antibody for each site that needs to be examined, and hence lacks the capability to detect co-occurring histone PTMs (e.g. H3-K9me3_K14ac). Recent findings show that antibodies exhibit promiscuity when two or more histone PTMs are placed too close together [87]. Many antibodies also lack the desired specificity to distinguish subtle differences in the chemical composition of PTMs (e.g. K9me2 and K9me3). In addition, antibodies are limited in their ability to determine PTM stoichiometries, which have proven valuable in monitoring PTM dynamics in various biological processes [166, 167].

Meanwhile, in the last decade, mass spectrometry (MS) has emerged as the most powerful analytical tool for identification and quantification of histone PTMs. Many high-throughput MS-based approaches have been developed for this purpose, and their combined effort has dramatically improved our understanding of the role histone PTMs play in various chromatin-associated processes. Histone PTM analyses using MS have so far been performed in three

different approaches: bottom-up, middle-down and top-down (Fig. 1.11) [168, 169]. The most widely used procedure is the bottom-up approach. In this approach, histone proteins are derivatized using deuterated acetic anhydride or propionic anhydride, digested into smaller peptides (usually less than 30 amino acids) using trypsin and analyzed by reverse-phase (RP)-LC-MS/MS. Identification of histone PTMs is achieved with the help of database search tools via peptide-spectrum matching processes [170, 171]. This approach is simple, straightforward and provides a relatively better quantification of site-specific changes in histone PTMs. However, the method is limited in terms of the length of peptide that can be analyzed and as a result long-range combinatorial PTMs and potential crosstalk between them cannot be investigated.

In the middle-down approach, histone proteins are digested into relatively longer peptide segments (upto 50 amino acids length) using proteases such as Glu-C and Asp-N, and separated on hydrophilic or reverse-phase analytical columns and analyzed by ETD-MS/MS [169]. The application of ETD in this approach improves the dissociation of the large and highly charged peptides, thereby increasing the number of PTMs per peptide that can be investigated. One major advantage of the middle-down approach is that it enables identification of multiple co-existing PTMs on the N-terminal tails of histones. However, due to the complexity of the ETD-MS/MS spectra, site-specific quantification of the individual PTMs is very challenging. Several computational tools have so far been developed to aid the deconvolution of such MS/MS spectra [172] but they often require extensive manual validation. Unlike the two previous approaches, the top-down approach does not require digestion of histones into peptides; instead the intact protein is fragmented in the MS by ETD. One major advantage of this technique is that it allows investigation of combinatorial histone PTMs at the proteome level, which is perhaps the ideal way to study the role of histone PTMs on chromatin architecture. However, the full application of this technique still demands major improvements in the following three areas. First, just like the middle-down approach, fragmentation of the intact histones generates very complex and composite MS/MS spectra and there are only few computational tools that can deal with such spectra. Even, the existing tools require rigorous manual validation and often a complete fragmentation spectra is rarely achieved. Second, very poor separation of histones into individual proteoforms is achieved using currently available chromatographic technologies. For this reason, some investigators choose to directly infuse the intact protein into the MS without chromatographic separation. This approach, however, generates MS/MS spectra containing fragments of a

variety of proteoforms, which complicates the data analysis. Third, some of the most widely used mass spectrometers are better suited for analysis of smaller polypeptides (usually less than 5 KDa) and they have lower performances when it comes to larger polypeptides. In addition, a complete dissociation of the proteoforms using the available fragmentation techniques is so far not achievable.

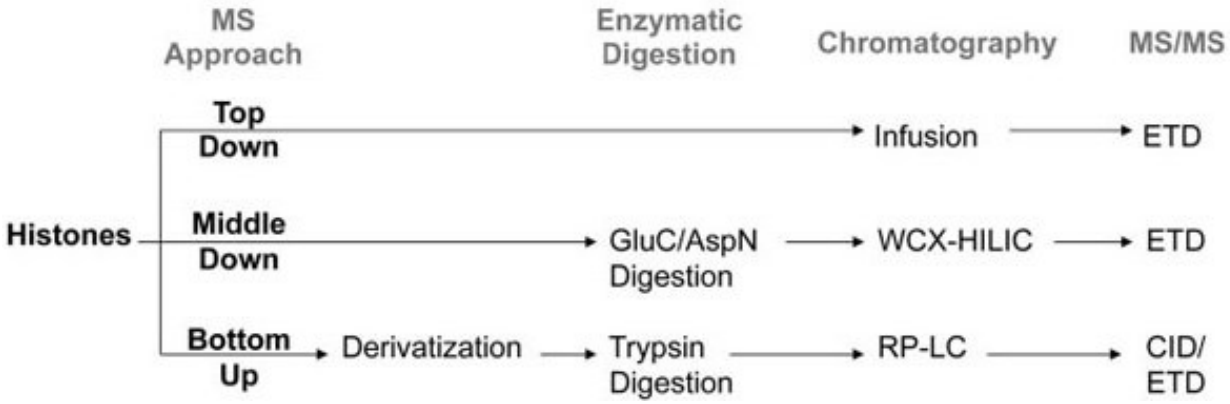


Figure 1.11 Bottom up, middle down, and top down MS approaches.
From reference [169] with permission from John Wiley and Sons.

WCX-HILIC: weak cation exchange hydrophilic interaction liquid chromatography
RP-LC: Reverse phase liquid chromatography; ETD: Electron transfer dissociation; CID: Collision induced dissociation

Despite their inherent limitations, the three aforementioned MS-based approaches have so far enabled the characterization of over fifteen different types of PTMs on the core histones [57]. The majority of these PTMs are localized on the unstructured N-terminal histone tails which extend outside the nucleosome core. Moreover, we can now investigate combinatorial PTM marks, as well as short- and long-range PTM crosstalk within the same histone molecule or even across different histones. These approaches offer several advantages over the traditional antibody-based analysis including good reproducibility, higher specificity, and ability to rapidly analyze numerous histone PTMs in a single LC-MS/MS experiment.

1.3.1. Basic steps in histone PTM analysis by mass spectrometry

1.3.1.1. Proteolytic digestion

Enzymes that cleave peptide bonds in proteins are called proteases [173]. Trypsin is the most commonly used enzyme for protein digestion in the bottom-up approach. It cleaves proteins on the C-terminal side of lysine or arginine residues, except when either amino acid is followed by proline or chemically modified. Trypsin has the advantage of high proteolytic specificity, wide sequence coverage and relatively lower auto-proteolysis (self-cleavage). It generates relatively short peptides, usually 5 to 30 amino acids long. Since histone proteins contain multiple arginine and lysine residues (20 to 25 %), their cleavage using trypsin generates peptides too short to be efficiently analyzed by MS. Moreover, tryptic digestion of histones produces varying lengths of heterogeneously modified peptides. A single lysine residue may end up being present in more than one type of tryptic peptide. This complicates the assignment and quantitation of histone PTMs. To circumvent these problems researchers have developed a chemical derivatization technique which helps to generate longer and more hydrophobic tryptic histone peptides. The derivatization of histones is commonly performed using a propionic anhydride ($(CH_3CH_2CO)_2O$) reagent which serves to transfer a propionyl group (CH_3CH_2CO-) to the ε-amino group of unmodified or mono-methylated lysine residues. This modification results in a shift in the mass of lysine residues by +56.02 Da. Moreover, propionylation of lysines effectively blocks the activity of trypsin at these sites, rendering peptides suitable for analysis by MS.

Alternatively, peptides of longer lengths can be obtained using endoproteinases such as Asp-N and Glu-C. Asp-N hydrolyzes peptide bonds mainly on the N-terminal side of aspartic acid, whereas Glu-C cleaves on the N-terminal side of mainly glutamic acid. Asp-N digestion of, for example, histone H4 generates an N-terminal peptide that is 23 amino acids long consisting of five internal lysines and one arginine residue (2-GRGKGGKGLGKGGAKRHRKVLRD-24). On the other hand, Glu-C digestion of histone H3 generates an N-terminal peptide 49 amino acid long and consisting of eight lysines and seven arginine residues (2-ARTKQTARKSTGGKAPRKQLATKAARKSA PATGGVKKPH RYRPGTVALRE-50). These two proteases are widely used in studies involving the characterization of combinatorial histone modifications and long-range PTM crosstalk. However, the complexity of PTM isoforms generated from Asp-N or Glu-C digestion of histones presents a major challenge for PTM identification using the existing bioinformatics tools.

There are two major approaches of proteolytic digestion [173]: In-gel and in-solution digestions. In the former approach, histone proteins are first resolved by gel electrophoresis then stained using MS compatible reagents such as coomassie blue or silver staining reagent. Then the gel bands corresponding to the proteins of interest are excised and cut into small pieces for proteolytic digestion using trypsin [174]. The peptide digests are extracted from the gel bands using high concentrations of organic solvents such as methanol or acetonitrile. This approach is not widely used for histone analysis because digestion of histones using trypsin without derivatization of the unmodified and mono-methylated lysines results in peptides too short to be detected by MS. On the other hand, the derivatization of histones inside a gel band is relatively less efficient.

In the later approach, the protease is directly added to the protein solution [175]. In typical large scale experiments, tryptic digestion is preceded by several steps including (1) protein denaturation using MS-compatible surfactants (e.g. urea, SDC (Sodium deoxycholate), ProteaseMAX) or by heating to open up globular proteins; (2) reduction using reagents such as DTT (dithiothreitol), TCEP (tris(2-carboxyethyl)phosphine)), or β-mercaptoethanol to break disulfide bridges (or S-S bonds); (3) alkylation using IAA (iodoacetamide) or CAA (chloroacetamide) reagent to block the reformation of disulfide bonds. However, histone proteins are relatively small and do not have that many cysteine residues that could form disulfide bridges, and as a result the denaturation, reduction, and alkylation steps are not required in histones analyses.

1.3.1.2. RP-HPLC separation

Reverse-phase high performace liquid chromatography (RP-HPLC) is the most commonly used separation technique for histone peptide or protein mixtures. RP-HPLC separates molecules based on their hydrophicity [176, 177]. This is achieved through continuous partitioning of analyte molecules between two different phases: a mobile phase and an immobilized hydrophobic stationary phase. The mobile phase is a binary mixture of aqueous and polar organic solvents. Both solvents are maintained at low pH by adding a small amount of organic acid (commonly 0.2% formic or acetic acid). The low pH aids in the separation and protonation of peptides or proteins. Acetonitrile is the most frequently used mobile phase organic solvent.

On the other hand, the stationary phase is usually packed in a column made from narrow bore stainless steel tube or fused silica capillary. Most RP columns are packed with hydrophobic alkyl chains containing four (C_4), eight (C_8), or eighteen (C_{18}) carbons [178]. The alkyl

chains are chemically linked to porous, spherical silica particles with internal diameter (i.d.) of 3 to 10 μm and pores of about 100 or 300Å diameter. For high grade separations, the stationary phase is packed in a column with a very narrow i.d, usually between 75 and 200 μm. The ultra small diameter of the columns helps to improve sensitivity by reducing excess dilution of the analyte by the mobile phase solvents. High performance and durable RP columns are obtained when the stationary phase is uniformly distributed, since this will create a homogeneous hydrophobic surface on which peptide and protein molecules are efficiently adsorbed [179]. For separation of intact histone proteins, analytical columns packed with shorter hydrocarbon chains (C_4 or C_8) and larger particle pore sizes (300 Å) are highly recommended (see Fig. 1.12). This is because large proteins form very strong hydrophobic interactions with the stationary phase, and thus are difficult to elute from the column. Such proteins are also too large to enter pores of smaller diameter. In contrast, separation of histone peptides is most efficient using RP columns packed with the *n*-octyldecyl (C_{18}) hydrocarbon chain.

Elution of peptides or proteins from the RP column can be achieved via isocratic or gradient approach [180, 181]. In the isocratic approach, the concentration of the organic solvent in the mobile phase is kept constant, whereas in the later approach the amount of the organic solvent is slowly increased till it reaches a desired level. Moreover, in isocratic elution, the same mobile solvent strength is used for both the less hydrophobic and strongly hydrophobic molecules, and as a result efficient elution of mainly one of the molecules is achieved. This approach is less common in histone analyses because peptides generated from tryptic digestion of histones exhibit wide variations in their hydrophobicity. Comparatively, in gradient elution, as the amount of the organic solvent increases, the adsorbed proteins or peptides are eluted in order of least to most hydrophobic molecules. Most researches use the gradient approach as it allows efficient separation, even for highly complex protein and peptide mixtures.

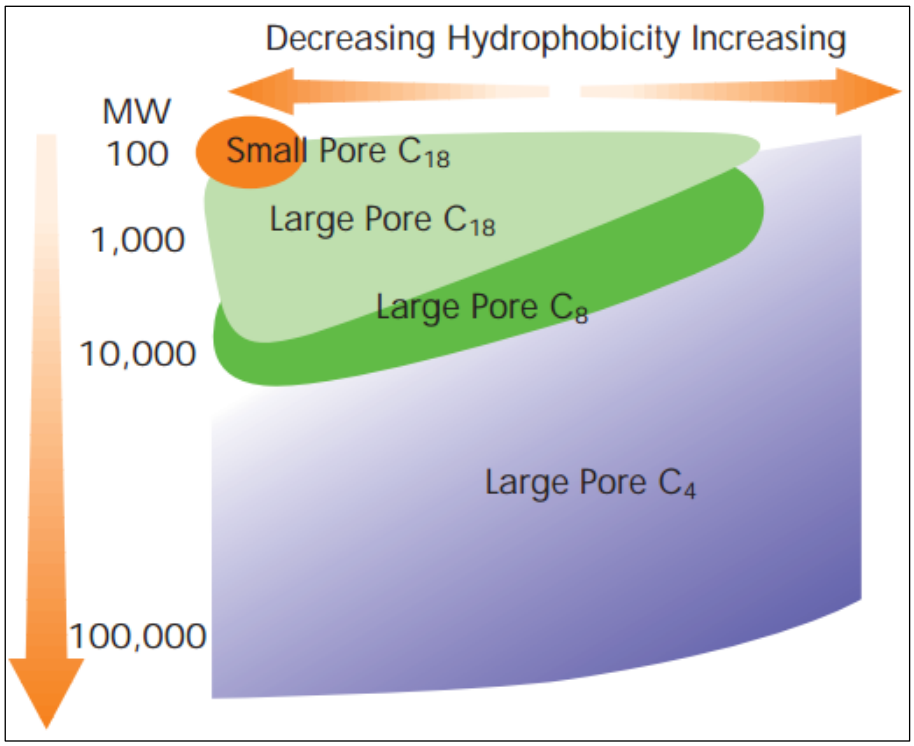


Figure 1.12 Recommended RP particle size versus protein MW.
From reference [182]

1.3.1.3. Sample ionization

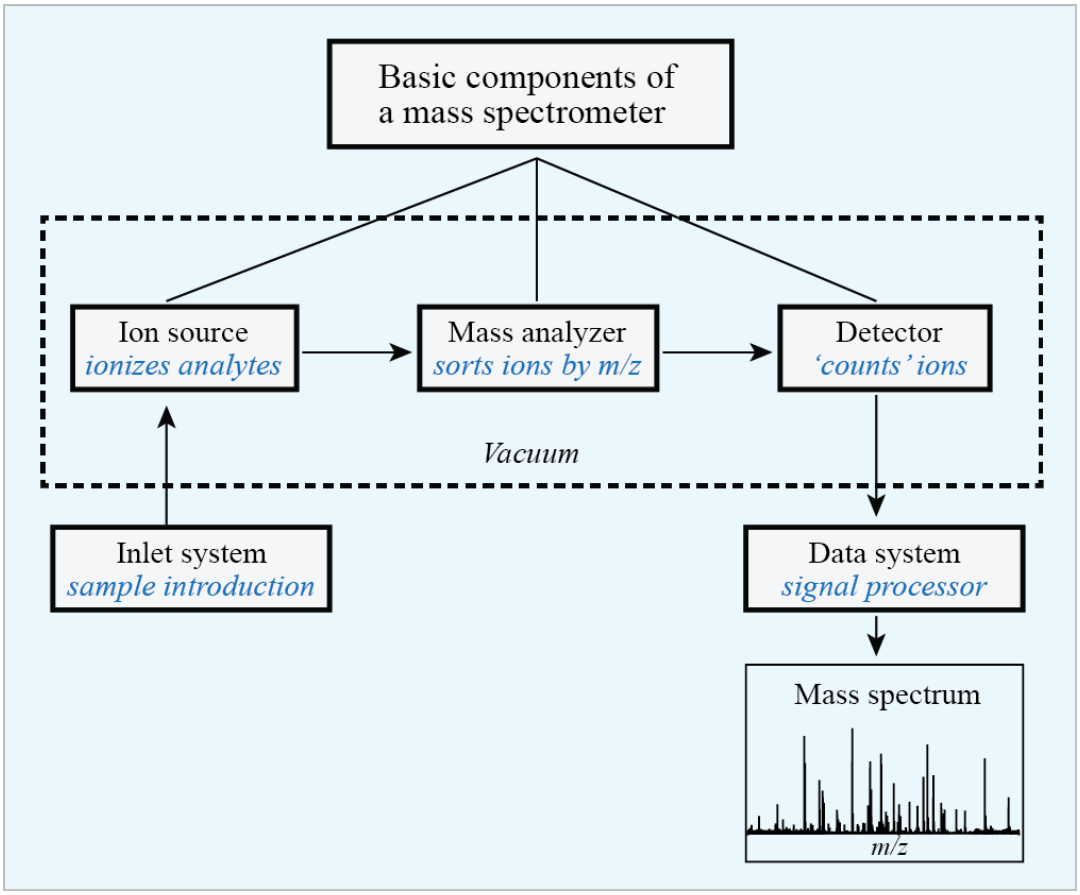
Any mass spectrometer consists of three basic components: (1) an ion source, (2) a mass analyzer, and (3) a detector. An illustration of the three components is shown in Figure 1.13a. The ion source is the section where ionization of analyte molecules takes place. Based on their effects on the analyte currently available ionization techniques are generally grouped into two major categories: hard and soft ionizations [183]. The hard ionization technique, in addition to forming ions of the analyte, results in breaking of chemical bonds. One common example of such technique is the electron ionization (EI) where gaseous analyte molecules are bombarded with a beam of electrons to generate radical cations [184]. This technique is commonly used in analysis of volatile and thermally stable organic compounds. The EI is a standard ionization technique in GC-MS instruments.

Comparatively, the soft ionization technique allows formation of ions of analyte molecules with little to no breakage of chemical bonds. The two most common soft ionization techniques used in analysis of biomolecules are MALDI and ESI. In MALDI, ionization of analyte molecules is achieved via laser irradiation of co-crystallized analyte-matrix mixtures [185]. The matrix serves to absorb the laser energy and transfers proton to the analyte,

thereby ionizing the analyte. This technique, in addition to low-mass ions, allows ionization of intact proteins larger than 100 kDa [186]. MALDI is widely used in mass fingerprinting analysis of peptides, nucleic acids and lipids.

The generation of analyte ions using the ESI technique was first described by J. B. Fenn [187]. It is the most commonly used ionization technique for peptides and proteins analyses. The exact mechanism of the ESI process is not fully understood; yet, several studies have postulated on the general behavior of analyte ions during the ESI [188], which is illustrated in Figure 1.13b . Samples are dissolved in a buffer similar to the mobile phase solvents and are directly sprayed across an electric field generated between the tip of the spray needle and a counter-electrode. The electric field helps to form highly charged liquid droplets consisting of the analyte molecules (denoted as 'M' in Figure 1.13b). A beam of nebulizing gas, usually nitrogen, passing across the tip of the spray needle evaporates the excess solvent leaving behind a solvent-free protonated analyte droplet with a diminished size. The decrease in droplet size results in increased repulsive force between neighboring like charged analytes. When the repulsive force reaches the Rayleigh limit (the maximum repulsive force that can exist between like charges before they overcome cohesive forces), the large analyte droplet disintegrates into smaller droplets due to the effect known as 'coulombic explosion'. The analyte ions formed in this process carry multiple charges, and as a result they have a small mass-to-charge ratios. This property of the ESI method allows analysis of molecules that have a larger mass than the upper mass limit of the mass spectrometer.

a)



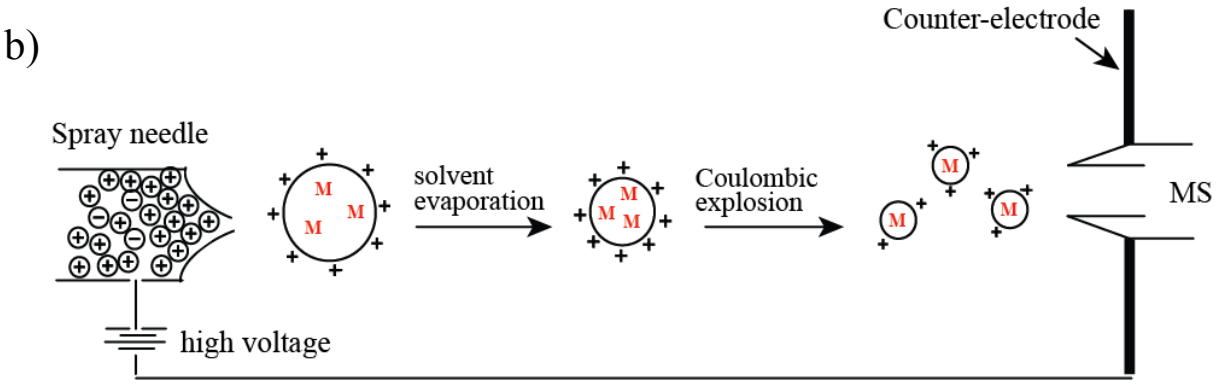


Figure 1.13 Basic components of mass spectrometry.

(a) Basic components of a mass spectrometer, (b) the ESI ionization process (adapted from Reference [188] with permission from John Wiley and Sons)

1.3.1.4. Mass analysis

A mass spectrometer separates the mass-to-charge ratio (*m/z*) of ionized (charged) analytes following their transmission through a mass analyzer. Most MS instruments are named after their mass analyzer. Some of the most common mass analyzers used in peptide/protein analyses include Quadrupole (Q), Quadrupole Ion Trap (QIT), Time-of-Flight (ToF), Fourier Transform Ion Cyclotron Resonance (FTICR) and Orbitrap [189, 190]). Each mass analyzer has its own distinct features and its own advantages and limitations. There are a number of characteristics that are used as a basis to compare the performances of the different mass analyzers. Some of these characteristics are mass resolving power, mass accuracy, dynamic range, and mass range. The mass resolving power is defined as M/ΔM, where M represents the ion m/z, and ΔM is the difference in m/z between two adjacent peaks in a mass spectrum [188]. For example, a mass analyzer with a resolving power of 1000 can effectively resolve ions of m/z 200.0 and m/z 200.2. The resolving power is also defined as the peak width at full width half maximum (FWHM). This definition is based on a single well-resolved peak where the M in the M/ΔM equation represents the m/z at the apex of the peak and ΔM represents the peak width at half maximum height of the peak. The mass accuracy represents the mass measurement error expressed in Dalton or part per million (ppm). It is determined by dividing the difference between the measured and the theoretical m/z by the theoretical m/z. The dynamic range is the number of orders of magnitude of analyte concentration over which the MS signal response is linear. The mass range is the range of m/z values over which a mass analyzer is able to detect ions or record a mass spectrum.

The Quadrupole mass analyzer is composed of four parallel, circular electric rods (Fig. 1.14 a) [191, 192]. Ions are filtered in the quadrupole based on the stability of their trajectory in a combined DC (direct current) and RF (radio frequency) potential applied to each of the opposing pair of rods. The applied potential can be adjusted to pass only a selected m/z ratio or mass range. This property of the quadrupole makes it particularly suitable for targeted MS analysis. However, the Quadrupole is limited to low resolution (usually unit resolution) and low mass accuracy MS analysis. On the other hand, the QIT (Quadrupole Ion Trap) analyzer is distinct in that it allows dynamic trapping and accumulation of large amounts of ions in a quadrupole device (Fig. 1.14b) [193]. The trapped ions can be manipulated by varying the applied RF potential; ions of a particular m/z range can be selectively ejected for detection or allowed to react with inert gases for the purpose of fragmentation. Despite such capabilities, the performance of the QIT is often limited by the phenomenon called ‘space-charge effects’

(characterized by ion coalescence, ion-ion repulsion), which is due to the accumulation of a large amount of ions in the trap. The space-charge effect severely affects the dynamic range and the mass accuracy of ion trap analyzers. To circumvent this problem, modern MS instruments are designed to have automatic ion-gauging mechanisms to control the amount of ions entering into the trap.

The TOF mass analyzer uses the differences in transit time through a high vacuum field-free region to separate ions of different m/z (Fig. 1.14c) [194]. Lighter ions travel faster than heavy ions, and thus reach the detector sooner. Unlike the quadrupole analyzers, all ions entering the TOF field-free region reach the detector. Moreover, the TOF analyzer has a unique capability of detecting ions with high mass range, even up to 10^6 Da [195, 196].

The FTICR analyzer is famous for its capability to provide superior resolving power (usually $> 10^6$ FWHM) and very high mass accuracy (sub-ppm errors are achievable) [197]. Mass analysis in the FTICR-MS is conducted in a very small closed analyzer cell (usually cube-shaped) which is kept under high vacuum and inside a strong magnetic field produced by a superconducting magnet (Fig. 1.14d). A combination of DC and RF potentials applied to the analyzer cell allows the trapping of ions at the center of the cell. For detection, the trapped ions are excited simultaneously by a pulse of RF electric field applied perpendicularly to the magnet. This causes the ions to move in a circular motion at a frequency known as the cyclotron frequency, which is inversely proportional to the m/z ratio. Each ion of a certain m/z is characterized by a particular cyclotron frequency, which is the basis for the separation and detection of ions in the FTICR-MS. Although the FTICR provides ultra-high resolution data, its capabilities are limited by the relatively slow acquisition rate and low dynamic range. The analyzer cell is also susceptible to the 'space-charge effect' due to the accumulation of large amounts of ions in a very limited space.

The Orbitrap is one of the latest and the most widely used mass analyzer in current proteomics studies. It has undergone a dramatic improvement since its introduction by Alexander Makarov in 2000 [198]. Like the FTICR and QIT analyzers, the Orbitrap functions as an ion trapping device. In contrast, however, the Orbitrap does not require a magnetic field or an RF potential for trapping ions or mass analysis. Instead, ion separation is achieved based on the motion of ions in an electric field applied between an outer barrel-like electrode and a central electrode (Fig. 1.14e). The ions rotate around the central electrode and at the same time oscillate left and right along the horizontal axis of the electrode. These

motions of the ions generate three characteristic frequencies: a rotational frequency, a radial frequency, and an axial frequency. Of the three, the axial frequency is used in mass resolving, since this frequency is dependent only on the m/z ratio of the oscillating ions and the applied potential between the central and outer electrode (which is held constant). The axial frequency is inversely proportional to the square-root of the ion m/z. The oscillatory motion of ions along the horizontal axis of the central electrode induces an image current on the outer electrode. This image current is subsequently converted into an m/z spectrum via Fourier transformation. At the current stage, the Orbitrap allows mass analysis at resolutions of several hundred thousand (the latest Orbitrap Fusion, Lumos, has a resolving power up to 500,000 (FWHM) at *m/z* 200). Mass accuracies in the low ppm to sub-ppm ranges are routinely achieved using the Orbitrap.

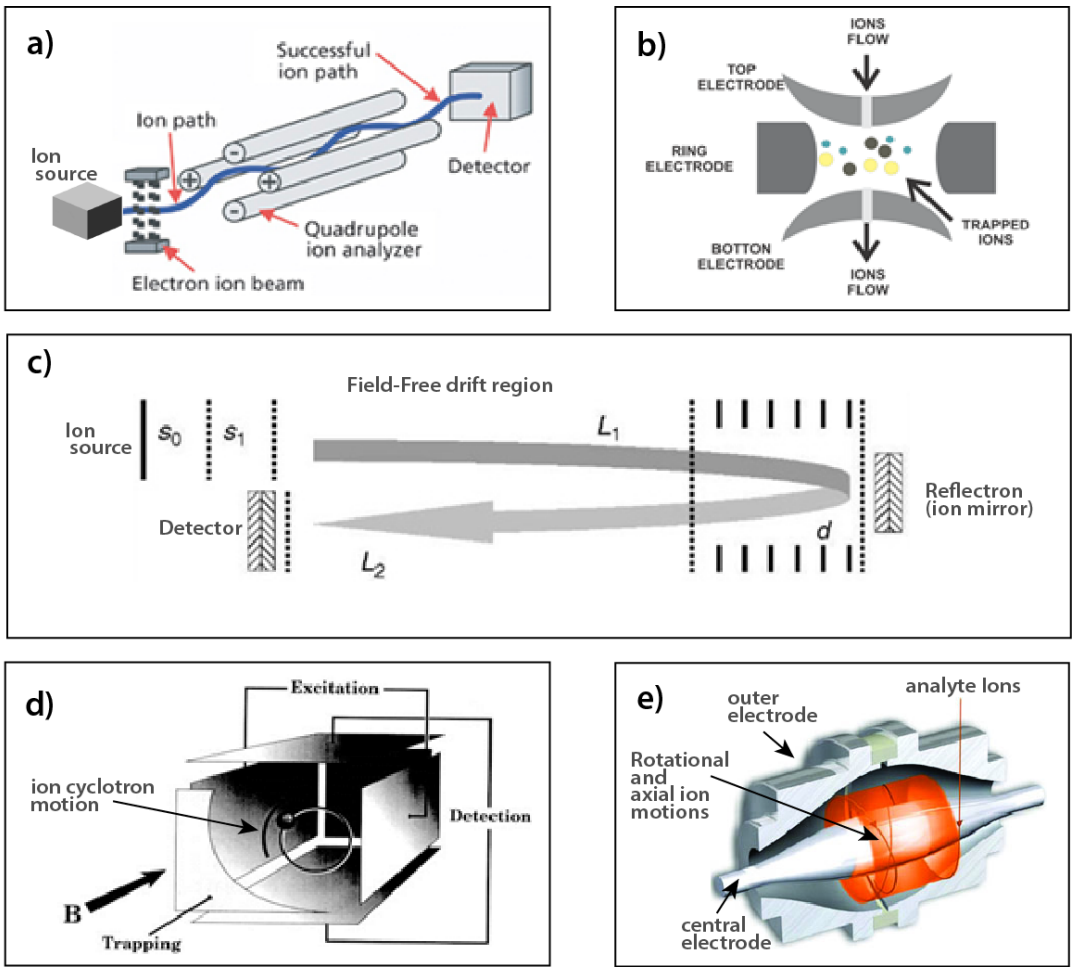


Figure 1.14 Schematics of some of the most common mass analyzers.

(a) Linear quadrupole, (b) Ion trap quadrupole, (c) TOF, (d) FTICR, and (e) Orbitrap mass analyzers. Adapted from reference [199]

1.3.1.5. Tandem mass analysis

Multiple stages of mass analysis (MSn) can be performed either in a single mass analyzer with each stage of analysis separated in time (tandem-in-time) or in more than one mass analyzer aligned in tandem (tandem-in-space) [189]. Tandem-in-time mass spectrometers have the capability to store ions and analyze them selectively at different times. The FTICR and QIT are equipped with such capability. In FTICR, in the first stage of mass analysis (MS1), ions of a certain *m/z* range are simultaneously selected and excited by applying an RF potential. The excited ions are detected as they undergo a cyclotron motion in the ICR cell. In the second stage of mass analysis (MS2), the ions are selectively excited and subjected to fragmentation via collision induced dissociation [197]. The MS1 analysis generates precursor or parent ion mass spectrum, whereas the MS2 (also denoted as MS/MS) analysis produces fragment or product ion mass spectrum.

On the other hand, tandem-in-space instruments have two or more mass analyzers configured in tandem. In one type of configuration, multiple identical mass analyzers are arranged in tandem. Typical examples of such instruments are the Triple Quadrupole (QQQ) and the TOF/TOF analyzers. In the second type of configuration, different mass analyzers are arranged in tandem, thereby forming hybrid mass spectrometers. Few examples of such instruments are Q-TOF, LTQ-Orbitrap, and Q-Exactive Plus, which are hybrid of, respectively, Quadrupole and TOF, Ion Trap and Orbitrap, and Quadrupole and Orbitrap analyzers. One important advantage of the tandem-in-space instruments is that they allow dissociation of the precursor ions in a specialized sector. This is particularly useful for the analysis of complex mixtures of peptides because the division of labor between the different analyzers improves the acquisition rate and minimizes the adverse effect of ion coalescence and the space-charge phenomenon observed mainly in tandem-in-time instruments.

Mass spectrometers are equipped with specialized techniques for fragmentation of peptides or proteins. Some of the common techniques are collision-induced dissociation (CID), High-energy collision dissociation (HCD), and electron transfer dissociation (ETD). Peptide dissociation using these techniques generates two categories of fragment ions (see Fig. 1.15): those that are charged on the N-terminal side are denoted as a-, b- and c-ions, whereas C-terminal charged fragments are termed x-, y- and z-ions.

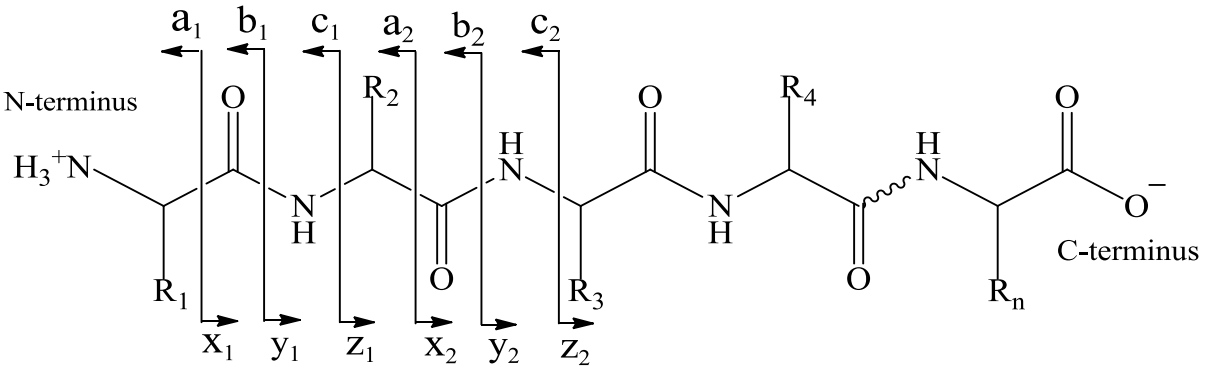


Figure 1.15 The Roepstorff nomenclature for peptide fragment ions.
From reference [200] with permission from John Wiley and Sons.

CID is by far the most commonly used method in peptide sequencing by MS/MS. In the CID mode, multiple collisions of a precursor peptide with a neutral gas molecule (e.g. Helium) induce a vibrational energy which redistributes along the backbone of the peptide ion [201]. The vibrational energy rapidly increases the internal energy of the peptide ion, which subsequently results in dissociation of the amide bonds along the peptide backbone. CID fragmentation primarily generates b- and y-ion fragments. Occasionally, the b-ion fragments are subject to lose of a carbonyl group (C=O, ≈ 28 Da), thereby generating a-ion fragments. Despite its widespread use, the CID mode is limited to analysis of relatively short peptide segments (usually less than 30 amino acids) and low-charged peptides (+2 and +3 charge states). Moreover, abundant internal fragmentation and loss of labile PTMs such as phosphorylation are frequently observed using this technique. One very common alternative fragmentation technique is the HCD [202, 203], which is basically a beam-type, high energy CID. Fragmentation by HCD also generates b- and y-ion fragments. Meanwhile, due to the high energy employed in this technique, HCD fragment spectra are dominated by y-ions; and the b-ions are usually dissociated into a-ions and other internal fragments. The HCD MS/MS spectra are analyzed at high resolution, high mass accuracy in the Orbitrap analyzers, and as such they provide higher quality peptide sequence information.

ETD is different from the CID-based methods in many ways. In this mode, first, radical anions are generated from a neutral molecule (e.g. fluoranthene, anthracene) in a separate chemical ionization source attached externally to the mass spectrometer [204]. The radical anions are then transferred to the fragmentation cell where the peptides are stored. In the cell,

the anions transfer an electron to multiply protonated peptides and induce fragmentation of the peptide backbone. This process leads to cleavage of the N─Cα bonds, thereby generating c- and z-ion fragments. ETD is particularly suited for fragmentation of highly charged and large peptides, even intact proteins. Moreover, in contrast to CID, ETD enables efficient analysis of labile PTMs such as phosphorylation and glycosylation.

1.3.1.6. Peptide identification

Classical approaches for histone PTM identification or detection involves Edman degradation, amino acid analysis coupled with thin-layer chromatography, and site-specific antibody-based assays. These approaches, however, are hampered by several technical limitations including the requirement for large amount of starting materials, inability to provide comprehensive and quantitative information, and due to the time-consuming, labor-intensive nature of the methods. On the other hand, MS-based approaches use database search algorithms to interpret the rich sequence information available in peptide fragment (MS/MS) spectra to identify site-specific modifications. In this approach, every MS/MS spectrum obtained from the collisional dissociation is compared with fragment ion mass values of *in silico* peptide sequences generated by applying appropriate cleavage rules in the sequence database [205, 206]. Typical peptide-spectrum matching (PSM) using search algorithms starts by setting appropriate search parameters on the user interface. As an example, the Mascot MS/MS ions search user interface is shown in Figure 1.16. The search begins by selecting a database (for instance - SWISS-PROT, NCBInr or the protein sequence, if known) to be searched. Prior knowledge of the protein sequence greatly simplifies the identification process by eliminating the need to look for possible matches in a database containing a large number of protein sequences. Next, the enzyme used for cleavage (e.g. trypsin, chymotrypsin etc.) is selected and maximum values for the number of missed cleavages that may result from incomplete digestion are specified. The variable and fixed (if any) PTMs that need to be considered in the search can be selected. PTMs result in a shift in mass of fragment ions consisting of the site(s) of modification compared to the corresponding fragment ions of the unmodified counterpart. This mass difference can be used to assign the identity of a PTM in a peptide. For example, acetylation of lysine side chains in histones result in an increase in mass of the residue by 42.02 Da, and mono-, di-, and tri-methylation result in an increase in mass by, respectively, 14.01, 28.02, or 42.06 Da. Some ambiguity can be encountered when identifying a trimethylated and an acetylated lysine residue since the mass difference between these two PTMs is only 0.036 Da. Meanwhile, this difference can

easily be resolved by analyzing the MS/MS spectra of peptides consisting of these PTMs at high mass accuracy and high resolution. Then, a desirable mass tolerance (in ppm or Da) for both the precursor peptide and the fragment ions is set based on the known mass accuracy of the MS analyzer. Finally, the raw MS/MS data are uploaded and submitted for database searching. In general, the success of peptide/PTM identification by the PSM approach depends on the quality of the MS/MS data, the performance of the search algorithm used by the search engine and the database accuracy.

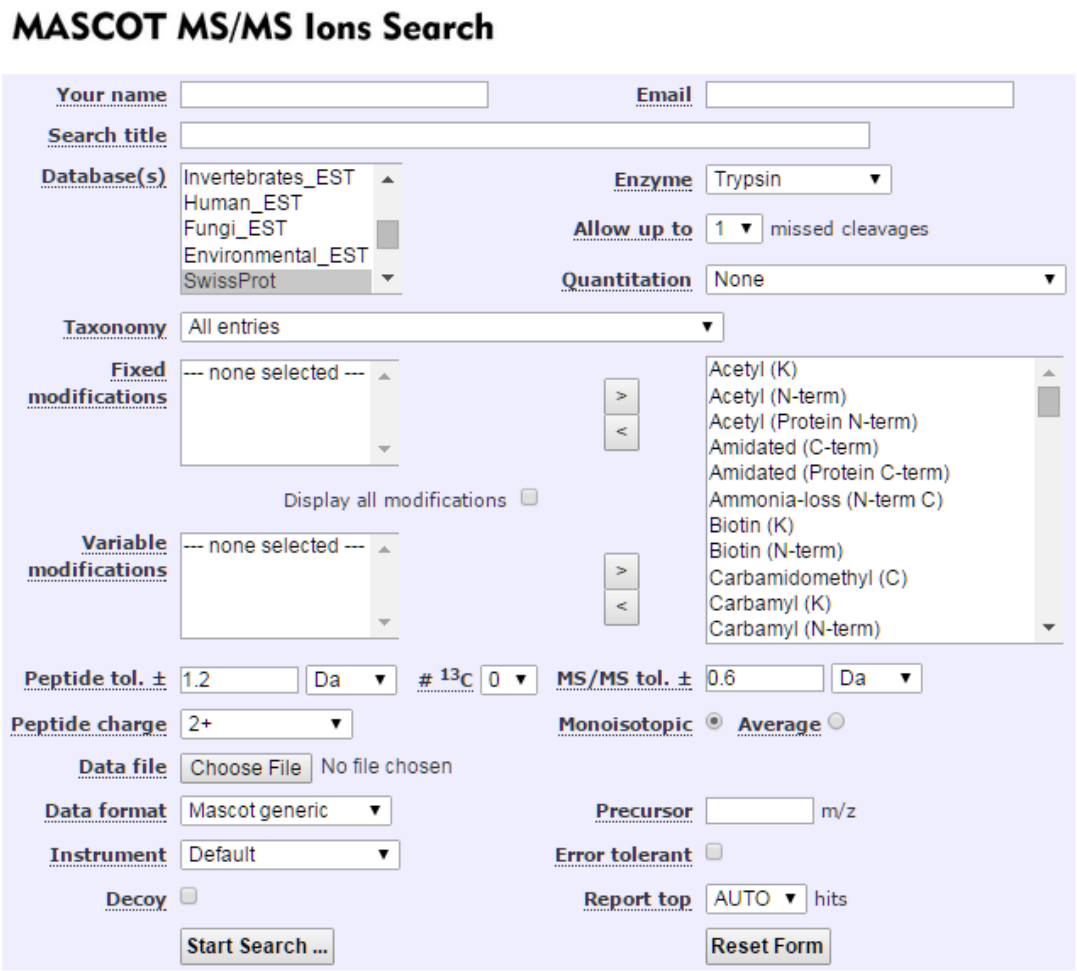


Figure 1.16 The user interface for Mascot MS/MS database search engine.
(Source: http://www.matrixscience.com/cgi/search_form.pl?FORMVER=2&SEARCH=MIS, July 2015.)

Most MS-based approaches for histone PTM analysis employ database search engines for the identification of modified residues. However, due to their complex modification patterns, histones are often difficult to analyze using conventional search engines. Yuan et al., have recently investigated the performance of eight different search engines (pFind, Mascot, SEQUEST, ProteinPilot, PEAKS, OMSSA in COMPASS, X!Tandem in TPP, and Andromeda in MaxQuant) for analysis of histone PTMs [171]. A comparison was made based on proportion of confident identifications, proportion of overlapp of confident identifications (to evaluate the similarity of search engines), search time and data storage space. The results from this study found Mascot and PFind to be the best search engines for the analysis of histone modifications as both identified relatively higher number of confident MS/MS spectra. However, almost all of the above search engines performed poorly in identifying combinatorial histone modifications. The combinatorial search consisting of PTMs - propionylation, acetylation, phosphorylation, and mono-, di-, tri-methylation resulted in large data storage space and also more false positive identifications. In addition, the search engines allowed limited number of PTMs that can be searched, for example - SEQUEST and Mascot could not identify more than six and nine modifications, respectively, in a single search although it is known that histones can be modified by more than fifteen different types of PTMs.

Also, the above search engines have poor performance in identifying PTMs present in isomeric histone peptides. Most histone tryptic peptides consist of multiple internal lysine residues (e.g. H3 peptide 27-KSAPATGGVKKPHR-40 and H4 peptide 4-GKGGKGLGKGGAKR-17). Often the same peptide can be modified on the side chain of different lysine residues thereby forming positional isomers. Such isomers have virtually the same m/z and they produce composite MS/MS spectra consisting of multiple shared fragment ions. Due to the lack of sufficient distinguishing fragment ions, conventional search engines often misassign sites of modification. To resolve the above problems some research labs have developed custom software tools that can aid the deconvolution of composite MS/MS spectra generated from mixtures of isomeric peptides. However, these tools also suffer from several limitations. For example, the algorithm called ByOnic [207] was able to deconvolute MS/MS spectra of an isomeric mixture of only two peptides. The recently developed algorithms isoScale [172] and EpiProfile [208] are only applicable for composite MS/MS spectra consisting of unique fragment ions representing each of the isomers in the mixture. Therefore, the development of a more comprehensive algorithm that does not necessarily rely

on the presence of unique fragment ions is much needed. Our lab has recently developed such algorithm, which is described in detail in Chapter 3 of the Thesis.

1.3.1.7. Peptide quantification

So far, numerous strategies have been developed for quantification of peptides or proteins in large scale proteomics studies. These strategies can be generally grouped into two categories: isotopic labeling and label-free approaches [209, 210]. Each approach has its own advantages and its own drawbacks and limitations. In the labeling approach, changes in the relative quantity of a given peptide/protein in response to a particular treatment can be evaluated by comparing its abundance in the untreated versus treated within a single LC-MS/MS run. In this approach, the two conditions are mixed into a single sample, and in order for the MS to distinguish between them one of the conditions, either the untreated or treated, is isotopically labeled. The labeling of peptides or proteins can be achieved via two approaches. One is by chemically incorporating isotope-coded mass tags on a specific amino acid comprising a highly reactive side chain (e.g. the ε-amino group of lysines). A few examples of methods using this labeling technique include the isotope-coded protein label (ICPL) [211], isotope tags for relative and absolute quantification (iTRAQ) [212], and tandem mass tags (TMT) [213].

The other labeling approach is the so-called metabolic labeling, which is basically performed *in vivo* via stable isotope labeling by amino acids in cell culture (SILAC) [214]. Here, cells are treated with amino acids containing heavy isotopes (e.g. nitrogen (^{15}N) or carbon (^{13}C)). The mass shift introduced as a result of incorporation of the heavy isotope can be recognized by the MS and at the same time provides the basis for quantification. The advantage of the SILAC approach is that the differentially treated samples can be mixed at the level of intact cells, and as a result it is unaffected by experimental errors introduced during subsequent sample preparation procedures, which is critical for unbiased and accurate quantification of peptides. However, this method is expensive, time consuming, and often less than 100% labeling is achieved, particularly in higher eukaryotes. Moreover, the SILAC approach allows comparison of only few conditions (usually three) due to the limited number of commercially available labels.

In contrast, the label-free approach does not require incorporation of specific mass tags or labels for quantification purpose. Here, quantitative information can be obtained by comparing the LC/MS signal intensity (or extracted ion chromatogram, XIC, intensity) of a

given precursor ion across several differentially treated samples [215]. Unlike the SILAC approach, however, the label-free approach is susceptible to experimental errors since each sample is independently processed and analyzed. This may cause unexpected variations in the overall intensity of peptides among the different conditions or among replicates within the same experimental condition. In such cases, the MS data are normalized for comparison. Global normalization is one of the most common normalization techniques used in label-free quantification [216]. In global normalization, it is assumed that the total MS abundances summed over all detected peptides should be the same in each sample, provided that equal amounts of starting materials are used. Given this assumption, a normalization factor can be calculated from the ratio of total abundances of, for example, untreated to treated samples. Then, by choosing the untreated sample as a reference, the individual abundance values in the treated samples can be adjusted by multiplying each value by the normalization factor. Alternatively, instead of the total abundance, the abundances of an unmodifiable peptide(s) reproducibly and abundantly detected in both the untreated and treated samples can be used to calculate the normalization factor. The label-free approach in principle does not have limitations in the number of conditions that can be compared, which is one of the major advantages of this method over stable isotope labeling techniques. Moreover, this approach is relatively less expensive, less time consuming and results in lower mass spectral complexity since the MS do not have to acquire labeled peptides.

1.4. Thesis objectives

A number of biochemical and/or mass-spectrometry-based methods have previously been developed to identify and characterize various types of histone modifications. While these methods have played an important role in advancing our current view of the biological functions of the modifications, there is still much room for improvement. In the MS-based proteomics field, there are three potential areas of improvement which are chromatography, mass spectrometry and bioinformatics. In this regard, the cutting-edge capabilities of the latest mass spectrometry technology platforms need be exploited to address some of the limitations observed in the currently available biochemical and analytical proteomics approaches.

Thus, the aim of this thesis was to develop improved quantitative MS-based methods that can be used for (1) measuring global changes in histone PTMs; (2) identifying and characterizing individual and combinatorial histone PTMs; and (3) defining the site specificity of HATs and HDACs toward histone substrates. This was achieved through the following three research objectives.

The first objective involved developing an improved MS-based strategy for determining the site specificity of *S. cerevisiae* HAT Rtt109 toward histones H3 and H4. Rtt109 is a fungal-specific HAT that was previously known to acetylate histone H3. Interestingly, early structural and biochemical studies showed that the specificity of Rtt109 toward H3 sites is modulated by the chaperone proteins Asf1 and Vps5. However, there were conflicting reports as to which sites are most preferred when Rtt109 binds to Asf1 versus Vps75. Therefore, in this particular study our objective was to (1) develop an improved MS-based workflow that can be used for measuring the global and site specific acetylation of histones; and (2) determine the role of Asf1 and Vps75 in directing the site specificity of Rtt109 during *in vitro* assay with histones H3 and H4.

The second objective involves developing a novel MS-based approach for deconvolution and quantification of isomeric histone peptides. Histone peptides suitable for bottom-up analysis are typically generated by tryptic digestion, which cleaves at the C-terminal side of arginine and unmodified or mono-methylated lysine residues. Unfortunately, peptides produced using this approach bear more than one internal lysine residue. These peptides can be acetylated at a single or multiple sites. Often a single acetylation can occur on different lysine residues of

the same peptide thereby forming positional isomers. Such isomers cannot be efficiently resolved using conventional RP-LC/MS, and appear as a single chromatographic peak. Moreover, CID of isomeric peptides generates composite MS/MS spectra. Conventional database search engines such as Mascot have difficulties assigning peptide sequences to this type of spectra because of lack of distinguishing fragment ions representing each isomer. Therefore, in order to address these problems we set out three objectives: (1) Using synthetic histone peptides, examine the detection efficiencies, fragmentation efficiencies and patterns, and chromatographic features of the individual and mixed isomers; (2) based on composite MS/MS spectra of isomeric synthetic peptides,develop a bioinformatics tool that can aid the deconvolution of both the composite MS/MS spectra and the co-eluting precursor chromatographic peaks; (3) validate the applicability of the method using different biological problems.

The third objective involves, using the aforementioned methods to determine the role of several fission yeast HATs and HDACs in regulating the steady-state levels of histone acetylation. A number of previous studies used biochemical approaches such as Western blot and ChIP in order to characterize the site specificity of fission yeast HATs and HDACs. These approaches, however, suffer from several limitations. First, the antibody-based approach requires prior knowledge of the sites of acetylation, and thus cannot be used to identify novel sites. Second, many antibodies used in studies of histone acetylation lack the desired specificity, particularly when two or more acetylation sites are too close together. Third, the antibody-based approach would only allow semi-quantitative assessment of histone acetylation. Therefore, we employed the above validated MS-based methods to investigate the global and site-specific changes of histone acetylation in fission yeast mutants lacking major HAT and HDAC activities. We were particularly interested in finding out the enzymes' distinct roles in acetylating or deacetylating histones H3 and H4. In addition, we wanted to shed light on the mechanisms underlying the dynamic interplay between HATs and HDACs in establishing the steady-state level of histone acetylation.

1.5. Thesis outline

The second chapter of the thesis presents a quantitative MS-based method that can be used for determining the global and site-specific acetylation of histones. The method was used in investigating the acetylation of histones H3 and H4 during *in vitro* acetylation assays with Rtt109-Vps75 or Rtt109-Asf1. Our results showed that Rtt109–Vps75 preferentially acetylates H3K9 and H3K23, whereas Rtt109–Asf1 showed a far greater specificity for H3K56 acetylation. Moreover, Rtt109–Vps75 exhibited moderate activity toward H4 sites, whereas Rtt109–Asf1 showed no detectable activity toward H4.

The third chapter presents a novel MS-based approach that is useful for analysis of isomeric histone peptides. The approach employs an in-house developed algorithm called Iso-PeptidAce to deconvolute composite MS/MS spectra of mixtures of isomers. The algorithm was built based on several chromatographic and fragmentation features of synthetic isomeric histone peptides. The applicability of the method was demonstrated using different biological problems including the identification of site-specific acetylation patterns in affinity-purified histones bound to CAF1 and profiling of histone acetylation in cells treated with different HDAC inhibitors.

In the fourth chapter, a study aimed at investigating the substrate specificity of several fission yeast HATs and HDACs is described. Bulk histones isolated from wild type or mutant fission yeast strains were fractionated into individual core histones using offline reverse-phase HPLC. Tryptic digests of derivatized H3/H4 were analyzed by a multiplexed parallel reaction monitoring (PRM) method on a QExactive Plus instrument. Data analyses were performed manually using Thermo Xcalibure Qual Browser software and the in-house built algorithm, Iso-PeptidAce. The site-specificity of each HAT and HDAC as determined based on the fold change ratios of mutant-to-control acetylation site occupancies of lysine residues on the N-terminal tails of H3 and H4. Our analysis identified several known and novel HAT and HDAC target sites.

The fifth chapter summarizes the different analytical tools developed and the results obtained in the above studies. A perspective on both the potential of our method and the remaining challenges in MS-based histone PTM analysis is also described.

1.6. References

1. Bruce Alberts, A.J., Julian Lewis, Martin Raff, Keith Roberts, and Peter Walter, *Molecular Biology of the Cell*. 2002, Garland Science: New York. p. 263 - 399.
2. Watson, J.D. and F.H.C. Crick, *Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid.* Nature, 1953. **171**(4356): p. 737-738.
3. Kornberg, R.D. and Y. Lorch, *Twenty-five years of the nucleosome, fundamental particle of the eukaryote chromosome.* Cell, 1999. **98**(3): p. 285-294.
4. Kornberg, R.D., *Chromatin structure: a repeating unit of histones and DNA.* Science, 1974. **184**(4139): p. 868-871.
5. Richmond, T. and J. Finch, *The structure of the nucleosome core particle.* Nature (London), 1984. **311**: p. 532-537.
6. Olins, D.E. and A.L. Olins, *Chromatin history: our view from the bridge.* Nature Reviews Molecular Cell Biology, 2003. **4**(10): p. 809-814.
7. Woodcock, F., *Structure of the 30 nm chromatin fiber.* 1986.
8. Bednar, J., et al., *Nucleosomes, linker DNA, and linker histone form a unique structural motif that directs the higher-order folding and compaction of chromatin.* Proceedings of the National Academy of Sciences, 1998. **95**(24): p. 14173-14178.
9. Thomas, J.O., *Histone H1: location and role.* Current opinion in cell biology, 1999. **11**(3): p. 312-317.
10. Hsu, T. and F.E. Arrighi, *Distribution of constitutive heterochromatin in mammalian chromosomes.* Chromosoma, 1971. **34**(3): p. 243-253.
11. Trojer, P. and D. Reinberg, *Facultative heterochromatin: is there a distinctive molecular signature?* Molecular cell, 2007. **28**(1): p. 1-13.
12. Felsenfeld, G. and M. Groudine, *Controlling the double helix.* Nature, 2003. **421**(6921): p. 448-453.
13. Dahm, R., *Friedrich Miescher and the discovery of DNA.* Developmental Biology, 2005. **278**(2): p. 274-288.
14. Doenecke, D. and P. Karlson, *Albrecht Kossel and the discovery of histones.* Trends in Biochemical Sciences. **9**(9): p. 404-405.
15. Kossel, A., *Protamines and histones.* 1928.
16. Rasmussen, P.S., K. Murray, and J.M. Luck, *On the Complexity of Calf Thymus Histone.* Biochemistry, 1962. **1**(1): p. 79-89.
17. Kornberg, R.D., *Structure of Chromatin.* Annual Review of Biochemistry, 1977. **46**(1): p. 931-954.
18. Oudet, P., M. Gross-Bellard, and P. Chambon, *Electron microscopic and biochemical evidence that chromatin structure is a repeating unit.* Cell, 1975. **4**(4): p. 281-300.
19. Thomas, J.O. and R.D. Kornberg, *An octamer of histones in chromatin and free in solution.* Proceedings of the National Academy of Sciences, 1975. **72**(7): p. 2626-2630.
20. Isenberg, I., *HISTONES* Ann. Rev. Biochem. , 1979. **48**: p. IS9-91
21. Khare, S.P., et al., *HIstome—a relational knowledgebase of human histone proteins and histone modifying enzymes.* Nucleic Acids Research, 2011.
22. Henikoff, S. and Y. Dalal, *Centromeric chromatin: what makes it unique?* Current Opinion in Genetics & Development, 2005. **15**(2): p. 177-184.
23. Tagami, H., et al., *Histone H3.1 and H3.3 Complexes Mediate Nucleosome Assembly Pathways Dependent or Independent of DNA Synthesis.* Cell, 2004. **116**(1): p. 51-61.
24. Jackson, V. and R. Chalkley, *Histone synthesis and deposition in the G1 and S phases of hepatoma tissue culture cells.* Biochemistry, 1985. **24**(24): p. 6921-6930.
25. NAKATANI, Y., et al., *Two Distinct Nucleosome Assembly Pathways: Dependent or Independent of DNA Synthesis Promoted by Histone H3.1 and H3.3 Complexes.* Cold Spring Harbor Symposia on Quantitative Biology, 2004. **69**: p. 273-280.

26. Corpet, A. and G. Almouzni, *Making copies of chromatin: the challenge of nucleosomal organization and epigenetic information.* Trends in Cell Biology, 2009. **19**(1): p. 29-41.
27. Rogakou, E.P., et al., *DNA Double-stranded Breaks Induce Histone H2AX Phosphorylation on Serine 139.* Journal of Biological Chemistry, 1998. **273**(10): p. 5858-5868.
28. Nakagawa, T., et al., *Deubiquitylation of histone H2A activates transcriptional initiation via trans-histone cross-talk with H3K4 di- and trimethylation.* Genes & Development, 2008. **22**(1): p. 37-49.
29. Chadwick, B.P. and H.F. Willard, *Histone H2A variants and the inactive X chromosome: identification of a second macroH2A variant.* Human Molecular Genetics, 2001. **10**(10): p. 1101-1113.
30. Eirín-López, J.M., T. Ishibashi, and J. Ausió, *H2A.Bbd: a quickly evolving hypervariable mammalian histone that destabilizes nucleosomes in an acetylation-independent way.* The FASEB Journal, 2008. **22**(1): p. 316-326.
31. Sharma, A., K. Singh, and A. Almasan, *Histone H2AX Phosphorylation: A Marker for DNA Damage*, in *DNA Repair Protocols*, L. Bjergbæk, Editor. 2012, Humana Press. p. 613-626.
32. Revet, I., et al., *Functional relevance of the histone γH2Ax in the response to DNA damaging agents.* Proceedings of the National Academy of Sciences, 2011. **108**(21): p. 8663-8667.
33. Xu, Y., et al., *Histone H2A.Z Controls a Critical Chromatin Remodeling Step Required for DNA Double-Strand Break Repair.* Molecular Cell, 2012. **48**(5): p. 723-733.
34. DeLange, R.J., et al., *Calf and Pea Histone IV: II. The complete amino acid sequence of calf thymus histone IV; presence of ε-N-acetyllysine.* Journal of Biological Chemistry, 1969. **244**(2): p. 319-334.
35. DeLange, R.J., et al., *Calf and Pea Histone IV: III. Complete amino acid sequence of pea seedling histone IV; comparison with the hologous calf thymus histone.* Journal of Biological Chemistry, 1969. **244**(20): p. 5669-5679.
36. Delange, R.J., J.A. Hooper, and E.L. Smith, *Complete Amino-Acid Sequence of Calf-Thymus Histone III.* Proceedings of the National Academy of Sciences, 1972. **69**(4): p. 882-884.
37. Kornberg, R.D. and J.O. Thonmas, *Chromatin Structure: Oligomers of the Histones.* Science, 1974. **184**(4139): p. 865-868.
38. D'Anna, J.A. and I. Isenberg, *Histone cross-complexing pattern.* Biochemistry, 1974. **13**(24): p. 4992-4997.
39. D'Anna, J.A. and I. Isenberg, *Interactions of histone LAK (f2a2) with histones KAS (f2b) and GRK (f2a1).* Biochemistry, 1974. **13**(10): p. 2098-2104.
40. Li, H.J., et al., *Circular dichroism of histone-bound regions in chromatin.* Biopolymers, 1975. **14**(1): p. 211-226.
41. Noll, M., *Subunit structure of chromatin.* Nature, 1974. **251**(5472): p. 249-251.
42. Eickbush, T.H. and E.N. Moudrianakis, *The histone core complex: an octamer assembled by two sets of protein-protein interactions.* Biochemistry, 1978. **17**(23): p. 4955-4964.
43. Klug, A., et al., *A low resolution structure for the histone core of the nucleosome.* Nature, 1980. **287**(5782): p. 509-516.
44. Finch, J.T., et al., *Structure of nucleosome core particles of chromatin.* Nature, 1977. **269**(5623): p. 29-36.
45. Richmond, T.J., et al., *Structure of the nucleosome core particle at 7 Å resolution.* Nature, 1984. **311**(5986): p. 532-537.
46. Burlingame, R., et al., *Crystallographic structure of the octameric histone core of the nucleosome at a resolution of 3.3 A.* Science, 1985. **228**(4699): p. 546-553.
47. Bentley, G.A., et al., *Crystal structure of the nucleosome core particle at 16 Å resolution.* Journal of Molecular Biology, 1984. **176**(1): p. 55-75.
48. Arents, G., et al., *The Nucleosomal Core Histone Octamer at 3.1 □ A Resolution: A Tripartite Protein Assembly and a Left-Handed Superhelix.* Proceedings of the National Academy of Sciences of the United States of America, 1991. **88**(22): p. 10148-10152.

49. Luger, K., et al., *Crystal structure of the nucleosome core particle at 2.8 Å resolution.* Nature, 1997. **389**(6648): p. 251-260.
50. Tsunaka, Y., et al., *Alteration of the nucleosomal DNA path in the crystal structure of a human nucleosome core particle.* Nucleic Acids Research, 2005. **33**(10): p. 3424-3434.
51. Ramakrishnan, V., *Histone structure and the organization of the nucleosome.* Annual Review of Biophysics and Biomolecular Structure, 1997. **26**(1): p. 83-112.
52. Walsh, C.T., S. Garneau-Tsodikova, and G.J. Gatto, *Protein Posttranslational Modifications: The Chemistry of Proteome Diversifications.* Angewandte Chemie International Edition, 2005. **44**(45): p. 7342-7372.
53. Knorre, D., N. Kudryashova, and T. Godovikova, *Chemical and functional aspects of posttranslational modification of proteins.* Acta naturae, 2009. **1**(3): p. 29.
54. Duan, G. and D. Walther, *The Roles of Post-translational Modifications in the Context of Protein Interaction Networks.* PLoS Comput Biol, 2015. **11**(2): p. e1004049.
55. Gu, Y., J. Rosenblatt, and D. Morgan, *Cell cycle regulation of CDK2 activity by phosphorylation of Thr160 and Tyr15.* The EMBO journal, 1992. **11**(11): p. 3995.
56. Spencer, V.A. and J.R. Davie, *Role of covalent modifications of histones in regulating gene expression.* Gene, 1999. **240**(1): p. 1-12.
57. Huang, H., et al., *SnapShot: Histone Modifications.* Cell, 2014. **159**(2): p. 458-458.e1.
58. Ito, K., P. J. Barnes, and I. M. Adcock, *Histone Acetylation and Deacetylation*, in *Asthma*, K. Fan Chung and I. Adcock, Editors. 2000, Humana Press. p. 309-319.
59. Tanner, K.G., et al., *Catalytic Mechanism and Function of Invariant Glutamic Acid 173 from the Histone Acetyltransferase GCN5 Transcriptional Coactivator.* Journal of Biological Chemistry, 1999. **274**(26): p. 18157-18160.
60. Bird, A.W., et al., *Acetylation of histone H4 by Esa1 is required for DNA double-strand break repair.* Nature, 2002. **419**(6905): p. 411-415.
61. Vogelauer, M., et al., *Histone Acetylation Regulates the Time of Replication Origin Firing.* Molecular Cell, 2002. **10**(5): p. 1223-1233.
62. Shogren-Knaak, M., et al., *Histone H4-K16 Acetylation Controls Chromatin Structure and Protein Interactions.* Science, 2006. **311**(5762): p. 844-847.
63. Hebbes, T.R., A.W. Thorne, and C. Crane-Robinson, *A direct link between core histone acetylation and transcriptionally active chromatin.* The EMBO journal, 1988. **7**(5): p. 1395.
64. Davie, J.R. and D.N. Chadee, *Regulation and regulatory parameters of histone modifications.* Journal of Cellular Biochemistry, 1998. **72**(S30–31): p. 203-213.
65. Taverna, S.D., et al., *How chromatin-binding modules interpret histone modifications: lessons from professional pocket pickers.* Nature Structural & Molecular Biology, 2007. **14**(11): p. 1025-1040.
66. Bannister, A.J. and T. Kouzarides, *Regulation of chromatin by histone modifications.* Cell research, 2011. **21**(3): p. 381-395.
67. Robert, F., et al., *Global position and recruitment of HATs and HDACs in the yeast genome.* Molecular Cell, 2004. **16**(2): p. 199-209.
68. Pokholok, D.K., et al., *Genome-wide map of nucleosome acetylation and methylation in yeast.* Cell, 2005. **122**(4): p. 517-527.
69. Wang, Z., et al., *Combinatorial patterns of histone acetylations and methylations in the human genome.* Nature genetics, 2008. **40**(7): p. 897-903.
70. Deckert, J. and K. Struhl, *Histone Acetylation at Promoters Is Differentially Affected by Specific Activators and Repressors.* Molecular and Cellular Biology, 2001. **21**(8): p. 2726-2735.
71. Black, Joshua C., C. Van Rechem, and Johnathan R. Whetstine, *Histone Lysine Methylation Dynamics: Establishment, Regulation, and Biological Impact.* Molecular Cell, 2012. **48**(4): p. 491-507.

72. Sims Iii, R.J., K. Nishioka, and D. Reinberg, *Histone lysine methylation: a signature for chromatin function.* Trends in Genetics, 2003. **19**(11): p. 629-639.
73. Vaissière, T., C. Sawan, and Z. Herceg, *Epigenetic interplay between histone modifications and DNA methylation in gene silencing.* Mutation Research/Reviews in Mutation Research, 2008. **659**(1–2): p. 40-48.
74. Zhang, Y. and D. Reinberg, *Transcription regulation by histone methylation: interplay between different covalent modifications of the core histone tails.* Genes & Development, 2001. **15**(18): p. 2343-2360.
75. Schmitz, R.J., S. Sung, and R.M. Amasino, *Histone arginine methylation is required for vernalization-induced epigenetic silencing of FLC in winter-annual Arabidopsis thaliana.* Proceedings of the National Academy of Sciences, 2008. **105**(2): p. 411-416.
76. Balint, B.L., et al., *Arginine Methylation Provides Epigenetic Transcription Memory for Retinoid-Induced Differentiation in Myeloid Cells.* Molecular and Cellular Biology, 2005. **25**(13): p. 5648-5663.
77. Helin, K. and D. Dhanak, *Chromatin proteins and modifications as drug targets.* Nature, 2013. **502**(7472): p. 480-488.
78. Liang, G., et al., *Distinct localization of histone H3 acetylation and H3-K4 methylation to the transcription start sites in the human genome.* Proceedings of the National Academy of Sciences of the United States of America, 2004. **101**(19): p. 7357-7362.
79. Pokholok, D.K., et al., *Genome-wide Map of Nucleosome Acetylation and Methylation in Yeast.* Cell, 2005. **122**(4): p. 517-527.
80. Barski, A., et al., *High-Resolution Profiling of Histone Methylations in the Human Genome.* Cell, 2007. **129**(4): p. 823-837.
81. Roh, T.-Y., et al., *The genomic landscape of histone modifications in human T cells.* Proceedings of the National Academy of Sciences, 2006. **103**(43): p. 15782-15787.
82. Bannister, A.J., et al., *Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain.* Nature, 2001. **410**(6824): p. 120-124.
83. Schotta, G., et al., *A silencing pathway to induce H3-K9 and H4-K20 trimethylation at constitutive heterochromatin.* Genes & Development, 2004. **18**(11): p. 1251-1262.
84. Loyola, A., et al., *PTMs on H3 Variants before Chromatin Assembly Potentiate Their Final Epigenetic State.* Molecular Cell, 2006. **24**(2): p. 309-316.
85. Yuan, J., et al., *Histone H3-K56 acetylation is important for genomic stability in mammals.* Cell cycle, 2009. **8**(11): p. 1747-1753.
86. Lee, J.-S., E. Smith, and A. Shilatifard, *The language of histone crosstalk.* Cell, 2010. **142**(5): p. 682-685.
87. Fuchs, S.M., et al., *Influence of Combinatorial Histone Modifications on Antibody and Effector Protein Recognition.* Current Biology, 2011. **21**(1): p. 53-58.
88. Schwämmle, V., et al., *Large Scale Analysis of Co-existing Post-translational Modifications in Histone Tails Reveals Global Fine Structure of Cross-talk.* Molecular & Cellular Proteomics, 2014. **13**(7): p. 1855-1865.
89. Cheung, P., et al., *Synergistic coupling of histone H3 phosphorylation and acetylation in response to epidermal growth factor stimulation.* Molecular cell, 2000. **5**(6): p. 905-915.
90. Lo, W.-S., et al., *Phosphorylation of serine 10 in histone H3 is functionally linked in vitro and in vivo to Gcn5-mediated acetylation at lysine 14.* Molecular cell, 2000. **5**(6): p. 917-926.
91. Portela, A. and M. Esteller, *Epigenetic modifications and human disease.* Nat Biotech, 2010. **28**(10): p. 1057-1068.
92. Sun, Z.-W. and C.D. Allis, *Ubiquitination of histone H2B regulates H3 methylation and gene silencing in yeast.* Nature, 2002. **418**(6893): p. 104-108.
93. Kim, J., et al., *RAD6-Mediated transcription-coupled H2B ubiquitylation directly stimulates H3K4 methylation in human cells.* Cell, 2009. **137**(3): p. 459-471.

94. Guillemette, B., et al., *H3 lysine 4 is acetylated at active gene promoters and is regulated by H3 lysine 4 methylation.* 2011.
95. Strahl, B.D. and C.D. Allis, *The language of covalent histone modifications.* Nature, 2000. **403**(6765): p. 41-45.
96. Zeng, L. and M.-M. Zhou, *Bromodomain: an acetyl-lysine binding domain.* FEBS letters, 2002. **513**(1): p. 124-128.
97. Filippakopoulos, P. and S. Knapp, *The bromodomain interaction module.* FEBS letters, 2012. **586**(17): p. 2692-2704.
98. Kasten, M., et al., *Tandem bromodomains in the chromatin remodeler RSC recognize acetylated histone H3 Lys14.* The EMBO journal, 2004. **23**(6): p. 1348-1359.
99. Jacobson, R.H., et al., *Structure and function of a human TAFII250 double bromodomain module.* Science, 2000. **288**(5470): p. 1422-1425.
100. Florence, B. and D.V. Faller, *You bet-cha: a novel family of transcriptional regulators.* Frontiers in bioscience: a journal and virtual library, 2001. **6**: p. D1008-18.
101. LeRoy, G., B. Rickards, and S. Flint, *The double bromodomain proteins Brd2 and Brd3 couple histone acetylation to transcription.* Molecular cell, 2008. **30**(1): p. 51-60.
102. Morinière, J., et al., *Cooperative binding of two acetylation marks on a histone tail by a single bromodomain.* Nature, 2009. **461**(7264): p. 664-668.
103. Nielsen, P.R., et al., *Structure of the HP1 chromodomain bound to histone H3 methylated at lysine 9.* Nature, 2002. **416**(6876): p. 103-107.
104. Rea, S., et al., *Regulation of chromatin structure by site-specific histone H3 methyltransferases.* Nature, 2000. **406**(6796): p. 593-599.
105. Carrozza, M.J., et al., *The diverse functions of histone acetyltransferase complexes.* Trends in Genetics, 2003. **19**(6): p. 321-329.
106. Taylor, G.C.A., et al., *H4K16 acetylation marks active genes and enhancers of embryonic stem cells, but does not alter chromatin compaction.* Genome Research, 2013. **23**(12): p. 2053-2065.
107. Papait, R., et al., *Genome-wide analysis of histone marks identifying an epigenetic signature of promoters and enhancers underlying cardiac hypertrophy.* Proceedings of the National Academy of Sciences, 2013. **110**(50): p. 20164-20169.
108. Wang, Z., et al., *Genome-wide Mapping of HATs and HDACs Reveals Distinct Functions in Active and Inactive Genes.* Cell. **138**(5): p. 1019-1031.
109. Struhl, K., *Histone acetylation and transcriptional regulatory mechanisms.* Genes & Development, 1998. **12**(5): p. 599-606.
110. Utley, R.T., et al., *Transcriptional activators direct histone acetyltransferase complexes to nucleosomes.* Nature, 1998. **394**(6692): p. 498-502.
111. Pokholok, D., et al., *Genome-wide map of nucleosome acetylation and methylation in yeast.* Cell, 2005. **122**(4): p. 517 - 527.
112. Singh, B.N., et al., *Nonhistone protein acetylation as cancer therapy targets.* 2010.
113. Friedmann, D.R. and R. Marmorstein, *Structure and mechanism of non-histone protein acetyltransferase enzymes.* FEBS Journal, 2013. **280**(22): p. 5570-5581.
114. Roth, S.Y., J.M. Denu, and C.D. Allis, *HISTONE ACETYLTRANSFERASES.* Annual Review of Biochemistry, 2001. **70**(1): p. 81-120.
115. Lee, K.K. and J.L. Workman, *Histone acetyltransferase complexes: one size doesn't fit all.* Nat Rev Mol Cell Biol, 2007. **8**(4): p. 284-295.
116. Craig, J.M. and N.C. Wong, *Epigenetics: a reference manual*. Concepts in Histone Acetyltransferase Biology. 2011, Norfolk, UK: Horizon Scientific Press.
117. Hayakawa, T. and J.-i. Nakayama, *Physiological roles of class I HDAC complex and histone demethylase.* BioMed Research International, 2010. **2011**.

118. Kasten, M.M., S. Dorland, and D.J. Stillman, *A large protein complex containing the yeast Sin3p and Rpd3p transcriptional regulators.* Molecular and Cellular Biology, 1997. **17**(8): p. 4852-8.
119. Zhang, Y., et al., *Histone Deacetylases and SAP18, a Novel Polypeptide, Are Components of a Human Sin3 Complex.* Cell, 1997. **89**(3): p. 357-364.
120. Rundlett, S.E., et al., *Transcriptional repression by UME6 involves deacetylation of lysine 5 of histone H4 by RPD3.* Nature, 1998. **392**(6678): p. 831-835.
121. Kadosh, D. and K. Struhl, *Repression by Ume6 Involves Recruitment of a Complex Containing Sin3 Corepressor and Rpd3 Histone Deacetylase to Target Promoters.* Cell, 1997. **89**(3): p. 365-371.
122. Shirakawa, K., et al., *Reactivation of latent HIV by histone deacetylase inhibitors.* Trends in Microbiology, 2013. **21**(6): p. 277-285.
123. Grozinger, C.M. and S.L. Schreiber, *Deacetylase enzymes: biological functions and the use of small-molecule inhibitors.* Chemistry & biology, 2002. **9**(1): p. 3-16.
124. Johnsson, A., et al., *HAT–HDAC interplay modulates global histone H3K14 acetylation in gene-coding regions during stress*. Vol. 10. 2009. 1009-1014.
125. Yamagoe, S., et al., *Interaction of Histone Acetylases and Deacetylases In Vivo.* Molecular and Cellular Biology, 2003. **23**(3): p. 1025-1033.
126. Che, J., et al., *Hyper-Acetylation of Histone H3K56 Limits Break-Induced Replication by Inhibiting Extensive Repair Synthesis.* PLoS Genet, 2015. **11**(2): p. e1004990.
127. Taylor, J.P., et al., *Aberrant histone acetylation, altered transcription, and retinal degeneration in a Drosophila model of polyglutamine disease are rescued by CREB-binding protein.* Genes & Development, 2003. **17**(12): p. 1463-1468.
128. Yang, F., et al., *Histone hyperacetylation during meiosis interferes with large-scale chromatin remodeling, axial chromatid condensation and sister chromatid separation in the mammalian oocyte.* Int J Dev Biol, 2012. **56**: p. 889-899.
129. Song, J.S., et al., *Global histone modification pattern associated with recurrence and disease-free survival in non-small cell lung cancer patients.* Pathology International, 2012. **62**(3): p. 182-190.
130. Seligson, D.B., et al., *Global histone modification patterns predict risk of prostate cancer recurrence.* Nature, 2005. **435**(7046): p. 1262-1266.
131. Barlési, F., et al., *Global Histone Modifications Predict Prognosis of Resected Non–Small-Cell Lung Cancer.* Journal of Clinical Oncology, 2007. **25**(28): p. 4358-4364.
132. Elsheikh, S.E., et al., *Global Histone Modifications in Breast Cancer Correlate with Tumor Phenotypes, Prognostic Factors, and Patient Outcome.* Cancer Research, 2009. **69**(9): p. 3802-3809.
133. Fraga, M.F., et al., *Loss of acetylation at Lys16 and trimethylation at Lys20 of histone H4 is a common hallmark of human cancer.* Nat Genet, 2005. **37**(4): p. 391-400.
134. Sharma, G.G., et al., *MOF and Histone H4 Acetylation at Lysine 16 Are Critical for DNA Damage Response and Double-Strand Break Repair.* Molecular and Cellular Biology, 2010. **30**(14): p. 3582-3595.
135. Vaquero, A., et al., *Human SirT1 Interacts with Histone H1 and Promotes Formation of Facultative Heterochromatin.* Molecular Cell, 2004. **16**(1): p. 93-105.
136. Song, J., et al., *Increased expression of histone deacetylase 2 is found in human gastric cancer.* APMIS, 2005. **113**(4): p. 264-268.
137. Halkidou, K., et al., *Upregulation and Nuclear Recruitment of HDAC1 in Hormone Refractory Prostate Cancer.* The Prostate, 2004. **59**(2): p. 177-189.
138. Minucci, S., et al., *Histone deacetylases: a common molecular target for differentiation treatment of acute myeloid leukemias?* Oncogene, 2001. **20**(24): p. 3110-3115.

139. Li, Z. and W.-G. Zhu, *Targeting histone deacetylases for cancer therapy: from molecular mechanisms to clinical implications.* International journal of biological sciences, 2014. **10**(7): p. 757.
140. Dokmanovic, M., C. Clarke, and P.A. Marks, *Histone Deacetylase Inhibitors: Overview and Perspectives.* Molecular Cancer Research, 2007. **5**(10): p. 981-989.
141. Mann, B.S., Johnson, John R., Cohen, Martin H., Justice, Robert, Pazdur, Richard, *FDA Approval Summary: Vorinostat for Treatment of Advanced Primary Cutaneous T-Cell Lymphoma.* The Oncologist, 2007. **12**(10): p. 1247-1252.
142. Drogaris, P., et al., *Histone Deacetylase Inhibitors Globally Enhance H3/H4 Tail Acetylation Without Affecting H3 Lysine 56 Acetylation.* Sci. Rep., 2012. **2**.
143. Li, C.T., et al., *Vorinostat, SAHA, represses telomerase activity via epigenetic regulation of telomerase reverse transcriptase in non-small cell lung cancer cells.* Journal of cellular biochemistry, 2011. **112**(10): p. 3044-3053.
144. Butler, L.M., et al., *The histone deacetylase inhibitor SAHA arrests cancer cell growth, up-regulates thioredoxin-binding protein-2, and down-regulates thioredoxin.* Proceedings of the National Academy of Sciences, 2002. **99**(18): p. 11700-11705.
145. Butler, L.M., et al., *Suberoylanilide Hydroxamic Acid, an Inhibitor of Histone Deacetylase, Suppresses the Growth of Prostate Cancer Cells in Vitro and in Vivo.* Cancer Research, 2000. **60**(18): p. 5165-5170.
146. VanderMolen, K.M., et al., *Romidepsin (Istodax, NSC 630176, FR901228, FK228, depsipeptide): a natural product recently approved for cutaneous T-cell lymphoma.* J Antibiot, 2011. **64**(8): p. 525-531.
147. Grant, C., et al., *Romidepsin: a new therapy for cutaneous T-cell lymphoma and a potential therapy for solid tumors.* Expert Review of Anticancer Therapy, 2010. **10**(7): p. 997-1008.
148. DiGiulio, S., *FDA Approves Farydak (Panobinostat) for Multiple Myeloma.* Oncology Times, 2015.
149. Tate, C.R., et al., *Targeting triple-negative breast cancer cells with the histone deacetylase inhibitor panobinostat.* Breast Cancer Res, 2012. **14**(3): p. R79.
150. Beckers, T., et al., *Distinct pharmacological properties of second generation HDAC inhibitors with the benzamide or hydroxamate head group.* International Journal of Cancer, 2007. **121**(5): p. 1138-1148.
151. Minucci, S. and P.G. Pelicci, *Histone deacetylase inhibitors and the promise of epigenetic (and more) treatments for cancer.* Nat Rev Cancer, 2006. **6**(1): p. 38-51.
152. McGee-Lawrence, M.E., et al., *Suberoylanilide hydroxamic acid (SAHA; vorinostat) causes bone loss by inhibiting immature osteoblasts.* Bone, 2011. **48**(5): p. 1117-1126.
153. Pratap, J., et al., *The Histone Deacetylase Inhibitor, Vorinostat, Reduces Tumor Growth at the Metastatic Bone Site and Associated Osteolysis, but Promotes Normal Bone Loss.* Molecular Cancer Therapeutics, 2010. **9**(12): p. 3210-3220.
154. Duvic, M. and C. Zhang, *Clinical and laboratory experience of vorinostat (suberoylanilide hydroxamic acid) in the treatment of cutaneous T-cell lymphoma.* Br J Cancer, 0000. **95**(S1): p. S13-S19.
155. Andreu-Vieyra, C.V. and J.R. Berenson, *The potential of panobinostat as a treatment option in patients with relapsed and refractory multiple myeloma.* Therapeutic Advances in Hematology, 2014. **5**(6): p. 197-210.
156. Phillips, D., *The presence of acetyl groups in histones.* Biochemical Journal, 1963. **87**(2): p. 258.
157. Phillips, D., *A micromethod for the determination of N-and O-acetyl groups.* Biochemical Journal, 1963. **86**(3): p. 397.
158. Gershey, E., G. Vidali, and V. Allfrey, *Chemical Studies of Histone Acetylation: The occurrence of ε-N-acetyllysine in the f2a1 histone* Journal of Biological Chemistry, 1968. **243**(19): p. 5018-5022.

159. Vidali, G., E. Gershey, and V. Allfrey, *Chemical Studies of Histone Acetylation: The distribution of ε-N-acetyllysine in calf thymus histones.* Journal of Biological Chemistry, 1968. **243**(24): p. 6361-6366.
160. BODE, J., K. HENCO, and E. WINGENDER, *Modulation of the nucleosome structure by histone acetylation.* European Journal of Biochemistry, 1980. **110**(1): p. 143-152.
161. BODE, J., M. GÓMEZ-LIRA, and H. SCHRÖTER, *Nucleosomal particles open as the histone core becomes hyperacetylated.* European Journal of Biochemistry, 1983. **130**(3): p. 437-445.
162. Garcia-Ramirez, M., C. Rocchini, and J. Ausio, *Modulation of Chromatin Folding by Histone Acetylation.* Journal of Biological Chemistry, 1995. **270**(30): p. 17923-17928.
163. Chahal, S.S., H.R. Matthews, and E.M. Bradbury, *Acetylation of histone H4 and its role in chromatin structure and function.* Nature, 1980. **287**(5777): p. 76-79.
164. Muller, S., et al., *Specificity of antibodies raised against triacetylated histone H4.* Molecular immunology, 1987. **24**(7): p. 779-789.
165. Pfeffer, U., N. Ferrari, and G. Vidali, *Availability of hyperacetylated H4 histone in intact nucleosomes to specific antibodies.* Journal of Biological Chemistry, 1986. **261**(6): p. 2496-2498.
166. Baeza, J., et al., *Stoichiometry of site-specific lysine acetylation in an entire proteome.* Journal of Biological Chemistry, 2014.
167. Weinert, B.T., et al., *Acetylation dynamics and stoichiometry in Saccharomyces cerevisiae*. Vol. 10. 2014.
168. Zee, B.M., N.L. Young, and B.A. Garcia, *Suppl 1: Quantitative Proteomic Approaches to Studying Histone Modifications.* Current chemical genomics, 2011. **5**: p. 106.
169. Moradian, A., et al., *The top-down, middle-down, and bottom-up mass spectrometry approaches for characterization of histone variants and their post-translational modifications.* PROTEOMICS, 2014. **14**(4-5): p. 489-497.
170. Guan, S. and A.L. Burlingame, *Data Processing Algorithms for Analysis of High Resolution MSMS Spectra of Peptides with Complex Patterns of Posttranslational Modifications.* Molecular & Cellular Proteomics, 2010. **9**(5): p. 804-810.
171. Yuan, Z.-F., et al., *Evaluation of Proteomic Search Engines for the Analysis of Histone Modifications.* Journal of Proteome Research, 2014. **13**(10): p. 4470-4478.
172. Sidoli, S., et al., *Middle-down hybrid chromatography/tandem mass spectrometry workflow for characterization of combinatorial post-translational modifications in histones.* PROTEOMICS, 2014. **14**(19): p. 2200-2211.
173. Walker, J.M., *The Protein Protocols Handbook*. 2 ed. 2002, Totowa, NJ Humana Press Inc.
174. Shevchenko, A., et al., *Mass Spectrometric Sequencing of Proteins from Silver-Stained Polyacrylamide Gels.* Analytical Chemistry, 1996. **68**(5): p. 850-858.
175. Riviere, L.R. and P. Tempst, *Current Protocols in Protein Science*. June, 1995, Brooklyn: John Wiley & Sons, Inc.
176. Aguilar, M.-I., *Reversed-Phase High-Performance Liquid Chromatography #*, in *T HPLC of Peptides and Proteins*. 2003. p. 9-22.
177. Vailaya, A. and C. Horváth, *Retention in reversed-phase chromatography: partition or adsorption?* Journal of Chromatography A, 1998. **829**(1): p. 1-27.
178. Pearson, J.D. and F.E. Regnier, *The Influence of Reversed-Phase n-Alkyl Chain Length on Protein Retention, Resolution and Recovery: Implications for Preparative HPLC.* Journal of Liquid Chromatography, 1983. **6**(3): p. 497-510.
179. Engelhardt, H. and H. Müller, *Optimal conditions for the reversed-phase chromatography of proteins.* Chromatographia, 1984. **19**(1): p. 77-84.
180. Shrivastava, A. and V.B. Gupta, *HPLC: Isocratic or Gradient Elution and Assessment of Linearity In Analytical Methods.* Journal of Advanced Scientific Research, 2012. **3**(2).
181. McGuffin, V.L., *High-Performance Gradient Elution: The Practical Application of the Linear-Solvent-Strength Model By Lloyd R. Snyder (LC Resources, Inc., Orinda, CA) and John W.*

Dolan (LC Resources, Inc., Amity, OR). J. Wiley & Sons, Inc.: Hoboken, NJ. 2007. xxviii+ 462 pp. $115.00. ISBN 0-471-70646-9. Journal of the American Chemical Society, 2007. **129**(27): p. 8669-8670.

182. Carr, D., *The handbook of analysis and purification of peptides and proteins by reversed-phase HPLC.* Hesperia, CA, USA: Grace Vydac, 2002.
183. Vestal, M.L., *Methods of ion generation.* Chemical reviews, 2001. **101**(2): p. 361-376.
184. Aiken, A.C., P.F. DeCarlo, and J.L. Jimenez, *Elemental analysis of organic species with electron ionization high-resolution mass spectrometry.* Analytical Chemistry, 2007. **79**(21): p. 8350-8358.
185. Knochenmuss, R. and R. Zenobi, *MALDI ionization: the role of in-plume processes.* Chemical reviews, 2003. **103**(2): p. 441-452.
186. Signor, L. and E.B. Erba, *Matrix-assisted Laser Desorption/Ionization Time of Flight (MALDI-TOF) Mass Spectrometric Analysis of Intact Proteins Larger than 100 kDa.* Journal of visualized experiments: JoVE, 2013(79).
187. Fenn, J., et al., *Electrospray ionization for mass spectrometry of large biomolecules.* Science, 1989. **246**(4926): p. 64-71.
188. Watson, J.T. and O.D. Sparkman, *Introduction to mass spectrometry: instrumentation, applications, and strategies for data interpretation*. 2007: John Wiley & Sons.
189. McLuckey, S.A. and J.M. Wells, *Mass analysis at the advent of the 21st century.* Chemical reviews, 2001. **101**(2): p. 571-606.
190. Scigelova, M. and A. Makarov, *Orbitrap mass analyzer-overview and applications in proteomics.* Proteomics, 2006. **6**(Suppl 2): p. 16-21.
191. Dawson, P., *Quadrupole mass analyzers: performance, design and some recent applications.* Mass Spectrometry Reviews, 1986. **5**(1): p. 1-37.
192. Miller, P.E. and M.B. Denton, *The quadrupole mass filter: basic operating concepts.* Journal of chemical education, 1986. **63**(7): p. 617.
193. March, R.E., *An introduction to quadrupole ion trap mass spectrometry.* Journal of mass spectrometry, 1997. **32**(4): p. 351-369.
194. Guilhaus, M., *Principles and instrumentation in time-of-flight mass spectrometry. Physical and instrumental concepts.* Journal of Mass Spectrometry, 1995. **30**(11): p. 1519-1532.
195. Van Berkel, W.J., et al., *Detection of intact megaDalton protein assemblies of vanillyl-alcohol oxidase by mass spectrometry.* Protein Science, 2000. **9**(3): p. 435-439.
196. Wenzel, R.J., et al., *Analysis of megadalton ions using cryodetection MALDI time-of-flight mass spectrometry.* Analytical chemistry, 2005. **77**(14): p. 4329-4337.
197. Marshall, A.G., C.L. Hendrickson, and G.S. Jackson, *Fourier transform ion cyclotron resonance mass spectrometry: a primer.* Mass spectrometry reviews, 1998. **17**(1): p. 1-35.
198. Makarov, A., *Electrostatic axially harmonic orbital trapping: a high-performance technique of mass analysis.* Analytical chemistry, 2000. **72**(6): p. 1156-1162.
199. Demartini, D.R., *A Short Overview of the Components in Mass Spectrometry Instrumentation for Proteomics Analyses.* 2013.
200. Roepstorff, P. and J. Fohlman, *Proposal for a common nomenclature for sequence ions in mass spectra of peptides.* Biomedical mass spectrometry, 1984. **11**(11).
201. Tang, X.J., P. Thibault, and R.K. Boyd, *Fragmentation reactions of multiply-protonated peptides and implications for sequencing by tandem mass spectrometry with low-energy collision-induced dissociation.* Analytical Chemistry, 1993. **65**(20): p. 2824-2834.
202. McAlister, G.C., et al., *Higher-energy Collision-activated Dissociation Without a Dedicated Collision Cell.* Molecular & Cellular Proteomics, 2011. **10**(5).
203. Olsen, J.V., et al., *Higher-energy C-trap dissociation for peptide modification analysis.* Nat Meth, 2007. **4**(9): p. 709-712.

204. Syka, J.E.P., et al., *Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry.* Proceedings of the National Academy of Sciences of the United States of America, 2004. **101**(26): p. 9528-9533.
205. Fenyö, D., *Identifying the proteome: software tools.* Current Opinion in Biotechnology, 2000. **11**(4): p. 391-395.
206. Eng, J., A. McCormack, and J. Yates, *An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database.* Journal of the American Society for Mass Spectrometry, 1994. **5**(11): p. 976-989.
207. Chen, X., P. Drogaris, and M. Bern, *Identification of Tandem Mass Spectra of Mixtures of Isomeric Peptides.* Journal of Proteome Research, 2010. **9**(6): p. 3270-3279.
208. Yuan, Z.-F., et al., *EpiProfile Quantifies Histone Peptides With Modifications by Extracting Retention Time and Intensity in High-resolution Mass Spectra.* Molecular & Cellular Proteomics, 2015. **14**(6): p. 1696-1707.
209. Bantscheff, M., et al., *Quantitative mass spectrometry in proteomics: a critical review.* Analytical and Bioanalytical Chemistry, 2007. **389**(4): p. 1017-1031.
210. Patel, V.J., et al., *A Comparison of Labeling and Label-Free Mass Spectrometry-Based Proteomics Approaches.* Journal of Proteome Research, 2009. **8**(7): p. 3752-3759.
211. Schmidt, A., J. Kellermann, and F. Lottspeich, *A novel strategy for quantitative proteomics using isotope-coded protein labels.* PROTEOMICS, 2005. **5**(1): p. 4-15.
212. Ross, P.L., et al., *Multiplexed Protein Quantitation in Saccharomyces cerevisiae Using Amine-reactive Isobaric Tagging Reagents.* Molecular & Cellular Proteomics, 2004. **3**(12): p. 1154-1169.
213. Thompson, A., et al., *Tandem Mass Tags: A Novel Quantification Strategy for Comparative Analysis of Complex Protein Mixtures by MS/MS.* Analytical Chemistry, 2003. **75**(8): p. 1895-1904.
214. Ong, S.-E., et al., *Stable Isotope Labeling by Amino Acids in Cell Culture, SILAC, as a Simple and Accurate Approach to Expression Proteomics.* Molecular & Cellular Proteomics, 2002. **1**(5): p. 376-386.
215. Neilson, K.A., et al., *Less label, more free: Approaches in label-free quantitative mass spectrometry.* PROTEOMICS, 2011. **11**(4): p. 535-553.
216. Listgarten, J. and A. Emili, *Statistical and Computational Methods for Comparative Proteomic Profiling Using Liquid Chromatography-Tandem Mass Spectrometry.* Molecular & Cellular Proteomics, 2005. **4**(4): p. 419-434.

CHAPTER 2 : Chaperone-mediated Acetylation of Histones by Rtt109 Identified by Quantitative Proteomics

Nebiyu Abshiru[1, 2], Kevin Ippersiel[2], Yong Tang[4, 5], Hua Yuan[4, 5], Ronen Marmorstein[4,5], Alain Verreault[1, 3,] , Pierre Thibault[1, 2]

[1] Institute for Research in Immunology and Cancer, Université de Montréal, Montréal, Canada

[2] Department of Chemistry, Université de Montréal, Montréal, Canada

[3] Department of Pathology and Cell Biology, Faculty of Medicine, Université de Montréal, Canada

[4] The Wistar Institute, Philadelphia, PA 19104, USA

[5] Department of Chemistry, University of Pennsylvania, Philadelphia, PA 19104, USA

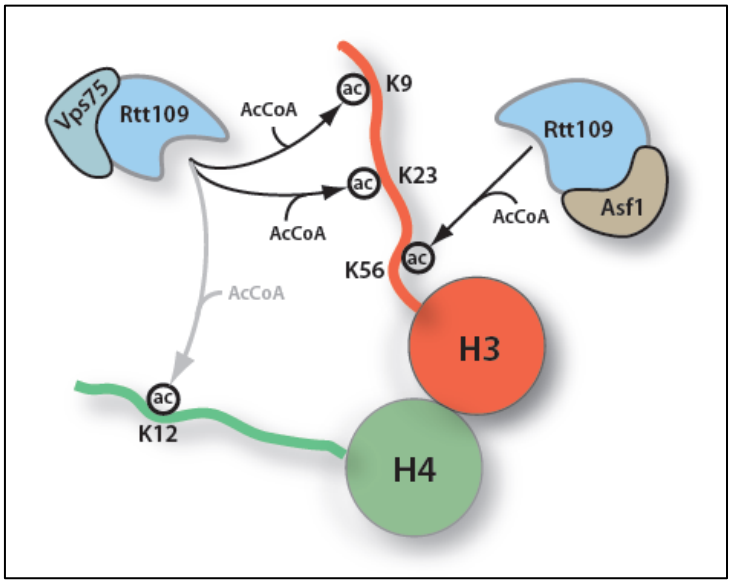
Published: J. Proteomics. 2013 April 9; 81: 80–90.

Author contributions:

N.A. designed and carried out the *in vitro* enzyme assay and the proteomics experiment. K.I. assisted in the proteomics experiment. Y.T, H.Y and R.M provided recombinant enzymes. N.A. analyzed data, wrote the first draft of the manuscript, and generated figures. P.T and A.V. conceived and led the project. N.A., P.T. and A.V. wrote the final draft of the manuscript, with input from all authors.

2.1. Abstract

Rtt109 is a fungal-specific histone acetyltransferase (HAT) that associates with either Vps75 or Asf1 to acetylate histone H3. Recent biochemical and structural studies suggest that site-specific acetylation of H3 by Rtt109 is dictated by the binding chaperone where Rtt109-Asf1 acetylates K56, while Rtt109-Vps75 acetylates K9 and K27. To gain further insights into the roles of Vps75 and Asf1 in directing site-specific acetylation of H3, we used quantitative proteomics to profile the global and site-specific changes in H3 and H4 during in vitro acetylation assays with Rtt109 and its chaperones. Our analyses showed that Rtt109-Vps75 preferentially acetylates H3 K9 and K23, the former residue being the major acetylation site. At high enzyme to substrate ratio, Rtt109 also acetylated K14, K18, K27 and to a lower extent K56 of histone H3. Importantly, this study revealed that in contrast to Rtt109-Vps75, Rtt109-Asf1 displayed a far greater site-specificity, with K56 being the primary site of acetylation. For the first time, we also report the acetylation of histone H4 K12 by Rtt109-Vps75, whereas Rtt109-Asf1 showed no detectable activity toward H4.



Graphical abstract

2.2. Introduction

Eukaryotic DNA is assembled into a highly compact and dynamic nucleoprotein structure known as chromatin. The basic repeating subunit of chromatin is the nucleosome core particle (NCP) that comprises 147 base pairs of DNA wrapped nearly twice around a histone octamer [1]. The octamer is formed of an $(H3–H4)_2$ tetramer (two molecules each of both H3 and H4) paired with two H2A/H2B dimers. Neighboring NCPs are joined by variable lengths of linker DNA (10–80 bp) [2] thatbinds to histone H1 family members to promote chromatin compaction into a 30-nm fiber [3]. The N-terminal tails of core histones protrude from the nucleosome, and contain multiple lysine residues that can be modified by post-translational modifications (PTMs) such as acetylation, methylation, ubiquitylation and sumoylation. These modifications can act individually or in a combinatorial fashion to mediate epigenetic regulation of chromatin structure [4,5]. The acetylation of lysine residues represents one of the most extensively studied histone modifications thus far. This modification can directly affect the chromatin structure and/or favor the recruitment of other binding partners by providing a docking platform for histone modifying enzymes or ATP-dependent chromatin remodeling machineries [6–9]. The acetylation of specific lysine residues can also provide a chromatin mark for the regulation of gene expression [10] or to identify sites of DNA damage. For instance, recent studies have shown that acetylation of H3K56 and H4K16 is significantly elevated in mammalian cells in response to DNA damage [11–13].

Histone acetylation is regulated by two enzyme families known as histone acetyltransferases (HATs) and histone deacetylases (HDACs) [14,15]. HATs catalyze the transfer of acetyl groups from acetyl-coenzyme A (CoA) to the ε-amino group of specific lysines on histone substrates while HDACs remove N-acetyl moieties on target lysine residues. These two enzyme families act antagonistically to modulate gene expression [16], and abnormal activity of some of these enzymes have been observed in different types of cancer [17]. For example, mutations in CBP (which encodes a HAT) and chromosomal translocation with MLL are hallmarks of different human diseases including specific types of leukemia [17–19].

The regulator of Ty1 transposition gene product 109 (Rtt109) is a fungal-specific HAT that acetylates newly synthesized histone H3 prior to its incorporation into chromatin. Rtt109 is involved in the response to genotoxic agents and the regulation of gene expression [20,21]. Rtt109 has a conserved core region that is structurally related to other HAT enzymes though they share limited sequence homology. Recent structural and biochemical analyses of Rtt109

revealed significant homology to the metazoan p300/CBP HAT domain, and also highlighted important differences in their catalytic properties [22]. The enzyme activity of Rtt109 is greatly enhanced by its association with either Vps75 or Asf1 histone chaperones [23,24]. These two chaperones are not structurally related but they both bind to new H3/H4 molecules [25,26]. While Rtt109 interacts only weakly with Asf1, crystal structures revealed the formation of a tight and stable complex with the chaperone Vps75 [27–29]. The chaperones are required to present the substrate in a distinct orientation that confers target specificity to Rtt109. A recent review by D'Arcy et al. indicated that Rtt109–Vps75 sequentially acetylates histone H3 at K9 and K27, whereas the Rtt109–Asf1 complex primarily acetylates K56 [29]. However, the exact sites of acetylation targeted by Rtt109–Vps75 are not entirely clear as an early study by Berndsen et al. [30] reported that this enzyme acetylated H3K9 and H3K23 but not H3K27.

In the present study, we used quantitative proteomics to determine the site-specific acetylation of histones by Rtt109 associated with either the Vps75 or the Asf1 histone chaperone. This is achieved by using a two-pronged approach that we previously developed to profile changes in histone modifications between wild-type and HAT mutant yeast strains [31] and for human cells incubated with different HDAC inhibitors [32]. First, temporal changes in H3–H4 acetylation following incubation with Rtt109–Vps75 and Rtt109–Asf1 are determined from the LC–MS mass profiles of intact histones. Second, the acetylated histones are subjected to propionylation, tryptic digestion and LC–MS/MS analyses to identify acetylated lysine residues and their respective changes in stoichiometry. This strategy enabled precise and comprehensive determination of acetylated lysine stoichiometry to unambiguously identify chaperone-mediated acetylation of histones H3 and H4 by Rtt109.

2.3. Methods

2.3.1. Reagents

Capillary LC columns for nano LC–MS were packed in-house using Jupiter C18 (3 μm particle Phenomenex), and fused silica tubing (Polymicro Technologies). Trapping columns were packed in-house using the same bulk material in Teflon tubing (Supelco). Water and acetonitrile (ACN) for chromatographic analysis were all HPLC grade (Fisher Scientific and in-house Milli-Q water). Trifluoroacetic acid (TFA), anhydrous methanol, ammonium bicarbonate, ammonium hydroxide, and propionic anhydride were all purchased from Sigma. Formic acid (FA) was obtained from EMD Science. Porcine modified trypsin (sequencing grade) was obtained from Promega.

2.3.2. Histone purification

Recombinant $(H3–H4)_2$ tetramers or monomeric H3 were purified from Escherichia coli as previously described [33]. The holoenzyme Rtt109–Vps75 was expressed and purified according to methods reported by Tang et al. [28]. Because Asf1 and Rtt109 do not form a stable complex, the two polypeptides were expressed and purified separately. The gene encoding full-length Rtt109 was cloned into a modified pET Duet vector (Novagen) with a TEV-cleavable GST tag, as previously described [22]. A gene encoding Saccharomyces cerevisiae 6His-Asf1N(1–154) was constructed into a modified pET Duet vector (Novagen) with a thrombin/TEV-cleavable 6His tag. Both recombinant proteins were expressed in E. coli BL21(DE3). After induction with 0.8 mM IPTG, cells were incubated overnight at 18 °C and then harvested for protein purification. The GST-Rtt109 protein was purified using a standard GST-affinity procedure with on-resin TEV protease digestion. The cleaved Rtt109 protein was eluted off the resin using PBS buffer and then flash frozen using liquid nitrogen and stored at −80 °C for future use. The 6His-tagged Asf1 protein was purified through Ni-NTA affinity and ion exchange chromatography before being subjected to TEV protease cleavage in solution and size exclusion chromatography.

2.3.3. HAT assay

In vitro HAT assays were performed using recombinant $(H3–H4)_2$ tetramer (or H3 monomer), acetyl-CoA and Rtt109–Vps75/Asf1 enzyme–chaperone complexes. The assays were performed byincubating the substrates and enzyme/chaperone complexes for different time periods (0 to 90 min) at 30 °C. Unless otherwise indicated, histone tetramers or

monomers were incubated with Rtt109–Vps75/Asf1 enzyme–chaperone complexes at a 1:1 enzyme-to-substrate ratio. Acetyl-CoA was added to all samples to a final concentration of 100 µM. The reaction buffer consisted of 1 mM DTT in 25 mM ammonium bicarbonate solution (pH 7.8). The enzymatic reactions were quenched with 5% TFA. A control sample (without Rtt109) was included to monitor assay conditions.

2.3.4. Sample preparation and protein digestion

Histone samples were desalted on a pre-equilibrated C18 ZipTip pipette tips (Millipore). Desalting was performed according to manufacturer's instructions. Each eluate was divided into two portions and dried in a speed-vac. The first portion of samples was resuspended in 0.2% FA for analysis of intact histones by LC/MS. The remaining samples were subjected to propionylation, tryptic digestion and peptide sequencing by LC–MS/MS. The propionylation reaction is performed by adding 200 µl of the propionylation reagent (de-ionized water: propionic anhydride, 2:1 (v/v)) to the histone samples and vortexing the mixture for 1 h at room temperature. Samples were then dried, resuspended in 0.1 M ammonium bicarbonate (pH 7.8), and digested overnight at 37 °C using trypsin (1:50, enzyme:substrate ratio). Digests were resuspended in 0.2 % FA in water (v/v) prior to LC–MS/MS analyses [31].

2.3.5. LC/MS analysis of intact histones

Intact histone analysis was performed on an Agilent QTOF 6520 LC/MS system equipped with nano-ESI source. The mass spectrometer was operated in positive ion mode and scanned from m/z 400 to 1600. Prior to injection, histone samples were diluted to 200 ng/µl using 0.2% FA in water. An aliquot of 5 µl was loaded on a custom C18 trap column (4 mm length, 360 µm i.d.) for 5 min at 15 µl/min using 0.2% FA. Histones were then eluted onto a C18 analytical column (10 cm length, 150 µm i.d.) using a gradient of 5 to 60% aqueous ACN (0.2% FA) in 60 min at 600 nl/min. The Agilent MassHunter Workstation Software (version B.01.04) was used to deconvolute the multiply-charged protein ions over a discrete time-segment of the LC–MS run.

2.3.6. LC–MS/MS analysis of tryptic peptides

Tryptic digests of propionylated histone samples were analyzed on LTQ-Orbitrap XL mass spectrometer (Thermo Fisher scientific) coupled to an Eksigent nano-LC system. The same column and solvent systems were used as described above for the LC/MS analysis of intact histones. The mass spectrometer was operated in a data-dependent acquisition mode with full scan (m/z 300–2000) resolution set to 60,000, with a target value of 1.0×10^{6} . The six most

abundant precursor ions were selected for fragmentation in the LTQ by CID at a normalized CE setting of 35. Fragment ions were analyzed in the LTQ ion trap over the mass range of m/z 200–2000 with a target value of 1.0×10^4.

2.3.7. Database searching and peptide clustering

MS data were analyzed using the Xcalibur software (version 2.0.7). Peak lists were generated using the Mascot distiller software (version 2.3.2.0, Matrix science) where MS processing was achieved using the LCQ_plus_zoom script. Database searches were performed using the search engine Mascot (version 2.2.0, Matrix Science, London, UK) with the concatenated forward and reverse Saccharomyces Genome Database (SGD, version 64.1). The tolerance window for experimental precursor mass values and fragment ion mass values were set to ±15 ppm and 0.5 Da, respectively. The number of allowed missed cleavage sites for trypsin was set to 5 and acetylation (K), propionylation (K), oxidation (M), and deamidation (NQ) were all selected as variable modifications. No fixed modification was included in the search. All peptide identifications were transferred to ProteoConnections [34], an in-house developed bioinformatics tool, to generate a .csv (comma separated value) file that can be opened and modified in Excel. The .csv output file contained only distinct peptides (peptide with maximum score is kept) with assigned Mascot score of greater than or equal to 20. Manual MS/MS spectra verification was performed on each modified peptide to confirm the sequence assignment.

The clustering of peptide abundances across different experimental conditions was performed using in-house tools as described previously [32]. Briefly, raw data files (.raw) from the Xcalibur software were converted into peptide map files representing all ions according to their corresponding m/z values, retention time, intensity, and charge state. An intensity threshold of 5000 counts was set as a cut-off for peptide detection. Peptide abundances were determined using the peak top intensity values. Clustering of peptide maps across different sample sets were performed on peptides associated to a Mascot entry using hierarchical clustering with tolerances of ±15 ppm and ±1 min for peptide mass and retention time, respectively. Normalization of retention time is performed on the initial peptide cluster list using a dynamic and nonlinear correction that confines the retention time distribution to less than 0.1 min (< 0.3% relative standard deviation, RSD) on average. Global intensities were normalized based on the average intensity of all peptide signals. Peptide clusters showing reproducible changes in abundance across conditions were manually inspected to validate identification and changes in abundance.

2.4. Results

2.4.1. nLC/MS analysis of histones

Our MS analysis of histones involves a two-pronged strategy to profile the distribution of molecular mass and identify sites of modification [31]. This approach utilizes both intact histones and tryptic peptides to profile global and site-specific changes in histone modifications. First, intact histones are analyzed to determine changes in their mass profiles that can be attributed to specific modifications. Second, peptide maps generated from nano LC–MS/MS analysis of the corresponding tryptic digests are compared to identify peptides showing statistically meaningful changes in abundance across conditions. In the present study, we extended the application of this approach to determine the kinetics and site-specificity of Rtt109 toward histones H3 and H4. The experimental workflow is illustrated in Fig. 2.1. Prior to MS analysis, recombinant histones H3 and H4 (monomeric or tetrameric forms) were incubated with either Rtt109–Vps75 or Rtt109–Asf1 HAT complex for different time periods (0 to 90 min). A portion of each intact histone sample was analyzed by LC/MS on a QTOF instrument, while the remaining sample was propionylated and digested by trypsin prior to peptide sequencing by LC–MS/MS on a LTQ-Orbitrap XL mass spectrometer.

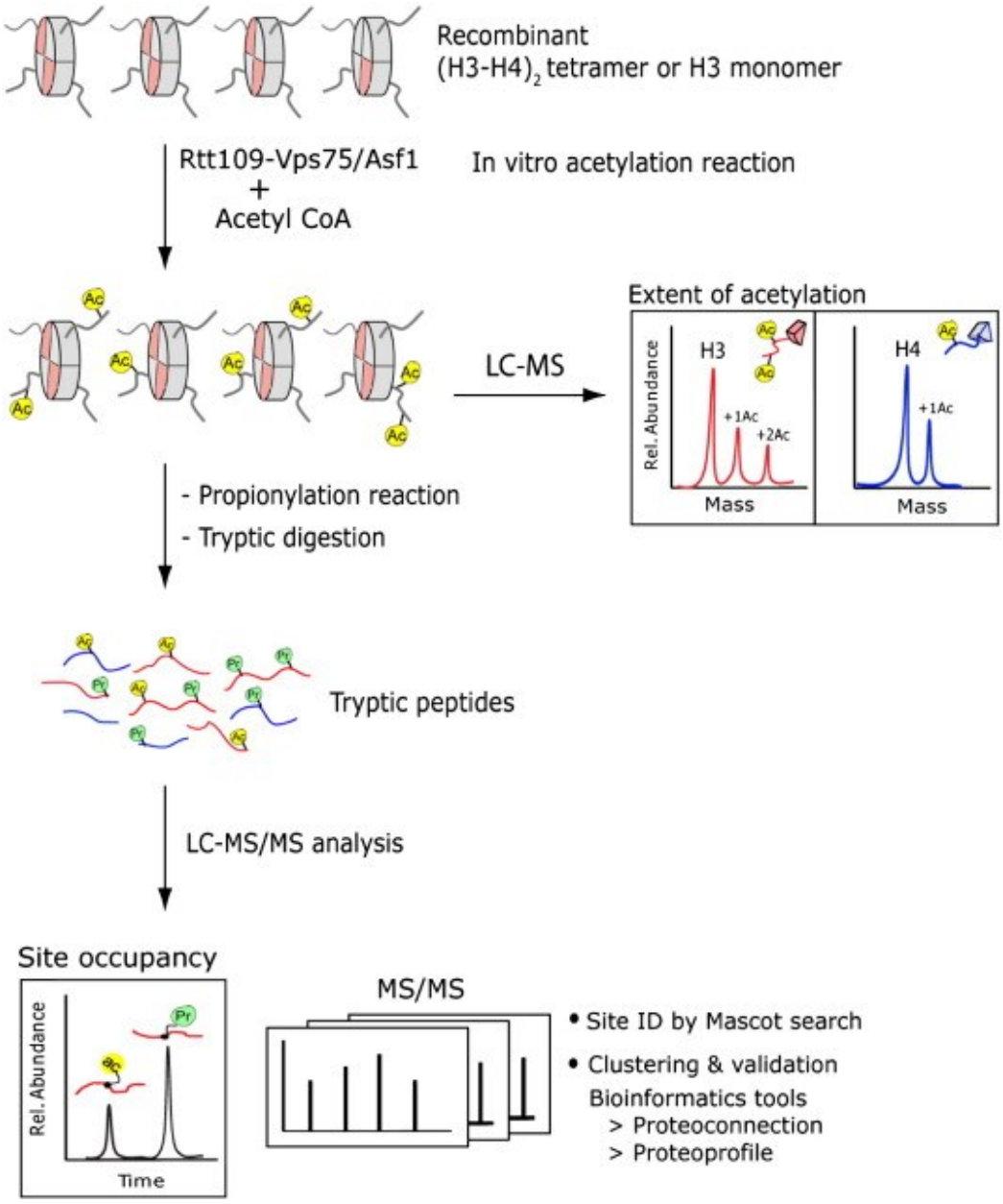


Figure 2.1 Overview of the analytical scheme for profiling histone modifications.

2.4.2. Histone acetylation by Rtt109–Vps75/Asf1 complexes

To profile the progressive changes in acetylation of the H3:H4 tetramers upon incubation with either Rtt109–Vps75 or Rtt109–Asf1 complexes, we first analyzed the enzymatic products using LC–MS. An enzyme-to-substrate ratio of 1:1 was maintained in all cases. H3 and H4 histones along with Rtt109 and its chaperones were chromatographically separated and displayed abundant multiply-charged ion spectra from which the molecular mass could be determined. Under the present chromatographic conditions, we identified H4 (Mobs: 11,237 Da), and H3 (Mobs: 15,225 Da), at 43.6 and 47.1 min respectively (Fig. S2.1). The

observed molecular mass was within ±1 Da of the value calculated from the protein sequence. To account for changes in histone acetylation over different incubation periods, we normalized the intensity of individual peak to that of histones from the control sample. The relative proportion of histones displaying different degrees of acetylation was determined from the abundance ratio of a specific acetylated form to the sum of all forms of identified histones.

Fig. 2.2 shows the intact mass profiles of acetylated histone H3 (M.theor. 15,225 Da) following the in vitro HAT assay with Rtt109–Vps75 and Rtt109–Asf1 complexes. As expected, a progressive increase in histone H3 acetylation is observed over the first 90 min of incubation. Interestingly, a broader distribution of acetylated products is observed for Rtt109–Vps75compared to Rtt109–Asf1 although identical enzyme:substrate ratios were used in all cases. Acetylation of H3 by Rtt109–Vps75 proceeded rapidly, and up to three acetyl groups were detected within the first minute of incubation (Fig. 2.2A). In contrast, acetylation of H3 by Rtt109–Asf1 progressed more slowly, and the first acetylated form of H3 was detected only after 5 min of incubation (Fig. 2.2B). We observed that incubation of H3 with Rtt109–Vps75 yielded up to 5 acetylation sites over the first 90 min of the in vitro HAT assay. The relative proportion of each acetylated form of H3 is shown in Fig. 2.2C. As indicated, H3 is rapidly acetylated by Rtt109–Vps75 within the first 15 min. The successive acetylation resulted in mono- to penta-acetylated H3 with the tetra acetylated form being predominant beyond 30 min of incubation. In contrast, Rtt109–Asf1 showed a narrower distribution of H3 acetylation, and no more than three sites were detected over the same incubation period (Fig. 2.2D). The relative proportion of unmodified H3 decreased very slowly, and the mono-acetylated H3 remained the most abundant form detected after 90 min of incubation.

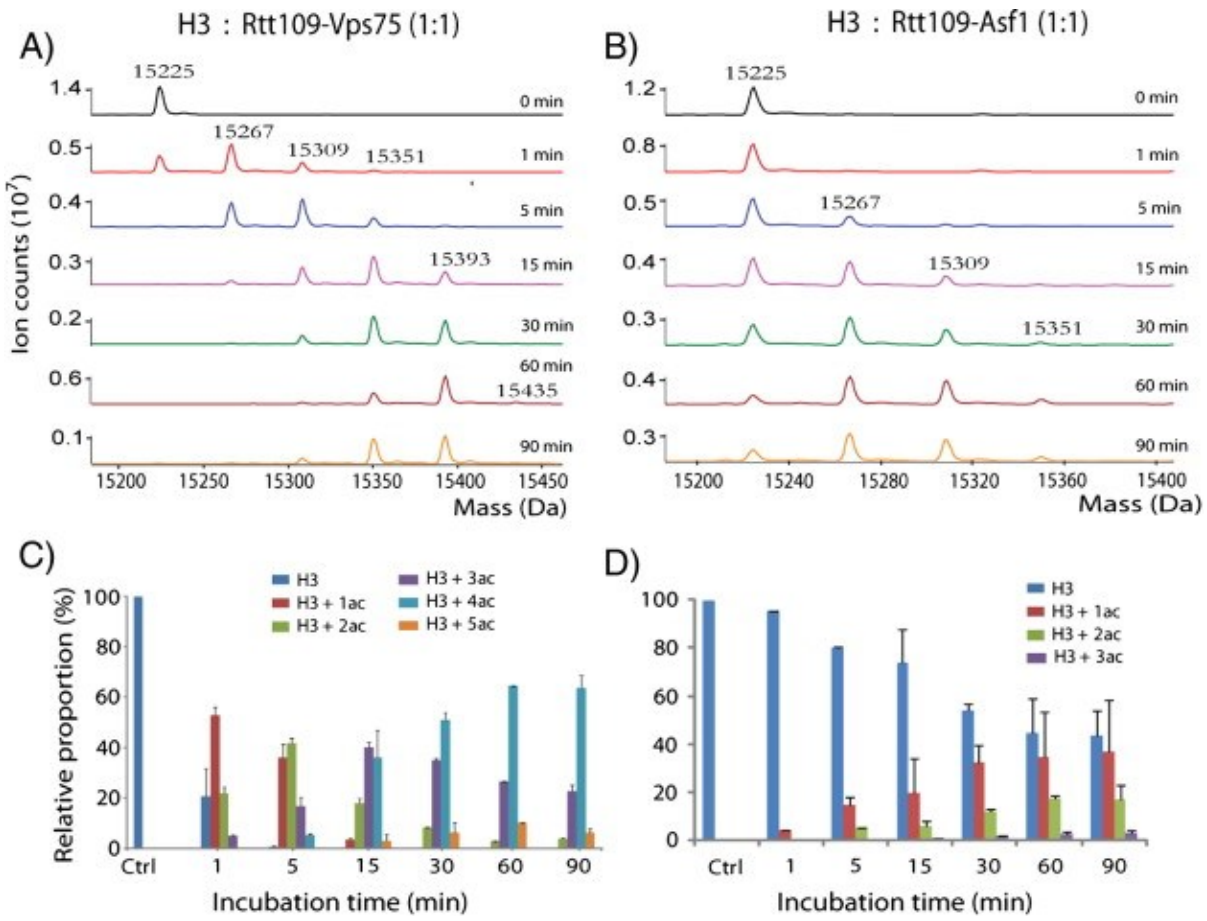


Figure 2.2 LC/MS analyses of intact histone H3 obtained from in vitro HAT assay.
(A) Intact mass profiles of histone H3 obtained from in vitro HAT assay with Rtt109–Vps75 and (B) with Rtt109–Asf1. Incubation times (from top to bottom) are 0, 1, 5, 15, 30, 60 and 90 min. (C) Bar graph represents the extent of H3 acetylation after HAT assay with Rtt109–Vps75 (raw intensity data: Table S2.1) and (D) with Rtt109–Asf1 (raw intensity data: Table S2.2).

Although Vps75 and Asf1 were previously reported to affect the acetylation status of histone H3, we also examined the activity of these chaperones on the acetylation of H4 by Rtt109. Similar to that described before, we determined the extent of acetylation of histone H4 following incubation of the H3:H4 tetramers with Rtt109–Vps75 and Rtt109–Asf1 (Fig. 2.3). Upon incubation with Rtt109–Vps75, no significant acetylation of histone H4 (Mobs: 11,237 Da) was observed until 5 min of incubation (Fig. 2.3A). Longer incubation periods led to increased abundance of the mono-, di-, and tri-acetylated forms of H4 at 11,279, 11,321 and 11,363 Da, respectively. However, the mono-acetylated form of H4 was the predominant reaction product of Rtt109–Vps75, while the original H4 substrate remained present over the

entire incubation period (Fig. 2.3B). In contrast, Rtt109–Asf1 did not show any activity towards histone H4 acetylation in spite of the high enzyme-to-substrate ratio used (Fig. 2.3C). To our knowledge, no previous study has reported the activity of Rtt109 on the acetylation of histone H4.

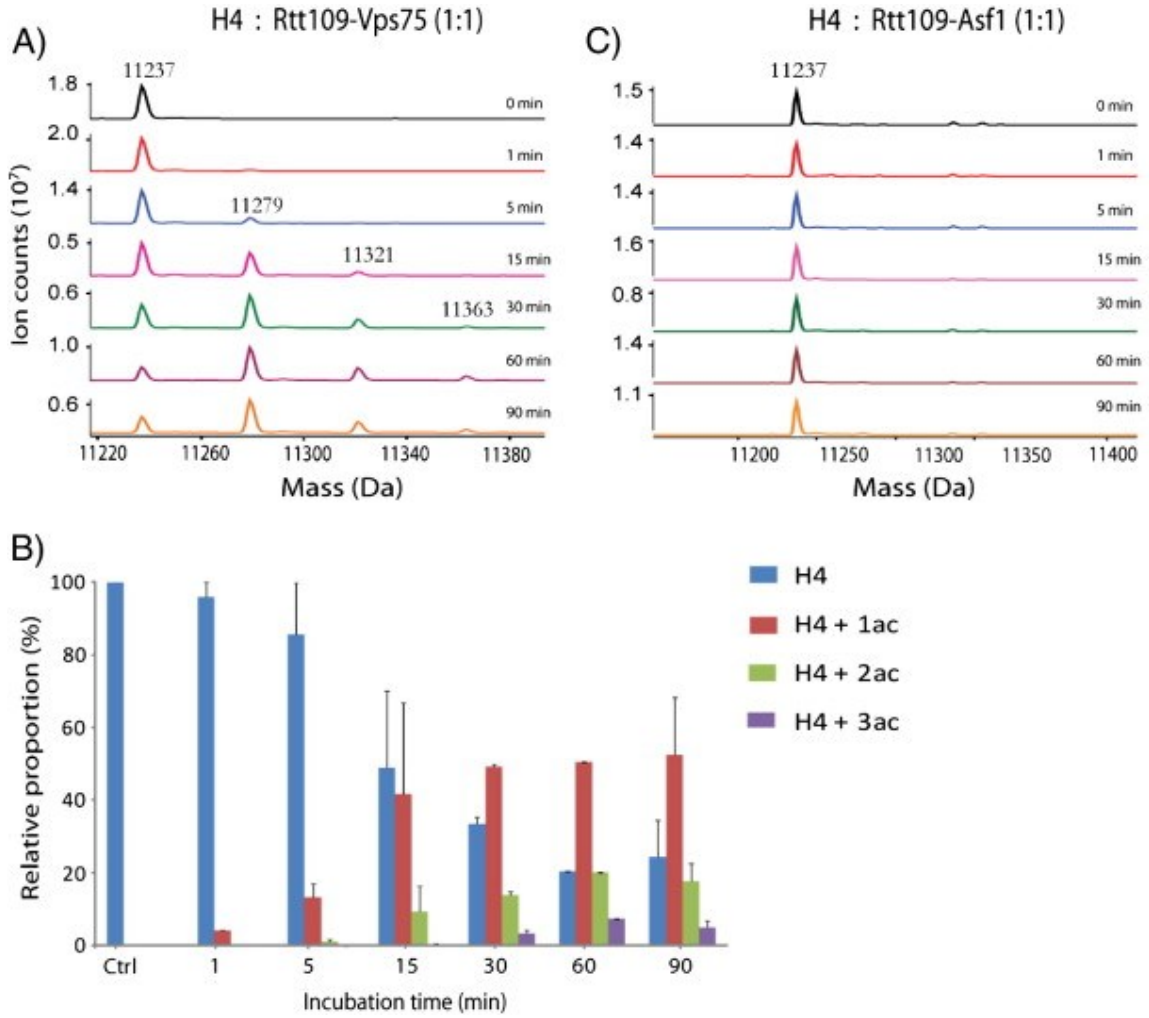


Figure 2.3 LC/MS analyses of intact histone H4 obtained from in vitro HAT assay.
(A) Intact mass profiles of histone H4 obtained from in vitro HAT assay with Rtt109–Vps75. (B) Bar graph showing the extent of H4 acetylation after assay with Rtt109–Vps75. (C) Intact mass profiles of histone H4 obtained from in vitro HAT assay with Rtt109–Asf1. Incubation times (top to bottom) are 0, 1, 5, 15, 30, 60, 90 min.

To address whether acetylation of histones H3 and H4 was dependent on the oligomeric state of the substrates, we performed in vitro HAT assays using tetrameric histones (H3–H4)2 or individual H3 or H4 histones. These experiments indicated that the chaperone-dependent activity of Rtt109 is not significantly affected by the oligomeric state of the histone substrates, and trends in acetylation observed for histone monomers were similar to those of tetramers (Fig. S2.2).

2.4.3. Effect of enzyme-substrate ratios on histone acetylation

Next, we examined the activity of Rtt109–Vps75 at different enzyme-to-substrate ratios. We performed in vitro HAT assays at enzyme-to-substrate molar ratios of 1:200, 1:100, 1:50 and 1:10 for a fixed incubation period of 15 min. The relative proportion of all detected acetylated histones at different enzyme:substrate ratios is shown in Fig. 2.4. At a molar ratio of 1:200 we observed two acetylated forms of H3, though their relative abundances were generally low (Fig. 2.4A). Further analyses of these tryptic digests using LC–MS/MS identified two acetylation sites at positions K9 and K23 of H3, with K9 being the most abundant (see Section 3.4). The relative proportion of the mono-acetylated form increased by 10-fold for an enzyme-to-substrate ratio of 1:100 and further increases in the mono and multiacetylated forms of H3 were observed for higher enzyme:substrate ratios. We also examined the acetylation profiles of H4 under the same incubation conditions (Fig. 2.4B). These analyses revealed that a mono-acetylated form of H4 is typically observed irrespective of the enzyme-to-substrate ratio used. These observations are consistent with those noted previously for an enzyme-to-substrate ratio of 1:1. LC–MS/MS analyses of the corresponding tryptic digests indicated that the primary site of H4 acetylation by Rtt109–Vps75 is K12 (see Section 3.4). It is noteworthy that only the mono-acetylated histone H3 was detected for incubation periods of 15 min and 60 min when experiments were performed using Rtt109 alone (Fig. 2.4C). Consistent with previous reports, we identified K9 as the mono-acetylated residue of H3 using LC–MS/MS of the corresponding tryptic digest [30].

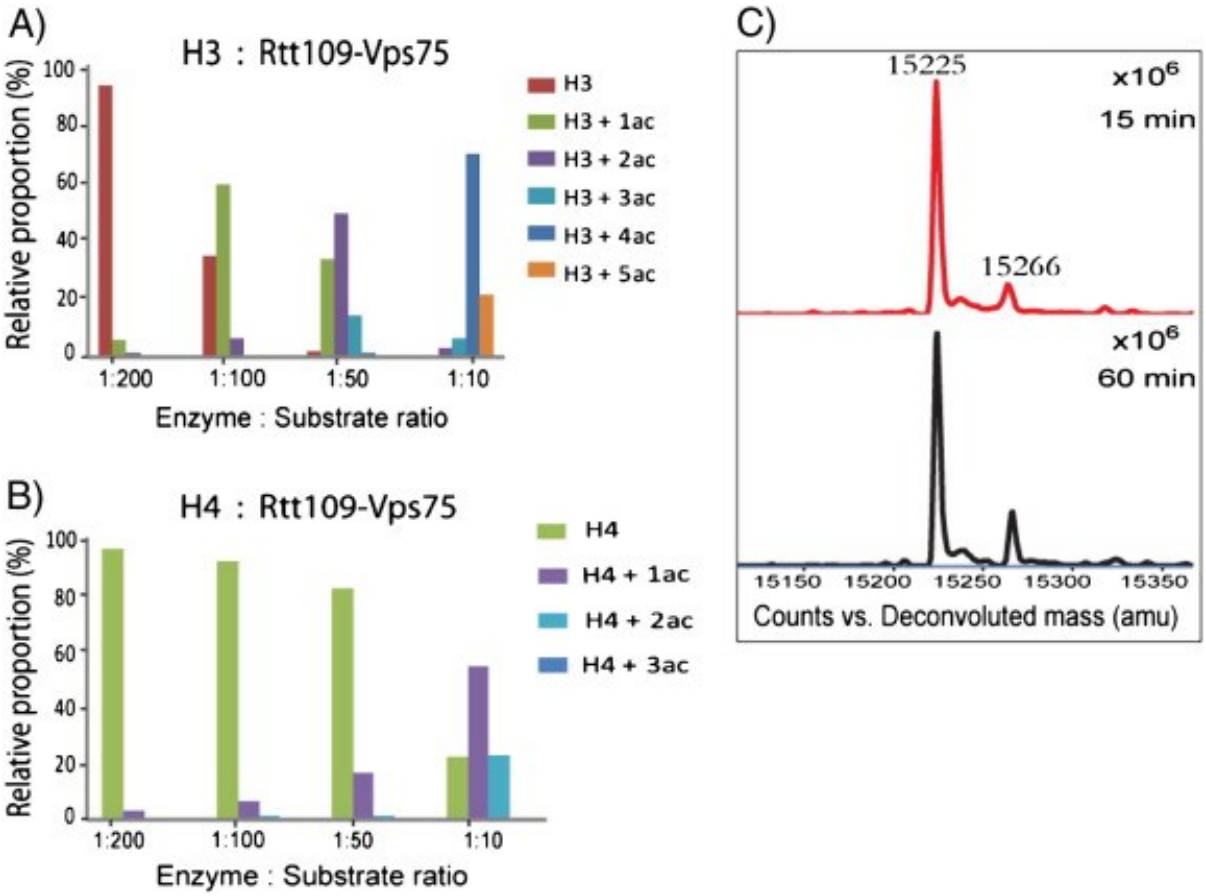


Figure 2.4 Relative proportion of histone acetylation by Rtt109–Vps75.
Distribution of unmodified, mono- (1ac), di-(2ac), tri-(3ac), and tetra-acetylated (4ac) histone (A) H3 and (B) H4 at different enzyme to substrate ratios. (C) In vitro acetylation of histone H3 acetylation by Rtt109 alone (without chaperone) after an incubation period of 15 min (upper panel) and 60 min (lower panel).

2.4.4. Peptide identification and determination of acetylation stoichiometry

To determine the extent and location of acetylation sites following in vitro HAT assay, all samples were subjected to propionylation, tryptic digestion and LC–MS/MS analyses. Propionylation of free Lys residues of histones was required to prevent the formation of small hydrophilic tryptic peptides that would be either too short for MS detection or could not be retained efficiently by reversed phase LC [31]. The LC–MS/MS analyses of the corresponding tryptic digests enabled the identification of at least 18 distinct peptides comprising 20 Lys residues, and these peptides typically contained one or two modified sites (see Tables S2.3–2.4). The extent of acetylation of a specific Lys residue was determined from the intensity ratio of the acetylated peptide to the sum of all forms of the corresponding tryptic peptide (acetylated and propionylated). For example, Fig. 2.5A shows representative extracted ion chromatograms of the doubly-protonated peptide ions corresponding to the

mono-acetylated (K9ac, K14pr) and di-propionylated (K9pr, K14pr) forms of the tryptic peptide 9 KSTGGKAPR17 at an incubation time of 1 min. The intensity ratio of the 9 KacSTGGKprAPR17 peptide to that of all forms of this peptide corresponded to 0.2 or an acetylation stoichiometry of 20%. The MS/MS spectrum of m/z 500.282+ assigned to the mono-acetylated tryptic peptide 9 K(ac) STGGK(pr)APR17 is shown in Fig. 2.5B. The incremental 42 Da mass shift of the b2–b5 fragment ions confirmed the acetylation of K9 residue while an additional mass shift of 56 Da on b6–b8 fragment ions was consistent with a propionylated K14 residue. The doubly-propionylated form of the same tryptic peptide displayed similar fragmentation patterns except that 56 Da shifts were observed for both K9 and K14 residues. Manual validation was performed on each MS/MS spectrum to confirm the identity of the modified peptides and to correlate the abundance of all peptides across the entire incubation period. Fig. 2.5C shows the site occupancy of Lys acetylation following in vitro incubation with Rtt109–Vps75. This graph indicates that Rtt109–Vps75 preferentially acetylates H3 at K9 and K23 residues with the former site displaying the most significant changes in acetylation during this period. Additional acetylation sites at K14, K18, K27 and K56 were also identified beyond 15 min suggesting that Rtt109–Vps75 can promiscuously catalyze reactions on other sites when long incubation periods or high enzyme-to-substrate ratios are used.

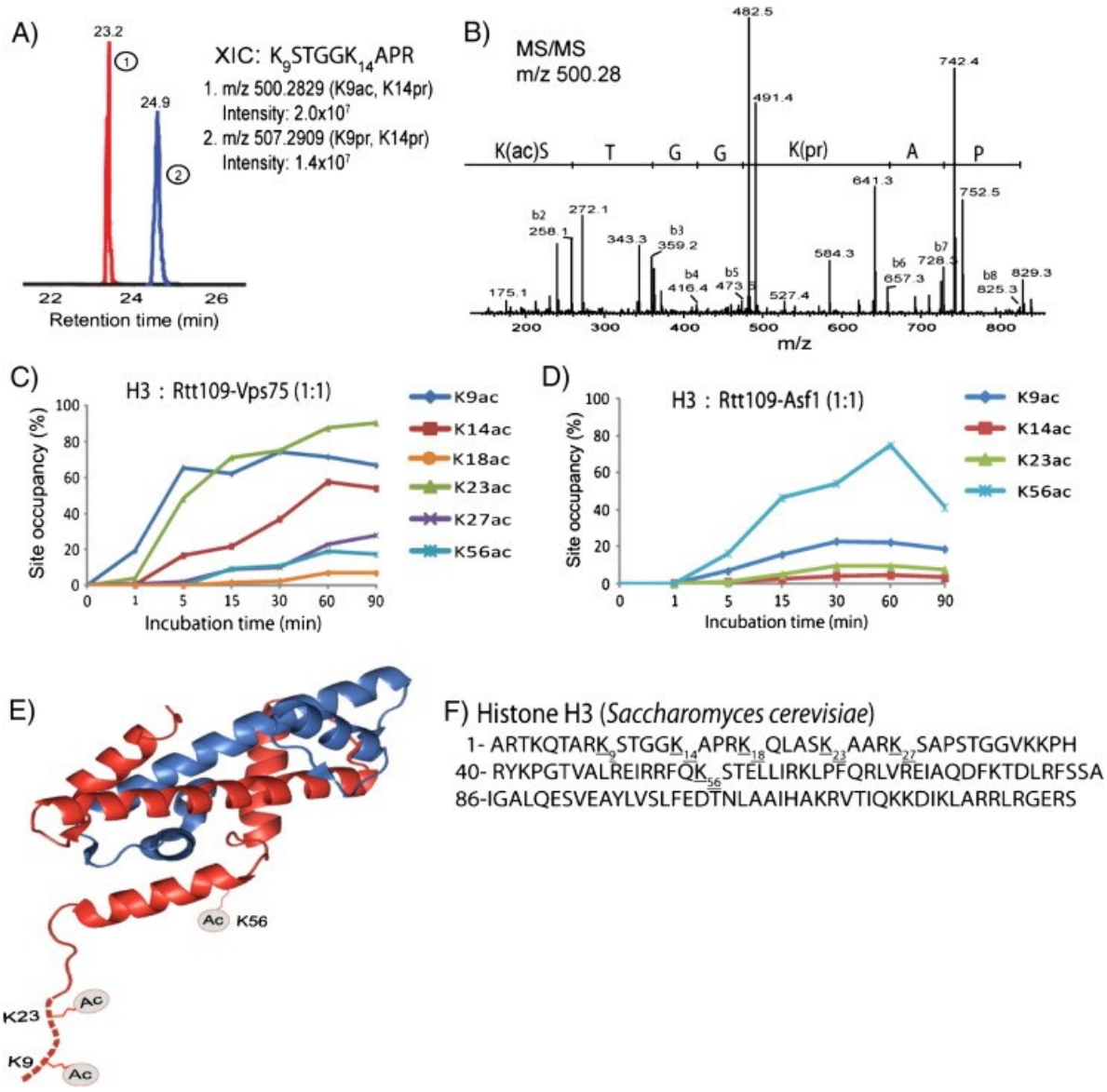


Figure 2.5 Analysis of acetylation stoichiometry of H3 incubated with Rtt109–Vps75.
(A) Extracted ion chromatograms of the doubly charged tryptic peptides corresponding to the acetylated and propionylated forms of 9 KSTGGKAPR17. (B) MS/MS spectrum of the acetylated tryptic peptide 9 KacSTGGKAPR17. Acetylation stoichiometry of H3 lysine residues following incubation with (C) Rtt109–Vps75 and (D) Rtt109–Asf1. (E) Primary acetylation sites of Rtt109–Vps75 (K9 and K23 located on the N-terminus tail of H3) and Rtt109–Asf1 (K56 located on the H3 fold domain). The cartoon representation of yeast histone H3 (red) and H4 (blue) are taken from nucleosome PDB 1ID3 (prepared using the Pymol Molecular Graphics system, version 1.2r1). (F) Amino acid sequence of yeast histone H3.

In contrast, in vitro reactions performed with histone H3 incubated with Rtt109–Asf1 showed strikingly different patterns with a limited number of acetylated residues (Fig. 2.5D). We noted that Rtt109–Asf1 displayed greater substrate specificity towards K56 acetylation, a

residue that is located in the histone fold domain unlike the N-terminal Lys residues (e.g. K9, K23) modified by Rtt109–Vps75 (Fig. 2.5E). Extended incubation periods with Rtt109–Asf1 led to the acetylation of K14 and K23 residues although these sites could arise from promiscuous enzymatic reactions.

We also examined the enzymatic activity of Rtt109–Vps75 and Rtt109–Asf1 on the acetylation of H4 also present in the histone tetramer. Fig. 2.6 shows the sites and extent of acetylation of histone H4 following in vitro incubation with Rtt109–Vps75. We identified a total of 3 acetylation sites on H4, consistent with that observed from intact mass measurements (Fig. 2.3). The LC–MS/MS analysis for the tryptic digest of histone tetramers after 15 min incubation with Rtt109–Vps75 is shown in Fig. 2.6A for the mono-acetylated (K12ac, K5pr, K8pr, K16pr) and the fully propionylated tryptic peptide 4 GKGGKGLGKGGAKR17. The MS/MS spectrum of m/z 740.942+ corresponding to the mono-acetylated tryptic peptide 4 GKGGKGLGKacGGAKR17 is shown in Fig. 2.6B. An incremental shift of 56 Da is observed for fragment ions y2–5 and y10–12 consistent with propionylation at K16 and K8 residues. The location of the acetylated K12 residue was confirmed by the observation of a mass shift of 42 Da for y6–9 fragment ions while b2 (m/z 242.0) and b3 (m/z 299.3) displayed a 56 Da shift, in agreement with a propionylated K5 residue. The acetylation profiles of K5, K8 and K12 residues are shown in Fig. 2.6C for the 90 min incubation period with Rtt109–Vps75. The acetylation of H4K12 proceeded more rapidly than that of K5 and K8 residues (Fig. 2.6C). An Rtt109–Vps75-dependent acetylation of histone H4 has not been reported in a literature, thus far. Interestingly, no acetylation of H4 was observed when the same reaction was performed using Rtt109–Asf1 suggesting again that this chaperone exerts more specific catalytic activity toward its histone substrate.

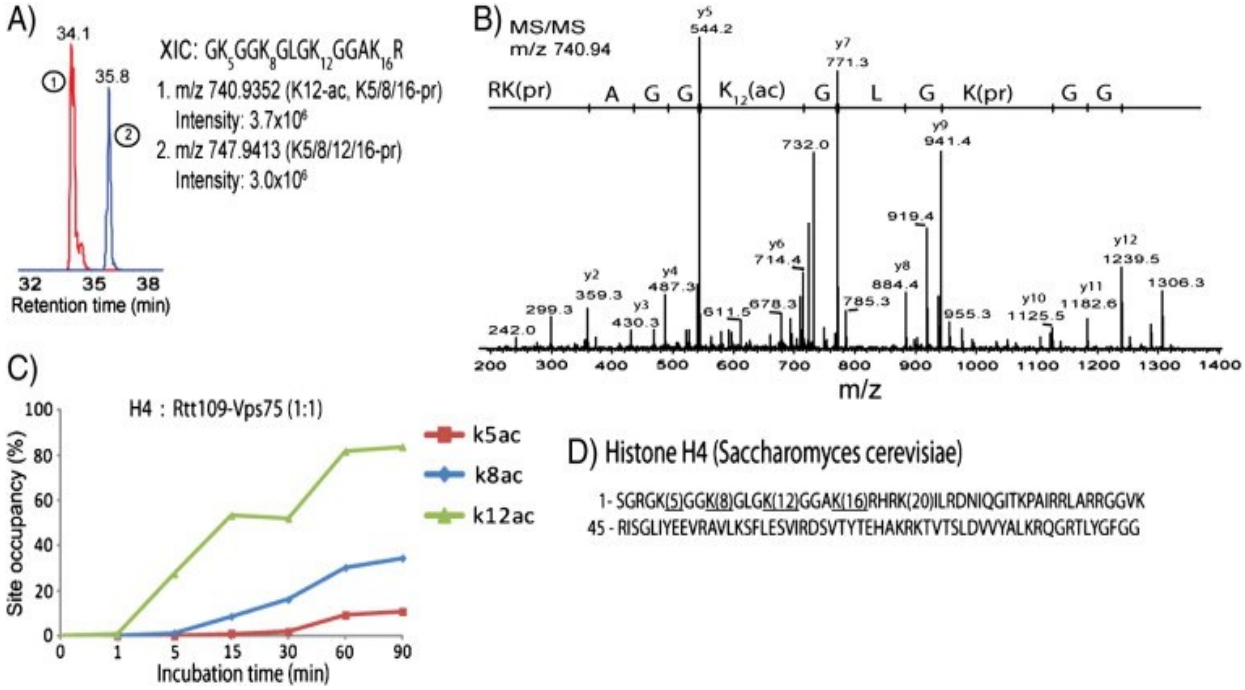


Figure 2.6 Analysis of acetylation stoichiometry of H4 incubated with Rtt109–Vps75.
(A) Extracted ion chromatograms of the doubly charged tryptic peptides corresponding to the acetylated and propionylated forms of 4 GKGGKGLGKGGAKR17. (B) MS/MS spectrum of the acetylated tryptic peptide 4 GKGGKGLGKacGGAKR17. (C) Acetylation stoichiometry of H4 lysine residues (K5, K8, K12) following incubation with Rtt109–Vps75. (D) Amino acid sequence of yeast histone H4.

2.5. Discussion

This study describes a comprehensive proteomics approach to determine the chaperone-dependent specificity of the Rtt109 enzyme. This approach is based on a two-pronged strategy used previously to profile global and site-specific changes in histone acetylation upon HDACi treatment [31]. Our global analyses revealed that the sites acetylated in histones H3 and H4 vary significantly depending on the type of chaperone associated with the Rtt109 HAT. While this enzyme can acetylate H3K9 on its own, its association with either Vps75 or Asf1 chaperones not only increases its catalytic activity but also confers additional substrate selectivity. In vitro, Rtt109–Vps75 catalyzed the acetylation of histone H3 and exhibited a broader substrate activity asevidenced by the number of acetylated forms detected (Figs. 2.2A and 2.3B). Our peptide sequencing analysis identified two preferential acetylation sites on H3 (K9 and K23) and only one on H4 (K12). It is known that H4K12 is acetylated in vivo. However, we cannot exclude the possibility that promiscuous acetylation of H4K12 by the Rtt109–Vps75 enzyme could take place in vivo as observed here for the in vitro HAT assay. The significance of this finding would clearly require further functional analyses given recent evidences indicating that acetylation of H4K12 was shown to regulate dynamic telomere properties in S. cerevisiae [35].

The identification of H3K23 acetylation using quantitative proteomics also provided clear information on the site-specific activity of Rtt109–Vps75. Although several studies agree on K9 as a primary H3 acetylation site, conflicting results were reported regarding a second acetylation site at positions K23 or K27 [24,30,36]. These sites were previously characterized using western blotting, and ambiguity on site assignment could possibly arise from cross reactivity of antibodies against H3 N-terminal tail acetylation sites that harbor similar sequences flanking acetylated residues [32,37]. By using quantitative proteomics we unambiguously identified H3K23 as a second acetylation site targeted by Rtt109–Vps75. It is noteworthy that additional acetylation sites (e.g. K14, K18, K27 and K56) were observed after extended incubation periods (or at high enzymeto-substrate ratio) showing that, at least in vitro, Rtt109–Vps75 can promiscuously acetylate sites other than K9 and K23. The physiological significance of the acetylation site preference exhibited by Rtt109–Vps75 is still unknown. Additional biochemical, structural and in vivo studies will be necessary to address the functional significance of this finding.

In contrast to Rtt109–Vps75, Rtt109–Asf1 catalyzed the acetylation of histone H3 but not H4 (Figs. 2.2D and 2.3C). This complex showed a striking specificity toward H3K56 acetylation, even at very high enzyme-to-substrate ratios. H3K56 is located at the C-terminal edge of an α-helix that is unique to histone H3 and is not part of the histone-fold domain. In addition, H3K56 cannot be acetylated by Rtt109 when H3 is incorporated into nucleosomes [38]. Rtt109 acetylation is mostly associated with non-nucleosomal, nascent histone H3 in yeast [39,40]. Rtt109 acetylation of newly synthesized H3 is consistent with recent evidences on its activity in the cytoplasm [41]. It is noteworthy that the orderly sequence by which H3 is acetylated or the events leading to the specific recruitment of Vps75 and Asf1 chaperones to Rtt109 are presently unknown. However, the possibility of determining enzyme-specificity in vitro or in vivo using quantitative proteomics opens up new perspectives to decipher the roles of chaperones in regulating HAT activity and substrate selectivity.

2.6. Acknowledgments

This work was carried out with a financial support from the Natural Sciences and Engineering Research Council (NSERC) and the Canadian Institute for Health Research (CIHR) to A.V and P.T. The Institute for Research in Immunology and Cancer (IRIC) receives infrastructure support from the Canadian Center of Excellence in Commercialization and Research, the Canadian Foundation for Innovation, and the Fonds de recherche du Québec-Santé (FRQS).

2.7. Supplementary data

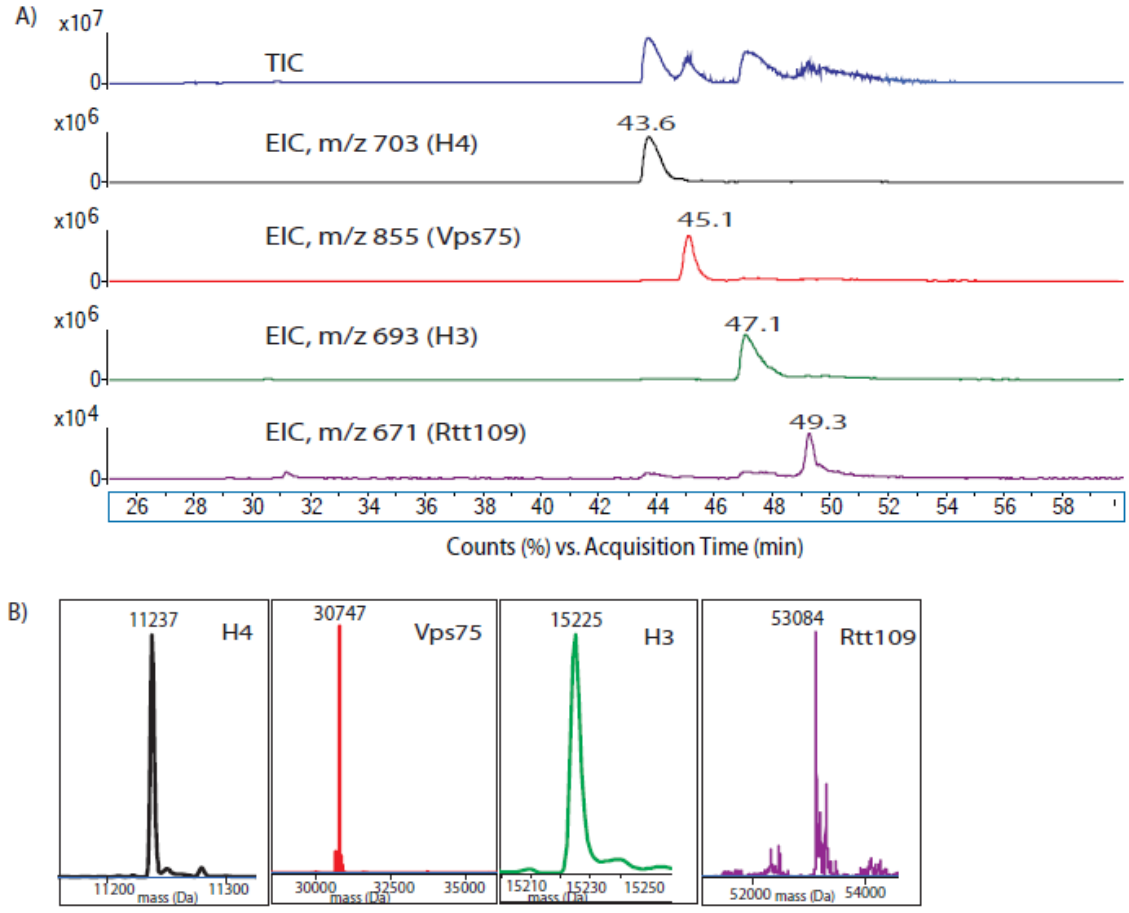


Figure S2.1 Intact mass analysis of intact proteins.
A) Total Ion Chromatogram (TIC) and Extracted Ion Chromatogram (EIC) of proteins obtained from ESI-LC/MS analysis of H3-H4/Rtt109-Vps75 assay, B) Shows the deconvoluted mass of the corresponding EIC of H4, Vps75, H3 and Rtt109.

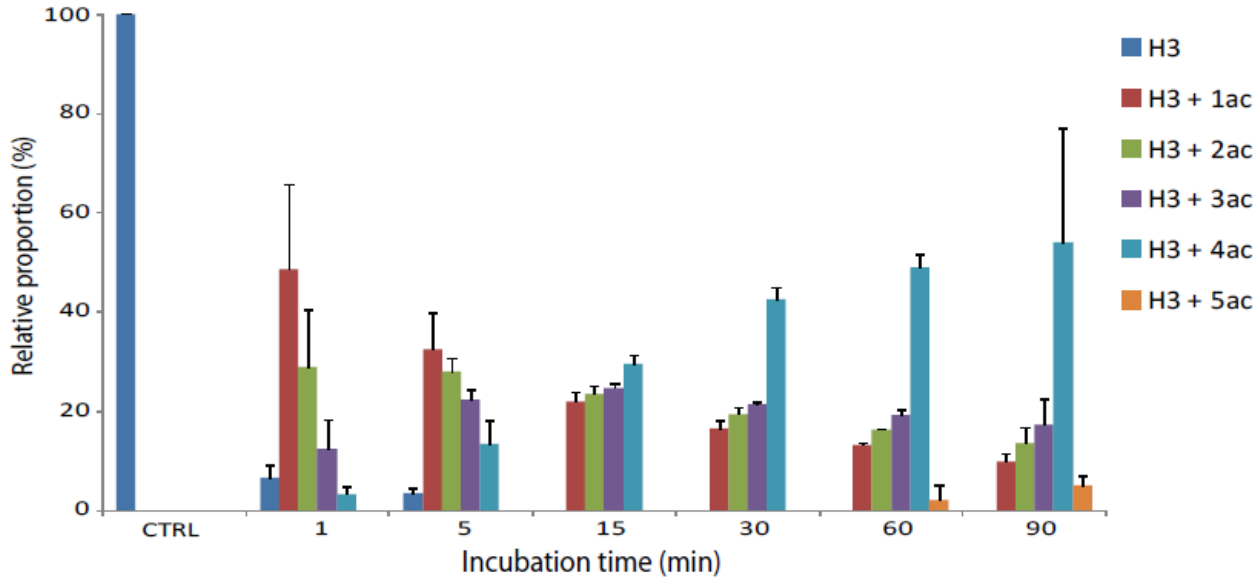


Figure S2.2 The proportion of histone H3 acetylation during 90min H3\Rtt109-Vps75 assay.

Table S2.1 H3/Rtt109-Vps75 assay intact histone intensity data (0 to 90 min)

CTRL

	Replicate	Mass (in Dalton)*	Height
H3	A	15225.11	5.15E+06
	B	15225.34	5.63E+06
Average	-	15225.23	5.39E+06
Standard deviation	-	0.2	3.40E+05
H3 + ac	A	ND	ND
	B	ND	ND
Average	-	-	-
Standard deviation	-	-	-
H3 + 2ac	A	ND	ND
	B	ND	ND
Average	-	-	-
Standard deviation	-	-	-
H3 + 3ac	A	ND	ND
	B	ND	ND
Average	-	-	-
Standard deviation	-	-	-
H3 + 4ac	A	ND	ND
	B	ND	ND
Average	-	-	-
Standard deviation	-	-	-
H3 +5ac	A	ND	ND
	B	ND	ND
Average	-	-	-
Standard deviation	-	-	-

1 min

	Replicate	Mass (in Dalton)	Height
H3	A	15225.31	1.10E+06
	B	15225.08	4.97E+05
Average	-	15225.20	8.00E+05
Standard deviation	-	0.2	4.28E+05
H3 + ac	A	15267.38	1.96E+06
	B	15267.39	2.14E+06
Average	-	15267.38	2.05E+06
Standard deviation	-	0	1.28E+05
H3 + 2ac	A	15309.17	7.83E+05
	B	15309.37	9.13E+05
Average	-	15309.27	8.48E+05
Standard deviation	-	0.1	9.25E+04
H3 + 3ac	A	15351.21	1.81E+05
	B	15351.20	2.01E+05
Average	-	15351.20	1.91E+05
Standard deviation	-	0	1.37E+04
H3 + 4ac	A	ND	ND
	B	ND	ND
Average	-	-	-
Standard deviation	-	-	-
H3 +5ac	A	ND	ND
	B	ND	ND
Average	-	-	-
Standard deviation	-	-	-

5 min

	Replicate	Mass (in Dalton)	Height
H3	A	15224.83	3.21E+04
	B	15224.82	2.54E+04
Average	-	15224.83	2.87E+04
Standard deviation	-	0.0	4.74E+03
H3 + ac	A	15267.25	1.27E+06
	B	15267.40	1.58E+06
Average	-	15267.32	1.43E+06
Standard deviation	-	0.1	2.17E+05
H3 + 2ac	A	15309.41	1.71E+06
	B	15309.39	1.61E+06
Average	-	15309.40	1.66E+06
Standard deviation	-	0.0	7.63E+04
H3 + 3ac	A	15351.25	7.67E+05
	B	15351.23	5.54E+05
Average	-	15351.24	6.60E+05
Standard deviation	-	0.0	1.50E+05
H3 + 4ac	A	15393.38	2.23E+05
	B	15393.32	1.85E+05
Average	-	15393.35	2.04E+05
Standard deviation	-	0.0	2.66E+04
H3 +5ac	A	ND	ND
	B	ND	ND
Average	-	-	-
Standard deviation	-	-	-

15 min

	Replicate	Mass (in Dalton)	Height
H3	A	ND	ND
	B	ND	ND
Average	-	-	-
Standard deviation	-	-	-
H3 + ac	A	15267.08	1.46E+05
	B	15267.04	1.04E+05
Average	-	15267.06	1.25E+05
Standard deviation	-	0.0	2.94E+04
H3 + 2ac	A	15309.36	7.41E+05
	B	15309.40	6.37E+05
Average	-	15309.38	6.89E+05
Standard deviation	-	0.0	7.38E+04
H3 + 3ac	A	15351.25	1.48E+06
	B	15351.49	1.60E+06
Average	-	15351.37	1.54E+06
Standard deviation	-	0.2	8.39E+04
H3 + 4ac	A	15393.30	1.10E+06
	B	15393.56	1.68E+06
Average	-	15393.43	1.39E+06
Standard deviation	-	0.2	4.13E+05
H3 +5ac	A	15435.38	4.58E+04
	B	15435.39	1.86E+05
Average	-	15435.38	1.16E+05
Standard deviation	-	0.0	9.90E+04

30 min

	Replicate	Mass (in Dalton)*	Height
H3	A	ND	ND
	B	ND	ND
Average	-	-	-
Standard deviation	-	-	-
H3 + ac	A	ND	ND
	B	ND	ND
Average	-	-	-
Standard deviation	-	-	-
H3 + 2ac	A	15309.36	2.82E+05
	B	15309.37	2.78E+05
Average	-	15309.37	2.80E+05
Standard deviation	-	0.0	2.83E+03
H3 + 3ac	A	15351.47	1.18E+06
	B	15351.49	1.22E+06
Average	-	15351.48	1.20E+06
Standard deviation	-	0.0	3.01E+04
H3 + 4ac	A	15393.55	1.82E+06
	B	15393.31	1.67E+06
Average	-	15393.43	1.75E+06
Standard deviation	-	0.2	1.03E+05
H3 +5ac	A	15435.34	1.16E+05
	B	15435.36	3.07E+05
Average	-	15435.35	2.11E+05
Standard deviation	-	0.0	1.35E+05

60 min

	Replicate	Mass (in Dalton)	Height
H3	A	ND	ND
	B	ND	ND
Average	-	-	-
Standard deviation	-	-	-
H3 + ac	A	ND	ND
	B	ND	ND
Average	-	-	-
Standard deviation	-	-	-
H3 + 2ac	A	15309.34	1.31E+05
	B	ND	ND
Average	-	-	-
Standard deviation	-	-	-
H3 + 3ac	A	15351.50	1.03E+06
	B	ND	ND
Average	-	-	-
Standard deviation	-	-	-
H3 + 4ac	A	15393.55	2.93E+06
	B	ND	ND
Average	-	-	-
Standard deviation	-	-	-
H3 +5ac	A	15435.41	4.54E+05
	B	ND	ND
Average	-	-	-
Standard deviation	-	-	-

90 min

	Replicate	Mass (in Dalton)	Height
H3	A	ND	ND
	B	ND	ND
Average	-	-	-
Standard deviation	-	-	-
H3 + ac	A	0	0.00E+00
	B	15226.61	8.09E+03
Average	-	7613.31	4.05E+03
Standard deviation	-	10766.84	5.72E+03
H3 + 2ac	A	15309.33	1.46E+05
	B	15309.30	1.26E+05
Average	-	15309.31	1.36E+05
Standard deviation	-	0.0	1.45E+04
H3 + 3ac	A	15351.27	1.02E+06
	B	15351.25	8.87E+05
Average	-	15351.26	9.52E+05
Standard deviation	-	0.0	9.25E+04
H3 + 4ac	A	15393.52	2.42E+06
	B	15393.52	2.17E+06
Average	-	15393.52	2.29E+06
Standard deviation	-	0.0	1.80E+05
H3 +5ac	A	15435.36	1.83E+05
	B	15435.36	2.59E+05
Average	-	15435.36	2.21E+05
Standard deviation	-	0.0	5.36E+04

ND: Not detected/determined
* Deconvoluted mass

Table S2.2 H3/Rtt109-Asf1 assay intact histone intensity data (0 to 90 min)

CTRL

	Replicate	Mass (in Dalton)*	Height
H3	A	15224.98	2.30E+06
	B	15225.54	2.54E+06
Average	-	15225.26	2.42E+06
Standard deviation	-	0.4	1.70E+05
H3 + ac	A	ND	ND
	B	ND	ND
Average	-	-	-
Standard deviation	-	-	-
H3 + 2ac	A	ND	ND
	B	ND	ND
Average	-	-	-
Standard deviation	-	-	-
H3 + 3ac	A	ND	ND
	B	ND	ND
Average	-	-	-
Standard deviation	-	-	-

1 min

	Replicate	Mass (in Dalton)	Height
H3	A	ND	ND
	B	15225.31	8.17E+05
Average	-	-	-
Standard deviation	-	-	-
H3 + ac	A	ND	ND
	B	15266.85	3.66E+04
Average	-	-	-
Standard deviation	-	-	-
H3 + 2ac	A	ND	ND
	B	ND	ND
Average	-	-	-
Standard deviation	-	-	-
H3 + 3ac	A	ND	ND
	B	ND	ND
Average	-	-	-
Standard deviation	-	-	-

5 min

	Replicate	Mass (in Dalton)	Height
H3	A	15225.02	6.31E+05
	B	15225.27	6.39E+05
Average	-	15225.15	6.35E+05
Standard deviation	-	0.2	5.59E+03
H3 + ac	A	15266.91	9.79E+04
	B	15267.22	1.36E+05
Average	-	15267.06	1.17E+05
Standard deviation	-	0.2	2.68E+04
H3 + 2ac	A	15309.41	4.22E+04
	B	15308.64	4.22E+04
Average	-	15309.02	4.22E+04
Standard deviation	-	0.5	0.0
H3 + 3ac	A	ND	ND
	B	ND	ND
Average	-	-	-
Standard deviation	-	-	-

15 min

	Replicate	Mass (in Dalton)	Height
H3	A	15225.0413	2.42E+06
	B	15225.3857	3.15E+06
Average	-	15225.2135	2.79E+06
Standard deviation	-	0.2	5.14E+05
H3 + ac	A	15267.17	1.13E+06
	B	15267.31	6.63E+05
Average	-	15267.24	8.97E+05
Standard deviation	-	0.1	3.31E+05
H3 + 2ac	A	15308.93	2.81E+05
	B	15309.20	1.42E+05
Average	-	15309.06	2.12E+05
Standard deviation	-	0.2	9.83E+04
H3 + 3ac	A	15348.29	3.87E+04
	B	15349.09	1.93E+04
Average	-	15348.69	2.90E+04
Standard deviation	-	0.6	1.37E+04

30 min

	Replicate	Mass (in Dalton)	Height
H3	A	15225.1786	1.18E+06
	B	15225.4	1.25E+06
Average	-	15225.2893	1.21E+06
Standard deviation	-	0.156553441	6.93E+04
H3 + ac	A	15267.1858	8.42E+05
	B	15267.2688	6.05E+05
Average	-	15267	7.23E+05
Standard deviation	-	0	1.68E+05
H3 + 2ac	A	15309.01	2.86E+05
	B	15309.20	2.39E+05
Average	-	15309.10	2.62E+05
Standard deviation	-	0.1	3.32E+04
H3 + 3ac	A	15349.71	4.40E+04
	B	15350.43	4.61E+04
Average	-	15350.07	4.50E+04
Standard deviation	-	0.5	1.51E+03

60 min

	Replicate	Mass (in Dalton)	Height
H3	A	15225.2157	9.91E+05
	B	15225.53	6.26E+05
Average	-	15225.37285	8.08E+05
Standard deviation	-	0.2	2.58E+05
H3 + ac	A	15267.2139	8.63E+05
	B	15267.3113	3.81E+05
Average	-	15267	6.22E+05
Standard deviation	-	0	3.40E+05
H3 + 2ac	A	15309.34	3.05E+05
	B	15309.29	3.29E+05
Average	-	1.53E+04	3.17E+05
Standard deviation	-	0.0	1.68E+04
H3 + 3ac	A	15348.33	4.81E+04
	B	15350.6	6.68E+04
Average	-	1.53E+04	5.75E+04
Standard deviation	-	1.6	1.32E+04

90min

	Replicate	Mass (in Dalton)	Height
H3	A	15225.2212	7.87E+05
	B	15225.4018	5.55E+05
Average	-	15225.3115	6.71E+05
Standard deviation	-	0.127703485	1.64E+05
H3 + ac	A	15267	8.05E+05
	B	15267.36	3.44E+05
Average	-	15267.28	5.74E+05
Standard deviation	-	0.11	3.26E+05
H3 + 2ac	A	15309.08	3.28E+05
	B	15309.31	1.94E+05
Average	-	15309.19	2.61E+05
Standard deviation	-	0.2	9.44E+04
H3 + 3ac	A	15349.56	5.63E+04
	B	15349.95	2.94E+04
Average	-	15349.76	4.28E+04
Standard deviation	-	0.3	1.90E+04

ND: Not detected
NA: Not available
* Deconvoluted mass

Table S2.3 H3/Rtt109-Vps75 assay peptide IDs and normalized intensities.

Peptide m/z	Peptide mass	Charge	UniProt ID	Protein Description	Sequence	Peptide Sequence	Pep Modification	CTRL	1 min	5 min	15 min	30 min	60 min	90 min
380.7194	759.4239	2	S000000214	Histone H3	3-8	TKQTAR	K4pr	1.88E+06	1.02E+06	6.43E+05	1.27E+06	1.12E+06	1.08E+06	1.16E+06
493.2744	984.5352	2	S000000214	Histone H3	9-17	KSTGGKAPR	K9ac, K14ac	2.94E+04	1.43E+05	1.90E+06	4.38E+06	6.30E+06	6.70E+06	5.74E+06
500.2829	998.5509	2	S000000214	Histone H3	9-17	KSTGGKAPR	K9ac, K14pr	8.08E+04	5.14E+06	5.53E+06	8.10E+06	6.40E+06	1.66E+06	1.38E+06
507.2931	1012.5670	2	S000000214	Histone H3	9-17	KSTGGKAPR	K9pr, K14pr	3.74E+07	2.17E+07	3.88E+06	7.49E+06	4.32E+06	3.27E+06	3.49E+06
528.8120	1055.6090	2	S000000214	Histone H3	18-26	KQLASKAAR	K18ac, K23ac	0.00E+00	0.00E+00	0.00E+00	1.93E+06	2.25E+06	6.46E+06	5.91E+06
535.8203	1069.6240	2	S000000214	Histone H3	18-26	KQLASKAAR	K18pr, K23ac	0.00E+00	3.37E+06	4.58E+07	7.92E+07	6.80E+07	7.46E+07	7.16E+07
542.8282	1083.6400	2	S000000214	Histone H3	18-26	KQLASKAAR	K18pr, K23pr	1.35E+08	8.67E+07	4.93E+07	3.27E+07	2.33E+07	1.15E+07	8.17E+06
539.9749	1616.9000	3	S000000214	Histone H3	27-40	KSAPSTGGVKKPHR	K27pr, K36pr, K36pr	1.05E+07	1.14E+07	5.73E+06	3.92E+06	6.75E+06	1.96E+06	8.83E+05
535.3008	1602.8840	3	S000000214	Histone H3	27-40	KSAPSTGGVKKPHR	K27ac, K36pr, K36pr	0.00E+00	1.10E+05	1.23E+05	3.91E+05	7.67E+05	5.77E+05	3.44E+05
717.3973	1432.7790	2	S000000214	Histone H3	28-40	SAPSTGGVKKPHR	K36pr, K36pr	4.55E+05	3.22E+05	2.35E+05	2.57E+05	1.60E+05	2.12E+05	2.21E+05
530.8120	1059.6080	2	S000000214	Histone H3	41-59	YKPGTVALR	K42pr	1.50E+07	8.05E+06	5.69E+06	7.34E+06	7.51E+06	6.82E+06	3.23E+06
638.8681	1275.7190	2	S000000214	Histone H3	54-63	FQKSTELLIR	K56ac	0.00E+00	0.00E+00	0.00E+00	3.32E+05	4.50E+05	8.73E+05	1.19E+06
645.8753	1289.7340	2	S000000214	Histone H3	54-63	FQKSTELLIR	K56pr	1.54E+07	6.27E+06	1.35E+06	3.27E+06	3.75E+06	3.72E+06	5.70E+06
422.7639	843.4967	2	S000000214	Histone H3	64-69	KLPFQR	K64pr	5.39E+07	2.36E+07	1.48E+07	1.89E+07	2.16E+07	1.98E+07	1.63E+07
696.3690	1390.7090	2	S000000214	Histone H3	73-83	EIAQDFKTDLR	K79pr	3.78E+07	1.67E+07	3.96E+06	1.31E+07	1.44E+07	1.36E+07	2.48E+07
592.8670	827.5116	2	S000000214	Histone H3	117-125	VTIQKKDIK	K121pr, K122pr, K125pr	6.52E+05	6.17E+05	2.17E+05	5.39E+05	4.41E+05	5.94E+05	2.68E+05
478.3003	954.5862	2	S000000214	Histone H3	122-128	KDIKLAR	K122pr, K125pr	3.12E+05	0.00E+00	0.00E+00	0.00E+00	0.00E+00	1.14E+05	0.00E+00
386.2400	770.4650	2	S000000214	Histone H3	123-128	DIKLAR	K125pr	6.17E+05	1.39E+05	0.00E+00	0.00E+00	0.00E+00	4.36E+05	0.00E+00

Table S2.4 H3/Rtt109-Asf1 assay peptide IDs and normalized intensities.

Peptide m/z	Peptide mass	Charge	UniProt ID	Protein Description	Sequence	Peptide Sequence	Pep Modification	CTRL	1 min	5 min	15 min	30 min	60 min	90 min
380.7213	759.4239	2	S000000214	Histone H3	3-8	TKQTAR	K4pr	2.30E+06	1.40E+07	4.23E+06	4.32E+06	5.02E+06	3.97E+06	3.06E+06
479.2773	956.5403	2	S000000214	Histone H3	9-17	KSTGGKAPR	K14pr	3.58E+03	0.00E+00	5.45E+04	2.53E+04	1.22E+04	4.14E+04	2.84E+04
493.2767	984.5352	2	S000000214	Histone H3	9-17	KSTGGKAPR	K9ac, K14ac	0.00E+00	0.00E+00	4.41E+05	1.68E+06	3.38E+06	2.09E+06	1.25E+06
500.2841	998.5509	2	S000000214	Histone H3	9-17	KSTGGKAPR	K9ac, K14pr	2.31E+05	0.00E+00	3.79E+06	1.03E+07	1.62E+07	8.05E+06	5.71E+06
507.2923	1012.5670	2	S000000214	Histone H3	9-17	KSTGGKAPR	K9pr, K14pr	7.77E+07	3.29E+07	5.53E+07	6.28E+07	6.48E+07	3.41E+07	2.94E+07
415.2309	828.4454	2	S000000214	Histone H3	10-17	STGGKAPR	K14pr	4.29E+05	7.92E+05	4.88E+05	3.89E+05	3.17E+05	5.73E+05	3.86E+05
365.7279	729.4385	2	S000000214	Histone H3	18-23	KQLASK	K18pr	8.77E+05	2.50E+06	1.42E+06	1.24E+06	8.65E+05	2.12E+06	1.11E+06
514.8135	1027.6140	2	S000000214	Histone H3	18-26	KQLASKAAR	K23pr	2.84E+04	0.00E+00	2.95E+04	2.00E+04	1.82E+04	4.94E+04	4.00E+04
535.8212	1069.6240	2	S000000214	Histone H3	18-26	KQLASKAAR	K18pr, K23ac	0.00E+00	0.00E+00	1.59E+05	2.46E+06	5.75E+06	3.59E+06	2.49E+06
542.8299	1083.6400	2	S000000214	Histone H3	18-26	KQLASKAAR	K18pr, K23pr	4.35E+07	1.09E+07	1.98E+07	4.74E+07	5.42E+07	3.28E+07	2.94E+07
539.9738	1616.9000	3	S000000214	Histone H3	27-40	KSAPSTGGVKKPHR	K27pr, K36pr, K36pr	2.84E+06	5.23E+05	2.70E+05	3.77E+05	3.72E+05	2.60E+05	3.15E+05
809.4565	1616.9000	2	S000000214	Histone H3	27-40	KSAPSTGGVKKPHR	K27pr, K36pr, K36pr	7.58E+05	1.35E+05	0.00E+00	6.44E+04	5.00E+04	5.38E+04	3.86E+04
717.3971	1432.7790	2	S000000214	Histone H3	28-40	SAPSTGGVKKPHR	K36pr, K36pr	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	5.67E+04	2.84E+04
530.8040	1060.6148	2	S000000214	Histone H3	41-49	YKPGTVALR	K42pr	4.11E+07	2.06E+05	3.78E+05	2.18E+06	1.84E+06	2.34E+06	1.04E+07
638.8670	1275.7190	2	S000000214	Histone H3	54-63	FQKSTELLIR	K56ac	0.00E+00	0.00E+00	1.33E+05	3.63E+05	4.30E+05	6.22E+05	6.62E+05
645.8761	1289.7340	2	S000000214	Histone H3	54-63	FQKSTELLIR	K56pr	4.08E+04	1.00E+06	6.80E+05	4.20E+05	3.63E+05	2.11E+05	9.44E+05
422.7567	843.4967	2	S000000214	Histone H3	64-69	KLPFQR	K64pr	1.31E+07	1.73E+05	1.26E+05	6.44E+05	2.58E+06	2.40E+06	8.04E+06
696.3654	1390.7090	2	S000000214	Histone H3	73-83	EIAQDFKTDLR	K79pr	2.43E+04	7.45E+05	7.07E+05	2.49E+06	1.01E+06	5.74E+05	2.23E+06
592.8686	1183.7180	2	S000000214	Histone H3	117-125	VTIQKKDIK	K121pr, K122pr, K125pr	0.00E+00	1.94E+05	1.30E+05	8.76E+04	1.20E+05	6.96E+04	1.89E+05

2.8. References

1. Luger K, Mader AW, Richmond RK, Sargent DF, Richmond TJ. Crystal structure of the nucleosome core particle at 2.8 Å resolution. Nature 1997;389:251-60.

2. Wang J-P, Fondufe-Mittendorf Y, Xi L, Tsai G-F, Segal E, Widom J. Preferentially quantized linker DNA lengths in Saccharomyces cerevisiae. PLoS Comput Biol 2008;4:e1000175.

3. Robinson PJJ, Rhodes D. Structure of the '30 nm' chromatin fibre: a key role for the linker histone. Curr Opin Struct Biol 2006;16:336-43.

4. Bannister AJ, Kouzarides T. Regulation of chromatin by histone modifications. Cell Res 2011;21:381-95.

5. Berger SL. The complex language of chromatin regulation during transcription. Nature 2007;447:407-12.

6. Kayne PS, Kim U-J, Han M, Mullen JR, Yoshizaki F, Grunstein M. Extremely conserved histone H4 N terminus is dispensable for growth but essential for repressing the silent mating loci in yeast. Cell 1988;55:27-39.

7. Durrin LK, Mann RK, Kayne PS, Grunstein M. Yeast histone H4 N-terminal sequence is required for promoter activation in vivo. Cell 1991;65:1023-31.

8. Cong P, Jie Z, Ying LH, Ming Z, Li WL, Hong ZQ, et al. The transcriptional regulation role of BRD7 by binding to acetylated histone through bromodomain. J Cell Biochem 2006;97:882-92.

9. Berger SL. Histone modifications in transcriptional regulation. Curr Opin Genet Dev 2002;12:142-8.

10. Grunstein M. Histone acetylation in chromatin structure and transcription. Nature 1997;389:349-52.

11. Vempati R. DNA damage in the presence of chemical genotoxic agents induce acetylation of H3K56 and H4K16 but not H3K9 in mammalian cells. Mol Biol Rep 2012;39:303-8.

12. Yuan J, Pu M, Lou Z. Histone H3-K56 acetylation is important for genomic stability in mammals. Cell Cycle 2009;8: 1747-53.

13. Fraga MF, Ballestar E, Villar-Garea A, Boix-Chornet M, Espada J, Schotta G, et al. Loss of acetylation at Lys16 and trimethylation at Lys20 of histone H4 is a common hallmark of human cancer. Nat Genet 2005;37:391-400.

14. Marmorstein R, Roth SY. Histone acetyltransferases: function, structure, and catalysis. Curr Opin Genet Dev 2001;11:155-61.

15. Thiagalingam SAM, Cheng K-H, Lee HJ, Mineva N, Thiagalingam A, Ponte JF. Histone deacetylases: unique players in shaping the epigenetic histone code. Ann N Y Acad Sci 2003;983:84–100.

16. Johnsson A, Durand-Dubief M, Xue-Franzen Y, Ronnerblad M, Ekwall K, Wright A. HAT-HDAC interplay modulates global histone H3K14 acetylation in gene-coding regions during stress. EMBO Rep 2009;10:1009-14.

17. Sobulo OM, Borrow J, Tomek R, Reshmi S, Harden A, Schlegelberger B, et al. MLL is fused to CBP, a histoneacetyltransferase, in therapy-related acute myeloid leukemia with a t(11;16)(q23;p13.3). Proc Natl Acad Sci 1997;94:8732-7.

18. Petrif F, Giles RH, Dauwerse HG, Saris JJ, Hennekam RCM, Masuno M, et al. Rubinstein-Taybi syndrome caused by mutations in the transcriptional co-activator CBP. Nature 1995;376:348-51.

19. Giles RH, Peters DJM, Breuning MH. Conjunction dysfunction: CBP/p300 in human disease. Trends Genet 1998;14:178-83.

20. Han J, Zhou H, Horazdovsky B, Zhang K, Xu R-M, Zhang Z. Rtt109 acetylates histone H3 lysine 56 and functions in DNA replication. Science 2007;315:653-5.

21. Driscoll R, Hudson A, Jackson SP. Yeast Rtt109 promotes genome stability by acetylating histone H3 on lysine 56. Science 2007;315:649-52.

22. Tang Y, Holbert MA, Wurtele H, Meeth K, Rocha W, Gharib M, et al. Fungal Rtt109 histone acetyltransferase is an unexpected structural homolog of metazoan p300/CBP. Nat Struct Mol Biol 2008;15:738-45.

23. Kolonko EM, Albaugh BN, Lindner SE, Chen Y, Satyshur KA, Arnold KM, et al. Catalytic activation of histone acetyltransferase Rtt109 by a histone chaperone. Proc Natl Acad Sci 2010;107:20275-80.

24. Fillingham J, Recht J, Silva AC, Suter B, Emili A, Stagljar I, et al. Chaperone control of the activity and specificity of the histone H3 acetyltransferase Rtt109. Mol Cell Biol 2008;28:4342-53.

25. English CM, Adkins MW, Carson JJ, Churchill ME, Tyler JK. Structural basis for the histone chaperone activity of Asf1. Cell 2006;127:495-508.

26. Bowman A, Ward R, Wiechens N, Singh V, El-Mkami H, Norman DG, et al. The histone chaperones Nap1 and Vps75 bind histones H3 and H4 in a tetrameric conformation. Mol Cell 2011;41:398-408.

27. Park Y-J, Sudhoff KB, Andrews AJ, Stargell LA, Luger K. Histone chaperone specificity in Rtt109 activation. Nat Struct Mol Biol 2008;15:957-64.

28. Tang Y, Holbert MA, Delgoshaie N, Wurtele H, Guillemette B, Meeth K, et al. Structure of the Rtt109–AcCoA/Vps75 complex and implications for chaperone-mediated histone acetylation. Structure 2011;19:221-31.

29. D'Arcy S, Luger K. Understanding histone acetyltransferase Rtt109 structure and function: how many chaperones does it take? Curr Opin Struct Biol 2011;21:728-34.

30. Berndsen CE, Tsubota T, Lindner SE, Lee S, Holton JM, Kaufman PD, et al. Molecular functions of the histone acetyltransferase chaperone complex Rtt109–Vps75. Nat Struct Mol Biol 2008;15:948-56.

31. Drogaris P, Wurtele H, Masumoto H, Verreault A, Thibault P. Comprehensive profiling of histone modifications using a label-free approach and its applications in determining structure−function relationships. Anal Chem 2008;80:6698-707.

32. Drogaris P, Villeneuve V, Pomiès C, Lee E-H, Bourdeau V, Bonneil É, et al. Histone deacetylase inhibitors globally enhance H3/H4 tail acetylation without affecting H3 lysine 56 acetylation. Sci Rep 2012;2.

33. Dyer PN, Edayathumangalam RS, White CL, Bao Y, Chakravarthy S, Muthurajan UM, et al. Reconstitution of nucleosome core particles from recombinant histones and DNA. Methods Enzymol 2004;375:23-44.

34. Courcelles M, Lemieux S, Voisin L, Meloche S, Thibault P. ProteoConnections: a bioinformatics platform to facilitate proteome and phosphoproteome analyses. Proteomics 2011;11:2654-71.

35. Zhou BO, Wang S-S, Zhang Y, Fu X-H, Dang W, Lenzmeier BA, et al. Histone H4 lysine 12 acetylation regulates telomeric heterochromatin plasticity in Saccharomyces cerevisiae. PLoS Genet 2011;7:e1001272.

36. Berndsen CE, Denu JM. Catalysis and substrate selection by histone/protein lysine acetyltransferases. Curr Opin Struct Biol 2008;18:682-9.

37. Fulzele A, Malgundkar SA, Govekar RB, D'Cruz AK, Chaturvedi P, Patil A, et al. Keratins in oral cancer: necessity of mass spectrometry for validation of antibody based identifications. J Proteomics 2012;75:2404-16.

38. Tsubota T, Berndsen CE, Erkmann JA, Smith CL, Yang L, Freitas MA, et al. Histone H3-K56 acetylation is catalyzed by histone chaperone-dependent complexes. Mol Cell 2007;25:703-12.

39. Masumoto H, Hawke D, Kobayashi R, Verreault A. A role for cell-cycle-regulated histone H3 lysine 56 acetylation in the DNA damage response. Nature 2005;436:294-8.

40. Kuo MH, Brownell JE, Sobel RE, Ranalli TA, Cook RG, Edmondson DG, et al. Transcription-linked acetylation by Gcn5p of histones H3 and H4 at specific lysines. Nature 1996;383:269-72.

41. Keck KM, Pemberton LF. Interaction with the histone chaperone Vps75 promotes nuclear localization and HAT activity of Rtt109 in vivo. Traffic 2011;12:826-39.

CHAPTER 3: Discovery of Protein Acetylation Patterns by Deconvolution of Peptide Isomer Mass Spectra

Nebiyu Abshiru[1,2], Olivier Caron-Lizotte[2], Roshan Elizabeth Rajan[2,3], Adil Jamai[2], Christelle Pomies[2], Alain Verreault[1,3], Pierre Thibault[1,2]

[1]Department of Chemistry, Université de Montréal, P.O. Box 6128, Station centre-ville, Montréal, Québec, H3C 3J7, Canada.

[2]Institute for Research in Immunology and Cancer, Université de Montréal, C.P. 6128, Succursale centre-ville, Montréal, Québec, H3C 3J7, Canada.

[3]Molecular Biology Programme, Université de Montréal, P.O. Box 6128, Station centre-ville, Montréal, Québec, H3C 3J7, Canada.

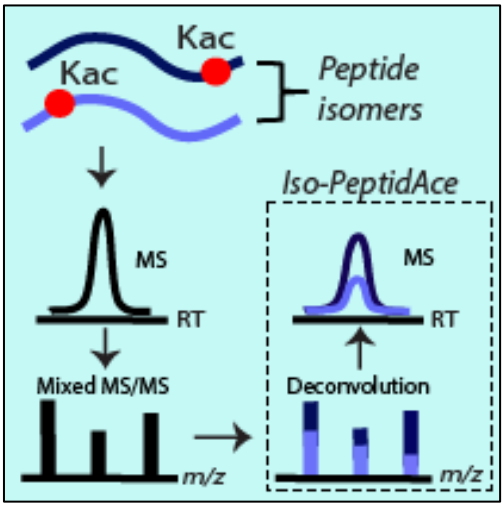
Manuscript: Published (Nature communications 6, 8648, 2015)

Author contributions:

N.A. designed and carried out the proteomics experiments. O.C.L. built and tested the Iso-PeptidAce deconvolution software. R.E.R and C.P prepared HDAC inhibitor-treated cells and performed histone extraction. A. J. isolated histones bound to chromatin modifying complex. N.A. analyzed data, wrote the first draft of the manuscript, and generated figures, with help from O.C.L. P.T and A.V. conceived and led the project. N.A., P.T. and A.V. wrote the final draft of the manuscript, with input from all authors.

3.1. Abstract

Protein post-translational modifications play an important role in the control of various biological processes including protein-protein interactions, epigenetics and cell cycle regulation. Mass spectrometry-based proteomics approaches enable comprehensive identification and quantitation of numerous types of PTMs. However, the analysis of PTMs is complicated by the presence of indistinguishable co-eluting isomeric peptides resulting in composite spectra with overlapping features that prevent the identification of individual components. In this study, we present Iso-PeptidAce, a novel software tool that enables deconvolution of composite MS/MS spectra of isomeric peptides based on features associated with their characteristic fragment ion patterns. We benchmark Iso-PeptidAce using dilution series prepared from mixtures of known amounts of synthetic acetylated isomers. We also demonstrate its applicability to different biological problems including the identification of site-specific acetylation patterns in affinity-purified histones bound to chromatin assembly factor-1 and profiling of histone acetylation in cells treated with different classes of HDAC inhibitors.



Graphical abstract

3.2. Introduction

Histone post-translational modifications (PTMs) control access to genetic information and, consequently, participate in several important cellular processes such as DNA repair and replication, nucleosome assembly and transcriptional regulation. [1, 2] It is now clear that many chromatin-modifying enzymes and enzymes that bind non-histone proteins, recognize "combinatorial patterns of PTMs", rather than single protein modification. For instance, ATP-dependent chromatin remodelers often recognize histones that contain more than one acetylated lysine residue that are located in close proximity to each other.[3,4] Although valuable to infer the functions of PTMs, accurately determining patterns and stoichiometries of PTMs located in isomeric and isobaric peptides, remains extremely difficult to achieve even by mass spectrometry (MS). This is because isomeric peptides, which are identical except for the position of the PTMs along the peptide chain, often co-elute during liquid chromatography and generate composite tandem mass spectra (MS/MS) containing indistinguishable fragment ions derived from two or more peptide isomers.

A few MS/MS methods were previously developed to quantify co-eluting isobaric histone peptides. Smith et al. were first to use a fragment ion-based approach to determine the fraction of histone H4 molecules acetylated at specific residues based on the normalized intensity ratio of peptide fragment ions containing either protiated or deuterated acetyl groups to distinguish lysine residues that were respectively acetylated or free *in vivo*.[5] However, this approach was limited by the lack of automated data analysis software and the difficulty to deconvolute co-eluting isomers. Recently, Feller et al. described a similar approach in which second generation fragment ions (known as MS^3) were necessary to obtain diagnostic fragment ions that can be used to estimate the abundance of each isomer by solving a set of linear equations.[6] A limitation of this approach is that singly charged MS^2 fragments, and particularly those with a nonmobile proton, fragment poorly and generate MS^3 spectra that are impossible to assign to specific isomers.[7, 8] Sidoli *et al.*, designed a software known as isoScale that can be used to determine the relative abundance of isobaric peptides, based on intensity ratios of fragment ions unique to each isomer.[9] isoScale relies on search engine outputs (e.g. Mascot.csv result files) to select isomer-specific fragments used in this analysis. The algorithm EpiProfile recently described by Yuan *et al.* quantifies isobaric histone peptides by solving a series of linear equations derived from peak heights of unique fragments representative of each isobaric species.[10] Several other approaches have also been

proposed to deconvolute mixed MS/MS spectra of peptide ions with different sequences [11-13], however, their application is limited when analyzing MS/MS spectra of co-fragmented isomeric ions that incorporate the same modification at different sites.

In this study, we present Iso-PeptidAce, a novel software tool that exploits high resolution LC-MS/MS data for deconvolution of mixed spectra and quantification of site-specific acetylation. The resolution of composite MS/MS spectra derived from isomeric peptides is based on features associated with their characteristic fragment ion patterns obtained from the corresponding synthetic peptides, and does not rely on unique fragment ions to distinguish isomers. The application of Iso-PeptidAce is presented here for peptide isomers that contain multiple acetylated lysine residues but can also be applied to other types of modifications such as isomeric phosphopeptides.[14] The use of Iso-PeptidAce is demonstrated to monitor temporal changes in acetylation patterns of histones H3 and H4 from human erythroleukemic (K562) cells treated with different classes of histone deacetylase inhibitors (HDACi), and for determining acetylation patterns in histones bound to *S. cerevisiae* chromatin assembly factor-1 (CAF1).

3.3. Methods

3.3.1. Synthetic peptide preparations

Synthetic histone H3 and H4 peptides were purchased from GenScript and include four variants of the H3 peptide 17-RKQLATKAAR-26: one with unmodified lysines, two peptides with one acetylated lysine and one peptide with both lysines acetylated (Table S3.1A). A total of sixteen H4 peptides were purchased. These included one unmodified, four mono-acetylated, six di-acetylated, four tri-acetylated and one tetra-acetylated form of peptide 1-SGRGKGGKGLGKGGAKRHR-19 (Table S3.1B). The concentration of each peptide was determined based on triplicate UV absorbance measurements at 205 nm using an Ultraspec 2100 Pro spectrophotometer (GE Healthcare). Each peptide solution was nominally in 1μg/μl range. We verified the purity and identity of each peptide by injecting directly 1 pmol of the peptides on a Q-Exactive Mass spectrometer (Thermo scientific). All MS files will be publicly available from PeptideAtlas database (data identifier: PASS00658, http://www.peptideatlas.org/PASS/PASS00658). For convenience MS/MS spectra of each isomeric peptide are provided in Appendix A.

3.3.2. Synthetic peptide standard dilutions

Standard curves for seven dilutions of mixtures of H4 peptides with peptide amounts ranging from 5 to 320 fmol were prepared (Table S3.2A-C). Two or three peptides from each group of isomers were selected and added to the mixtures at a fixed amount of 80 fmols. The peptide mixtures were subjected to propionylation, tryptic digestion and LC-MS/MS analysis on a Q Exactive mass spectrometer coupled to an EASY nLC II system (Thermo scientific).

3.3.3. Assessment of relative detection efficiency of H3 and H4 acetylated peptides

We prepared five mixtures of H3 or H4 synthetic peptides. The amount of each peptide in mixtures 1,2 3, 4 and 5 are 25, 50, 100, 200 and 400 fmols, respectively. These samples were subjected to propionylation, tryptic digestion and LC-MS/MS analysis on a Q Exactive Plus instrument. The signal intensity of each peptide was plotted against the amount of peptide injected as shown in Figs.S3.2 A and B.

3.3.4. Cell treatment with HDAC inhibitors

Suspension cultures of K-562 (human erythroleukemic) cells were grown in T-75 flasks at 37°C and 5% CO2 in RPMI-1640 medium (Gibco) supplemented with 10% FBS (Wisent) and 1% Penicillin/Streptomycin. Asynchronously growing cells were treated with 1 μM of SAHA, MS-275 or JNJ-26481585 (Selleck Chemicals) for 1, 6 and 24 hours. Control cells were treated with DMSO. Two biological replicates were prepared for each condition. After treatment, the cells were harvested, centrifuged at 1000 rpm at room temperature, washed with PBS, and flash frozen in liquid nitrogen. Histones were isolated using acid extraction as previously described.[37] Briefly, nuclei from 10^7 cells were isolated with a hypotonic lysis buffer, followed by histone extraction using 0.2 M H_2SO_4 and TCA precipitation. The concentration of histone samples was determined by a Micro-BCA assay (Pierce). Further purification of core histones was performed by reverse-phase HPLC.

3.3.5. Purification of yeast CAF1

The Cac2-TAP purification was adapted from Zhou *et. al*.[28] Briefly, exponentially growing yeast cells (4L at approximately 1.8 x 10^7 cells/ml in YPD medium) were harvested by centrifugation and washed twice with ice-cold water. All further steps were carried out on ice. The pellet was resuspended in an equal volume of lysis buffer (25mM Tris-HCl pH 7.5, 100mM NaCl, 10mM of MgCl2, 0.1% NP-40, 1mM EDTA, 10% glycerol, 1mM DTT, 1mM PMSF) containing a cocktail of protease inhibitors (5μM Leupeptin, 25μM Aprotinin, 5μM Pepstatin A) and HDAC inhibitors (10mM nicotinamide, 10mM sodium butyrate, 10 μM trichostatin A and 100μM SAHA). Cell suspensions were frozen in liquid nitrogen and lysed in a freezer mill (Spex Certiprep) using two repeats of 4 cycles with 2 min of grinding and 2 min (15 impacts per second) of cooling per cycle. The resulting cell lysate powder was thawed out on ice. Ethidium bromide (75μg/mL) was added and the lysate was incubated on ice for 1 h with benzonase (25U/mL). The extract was clarified by centrifugation at 25,000 rpm for 30 min. The supernatant was then incubated with 200 μL of IgG-Sepharose beads (SIGMA) for 2 h at 4°C. The beads were recovered by centrifugation, extensively washed with lysis buffer followed by several washes with TEV cleavage buffer (10mM Tris-HCl, pH 8, 100mM NaCl, 1mM DTT, 0.5mM EDTA) containing the protease inhibitors and HDAC inhibitor mixture. In order to elute the Cac2-TAP, the beads were incubated with TEV protease, at 4°C overnight in TEV cleavage buffer. A portion of the TEV eluate (30μl) was resolved by SDS-PAGE and the presence of Cac2-CBP and histones H3 and H4 was confirmed by immunoblotting. The remainder of the TEV eluate (170μL) was dried in a

Speed-Vac, resuspended in 1ml H_2O and applied to a C18 column equilibrated with 0.1% TFA in water. The C18 column was washed twice with 1ml of 0.1% TFA in 5% methanol and eluted with 1ml of 70% acetonitrile. The eluate was dried in a Speed-Vac and resuspended in 0.1% TFA for purification of intact histones by offline HPLC.

3.3.6. Fractionation of core histones by RP-HPLC

Separation of core histones was achieved using a narrow-bore Zorbax C8 reverse-phase column (2.1 x 150 mm, 5µm, 300Å) on an Agilent 1200 HPLC system. Solvent A was aqueous 0.1% TFA (Sigma) and solvent B was 0.1% TFA in 100% acetonitrile (ACN). Approximately 2 µg of acid extracted histones were loaded onto the C8 column at a flow rate of 150 µl/min. Histones were eluted from the column using a gradient of 5 – 80% solvent B in 60 minutes. Fractions were collected in a 96-well plate at a rate of one fraction per minute. Fractions containing histones H3 or H4 were pooled and dried completely in a Speed-vac concentrator.

3.3.7. Propionylation and trypsin digestion of histone peptides/proteins

In-solution tryptic digestion of synthetic peptides or core histones was performed as previously described.[38] Briefly, a total of 2 µg of HPLC purified histones H3 and H4 were subjected to propionylation by adding 200 µl of freshly prepared 2:1 (v/v) water: propionic anhydride (Sigma) mixture and vortexing the mixture for 1 h at room temperature. The samples were then dried in a Speed-vac at 4^oC. The dried samples were resuspended in 50 mM ammonium bicarbonate, vortexed for 2 minutes and subjected to a second round of evaporation at 4oC. The samples were collected and resuspended in 100 µl of 50 mM ammonium bicarbonate and vortexed for 5 minutes in order to re-dissolve the proteins. Our digestion solution was prepared by adding 200 µl of 50 mM ammonium bicarbonate in a vial containing 20 µg of lyophilized trypsin (Promega). About 0.5 µl of this solution was added to each histone sample and digested overnight at 37^oC. After digestion, samples were dried completely in a Speed-vac and then resuspended in 0.2% formic acid prior to LC-MS/MS analyses.

3.3.8. LC-MS/MS analyses of histone digests

MS data were acquired in duplicate on a Q Exactive Plus mass spectrometer (Thermo scientific) coupled to an EASY nLC II system (Thermo scientific). A total of 1µg of histone H3 and H4 digests generated from the control or HDACi-treated cells were first desalted on a Jupiter C18 (3µm particles, Phenomenex) trap column (4mm length, 360µm i.d.) for 5 min at

10 µl/min, prior to their elution onto a C18 analytical column (18 cm length, 150 µm i.d.). A linear gradient from 5 to 60% acetonitrile (containing 0.2% formic acid) at 600 nl/min over 60 min was used for peptide elution. The MS instrument was operated in positive ion mode, and capillary voltage of 1.6 kV. MS scans were acquired in the Orbitrap analyzer over the range of 300 – 1500 *m/z* at a resolution of 70,000 and automatic gain control target value of 1.0×10^6. An inclusion list containing *m/z*, charge state and collision energy CE values of H3 and H4 peptides were used to trigger MS/MS acquisition. Every precursor ion found in the inclusion list was automatically selected for fragmentation in the HCD cell at a normalized CE setting of 27. The fragments were analyzed in the Orbitrap at a resolution of 35,000 and a target value of 5.5×10^5. The dynamic exclusion setting was disabled in order to acquire multiple MS/MS spectra per peptide.

3.3.9. Data analysis

Data analysis was performed using the Iso-PeptidAce software. Raw MS and MS/MS files of individual isomeric peptides, peptide mixtures or histone digests and Fasta files with H3 and H4 protein sequences were submitted to Iso-PeptidAce. The default settings for deconvolution of mixture MS/MS spectra are: Precursor mass tolerance: 8 ppm, fragment mass tolerance: 0.05 Da, the minimum number of fragment ions (per isomer) considered for deconvolution: 5, types of fragment ions considered: b and y, digestion enzyme: trypsin, missed cleavages: Iso-PeptidAce uses no-enzyme in-silico protein digestion routine to parse the provided FASTA file for potential matches. PTMs included in peptide-spectrum matching: carbamidomethylation of cysteine (C, 57.0215 Da), oxidation of methionine residue (M, 15.9949), phosphorylation of serine, threonine or tyrosine (S/T/Y, 79.9663), deamidation of aspargine and glutamine (N/Q 0.98400.9847), acetylation of lysine (K, 42.0106), propionylation of lysine (K, 56.0262), methylation of lysine (K, 14.0157), di-methylation of lysine (K, 28.0313), tri-methylation of lysine (K, 42.0470), acetylation of protein N-terminus (42.0106). All the PTMs were considered as variable modifications. The output from Iso-PeptidAce is a combined result file (spreadsheet) containing intensity values for all the acetylated and non-acetylated forms of deconvoluted H3 and H4 peptides. A representative result file for the H4 peptide GKGGKGLGKGGAKR is shown in Table S3.7. A total of 16 different isoforms of this peptide were detected (shown in each column in Table S3.7). These peptides are divided into five groups (per sample) based on their m/z values (shown in each row in Table S3.7): one un-acetylated (0Ac, m/z 747.94), four mono-acetylated (1Ac, m/z 740.93), six di-acetylated (2Ac, 733.93), four tri-acetylated (3Ac, m/z

726.92) and one tetra-acetylated (4Ac, m/z 719.91). Each intensity value in Table S3.7 is normalized based on the peptide's response factor, which is determined from the slope of the lines shown in Fig.s S3.2 A and B. The slope of each line was determined from the linear equation that best-fit the MS signal responses. Acetylation site occupancy at a specific Lys residue was determined from the ratio of the sum of intensities of peptides bearing the ac-Lys to the sum of intensities of all the 16 peptide isoforms. Iso-PeptidAce can be downloaded from the website: http://proteomics.iric.ca/tools/Iso-PeptidAce. Detailed information on Iso-PeptidAce is also provided as Supplementary Method.

3.4. Results

3.4.1. Deconvolution of mixed MS/MS spectra by Iso-PeptidAce

We designed a new software tool named Iso-PeptidAce that deconvolutes mixed MS/MS spectra of known isomeric peptides. The software takes as input raw MS and MS/MS files of the individual and the mixed isomers, the FASTA sequence file of the protein of interest, and a set of parameters for peptide-spectrum matching (PSM) such as modifications, precursor and fragment ion tolerances (See the 'Data analysis' section of the Method). The raw files are processed to extract precursor intensity, MS1 and MS2 injection times, and MS2 peak lists containing m/z and intensity values (see Supplementary Method for a detailed section on spectral deconvolution and normalization of MS signal intensities).

Iso-PeptidAce computes the proportion of individual isomers in mixed MS/MS spectra based on their fragment ion patterns. In Iso-PeptidAce, ion patterns are reduced to a set of maximum network flow problems, for which a number of efficient algorithms are known.[15] This approach has previously been used in a wide-range of complex problems, such as in predicting molecular pathways in complex diseases and in selecting single nucleotide polymorphisms (SNPs) and their associated allele in case and control groups.[16-18] In our study, we implemented the Network Flows approach to deconvolute mixed fragment spectra of acetylated isomers. Although these isomers share multiple fragment ions, they produce distinct fragment ion patterns that can be transformed into a set of network flow problems. Fig. 3.1A shows a schematic overview of the deconvolution conducted by Iso-PeptidAce for a series of mixed MS/MS spectra generated from two hypothetical co-eluting isomeric peptides labeled X and Y. For every mixed spectrum acquired across the elution curve, fragment ion intensities (Fig. 3.1A, step 1) are modeled into maximum network flow problems (flow capacity shown as empty bar charts in Fig. 3.1A, step 2). The networks are filled with fragment ion patterns for isomers X and Y (Fig. 3.1A, step 3) and merged into a single network with excess flow (represented as overrunning color bars in Fig. 3.1A, step 4). The resulting network is processed iteratively by a multivariate optimization technique known as Gradient Descent[19] (Fig. 3.1A, step 5), to remove the excess flow in the network. Each iteration step converges towards the maximum flow, that refers to the optimal ratio of X and Y in the mixed spectrum (Fig. 3.1A, left inner circle). Finally, individual elution curves are generated based on the abundance ratios of MS/MS spectra (Fig. 3.1A, right inner circle) and the peak intensity or peak area of the precursor ions.

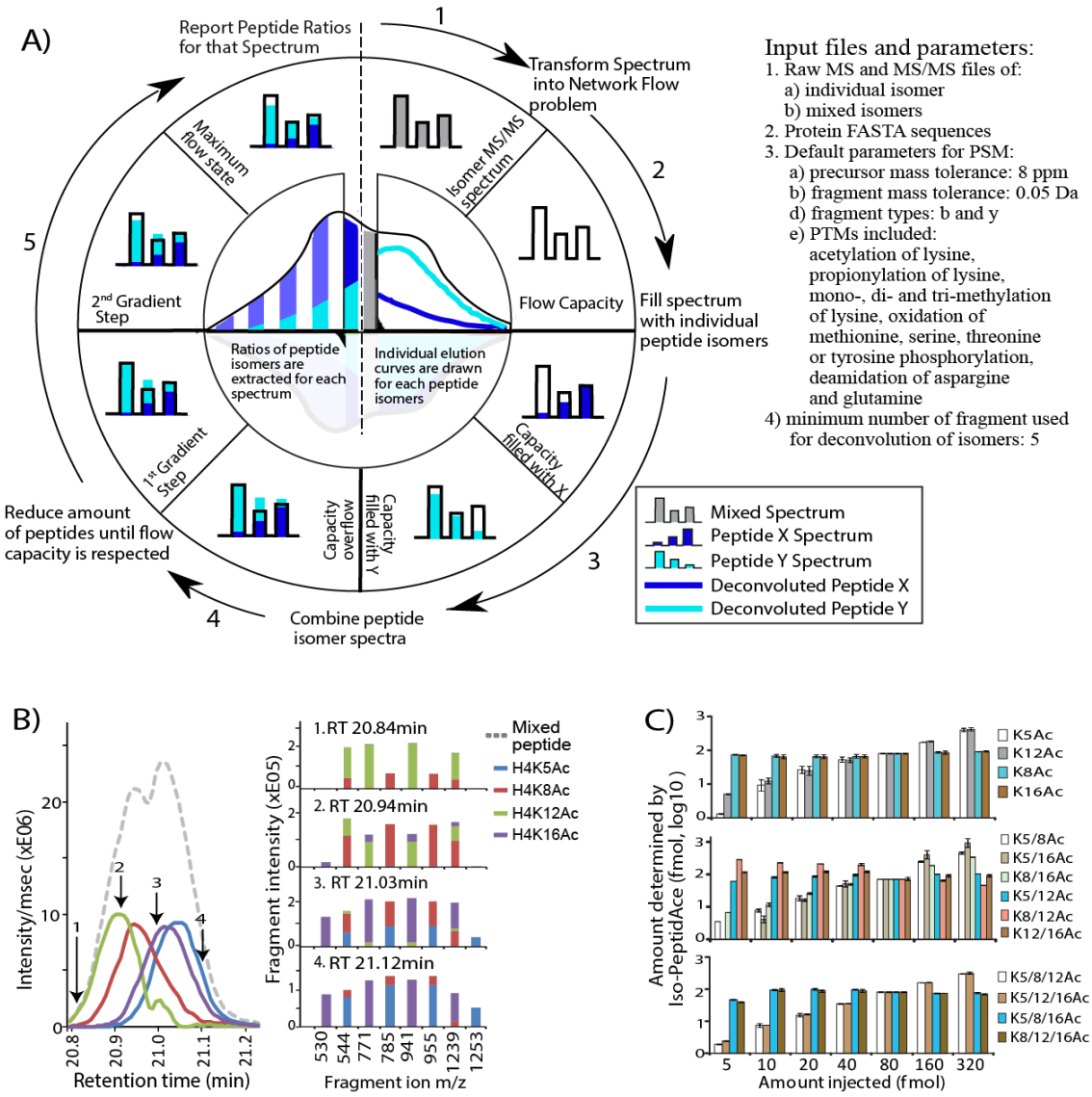


Figure 3.1 Deconvolution of co-eluting acetylated isomers by Iso-PeptidAce.

A) Schematic showing the process of spectral deconvolution by Iso-PeptidAce. The input files and parameters used for deconvolution of mixed isomers are listed on the right side of the panel. The process involves building an individual elution curve for X and for Y based on ratios computed for each MS/MS events. Isomeric peptide ratios are computed in 5 steps: 1) acquire and normalize fragment intensities of mixed spectra; 2) model fragment intensities as a network flow with fixed capacities; 3) find the maximum flow for peptide X and for peptide Y; 4) combine both X and Y network flow capacities into the same network flow model; 5) iteratively reduce the ratio of X and Y until the maximum flow of is reached. B) Left panel: Elution profiles of a mixture of acetylated isomers before (dashed line) or after (colored lines) deconvolution. Right panel: The relative intensity of representative fragment ions extracted from the deconvoluted isomers at four different retention times, RT 20.84, 20.94, 21.03 and 21.12 min. C) Bar graphs showing the amount of each peptide isomer determined by Iso-PeptidAce as a function of the amount of the same peptide injected as part of mixtures of

known amounts of mono-, di- or tri-acetylated groups (from top to bottom). The mean value and the standard deviation for each dilution point are shown in Table S3.3 A-C. The amount determined at each dilution point was normalized to the 80 fmol signal response. The amounts of peptides K5ac and K12ac (upper panel); K5_8ac, K5_16ac and K8_16ac (middle panel); K5_8_12ac and K5_12_16ac (lower panel) were varied from 5 to 320 fmol. Peptides K8ac and K16ac (upper panel); K5_12ac, K8_12ac and K12_16ac (middle panel); K5_8_16ac and K8_12_16ac (lower panel) were added to the mixture at fixed amount of 80 fmol. Each bar graph represents the mean of two technical replicates with error bars showing the relative distance of the maximum and the minimum values from the mean. Error bars indicate standard deviation of the mean of two technical replicates.

3.4.2. Extraction of elution profiles and fragment ion patterns

We first used Iso-PeptidAce for MS/MS spectra of synthetic peptides from acetylated isomers of histone H3 and H4 (Tables S3.1). All possible non-acetylated and acetylated forms of H4 peptide 1-SGRGKGGKGLGKGGAKRHR-19 were synthesized, and the purity of each peptide was confirmed by MS analyses. These peptides were propionylated, digested by trypsin and subjected to LC-MS/MS analysis on a Q Exactive Plus instrument.

We obtained three isomeric groups of H4 peptides acetylated at one, two or three lysine residues, and members of each group generally have very narrow retention time differences (Fig. S3.1). To normalize precursor intensity, we determined the signal intensity response for each peptide as a function of the amount injected. We observed that mono- or multiply acetylated isomers yielded a higher response than their propionylated counterparts (Fig. S3.2A). We then prepared equimolar mixtures of peptides for each group of isomers and analyzed them by LC-MS/MS. Although propionylated H4 peptides were readily separated into non-acetylated, mono-, di-, tri- and tetra-acetylated peaks (denoted as Ac0, 1, 2, 3, and 4 in Fig. S3.3A, top panel), we observed that peptides within the same isomeric group co-eluted and produced composite MS/MS spectra (Fig. S3.3 A-B). In addition, these peptides shared multiple fragment ions (Fig. S3.4 A-E). Figure S3.4E shows a three-dimensional view of the distribution of selected fragment ions derived from the four mono-acetylated isomers. H4K5ac and H4K16ac are the only H4 peptides that produced unique fragment ions at *m/z* 1253.7 and 530.3, respectively (Fig. S3.4E, see blue and purple bars). The peak intensities of these fragments can be used to infer relative abundances of H4K5ac and H4K16ac in a given sample. This approach has been previously implemented in existing software tools such as IsoScale.[5] However, the other two mono-acetylated isomers H4K8ac and H4K12ac did not produce unique fragments (Fig. S3.4E, red and green bars). These peptides share all of their

fragments with either H4K5ac or H4K16ac, and as a result fragment ion ratios cannot be used directly to measure the relative proportion of individual peptides. In Iso-PeptidAce, deconvolution of MS/MS spectra containing fragments of the four mono-acetylated isomers is achieved by exploiting the distinguishing features associated with isomer-specific fragment ion patterns. Here, the fragment ion pattern is defined by the retention time at which the peptides were eluted and selected for fragmentation, the peak height and m/z of each fragment ion, and the sites of modification. For example, the fragment ion pattern for H4K8ac is different from that of H4K12ac by the presence of additional y-ion fragments at m/z 785.5 and 955.6, and the absence of y-ion fragments at m/z 771.4 and 941.6 and b-ion at m/z 299.2. In addition, their shared fragments (e.g. m/z 242.1, 544.3 and 1239.7) differ slightly in peak intensities. Thus, by exploiting such isomer-specific differences, Iso-PeptidAce deconvoluted the co-eluting mono-acetylated peptides into four distinct elution profiles where the apex of individual isomer peaks were separated by a few seconds (Fig. 3.1B, left panel, colored solid lines). The elution curve prior to the deconvolution is shown in Figure 3.1B (gray dashed line). Using Iso-PeptidAce, we were also able to show the relative intensity of selected fragment ions extracted from the deconvoluted isomers (Fig. 3.1B, right panel). Most of these fragments are shared between the different isomers, but their distribution and proportion varies across peptides. Fragment patterns representative of H4K8ac and H4K12ac were more prominent at the beginning of elution profiles, whereas those of H4K5ac and H4K16ac were observed at later time points. Similarly, deconvolution of the di- and tri-acetylated isomers by Iso-PeptidAce enabled the resolution of each isomeric peptide (Fig. S3.5).

We then investigated the performance of Iso-PeptidAce using dilution series prepared from mixtures of known amounts of propionylated and acetylated isomeric H4 peptides each ranging from 5 to 320 fmols. Two to three peptides from each isomeric group were selected and added to the dilutions at a fixed amount of 80 fmols (Table S3.2A-C). Raw MS and MS/MS files from these samples were submitted to Iso-PeptidAce for deconvolution. The amount of peptide detected as a function of amount injected (in fmol) for the mono-, di- and tri-acetylated groups of isomers is shown in Fig. 3.1C (see Table S3.3A-C for the mean values and standard deviations). In the mono-acetylated mixtures we observed an increase in the relative proportion of peptides H4K5ac and H4K12ac whereas peptides H4K8ac and H4K16ac remained unchanged in each mixture (Fig. 3.1C, upper panel, and Table S3.3A). Similarly, we observed increasing levels of H4K5ac/K8ac, H4K5ac/K16ac and

H4K8ac/K16ac in the di-acetylated mixtures (Fig. 3.1C, middle panel, and Table S3.3B) and of H4K5ac/K8ac/K12ac and H4K5ac/K12ac/K16ac in the tri-acetylated mixtures (Fig. 3.1C, bottom panel, and Table S3.3C) in accordance to their expected abundance. Thus, Iso-PeptidAce successfully deconvoluted mixed spectra generated from different co-eluting acetylated isomers present at concentrations ranging over almost two orders of magnitude.

3.4.3. Temporal profiling of histone acetylation following HDAC inhibition

We tested our software using tryptic digests of histones isolated from human K562 cells treated with HDAC inhibitors (Fig. 3.2A). Cells were treated in duplicate for 1, 6 or 24 h with DMSO (control) or with the HDAC inhibitors MS-275, SAHA and JNJ-26481585. Histones were isolated by acid extraction and fractionated by liquid chromatography (Fig. S3.6). Fractions containing histones H3 or H4 were subjected to propionylation, tryptic digestion and LC-MS/MS analysis prior to data deconvolution using Iso-PeptidAce (Fig. 3.2A).

Our results show that, 24 h after treatment with SAHA or JNJ-26481585, approximately 50% of H4 peptides 4-GKGGKGLGKGGAKR-17 were acetylated at the four available sites, whereas only 20% of H4 molecules were tetra-acetylated in cells treated with MS-275 (Fig. 3.2B). Using Iso-PeptidAce we then determined acetylation site occupancies at positions H3K18, H3K23, H4K5, H4K8, H4K12 and H4K16 (Table S3.4A-B). The level of H3K23ac is higher than that of H3K18ac in both untreated and treated samples (Fig. 3.2C and Table S3.4A). Although previous studies have reported trace amounts of H3K18 methylation in mammalian cells[20], we did not detect this modification in our samples. Cells treated with all HDACi demonstrated similar increase of H4 acetylation with time. We observed more than 8-fold increase in acetylation at sites H4K5, H4K8 or H4K12 after 24 h treatment with MS275, SAHA or JNJ (Fig. 3.2D, and Table S3.4B). Up to 3-fold increase in acetylation was observed at position H4K16.

Although a number of studies have previously shown that HDACis MS-275 and SAHA induce global increases in histones H3 and H4 acetylation, there was no efficient technique to measure sitespecific acetylation stoichiometries of lysines in the N-terminal tail of H4 because some of the acetylated isomers cannot be resolved by nano-LC and do not generate isomer-specific fragments. In our study we observed that both MS-275 and SAHA caused a major increase in acetylation of the four lysines in the N-terminal tail of H4. In addition, for the first time, we report site-specific changes in histone acetylation caused by the more recently characterized pan-HDACi JNJ-26481585.[21] This drug, also known as Quisinostat, is

currently in phase II clinical trial and exhibits antitumor activity in human multiple myeloma and leukemic cells[22-25] , and recent studies have shown that JNJ-26481585 elicits global increases in acetylation of histones H3 and H4.[25, 26] However, due to the inherent limitations of immunoassays, these studies did not report the specific lysine residues affected by the drug treatment of the stoichiometries of acetylation the result from treatment with JNJ-26481585. Our approach allowed identification of several H3 and H4 sites that increased in acetylation after JNJ-26481585 treatment. When compared with MS-275 and SAHA, JNJ-26481585 caused a more rapid and higher fold-increase in acetylation at all the sites that we investigated (Figs.3.2B-D). Consistent with these results, JNJ-26481585 has been reported to be 500-fold more potent than SAHA at inhibiting HDAC1.[21]

These examples illustrate that Iso-PeptidAce provides an automated and rapid approach to quantify changes in acetylation site occupancies on histone lysine residues. Our method is implemented as a software tool that deconvolutes mixed spectra derived from two or more co-eluting isomeric acetylated peptides. In turn, this enables an accurate assessment of acetylation site occupancies at each of the lysines within tryptic peptides.

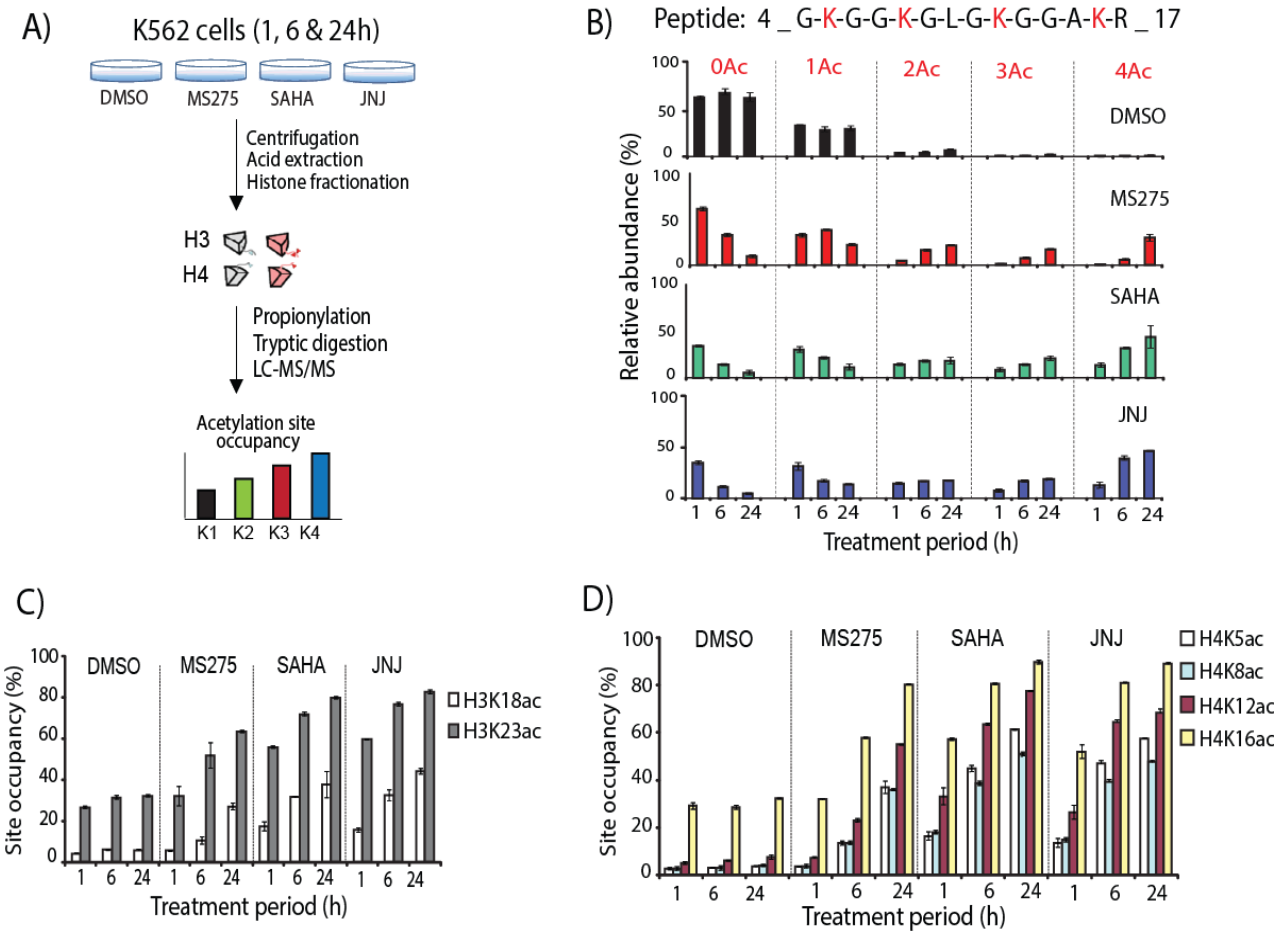


Figure 3.2 Acetylation site occupancies of H3 and H4 before or after HDACi treatment.

A) Histone isolation and MS analysis workflow. Total histones were isolated from human erythroleukemic (K562) cells treated for 1, 6 or 24 h with DMSO (control) or with the HDACi MS-275, SAHA and JNJ-26481585. Total histones were fractionated into individual core histones via off-line HPLC. Fraction containing histones H3 and H4 were subjected to propionylation, tryptic digestion and LC-MS/MS analysis. Acetylation site occupancy for each lysine residue was determined based on Iso-PeptidAce reported peptide intensities. B) Relative abundances of un-modified (Ac0), mono- (Ac1), di- (Ac2), tri- (Ac3), and tetra-acetylated (Ac4) peptide 4 -GKGGKGLGKGGAKR-17 before (DMSO) and after treatment with HDACis MS-275, SAHA, and JNJ-26481585. C) Acetylation site occupancies of histone H3 at positions K18 and K23, and D) histone H4 at positions K5, K8, K12 and K16. Each bar graph represents the mean of two technical replicates with error bars showing the relative distance of the maximum and the minimum values from the mean.

3.4.4. Identification of acetylation patterns in CAF1-bound histones

Acetylation of multiple residues in newly synthesized histones has been implicated in nucleosome assembly[27]. However, the molecular function and the biological implications of this acetylation have remained unclear because of redundancy among several acetylation sites in new H3 and H4 molecules. The capability of Iso-PeptidAce to determine acetylation patterns in situations where samples are limiting was evaluated for affinity-purified histones bound to chromatin assembly factor 1 (CAF1), a protein complex that deposits new H3/H4 molecules onto nascent DNA during replication[28].

As previously described[28], we purified the CAF1 complex from asynchronously growing yeast cells via a tandem affinity purification (TAP)-tagged Cac2 subunit. We also purified total histones from the same culture in order to compare their acetylation patterns with those of histones bound to CAF1 (Fig. 3.3A). The sites and stoichiometries of H3/H4 acetylation were clearly different in CAF1-bound versus total histones (Figs. 3.3B-D, Table S3.5). We also detected mono-, di- and tri-methylation of H3K36 in the total histones, however, these PTMs were absent in CAF1-bound samples (Table S3.5). In CAF1-bound H3 molecules, we found a high acetylation site occupancy (between 49 and 79%) at each of five lysine residues previously known to be acetylated in new histones from *S. cerevisiae*: H3K56, K9, K14, K23 and K27 but not at K18 (Fig. 3.3B). This high degree of acetylation at multiple sites demonstrates that essentially all the H3 molecules bound to CAF1 are acetylated. In contrast, a large fraction (64%, Fig. 3.3C) of H4 molecules bound to CAF1 were not acetylated at any of the four lysine residues located in their N-terminal tail. These results provide an explanation to a long-standing paradox. Hat1 is an enzyme that acetylates new H4 molecules at lysines 5 and 12, sites of acetylation that are conserved from yeast to humans.[27] Given this, the paradox was that *S. cerevisiae* cells lacking Hat1 or cells where lysines 5 and 12 were mutated showed essentially no phenotype unless several lysines of H3 were also mutated to block their acetylation.[29, 30] The fact that, in *S. cerevisiae*, the acetylation of new H3 molecules is far more abundant than that of new H4 molecules provides an explanation for this puzzle.

Our results also revealed an unexpected acetylation pattern in new H4 molecules where, in addition to lysines 5 and 12, lysine 16 is acetylated (Fig. 3.3C-D). Until now, the dogma was that new H4 molecules were predominantly acetylated at lysines 5 and 12, but not at lysine 16. This was based on pulse labeling of new histones with [^{3}H]-lysine and Edman sequencing to quantify PTH derivatives of acetyl-lysine and lysine[31]. The sensitivity of this technique rapidly decreases as a function of distance from the N-terminus and this may account for the fact that lysine 16 acetylation in new H4 molecules was previously missed. Based on mass spectrometry, the presence of K16 acetylation in new H4 molecules bound to *S. cerevisiae* CAF1 was previously reported[28] but its prevalence was unknown. For the first time, our experiments demonstrate that the most abundant form of H4 bound to *S. cerevisiae* CAF1 consists of molecules that are tri-acetylated at K5, K12 and K16, with significantly lower amounts of mono- and di-acetylated forms that also contain K16 acetylation (Fig. 3.3C). Consistent with this, the acetylation site occupancies at K5, K12 and K16 were roughly 20% each in CAF1-bound H4, a pattern strikingly different from that observed in total histones (Fig. 3.3D). In *S. cerevisiae*, the integrity of heterochromatin depends upon removal of H4K16 acetylation by the deacetylase Sir2[32]. Our results imply that deacetylation of H4K16 in new molecules deposited onto DNA during replication is likely important for propagation of heterochromatin structure in proliferating cells.

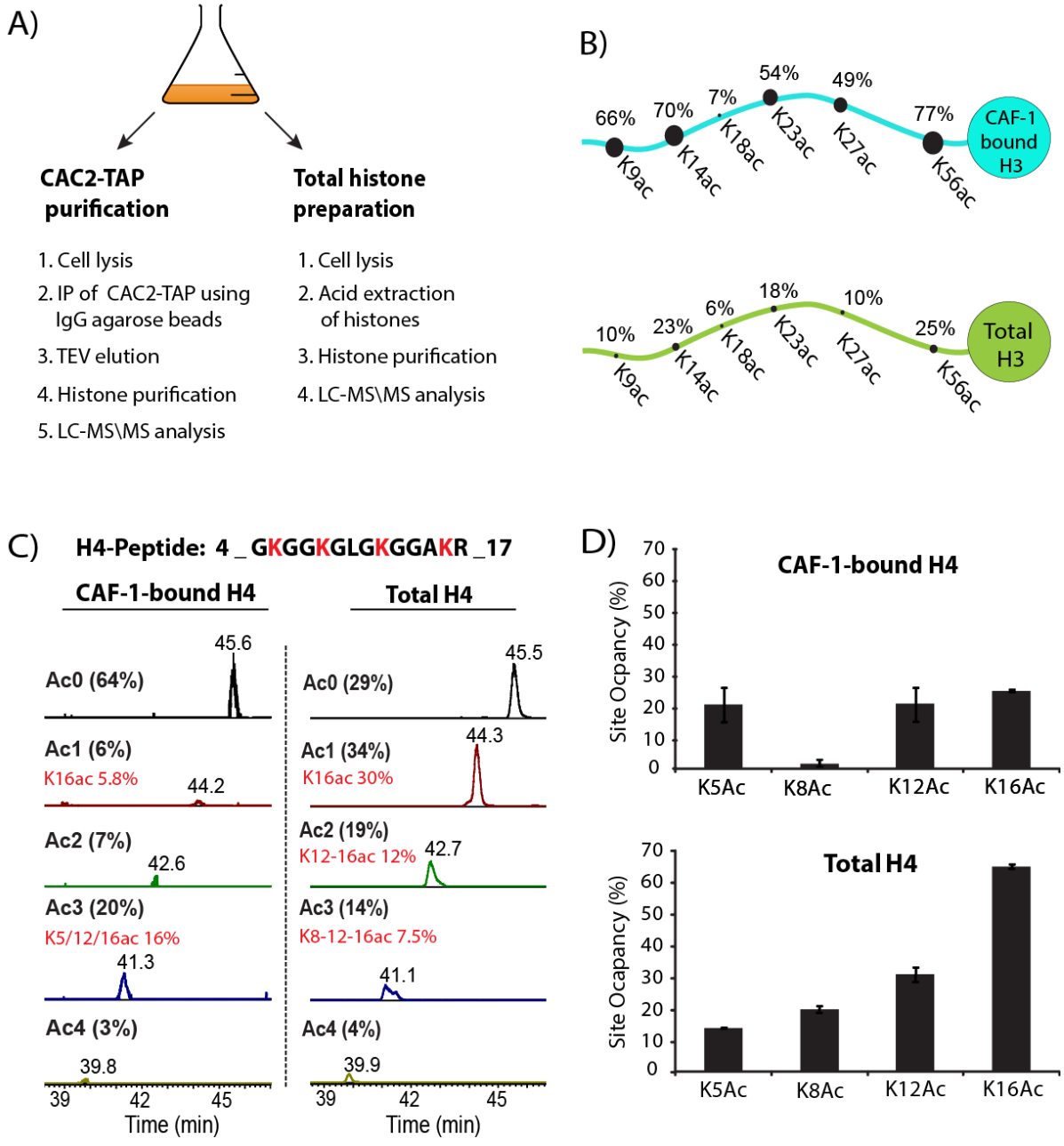


Figure 3.3 Acetylation site occupancies of CAF1-bound H3 and H4.

A) Schematic representation of the experimental workflow. Tandem Affinity Purification (TAP)-Cac2 was purified from 3.5 L of cell culture and the remainder (0.5 L) was processed for total histone extraction. Histones from both samples were fractionated by HPLC and analyzed by LC-MS/MS. B) Comparison of lysine acetylation site occupancies in total histone H3 (lower panel) and CAF1 bound H3 (upper panel). C) Extracted ion chromatograms of histone H4 peptide 4 -GKGGKGLGKGGAKR -17 acetylated at zero (Ac0), one (Ac1), two (Ac2), three (Ac3) or four (Ac4) lysine residues in CAF1 bound (left panel) or total histones (right panel). The predominant acetylation pattern in each isomeric group is highlighted in red. D) Comparison of acetylation site occupancies at positions H4K5,

H4K8, H4K12 and H4K16 in CAF1 bound H4 (upper panel) and total histone H4 (lower panel). Each bar graph represents the mean of two technical replicates with error bars showing the relative distance of the maximum and the minimum values from the mean.

3.5. Discussion

Iso-PeptidAce was initially created to solve a specific problem that had apparently been neglected, even though it was pointed out by Smith *et al*. more than twelve years ago.[5] The problem of determining the relative abundance of multiply acetylated forms derived from the N-terminal tail of histone H4, which contains four lysines that can be acetylated, is compounded by the fact that several acetylated isomers co-elute during liquid chromatography and cannot be distinguished from each other based on the masses of MS^2 fragments. The value of determining the relative abundance of acetylated isomers was illustrated by our studies of new histone molecules bound to CAF1. The predominant pattern of acetylation that we found consisted of H4 molecules tri-acetylated at K5, K12 and K16. Although further studies are necessary to assess its biological significance, the prevalence of this acetylation pattern has implications for both replication-coupled nucleosome assembly and the roles of histone deacetylases in the propagation of heterochromatin structures in proliferating cells.

We expect that Iso-PeptidAce will be an asset to investigate a broad range of biological problems related to protein lysine acetylation and other combinatorial patterns of PTMs. For instance, Iso-PeptidAce will be valuable to determine the *in vivo* substrate specificity of the multitude of bromodomain (BRD)-containing proteins that exist in model organisms and human cells[3]. Because large protein complexes contain multiple BRDs located in different polypeptide subunits, deciphering their physiologically relevant acetylated substrates remains a formidable task. Although our approach is limited to patterns of PTMs that are confined to single peptides, previous *in vitro* studies have clearly shown that a number of BRDs bind with higher affinity to peptides that contain acetylated lysines located in close proximity[3]. We therefore anticipate that, when coupled with methods to co-purify protein complexes with their acetylated substrates, Iso-PeptidAce will help identify combinatorial patterns of lysine acetylation that are recognized by protein complexes involved in several biological processes.

Proteomics studies performed in several species have identified a myriad of acetylation sites in proteins with a wide range of biological functions[33-35]. In many cases, several lysines located near each other have been reported but, without isomer deconvolution, it is not possible to determine exactly which lysines are acetylated or whether a predominant pattern of acetylation does exist. Table S3.6 shows seven examples of known acetylated peptides where isomers could now be distinguished based on their fragmentation patterns using Iso-

PeptidAce. In addition, when combined with recently developed multiplexed parallel reaction monitoring approaches[20, 36], we anticipate that our method will prove valuable for deconvolution of a large set of isomeric peptides bearing various combinations of PTMs that include, but are not restricted to lysine acetylation.

It is known that in addition to lysine acetylation and methylation, histones are also modified with numerous other types of PTMs including mono- or di-methylation of arginine, crotonylation, butyrylation and formylation of lysine residues. In the current study we did not determine the later type of PTMs as our method is targeted toward known peptides and PTM combinations. Therefore, based on our dataset, we cannot rule out the possibility that other type of PTM combinations might exist on the peptides we reported in this study. Meanwhile, when coupled with recently developed multiplexed parallel reaction monitoring approaches[20, 36] we anticipate that our method will prove valuable for deconvolution of a large set of isobaric peptides bearing variable combinations of PTMs.

We also expect that our method will prove an asset to investigate a broad range of biological problems related to protein lysine acetylation. For instance, our approach will be valuable to determine the *in vivo* substrate specificity of the multitude of bromodomain (BRD)-containing proteins that exist in model organisms and human cells[3]. Previous *in vitro* studies have clearly shown that BRDs derived from chromatin-modifying enzymes bind with higher affinity to peptides that contain acetylated lysines located in close proximity[3]. Moreover, because large protein complexes contain multiple BRDs located in different subunits of a given complex, deciphering their physiologically relevant acetylated substrates remains a formidable task. Proteomics studies performed in several species have identified a myriad of acetylation sites in proteins with a wide range of biological functions[33-35]. We therefore anticipate that, when coupled with methods to co-purify protein complexes with their acetylated substrates, the method described here will help identify combinatorial patterns of lysine acetylation that are recognized by protein complexes involved in numerous biological processes.

3.6. Acknowledgments

We thank Dr. Sébastien Lemieux (UdeM) for his invaluable advice on the use of different bioinformatics tools and approaches. We also thank Éric Bonneil (UdeM) for his assistance with MS experiments. This work was carried out with financial support from the Canadian Institute for Health Research to A.V. (CIHR, FRN 125916) and the Natural Sciences and Engineering Research Council (NSERC 311598) to P.T. The Institute for Research in Immunology and Cancer (IRIC) receives infrastructure support from IRICoR, the Canadian Foundation for Innovation, and the Fonds de Recherche du Québec - Santé (FRQS).

3.7. Supplementary data

Table S3.1 Lists of synthetic H3 and H4 peptides and their concentration in mM.

List of synthetic histone (A) H3 and (B) H4 peptides and their concentration in mM. The concentration of each peptide was determined based on triplicate UV absorbance measurements at 205 nm.

A) H3 peptides

Peptide	Concentration (in mM)
Peptide 1: H3K18_K23NoAc	4.74
Peptide 2: H3K18Ac	5.58
Peptide 3: H3K23Ac	5.42
Peptide 4: H3K18_K23Ac	5.60

B) H4 peptides

Peptide	Concentration (in mM)
Peptide 1: H4NoAc	0.60
Peptide 2: H4K5Ac	0.51
Peptide 3: H4K8Ac	0.57
Peptide 4: H4K12Ac	0.77
Peptide 5: H4K16Ac	0.79
Peptide 6: H4K5_8Ac	0.71
Peptide 7: H4K5_12Ac	0.68
Peptide 8: H4K5_16Ac	0.76
Peptide 9: H4K8_12Ac	0.70
Peptide 10: H4K8_16Ac	0.70
Peptide 11: H4K12_16Ac	0.71
Peptide 12: H4K5_8_12Ac	0.69
Peptide 13: H4K5_12_16Ac	0.75
Peptide 14: H4K8_12_16Ac	0.74
Peptide 15: H4K5_8_16Ac	0.82
Peptide 16: H4K5_8_12_16Ac	0.66

Table S3.2 Synthetic H4 peptide standard dilutions.

Standard curves for seven dilutions of mixtures of H4 (A) mono-, (B) di- and (C) tri acetylated peptides with peptide amounts ranging from 5 to 320 fmol were prepared. Two peptides from the mono- or tri-acetylated group and three peptides from the di-acetylated group were added to the mixture at fixed amount of 80 fmol.

A) Mono-acetylated isomers

Sample #	Amount injected. (fmol)			
	K5Ac	K8Ac	K12Ac	K16Ac
1	5	80	5	80
2	10	80	10	80
3	20	80	20	80
4	40	80	40	80
5	80	80	80	80
6	160	80	160	80
7	320	80	320	80

B) Tri-acetylated isomers

Sample #	Amount injected. (fmol)			
	K5/8/12Ac	K5/12/16Ac	K8/12/16Ac	K5/8/16Ac
1	5	80	5	80
2	10	80	10	80
3	20	80	20	80
4	40	80	40	80
5	80	80	80	80
6	160	80	160	80
7	320	80	320	80

C) Di-acetylated isomers

sample #	Amount injected. (fmol)					
	K5/8Ac	K5/12Ac	K5/16Ac	K8/12Ac	K8/16Ac	K12/16Ac
1	5	80	5	80	5	80
2	10	80	10	80	10	80
3	20	80	20	80	20	80
4	40	80	40	80	40	80
5	80	80	80	80	80	80
6	160	80	160	80	160	80
7	320	80	320	80	320	80

Table S3.3 Deconvolution of mixture spectra by Iso-PeptidAce.
Amount of each peptide isomer determined by Iso-PeptidAce after deconvolution of mixtures of known amounts of (A) four mono-, (B) four tri- and (C) six di-acetylated isomeric groups. Sample at 80 fmols was used for normalisation of peptide intensities.

A) mono-acetylated H4 peptides (Mean of two technical replicates ± range error bars)

Amount Injected	Acetylated synthetic peptide			
	K5ac	K12ac	K8ac	K16ac
5	1.33 ± 0.0	5.05 ± 0.0	74.28 ± 0.0	74.52 ± 0.0
10	9.27 ± 4.4	12.36 ± 3.1	69.09 ± 7.51	66.59 ± 9.6
20	24.24 ± 8.2	25.17 ± 8.6	68.51 ± 8.9	67.18 ± 9.8
40	50.52 ± 12.3	49.49 ± 9.3	69.13 ± 8.7	68.23 ± 7.7
80	80 ± 0.0	80 ± 0.0	80 ± 0.0	80 ± 0.0
160	168.61 ± 4.4	181.21 ± 2.4	84.47 ± 7.6	86.99 ± 9.9
320	370.1 ± 58.4	401.15 ± 55.7	89.28 ± 0.3	89.21 ± 6.9

B) Tri-acetylated H4 peptides (Mean of two technical replicates ± range error bars)

Amount Injected	Acetylated synthetic peptide			
	K5/8/12ac	K5/12/16ac	K5/8/16ac	K8/12/16ac
5	1.92 ± 0.0	2.44 ± 0.0	65.4 ± 0.	60.23 ± 0.0
10	7.47 ± 1.1	7.49 ± 0.3	93.9 ± 4.9	89.55 ± 10.1
20	15.38 ± 2.3	16.18 ± 1.2	99.13 ± 9.6	86.01 ± 7.6
40	35.1 ± 0.5	35.55 ± 1.0	95.25 ± 2.1	85.31 ± 9.7
80	80 ± 0.0	80 ± 0.0	80 ± 0.0	80 ± 0.0
160	155.82 ± 4.3	157.31 ± 8.7	76.56 ± 0.0	77.02 ± 1.5
320	295.55 ± 6.6	310.17 ± 20.9	78.86 ± 3.3	72.15 ± 3.4

C) Di-acetylated H4 peptides (Mean of two technical replicates ± range error bars)

Amount Injected	Amount determined by Iso-PeptidAce (fmol)					
	K5/8ac	K5/16ac	K8/16ac	K5/12ac	K8/12ac	K12/16ac
5	3.7 ± 0.0	1 ± 0.0	6.92 ± 0	70.98 ± 0.0	286.89 ± 0	114.68 ± 0.0
10	8.17 ± 1.0	4.31 ± 1.1	11.56 ± 2.0	85.64 ± 8.2	223.19 ± 8.7	110.13 ± 0.5
20	19.42 ± 4.8	17.03 ± 1.7	25.73 ± 1.8	96.52 ± 8.1	215.69 ± 5.2	114.52 ± 1.1
40	46.33 ± 2.9	52.58 ± 12.5	50.85 ± 4.1	101.69 ± 11.2	200.39 ± 26.2	115.94 ± 0.4
80	80 ± 0.0	80 ± 0.0	80 ± 0.0	80 ± 0.0	80 ± 0.0	80 ± 0.0
160	276.67 ± 21.2	459.33 ± 159.6	207.11 ± 3.9	104.26 ± 3.6	71.11 ± 3.6	98.01 ± 3.7
320	510.19 ± 63.7	1051.96 ± 500.8	372.77 ± 28.5	109.1 ± 0.6	65.78 ± 2.3	95.97 ± 1.37

* Values in this row represent signal response measurements from only one replicate.

Table S3.4 Calculated acetylation site occupancies.
Acetylation site occupancies of (A) H3 and (B) H4 peptides before or after HDACi treatment.Acetylation site occupancy of, for instance H4K5ac, was determined by dividing the sum of normalized intensities of all peptides containing acetylated Lys-5 by the sum ofnormalized intensities of all forms of acetylated and un-acetylated H4 peptide 4 GKGGKGLGKGGAKR-17.

A. H3_acetylation site occupancy (Mean of two technical replicates ± range error bars)

Samples	H3K18	H3K23
DMSO-1h	4.23 ± 0.14	26.56 ± 0.66
DMSO-6h	6.03 ± 0.14	31.45 ± 0.93
DMSO-24h	5.89 ± 0.74	32.06 ± 0.98
MS275-1h	5.74 ± 0.15	32.03 ± 4.85
MS275-6h	10.64 ± 1.89	51.95 ± 6.32
MS275-24h	27.11 ± 1.67	63.35 ± 0.81
SAHA-1h	17.19 ± 2.46	55.89 ± 0.89
SAHA-6h	31.94 ± 0.28	71.92 ± 1.22
SAHA-24h	37.70 ± 6.36	79.73 ± 0.77
JNJ-1h	15.58 ± 1.17	59.95 ± 0.02
JNJ-6h	32.76 ± 2.66	76.55 ± 1.39
JNJ-24h	44.17 ± 1.53	82.77 ± 1.21

B. H4_acetylation site occupancy (Mean of two technical replicates ± range error bars)

Samples	H4K5ac	H4K8ac	H4K12ac	H4K16ac
DMSO-1h	2.68 ± 0.59	2.63 ± 0.46	5.14 ± 0.53	29.27 ± 1.27
DMSO-6h	3.05 ± 0.07	2.86 ± 0.02	5.88 ± 0.04	28.42 ± 1.00
DMSO-24h	3.43 ± 0.60	3.79 ± 0.51	7.52 ± 0.98	32.27 ± 0.08
MS275-1h	3.42 ± 0.16	3.66 ± 0.35	7.12 ± 0.58	31.79 ± 0.52
MS275-6h	13.54 ± 0.78	13.43 ± 1.25	22.77 ± 1.21	57.81 ± 0.45
MS275-24h	36.79 ± 2.90	35.66 ± 1.47	55.16 ± 0.05	80.28 ± 0.23
SAHA-1h	16.24 ± 1.89	17.98 ± 1.65	33.03 ± 3.86	57.19 ± 0.79
SAHA-6h	44.96 ± 1.19	38.59 ± 1.91	63.45 ± 0.81	80.77 ± 0.17
SAHA-24h	61.35 ± 0.21	51.05 ± 3.30	77.49 ± 0.40	89.77 ± 1.16
JNJ-1h	13.51 ± 1.96	14.83 ± 1.43	26.48 ± 3.10	51.90 ± 3.30
JNJ-6h	47.15 ± 1.38	39.41 ± 0.06	64.59 ± 0.85	80.96 ± 0.49
JNJ-24h	57.67 ± 0.39	47.59 ± 1.18	68.25 ± 2.07	89.00 ± 0.80

Table S3.5 LC/MS intensity data for CAF-1-bound and total H3 and H4 peptides.

Histone H3 Peptide sequence	m/z	Charge	PTM	LC/MS Intensity	
				CAF-1 bound	Total histone
9-KSTGGKAPR-17	507.2906	2	H3K9prK14pr	860033.68 ± 231000	390383113.88 ± 11300000
	493.2748	2	H3K9acK14pr	255287.72 ± 17000	35576306.85 ± 12000000
	500.2836	2	H3K9prK14ac	415406.09 ± 2060	109502233.83 ± 2840000
		2	H3K9acK14ac	2253016.49 ± 418000	19487793.22 ± 5660000
18-KQLASKAAR-23	542.827	2	K18prK23pr	493253.48 ± 265000	703000000 ± 137000000
	535.819	2	K18acK23pr	28786.03 ± 26500	40300000 ± 14600000
		2	K18prK23ac	608652.18 ± 46000	149875000 ± 59300000
	528.811	2	K18acK23ac	58024.76 ± 34300	17545000 ± 6020000
27-KSAPSTGGVKKPHR-40	809.4568	2	K27prK36prK37pr	47760.38 ± 3200	46475000 ± 12900000
	816.4646	2	K27prK36pr+me1K37pr	0.00E+00	13272500 ± 4710000
	795.4595	2	K27prK36me2K37pr	0.00E+00	8742500 ± 300000
	530.642	3	K27prK36me2K37pr	0.00E+00	50900000 ± 2100000
	802.4673	2	K27prK36me3K37pr	0.00E+00	15470000 ± 600000
	535.3139	3	K27prK36me3K37pr	0.00E+00	80525000 ± 3000000
	802.4491	2	K27acK36prK37pr	43890.86 ± 6660	9195000 ± 300000
	809.4569	2	K27acK36pr+me1K37pr	0.00E+00	2970000 ± 800000
	788.4518	2	K27acK36me2K37pr	0.00E+00	574750 ± 100000
	525.9702	3	K27acK36me2K37pr	0.00E+00	2940000 ± 100000
	795.4596	2	K27acK36me3K37pr	0.00E+00	609250 ± 130000
	530.6421	3	K27acK36me3K37pr	0.00E+00	3507500 ± 1330000
	539.9736	3	K27prK36prK37pr	182908.65 ± 11300	168275000 ± 6040000
	544.6455	3	K27prK36pr+me1K37pr	0.00E+00	41650000 ± 15600000
	398.2333	4	K27prK36me2K37pr	0.00E+00	58900000 ± 21800000
	401.7373	4	K27prK36me3K37pr	0.00E+00	105650000 ± 38700000
	535.3018	3	K27acK36prK37pr	163800.96 ± 61100	29250000 ± 1380000
	539.9737	3	K27acK36pr+me1K37pr	0.00E+00	8445000 ± 1800000
	394.7295	4	K27acK36me2K37pr	0.00E+00	3520000 ± 1240000
	398.23334	4	K27acK36me3K37pr	0.00E+00	4467500 ± 1590000
54-FQK$_{56}$STELLIR-63	645.8743	2	K56pr	80480.47 ± 9580	47700000 ± 5940000
	638.8666	2	K56ac	197734.94 ± 39500	15050000 ± 1630000
53-RFQK$_{56}$STELLIR-63	723.9248	2	K56pr	23625.85 ± 20000	13207500 ± 2960000
	482.9523	3	K56pr	11242.59	11172500 ± 2580000
	478.2805	3	K56ac	49589.66	3972500 ± 1440000
	716.9171	2	K56ac	167034.28 ± 107000	4497500 ± 1250000

(continued…Table S3.5)

Histone H4 Peptide sequence	m/z	Charge	PTM	LC/MS Intensity	
				CAF-1 bound	Total histone
4-GKGGKGLGKGGAKR-17	747.9412	2	K5-8-12-16pr	438525.03 ± 2090	21069549.87 ± 178000
	740.9334	2	K5ac\K8-12-16pr	0.00E+00	21208.89 ± 3630
			K8ac\K5-12-16pr	0.00E+00	357774.80 ± 39300
			K12ac\K5-8-16pr	1744.11 ± 2470	2410070.69 ± 145000
			K16ac\K5-8-12pr	24985.21 ± 35300	21740122.97 ± 1810000
	733.9255	2	K5-8ac\K12-16pr	0.00E+00	31596.76 ± 2540
			K5-12ac\K8-16pr	0.00E+00	347870.13 ± 25700
			K5-16ac\K8-12pr	0.00E+00	1666742.07 ± 389000
			K8-12ac\K5-16pr	0.00E+00	815051.26 ± 64900
			K8-16ac\K5-12pr	0.00E+00	2629819.07 ± 255000
			K12-16ac\K5-8pr	0.00E+00	8159580.12 ± 323000
	726.9177	2	K5-8-12ac\K16pr	488.08 ± 690	254396.05 ± 58200
			K5-8-16ac\K12pr	976.16	2092262.66 ± 58900
			K5-12-16ac\K8pr	110061.69 ± 18200	1909117.03 ± 582000
			K8-12-16ac\K5pr	17472.77 ± 12400	4395165.79 ± 1010000
	719.9092	2	K5-8-12-16ac	0.00E+00	3553895.37 ± 593000

Note: Each intensity value is the Mean of two technical replicates + range error bars)

Table S3.6 List of known acetylation sites in selected bromodomains substrates.

Peptides containing multiple acetylated lysine residues generated from p53, GATA1, Cyclin T1, or MyoD are shown. List of publications for each of the previously identified acetylation sites are indicated.

Protein name	Interacting Bromodomain containing protein/s	Peptide segment containing multiple lysine residues	Known acetylation sites	Reference
p53	CREBBP, TAF1	290-RKKGEPHHELPPGSTKR-306	K292, K305	37- 43
		306-RALPNNTSSSPQPKKKPLDGEYFTLQIR-333	K319, K320, K321	
		363- RAHSSHLKSKKGQSTSRHKKLMFKTEGP-390	K370, K372, K373, K381, K382, K386	
GATA1	BRD3	220 - RTGHYLCNACGL YHKMNGQNRPLIRP KKR- 247	K233, K245, K246	44 - 46
		301- GIQTRNRKASGKGKKKRGSS -320	K308, K312, K314, K315, K316	
CyclinT1	BRD4	374-SQKQNSKSVPSAKVSLKEYR-393	K380, K386, K390	38, 47
MyoD	CREBBP or EP300	92- RCLLWACKACKRKTTNADR-110	K99, K102, K104	48 - 49

Table S3.7 Raw intensity data for control and HDACi treated samples.
For each sample Iso-PeptidAce reports five different groups of peptides: one un-acetylated (m/z 747.94), four mono-acetylated (m/z 740.93), six di-acetylated (m/z 733.93), four tri-acetylated (m/z 726.92) and one tetra-acetylated (m/z 719.91).

Sample	Precursor m/z	K5/8/12/16pr	K5ac	K8ac	K12ac	K16ac	K5-8ac	K5-12ac	K5-16ac	K8-12ac	K8-16ac	K12-16ac	K5-8-12ac	K5-8-16ac	K5-12-16ac	K8-12-16ac	K5-8-12-16ac
DMSO-1h	747.9412	5E+07	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
DMSO-1h	740.9334	0	1E+06	6E+05	2E+06	2E+07	0	0	0	0	0	0	0	0	0	0	0
DMSO-1h	733.9255	0	0	0	0	0	30356	1E+05	4E+05	2E+05	7E+05	2E+06	0	0	0	0	0
DMSO-1h	726.9177	0	0	0	0	0	0	0	0	0	0	0	68875	2E+05	3E+05	5E+05	0
DMSO-1h	719.9099	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2E+05
DMSO-6h	747.9412	3E+08	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
DMSO-6h	740.9334	0	7E+06	4E+06	1E+07	1E+08	0	0	0	0	0	0	0	0	0	0	0
DMSO-6h	733.9255	0	0	0	0	0	1E+05	9E+05	3E+06	1E+06	4E+06	1E+07	0	0	0	0	0
DMSO-6h	726.9177	0	0	0	0	0	0	0	0	0	0	0	6E+05	7E+05	2E+06	2E+06	0
DMSO-6h	719.9099	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1E+06
DMSO-24h	747.9412	3E+08	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
DMSO-24h	740.9334	0	7E+06	4E+06	1E+07	1E+08	0	0	0	0	0	0	0	0	0	0	0
DMSO-24h	733.9255	0	0	0	0	0	1E+05	8E+05	3E+06	1E+06	5E+06	1E+07	0	0	0	0	0
DMSO-24h	726.9177	0	0	0	0	0	0	0	0	0	0	0	7E+05	1E+06	1E+06	3E+06	0
DMSO-24h	719.9099	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2E+06
JNJ-1h	747.9412	1E+08	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
JNJ-1h	740.9334	0	4E+06	2E+06	9E+06	7E+07	0	0	0	0	0	0	0	0	0	0	0
JNJ-1h	733.9255	0	0	0	0	0	95795	2E+06	3E+06	1E+06	7E+06	2E+07	0	0	0	0	0
JNJ-1h	726.9177	0	0	0	0	0	0	0	0	0	0	0	9E+05	1E+06	6E+06	9E+06	0
JNJ-1h	719.9099	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2E+07
JNJ-6h	747.9412	2E+07	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
JNJ-6h	740.9334	0	2E+06	9E+05	3E+06	2E+07	0	0	0	0	0	0	0	0	0	0	0
JNJ-6h	733.9255	0	0	0	0	0	0	3E+06	3E+06	7E+05	5E+06	2E+07	0	0	0	0	0
JNJ-6h	726.9177	0	0	0	0	0	0	0	0	0	0	0	4E+05	3E+06	2E+07	1E+07	0
JNJ-6h	719.9099	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4E+07
JNJ-24h	747.9412	1E+07	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
JNJ-24h	740.9334	0	5E+06	2E+06	4E+06	3E+07	0	0	0	0	0	0	0	0	0	0	0
JNJ-24h	733.9255	0	0	0	0	0	0	5E+06	1E+07	2E+06	1E+07	4E+07	0	0	0	0	0
JNJ-24h	726.9177	0	0	0	0	0	0	0	0	0	0	0	1E+06	2E+07	4E+07	2E+07	0
JNJ-24h	719.9099	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1E+08
MS275-1h	747.9412	1E+08	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MS275-1h	740.9334	0	3E+06	1E+06	4E+06	5E+07	0	0	0	0	0	0	0	0	0	0	0
MS275-1h	733.9255	0	0	0	0	0	94043	4E+05	1E+06	4E+05	2E+06	5E+06	0	0	0	0	0
MS275-1h	726.9177	0	0	0	0	0	0	0	0	0	0	0	2E+05	3E+05	5E+05	1E+06	0
MS275-1h	719.9099	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4E+05
MS275-6h	747.9412	8E+07	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MS275-6h	740.9334	0	6E+06	2E+06	7E+06	7E+07	0	0	0	0	0	0	0	0	0	0	0
MS275-6h	733.9255	0	0	0	0	0	42800	2E+06	9E+06	1E+06	1E+07	2E+07	0	0	0	0	0
MS275-6h	726.9177	0	0	0	0	0	0	0	0	0	0	0	4E+05	3E+06	7E+06	8E+06	0
MS275-6h	719.9099	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5E+06
MS275-24h	747.9412	6E+07	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MS275-24h	740.9334	0	1E+07	4E+06	2E+07	1E+08	0	0	0	0	0	0	0	0	0	0	0
MS275-24h	733.9255	0	0	0	0	0	4E+05	1E+07	3E+07	1E+07	4E+07	8E+07	0	0	0	0	0
MS275-24h	726.9177	0	0	0	0	0	0	0	0	0	0	0	2E+06	2E+07	5E+07	5E+07	0
MS275-24h	719.9099	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1E+08
SAHA-1h	747.9412	8E+07	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
SAHA-1h	740.9334	0	3E+06	2E+06	7E+06	6E+07	0	0	0	0	0	0	0	0	0	0	0
SAHA-1h	733.9255	0	0	0	0	0	36592	2E+06	3E+06	1E+06	4E+06	2E+07	0	0	0	0	0
SAHA-1h	726.9177	0	0	0	0	0	0	0	0	0	0	0	9E+05	1E+06	6E+06	1E+07	0
SAHA-1h	719.9099	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2E+07
SAHA-6h	747.9412	6E+07	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
SAHA-6h	740.9334	0	5E+06	2E+06	9E+06	7E+07	0	0	0	0	0	0	0	0	0	0	0
SAHA-6h	733.9255	0	0	0	0	0	1E+05	1E+07	1E+07	2E+06	2E+07	7E+07	0	0	0	0	0
SAHA-6h	726.9177	0	0	0	0	0	0	0	0	0	0	0	1E+06	7E+06	5E+07	3E+07	0
SAHA-6h	719.9099	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1E+08
SAHA-24h	747.9412	2E+07	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
SAHA-24h	740.9334	0	5E+06	1E+06	5E+06	3E+07	0	0	0	0	0	0	0	0	0	0	0
SAHA-24h	733.9255	0	0	0	0	0	2E+05	7E+06	1E+07	5E+06	1E+07	5E+07	0	0	0	0	0
SAHA-24h	726.9177	0	0	0	0	0	0	0	0	0	0	0	3E+06	1E+07	6E+07	3E+07	0
SAHA-24h	719.9099	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1E+08

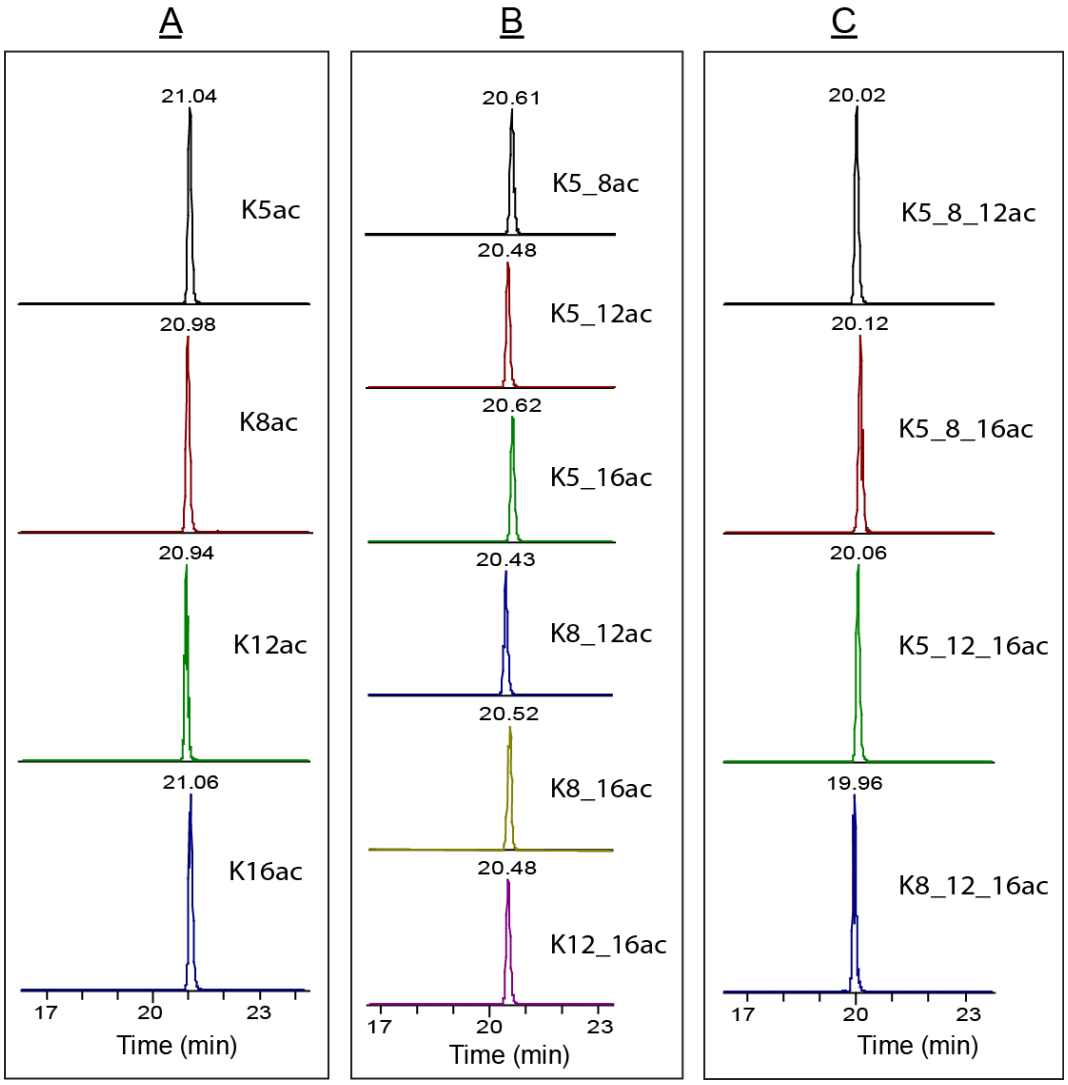
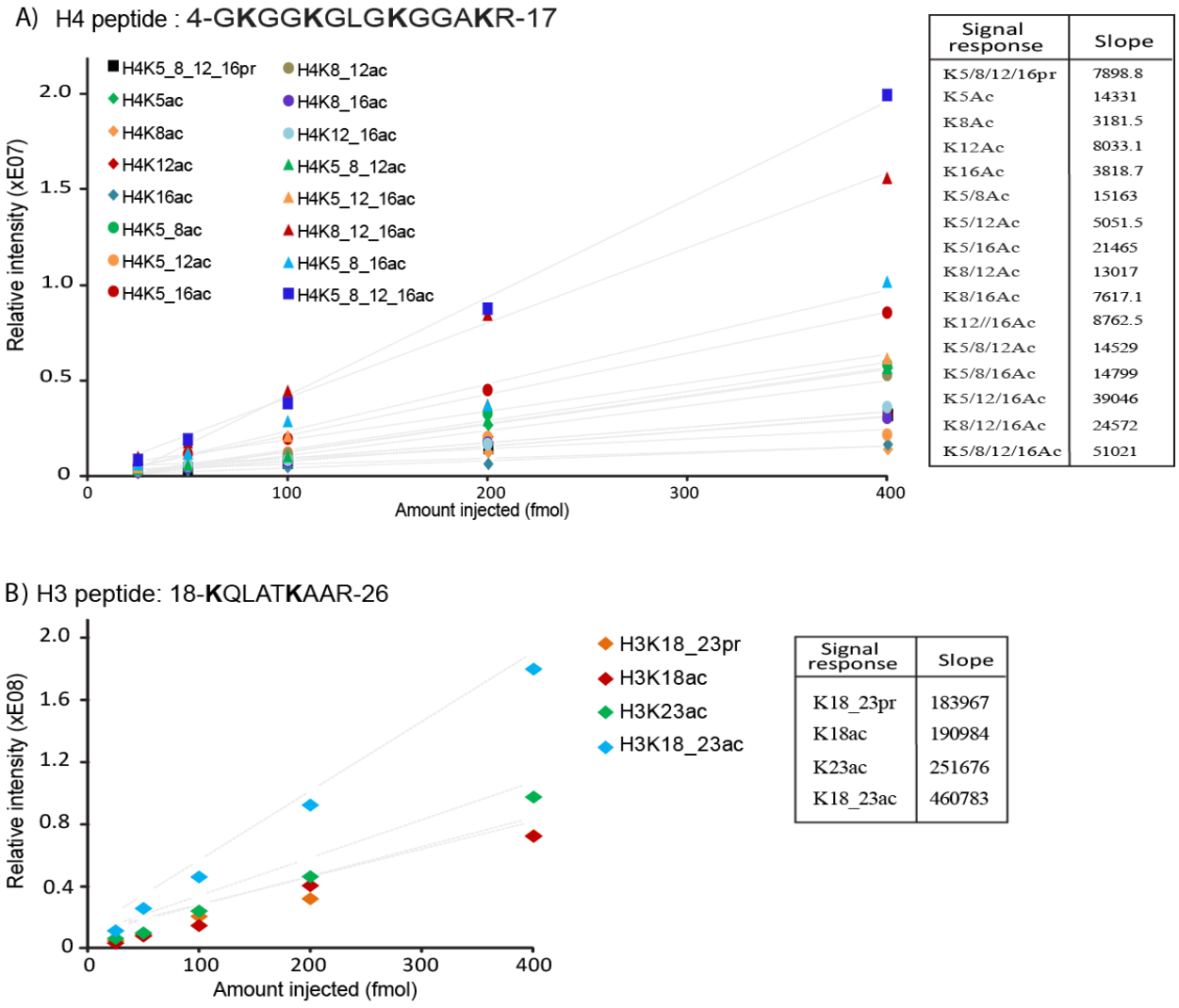


Figure S3.1 Peptides within the same isomeric groups exhibit very narrow retention.

Shown in each column are the extracted ion chromatograms of (A) four mono-acetylated (m/z 740.93, +2), (B) six di-acetylated (m/z 733.93, 2+), and (C) four tri-acetylated (m/z 726.92, +2) isomers derived from the H4 peptide 4-GKGGKGLGKGGAKR-17.



Signal response	Slope
K5/8/12/16pr	7898.8
K5Ac	14331
K8Ac	3181.5
K12Ac	8033.1
K16Ac	3818.7
K5/8Ac	15163
K5/12Ac	5051.5
K5/16Ac	21465
K8/12Ac	13017
K8/16Ac	7617.1
K12//16Ac	8762.5
K5/8/12Ac	14529
K5/8/16Ac	14799
K5/12/16Ac	39046
K8/12/16Ac	24572
K5/8/12/16Ac	51021

Signal response	Slope
K18_23pr	183967
K18ac	190984
K23ac	251676
K18_23ac	460783

Figure S3.2 MS signal responses of synthetic H3 and H4 peptides.

After i*n vitro* propionylation and trypsin digestion of synthetic peptides, the relative MS response of each peptide was determined based on the curve of ion intensity versus amount injected for the non-acetylated and acetylated forms of (A) H3 peptide 18-KQLATKAAR-26 and (B) H4 peptide 4-GKGGKGLGKGGAKR-17. The tables on the right show the slopes of the lines determined from the linear equation that best-fit (as measure by R2 values) the signal responses.

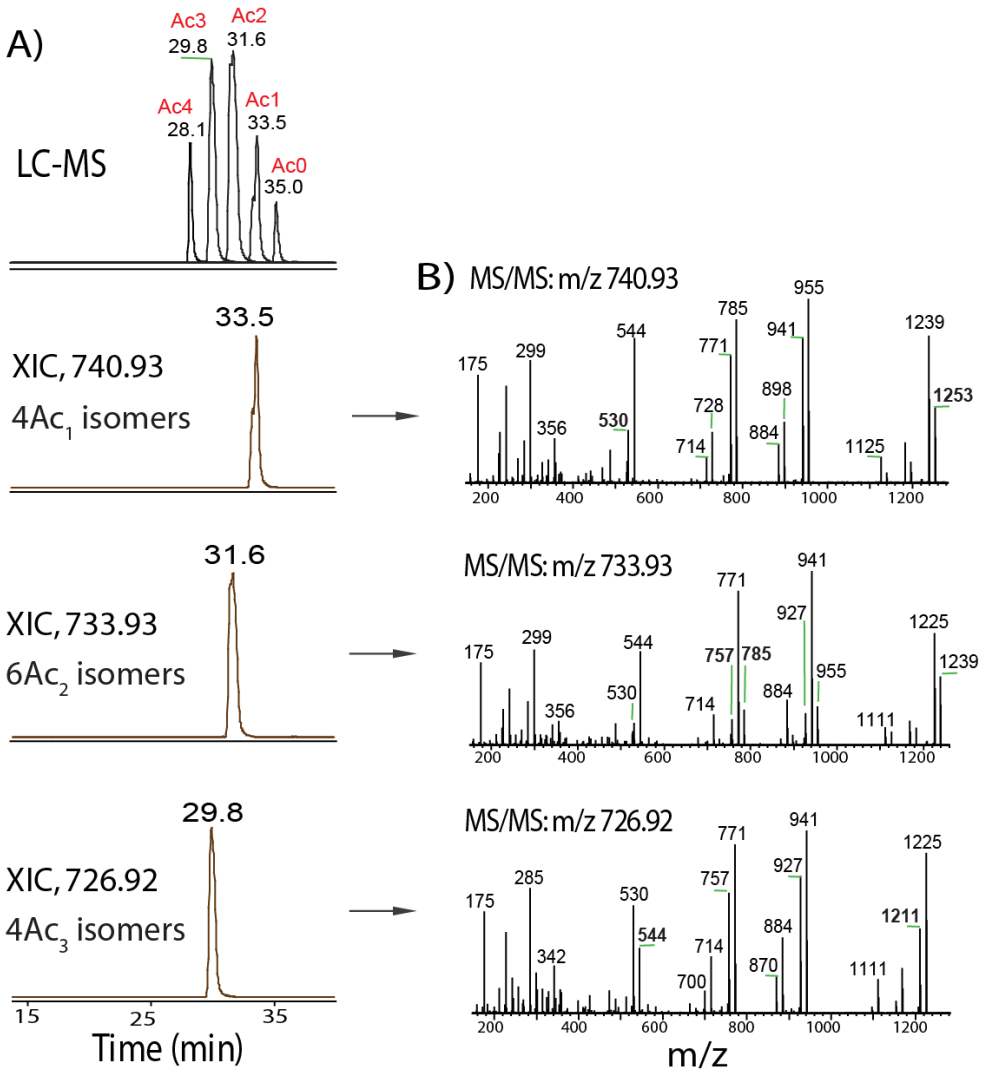


Figure S3.3 Co-elution and co-fragmentation of acetylated isomers.

A) LC-MS trace of a mixture of sixteen H4 peptides, and extracted ion chromatograms of co-eluting four mono-, six di- and four tri-acetylated isomers of peptide 4-GKGGKGLGKGGAKR-17 (top to bottom panels). B) Mixed MS/MS spectra of four mono-, six di- and four tri-acetylated isomers (from top to bottom panel).

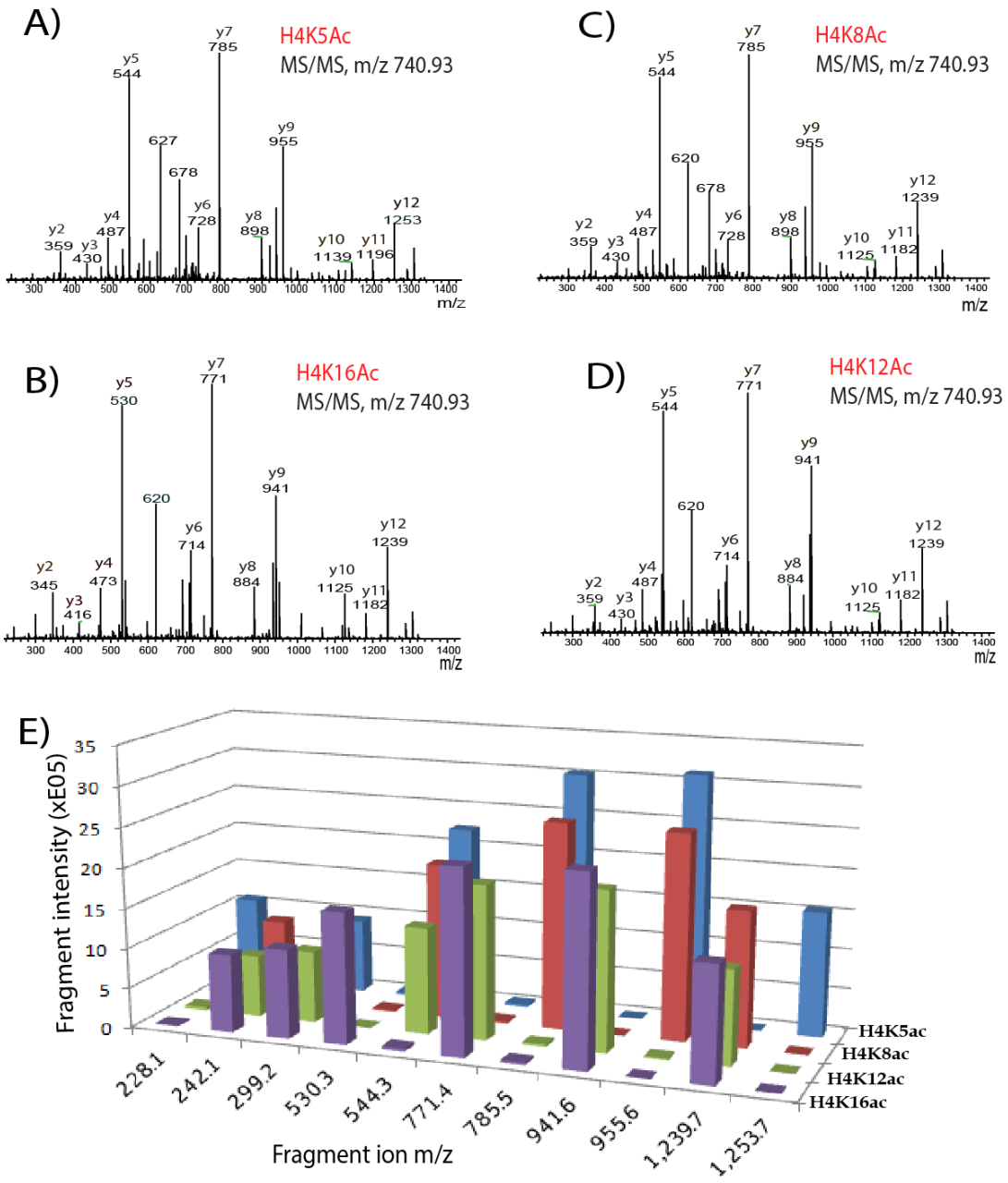


Figure S3.4 Isomeric peptides produce distinct fragment ion patterns.

Representative fragment ion spectra of four mono-acetylated isomers of the peptide 4-GKGGKGLGKGGAKR-17: (A) H4K5ac, (B) H4K16ac, (C) H4K8ac, and (D) H4K12ac. The 'y' type fragment ions are indicated in each spectrum. E) A three-dimensional view of the distribution of selected 'b' and 'y' type fragment ions derived from the individual mono-acetylated isomers. The 'b' type fragment ions are shown at m/z 228.1 (b2, K5ac), 242.1 (b2, K5pr) and 299.2 (b3, K5pr). The 'y' type fragment ions are m/z 530.3, 544.30, 771.4, 785.5, 941.6, 955.6, 1239.7 and 1253.7.

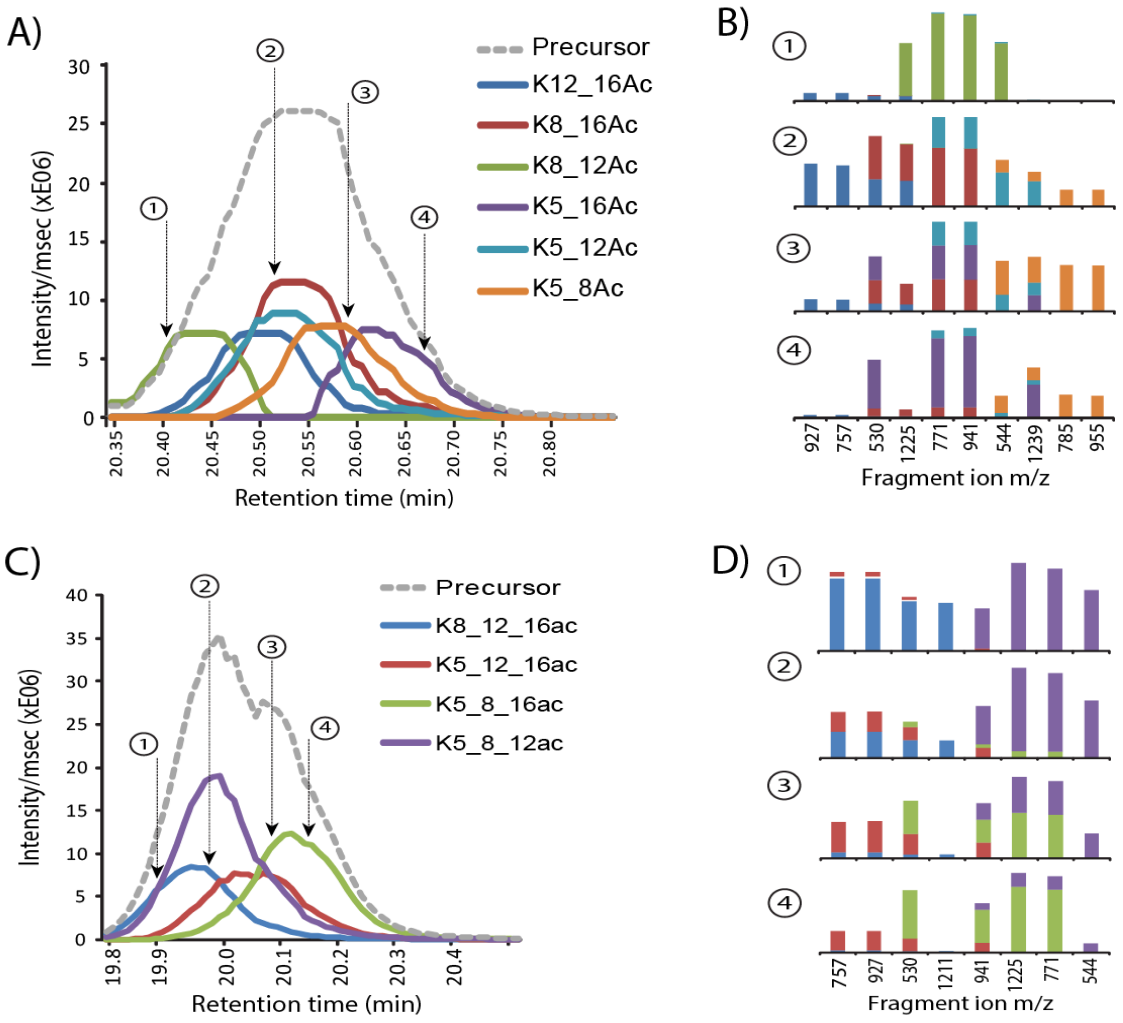


Figure S3.5 Deconvolution of di- and tri-acetylated isomers of histone H4.

Elution profiles before (dashed line) or after (colored lines) deconvolution of a mixture of (A) six di-acetylated, and (B) four tri-acetylated isomers of the peptide 4-GKGGKGLGKGGAKR-17. Distributions of fragment peaks of the deconvoluted di- (C) and tri-acetylated (D) isomers.

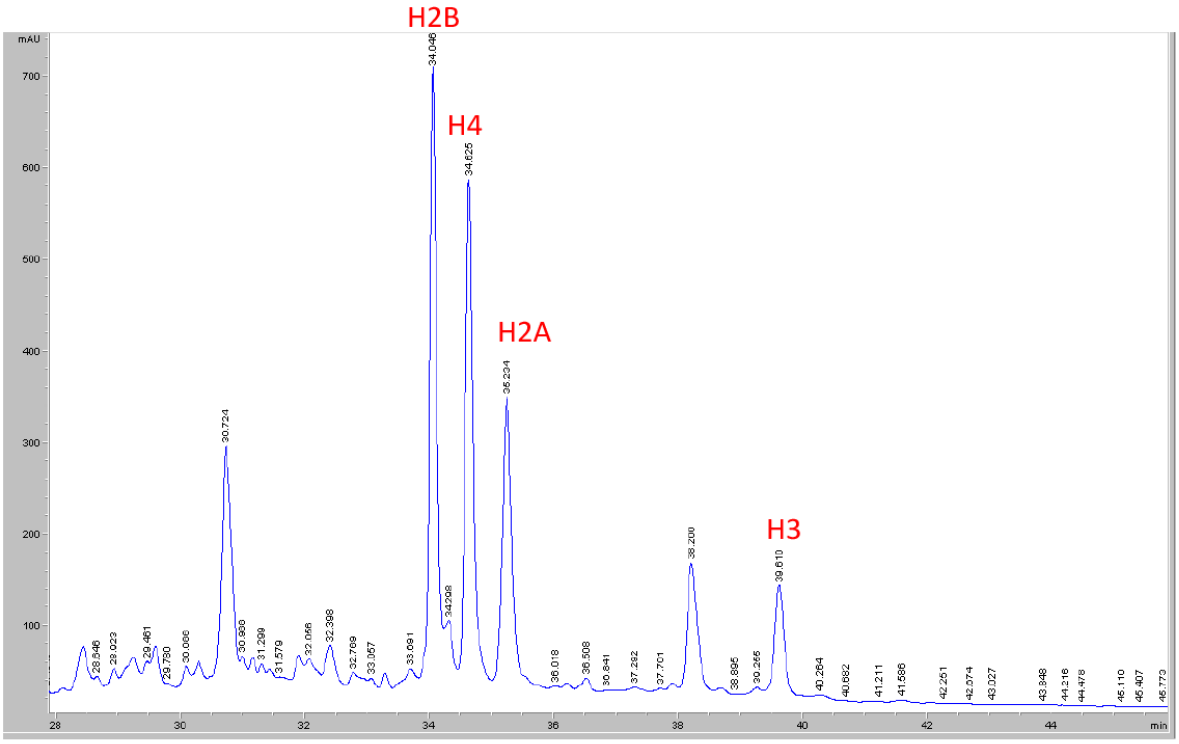


Figure S3.6 Fractionation of total histones by RP-HPLC.

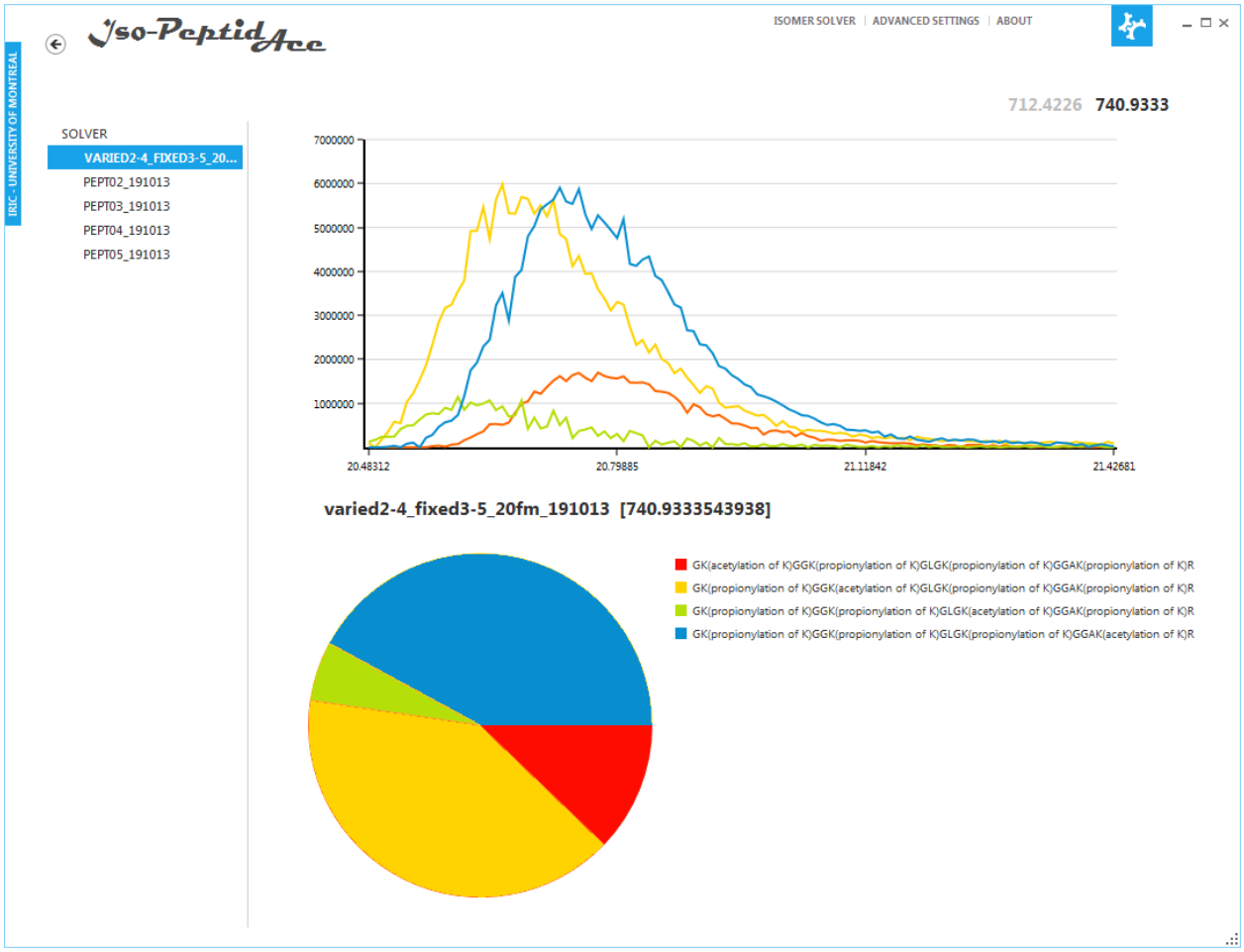
Total histones were baselineseparated as H2B, H4, H2A and H3.1 in increasing order of retention time. Histone H3.1 and H4 fractions were pooled and dried completely in a speed-vac concentrator. Pump A: 0.1% trifluoroacetic acid in 100% filtered water (v/v). Pump B: 0.1% trifluoroacetic acid in 100% acetonitrile (v/v)., flow rate:150μl/min, Absorbance wavelength: 214 nm, Column: Zorbax C8, 5µm, 2.1x150 mm.

3.8. Supplementary method

Iso-PeptidAce : In-silico separation of co-eluting peptides

Overview

Here we describe Iso-PeptidAce, a stand-alone software for the in-silico separation of co-eluting peptide isomers in high resolution mass spectrometry data. Peptide isomers of interest need to be synthesized and processed individually in a high resolution mass spectrometer, with dynamic exclusion turned off. Samples with unknown proportions of co-eluting peptide isomers processed in the same instrument, with a similar method will be separated by the software. Iso-PeptidAce can also quantify (relatively) co-eluting peptides from mixed samples. To our knowledge, this is the first software capable of such a feat.



List of histone H4 peptides used for testing Iso-PeptidAce

Peptide 1	GK(propionylation of K)GGK(propionylation of K)GLGK(propionylation of K)GGAK(propionylation of K)R
Peptide 2	GK(acetylation of K)GGK(propionylation of K)GLGK(propionylation of K)GGAK(propionylation of K)R
Peptide 3	GK(propionylation of K)GGK(acetylation of K)GLGK(propionylation of K)GGAK(propionylation of K)R
Peptide 4	GK(propionylation of K)GGK(propionylation of K)GLGK(acetylation of K)GGAK(propionylation of K)R
Peptide 5	GK(propionylation of K)GGK(propionylation of K)GLGK(propionylation of K)GGAK(acetylation of K)R
Peptide 6	GK(acetylation of K)GGK(acetylation of K)GLGK(propionylation of K)GGAK(propionylation of K)R
Peptide 7	GK(acetylation of K)GGK(propionylation of K)GLGK(acetylation of K)GGAK(propionylation of K)R
Peptide 8	GK(acetylation of K)GGK(propionylation of K)GLGK(propionylation of K)GGAK(acetylation of K)R
Peptide 9	GK(propionylation of K)GGK(acetylation of K)GLGK(acetylation of K)GGAK(propionylation of K)R
Peptide 10	GK(propionylation of K)GGK(acetylation of K)GLGK(propionylation of K)GGAK(acetylation of K)R
Peptide 11	GK(propionylation of K)GGK(propionylation of K)GLGK(acetylation of K)GGAK(acetylation of K)R
Peptide 12	GK(acetylation of K)GGK(acetylation of K)GLGK(acetylation of K)GGAK(propionylation of K)R
Peptide 13	GK(acetylation of K)GGK(acetylation of K)GLGK(propionylation of K)GGAK(acetylation of K)R
Peptide 14	GK(acetylation of K)GGK(propionylation of K)GLGK(acetylation of K)GGAK(acetylation of K)R
Peptide 15	GK(propionylation of K)GGK(acetylation of K)GLGK(acetylation of K)GGAK(acetylation of K)R
Peptide 16	GK(acetylation of K)GGK(acetylation of K)GLGK(acetylation of K)GGAK(acetylation of K)R

Symbols used

- *MS1i, MS2i* = Injection Time for MS1 and MS2 Scans, in milliseconds
- *iMS* = ions per milliseconds
- *#ion* = Number of monoisotopic precursor ions fragmented in a scan
- *psm* = Peptide Spectrum Match
- *fi* = Fragment Intensity
- *nfi* = Normalized Fragment Intensity
- *n* = Number of fragments considered per characterized peptide
- *rt* = Scan Time
- Δmz= Precursor mass error, in ppm
- *mif* = Matching fragment ion divided by the number of theoretical ion tested
- c_j= Number of ions detected for peptide *j*
- *aMz* = List of selected fragment masses
- *ac* = Area Under the Curve
- *pc* = Peptide Count (Estimation of the number of a certain peptide in a sample)

Description of how spectral deconvolution is carried out by Iso-PeptidAce

Raw file extraction

Information is extracted from Raw files using ProteoWizard[50] and MsFileReader. MS2 peak lists (masses and intensities) are stored per precursor mass and sorted by retention time. Spectrum precursor intensity, MS1 and MS2 injection times are also extracted.
Precursor intensity counts are converted to ions per millisecond (*iMS*):

$$iMS = Precursor\ Intensity\ Count/MS1\ injection\ time$$

For each different precursor mass in a file, a curve describing the number of ions flowing in the system at any given time is built using the computed *iMS* values. Linear interpolation is used to compute the area under the curve for a given time point. In particular, the number of ions that entered the CTrap and became fragmented (*#ion*) is estimated as the area under the curve between the beginning (*rt*) and the end of the scan (*rt* + *MS2i*).

$$\#ion = Area\ Under\ The\ Curve\ (from: rt, to: rt + MS2i)$$

These operations are made for each file (both mixed samples and synthetic peptide runs). For the synthetic peptide runs, spectrum fragment intensity counts (*fi*) are normalized (*nfi*) using the number of ions (*#ion*) that became fragmented.

$$nfi = fi/\#ion$$

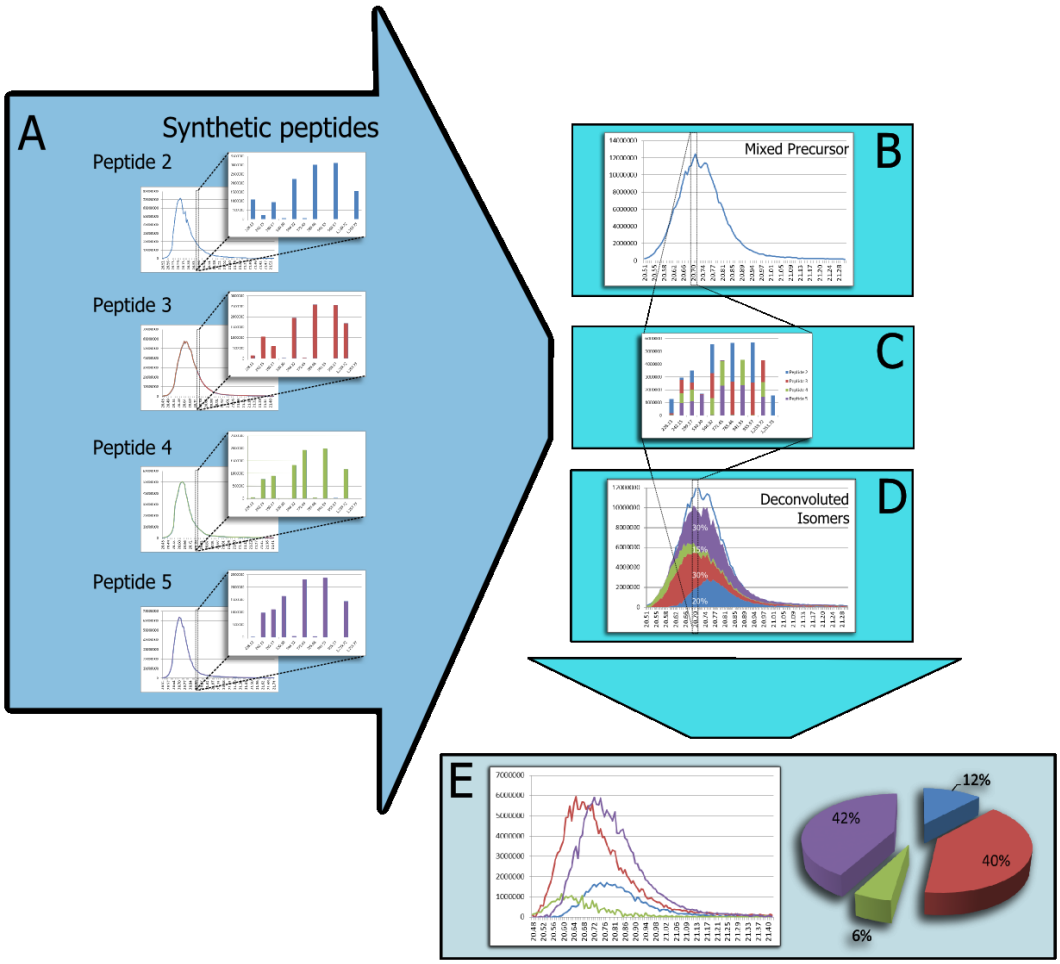
Peptide Spectrum Matching

Iso-PeptidAce uses PeptidAce peptide spectrum matching abilities (which is derived from Morpheus) to automatically identify peptide sequences and modifications across synthetic peptide runs. PeptidAce uses a no-enzyme in-silico protein digestion routine to parse the provided Fasta file for potential matches. All modifications (specified in the "Modifications.csv" user configurable Comma Separated Values file) are searched for. This automatic identification is useful to discover impurities in the peptide synthesis process.
PeptidAce matches peptide sequences to spectra by comparing theoretical ions (fragments *a*, *b*, *c*, *x*, *y* and *z*) within a specified ppm tolerance window. Peptide spectrum matches (*psm*) are scored by their cumulated normalized fragment intensities (*nfi*), number of matching ions versus unmatched theoretical ions (*mif*) and precursor accuracy (Δmz, normalized between [0,1]):

$$psm\ score = 0.33 \cdot \Delta mz + 0.33 \cdot mif + 0.33 \cdot nfi$$

In case of discrepancy (precursor with inconsistent *psm* across the elution curve), *psm* scores of similar peptides are summed together and the best peptide is associated to the precursor. Ambiguous spectra assigned to different peptides are discarded.

Describing the system as a Maximum Flow problem



The network flow was built from mixed spectra and spectra acquired from peptide isomers. The objective, finding the maximum flow of the system, is stated as such:

How many ions of peptide isomer does it take to optimally fill the mixed spectrum fragment intensity counts?

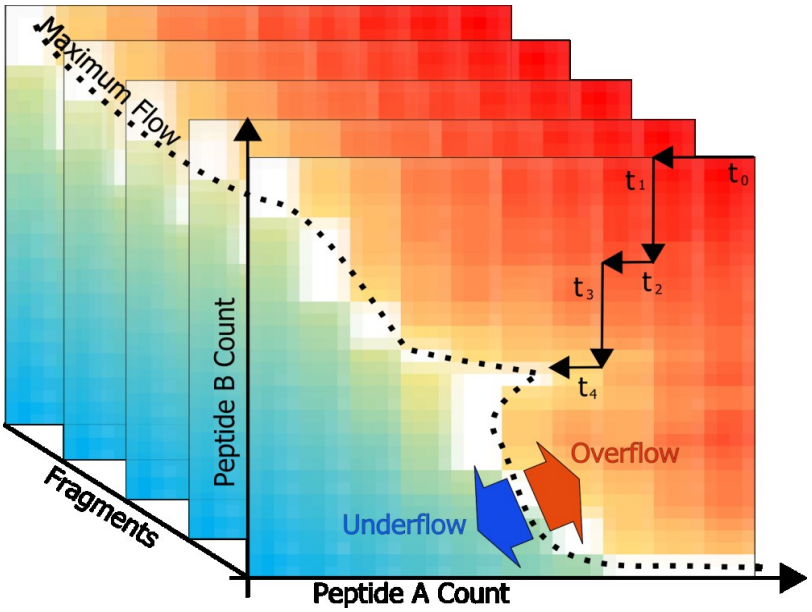
For each synthetic peptide found, normalized fragment intensities (*nfi*) of every spectrum were averaged. The *n* most intense fragments were chosen for each synthetic peptide. Only fragment masses with consistent normalized intensities (standard deviation above 50% of the mean) were considered reproducible enough. Unstable fragments were flagged and ignored in following analyses.

For synthetic peptides of similar masses, these *n* fragments were used to build a list of fragment masses (aMz), both with common and unique fragment masses. The *nfi* of their averaged spectrum was used to associate intensities to a list of fragment masses (aMz_j, one aMz for each potential isomer *j*). A similar list of intensity count (aMz_m) is built for the mixed spectrum *m* of matching precursor mass. For each mixed spectrum *m*, the Maximum Flow is defined by the sum of aMz_j that optimally fills the aMz_m fragment intensity counts (were c_j is the number of ions for the *j* isomer, and aMz_t is the theoretical list being built):

$$aMz_t = \sum_{isomers}^{j}(c_j \cdot aMz_j) \text{ and } aMz_t \leq aMz_m$$

Thus, the network flow can be described as the source being connected to peptide isomers *j* (edges of unknown capacity) which are themselves connected to every selected fragment *aMz* (edge capacity determined by c_j) linking to the sink (edge capacity determined by aMz_j). When the maximum flow is reached, the first set of edges from the source represents the ratio of each peptide isomer found in the mixed spectrum.

Gradient Descent



Because of the dependencies between fragment intensities within a synthetic peptide spectrum, traditional approaches (such as PreFlow Push) could not be used. Instead, a Gradient Descent was implemented.

The Gradient Descent is an iterative approach with each step being closer to the Maximum Flow than the previous iteration. The initial state of the gradient descent (*c* value for each *j* isomer) is computed as the number of times each *j* isomer can fit in the mixed spectrum.

$$c_j = aMz_m / aMz_j$$

This puts the system in an Overflow state (some fragment intensities within the theoretical aMz (aMz_t) are over their capacity). There might also be Underflow (fragments of aMz_t lower than their capacity in aMz_m).

Iteratively, until there is no Overflow in the network, the *j* isomer with the highest Overflow/Underflow score (objective function) has its c_j value reduced (ions are removed one unit at a time). Ambiguous states are resolved randomly.

Deconvolution

In PeptidAce, the peak tops and areas under the curve of precursors are both reported for quantification purposes. Iso-PeptidAce also reports area under the curve of the individual isomers populating a mixed precursor elution profile. For each isomer *j*, the number of ions populating the spectrum (deduced by each Gradient Descent result) is used to compute the number of ions per millisecond flowing in the system at different time points. The resulting, deconvoluted curve is traced through linear interpolation, which is in turn used to compute the area under the curve (ac_j) for that peptide.

Quantification

If peptide intensities are properly normalized based on the synthetic sample runs, it is possible to use this method for absolute quantification of peptides. When concentrations of spiked synthetic peptides are known, Iso-PeptidAce's area under the curve can be converted into absolute peptide counts. Assuming that synthetic peptide samples were processed exactly as the mixed spectrum samples, the number of peptides *j* in the sample *m* (pc_{mj}) is defined by the area under the curve (ac_{mj}) normalized by the synthetic peptide sample *s* information (peptide count in sample pc_{sj}, and area under the curve ac_{sj}).

$$pc_{mj} = ac_{mj} \cdot \left(pc_{sj} / ac_{sj}\right)$$

Results

Experimental design

To assess the precision of the approach, we designed an experiment to test deconvolution across different peptide isomer concentration ratios. Because of its high level of complexity, characterization of acetylation sites in the Histone H4 peptide **[1]** was selected. The software method was tested for the mono, di and tri-acetylated versions of the peptides while all other lysine sites were propyonylated. The next table depicts the concentrations normalized using the equimolar run (observed / expected):

	5fMol - 80fMol	10fMol - 80fMol	20fMol - 80fMol	40fMol - 80fMol	80fMol - 80fMol	160fMol - 80fMol	320fMol - 80fMol
Peptide 2	0.0 / 5	7.4 / 10	19.8 / 20	41.9 / 40	80.0 / 80	168.3 / 160	329.8 / 320
Peptide 3	72.7 / 80	74.0 / 80	72.9 / 80	73.6 / 80	80.0 / 80	74.8 / 80	83.7 / 80
Peptide 4	8.2 / 5	13.0 / 10	19.0 / 20	45.4 / 40	80.0 / 80	189.6 / 160	362.0 / 320
Peptide 5	75.7 / 80	75.9 / 80	78.1 / 80	74.8 / 80	80.0 / 80	72.9 / 80	81.9 / 80
Peptide 6	2.8 / 5	6.3 / 10	14.6 / 20	30.5 / 40	80.0 / 80	168.6 / 160	319.7 / 320
Peptide 7	53.0 / 80	60.6 / 80	67.8 / 80	69.6 / 80	80.0 / 80	74.1 / 80	82.2 / 80
Peptide 8	0.0 / 5	0.4 / 10	7.8 / 20	32.0 / 40	80.0 / 80	382.8 / 160	884.9 / 320
Peptide 9	215.8 / 80	177.4 / 80	150.4 / 80	165.4 / 80	80.0 / 80	44.4 / 80	27.1 / 80
Peptide 10	5.4 / 5	11.7 / 10	21.8 / 20	36.6 / 40	80.0 / 80	133.4 / 160	228.8 / 320
Peptide 11	78.5 / 80	81.9 / 80	82.8 / 80	84.6 / 80	80.0 / 80	66.6 / 80	66.4 / 80
Peptide 12	3.0 / 5	7.5 / 10	17.9 / 20	35.3 / 40	80.0 / 80	162.2 / 160	303.5 / 320
Peptide 13	93.8 / 80	92.4 / 80	90.4 / 80	95.1 / 80	80.0 / 80	67.8 / 80	71.6 / 80
Peptide 14	7.0 / 5	11.1 / 10	21.9 / 20	38.8 / 40	80.0 / 80	162.8 / 160	332.8 / 320
Peptide 15	79.5 / 80	81.8 / 80	81.0 / 80	78.8 / 80	80.0 / 80	73.7 / 80	65.2 / 80

Automatic identification of synthetic peptides

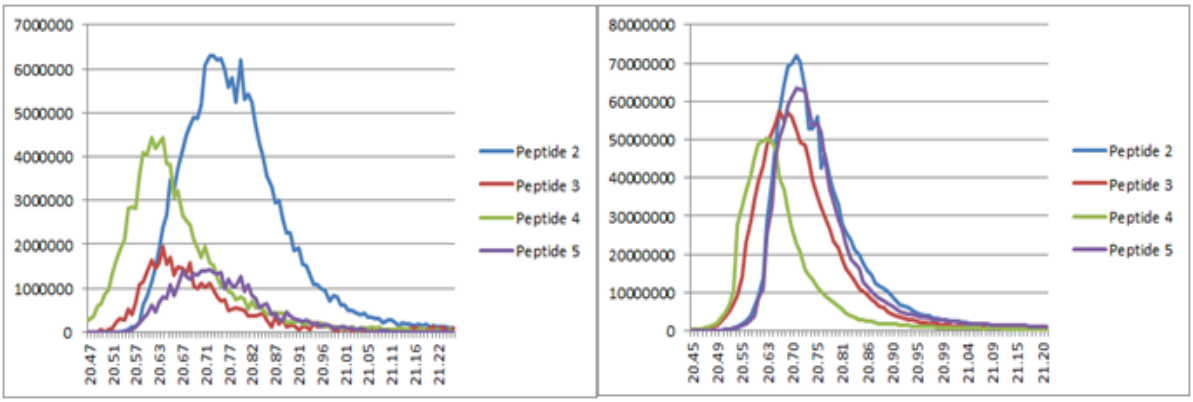
Automatic identification of peptide sequence and modification did not yield in any particular error. However, we did find impurities in the synthesized peptide samples. For instance, incomplete or non-acetylated forms of the peptides were seen in small quantities (less than 0.2% of total intensity). Cross contamination of peptide forms was tested by running Iso-

PeptidAce on sample containing only one peptide form. The resulting contamination intensity detected were too low to be worth accounting for.

Consistency of the Gradient Descent approach

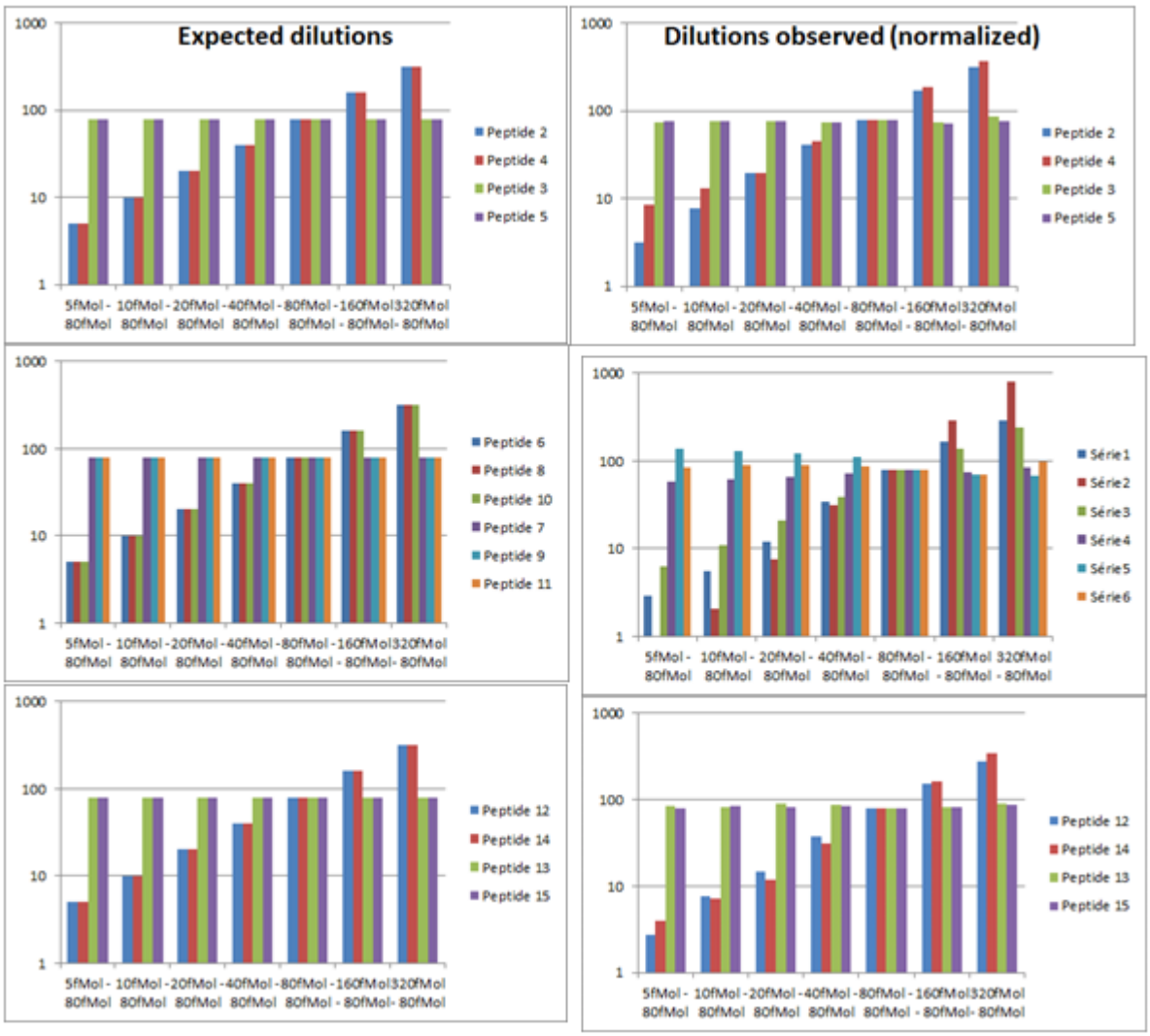
Despite the fact that each gradient descent is done independently of the others, the results are surprisingly consistent across the mixed elution profile. As shown in the figure below, the deconvoluted elution profile for the mixed peptide spectrum (left panel) is comparable to the profile obtained for the individual peptide spectrum (right panel). Both profiles have comparable elution time range (x-axis) and intensityvalues (y-axis)..

Elution profiles of mixed peptide spectra (left panel) and the individual synthetic peptide spectrum (right panel) after deconvolution by the gradient descent approach



Expected versus Observed Concentration

The precision of the method is shown in the next figure, where observed peptide concentration ratios are plotted against their expected values. The average difference between observed and expected concentration varies from 1.7% to 4.3% with a maximum recorded error of 35% (Di-Acetylation, 320fMol sample). These numbers are well within the error estimates of expected peptide count per sample (**80% purity** per synthetic peptide plus errors introduced during dilutions).



Conclusion

Iso-PeptidAce was successfully tested for three types of peptide positional isomers on a Q-Exactive. Peptide fragmentation was done in HCD, but in theory should work equally well with other fragmentation methods. All mixed spectra samples were resolved with accuracies below the errors that were expected from sample preparation and manipulation. Furthermore, the included user interface makes the tool readily accessible to all scientists, not just bioinformaticians.

About PeptidAce

PeptidAce is an open source C# library that includes a lot of tools for peptide spectrum matching. It was based on the Morpheus open source project, with the ultimate aim of improving no-enzyme in-silico digestion of huge, unannotated transcriptome databases. It excels at identifying non-tryptic peptides and can identify peptides with non-standard fragmentation patterns. It can also be used for deep spectrum annotation, matching peaks to all types of fragments and modifications. The source code and the documentation of the library are accessible from the github page:
https://github.com/olivierlizotte/PeptidAce

About Iso-PeptidAce

The user interface for Iso-PeptidAce is built using .Net 4. It is compatible with 64 bit versions of Windows 7 (and up). To work with Thermo instruments, MsFileReader (64 bit version) must be installed first. Iso-PeptidAce installer can be downloaded for free from IRIC's Proteomic Platform website:
http://proteomics.iric.ca/tools/Iso-PeptidAce

3.9. References

1. Unnikrishnan, A., Gafken, P. R. & Tsukiyama, T. Dynamic changes in histone acetylation regulate origins of DNA replication. *Nat Struct Mol Biol* 17, 430-437, (2010).

2. Zentner, G. E. & Henikoff, S. Regulation of nucleosome dynamics by histone modifications. *Nat Struct Mol Biol* 20, 259-266, (2013).

3. Filippakopoulos, P. & Knapp, S. The bromodomain interaction module. *FEBS Lett* 586, 2692-2704, (2012).

4. Filippakopoulos, P. *et al.* Histone recognition and large-scale structural analysis of the human bromodomain family. *Cell* 149, 214-231, (2012).

5. Smith, C. M. *et al.* Mass spectrometric quantification of acetylation at specific lysines within the amino-terminal tail of histone H4. *Analytical Biochemistry* 316, 23-33, (2003).

6. Feller, C., Forné, I., Imhof, A. & Becker, Peter B. Global and Specific Responses of the Histone Acetylome to Systematic Perturbation. *Molecular Cell* 57, 559-571, (2015).

7. Sidoli, S. *et al.* Middle-down hybrid chromatography/tandem mass spectrometry workflow for characterization of combinatorial post-translational modifications in histones. *PROTEOMICS* 14, 2200-2211, (2014).

8. Yuan, Z.-F. *et al.* EpiProfile Quantifies Histone Peptides With Modifications by Extracting Retention Time and Intensity in High-resolution Mass Spectra. *Molecular & Cellular Proteomics* 14, 1696-1707, (2015).

9. Bern, M. *et al.* Deconvolution of mixture spectra from ion-trap data-independent-acquisition tandem mass spectrometry. *Anal Chem* 82, 833-841, (2010).

10. Wang, J., Bourne, P. E. & Bandeira, N. MixGF: spectral probabilities for mixture spectra from more than one peptide. *Mol Cell Proteomics* 13, 3688-3697, (2014).

11. Zhang, B., Pirmoradian, M., Chernobrovkin, A. & Zubarev, R. A. DeMix workflow for efficient identification of cofragmented peptides in high resolution data-dependent tandem mass spectrometry. *Mol Cell Proteomics* 13, 3211-3223, (2014).

12. Courcelles, M., Bridon, G., Lemieux, S. & Thibault, P. Occurrence and detection of phosphopeptide isomers in large-scale phosphoproteomics experiments. *J Proteome Res* 11, 3753-3765, (2012).

13. Ahuja, R. K., Magnanti, T. L. & Orlin, J. B. *Network Flows: Theory, Algorithms, and Applications*. (PRENTICE HALL, 1993).

14. Cho, D.-Y., Kim, Y.-A. & Przytycka, T. M. Chapter 5: Network Biology Approach to Complex Diseases. *PLoS Comput Biol* 8, e1002820, (2012).

15. Albert, R. Scale-free networks in cell biology. *Journal of Cell Science* 118, 4947-4957, (2005).

16. Dutt, S., Dai, Y., Ren, H. & Fontanarosa, J. in *Bioinformatics and Computational Biology* Vol. 5462 *Lecture Notes in Computer Science* (ed Sanguthevar Rajasekaran) Ch. 21, 211-223 (Springer Berlin Heidelberg, 2009).

17. SNYMAN, J. A. *Practical Mathematical Optimization: An Introduction to Basic Optimization Theory and Classical and New Gradient-Based Algorithms*. (Springer Publishing, 2005).

18. Tang, H. *et al.* Multiplexed Parallel Reaction Monitoring Targeting Histone Modifications on the QExactive Mass Spectrometer. *Analytical Chemistry* 86, 5526-5534, (2014).

19. Li, Q., Burgess, R. & Zhang, Z. All roads lead to chromatin: multiple pathways for histone deposition. *Biochim Biophys Acta* 1819, 238-246 (2013).

20. Zhou, H., Madden, B. J., Muddiman, D. C. & Zhang, Z. Chromatin assembly factor 1 interacts with histone H3 methylated at lysine 79 in the processes of epigenetic silencing and DNA repair. *Biochemistry* 45, 2852-2861, (2006).

21. Kelly, T. J., Qin, S., Gottschling, D. E. & Parthun, M. R. Type B histone acetyltransferase Hat1p participates in telomeric silencing. *Mol Cell Biol* 20, 7051-7058 (2000).

22. Qin, S. & Parthun, M. R. Histone H3 and the histone acetyltransferase Hat1p contribute to DNA double-strand break repair. *Mol Cell Biol* 22, 8353-8365 (2002).

23. Sobel, R. E., Cook, R. G., Perry, C. A., Annunziato, A. T. & Allis, C. D. Conservation of deposition-related acetylation sites in newly synthesized histones H3 and H4. *Proc Natl Acad Sci U S A* 92, 1237-1241 (1995).

24. Rusche, L. N., Kirchmaier, A. L. & Rine, J. The establishment, inheritance, and function of silenced chromatin in Saccharomyces cerevisiae. *Annu Rev Biochem* 72, 481-516, (2003).

25. Arts, J. *et al.* JNJ-26481585, a Novel "Second-Generation" Oral Histone Deacetylase Inhibitor, Shows Broad-Spectrum Preclinical Antitumoral Activity. *Clinical Cancer Research* 15, 6841-6851, (2009).

26. Stühmer, T. *et al.* Preclinical anti-myeloma activity of the novel HDAC-inhibitor JNJ-26481585. *British Journal of Haematology* 149, 529-536, (2010).

27. Tong, W.-G. *et al.* Preclinical antileukemia activity of JNJ-26481585, a potent second-generation histone deacetylase inhibitor. *Leukemia Research* 34, 221-228, (2010).

28. Deleu, S. *et al.* The effects of JNJ-26481585, a novel hydroxamate-based histone deacetylase inhibitor, on the development of multiple myeloma in the 5T2MM and 5T33MM murine models. *Leukemia* 23, 1894-1903 (2009).

29. Venugopal, B. *et al.* A Phase I Study of Quisinostat (JNJ-26481585), an Oral Hydroxamate Histone Deacetylase Inhibitor with Evidence of Target Modulation and Antitumor Activity, in Patients with Advanced Solid Tumors. *Clinical Cancer Research* 19, 4262-4272, (2013).

30. Carol, H. *et al.* Initial testing (stage 1) of the histone deacetylase inhibitor, quisinostat (JNJ-26481585), by the Pediatric Preclinical Testing Program. *Pediatric Blood & Cancer* 61, 245-252, (2014).

31. Gallien, S. *et al.* Targeted Proteomic Quantification on Quadrupole-Orbitrap Mass Spectrometer. *Molecular & Cellular Proteomics* 11, 1709-1723, (2012).

32. Chen, Y. *et al.* Quantitative acetylome analysis reveals the roles of SIRT1 in regulating diverse substrates and cellular pathways. *Mol Cell Proteomics* 11, 1048-1062, (2012).

33. Downey, M. *et al.* Acetylome profiling reveals overlap in the regulation of diverse processes by sirtuins, gcn5, and esa1. *Mol Cell Proteomics* 14, 162-176, (2015).

34. Weinert, B. T. *et al.* Proteome-wide mapping of the Drosophila acetylome demonstrates a high degree of conservation of lysine acetylation. *Sci Signal* 4, ra48, (2011).

35. Shechter, D., Dormann, H. L., Allis, C. D. & Hake, S. B. Extraction, purification and analysis of histones. *Nat. Protocols* 2, 1445-1457 (2007).

36. Abshiru, N. *et al.* Chaperone-mediated acetylation of histones by Rtt109 identified by quantitative proteomics. *Proteomics* 81, 80-90 (2013).

37. Zeng, L., Zhang, Q., Gerona-Navarro, G., Moshkina, N. & Zhou, M.M. Structural Basis of Site-Specific Histone Recognition by the Bromodomains of Human Coactivators PCAF and CBP/p300. *Structure* 16, 643-652 (2008).

38. Choudhary, C. et al. Lysine Acetylation Targets Protein Complexes and Co-Regulates Major Cellular Functions. *Science* 325, 834-840 (2009).

39. Wu, S.Y. & Chiang, C.M. Crosstalk between sumoylation and acetylation regulates p53-dependent chromatin transcription and DNA binding, Vol. 28. (2009).

40. Joubel, A., Chalkley, R.J., Medzihradszky, K.F., Hondermarck, H. & Burlingame, A.L. Identification of New p53 Acetylation Sites in COS-1 Cells. *Molecular & Cellular Proteomics* 8, 1167-1173 (2009).

41. Sen, N., Kumari, R., Singh, Manika I. & Das, S. HDAC5, a Key Component in Temporal Regulation of p53-Mediated Transactivation in Response to Genotoxic Stress. *Molecular Cell* 52, 406-420 (2013).

42. Li, A.G. et al. An Acetylation Switch in p53 Mediates Holo-TFIID Recruitment. *Molecular Cell* 28, 408-421 (2007).

43. Tang, Y., Zhao, W., Chen, Y., Zhao, Y. & Gu, W. Acetylation Is Indispensable for p53 Activation. *Cell* 133, 612-626 (2008).

44. Boyes, J., Byfield, P., Nakatani, Y. & Ogryzko, V. Regulation of activity of the transcription factor GATA-1 by acetylation. *Nature* 396, 594-598 (1998).

45. Lamonica, J.M., Vakoc, C.R. & Blobel, G.A. Acetylation of GATA-1 is required for chromatin occupancy, Vol. 108. (2006).

46. Lamonica, J.M. et al. Bromodomain protein Brd3 associates with acetylated GATA1 to promote its chromatin occupancy at erythroid target genes. *Proceedings of the National Academy of Sciences* 108, E159-E168 (2011).

47. Schröder, S. et al. Two-pronged Binding with Bromodomain-containing Protein 4 Liberates Positive Transcription Elongation Factor b from Inactive Ribonucleoprotein Complexes. *Journal of Biological Chemistry* 287, 1090-1099 (2012).

48. Sartorelli, V. et al. Acetylation of MyoD Directed by PCAF Is Necessary for the Execution of the Muscle Program. *Molecular Cell* 4, 725-734 (1999).

49. Wei, L., Jamonnak, N., Choy, J., Wang, Z. & Zheng, W. Differential binding modes of the bromodomains of CREB-binding protein (CBP) and p300 with acetylated MyoD. *Biochemical and Biophysical Research Communications* 368, 279-284 (2008).

50. Kessner, D., Chambers, M., Burke, R., Agus, D. & Mallick, P. Bioinformatics 24, 2534-2536 (2008).

CHAPTER 4: Unraveling site-specific and combinatorial acetylations of histones using high-resolution mass spectrometry in HAT and HDAC mutants of fission yeast

Nebiyu Abshiru[1,2,4], Roshan Elizabeth Rajan[2,3,4], Alain Verreault[1,3], Pierre Thibault[1,2]

[1]Department of Chemistry, Université de Montréal, P.O. Box 6128, Station centre-ville, Montréal,
Québec, H3C 3J7, Canada.
[2]Institute for Research in Immunology and Cancer, Université de Montréal, C.P. 6128, Succursale centre
ville, Montréal, Québec, H3C 3J7, Canada.
[3]Molecular Biology Programme, Université de Montréal, P.O. Box 6128, Station centre-ville, Montréal,
Québec, H3C 3J7, Canada.

Manuscript: In revision

Author contributions:

N.A. designed and carried out the proteomics experiment. R.E.R. isolated histones from control and mutant fission yeast strains. N.A. analyzed data, wrote the first draft of the manuscript, and generated figures. P.T and A.V. conceived and led the project. N.A., R.E.R., P.T. and A.V. wrote the final draft of the manuscript.

4.1. Abstract

Histone acetylation is regulated by the opposing activities of two enzyme families known as histone acetyltransferases (HATs) and histone deacetylases (HDACs). The dynamic interplay between these two enzymes regulates global histone acetylation levels. This process is crucial in the maintenance of chromatin reorganization and gene expression as dysregulation of either enzymatic activity can cause severe defects in the genome integrity of cells. The fission yeast, *Schizosaccharomyces pombe,* genome encodes for few HATs and HDACs. The substrate specificities of most of these enzymes were previously determined by antibody-based methods whereby only a limited number of sites were tested. Even less is known about their role in the establishment and regulation of combinatorial acetylation codes, which are valuable in mediating functional chromatin-effector protein interactions. Here, we employed a targeted quantitative mass spectrometry (MS)-based approach to conduct an in-depth analysis of site-specific and combinatorial acetylations of histones H3 and H4 in *S. pombe* lacking major HAT and HDAC activities. Our analysis identified several known and novel HAT and HDAC target sites. We also found that the *S. pombe* Class I HDAC known as Clr6 has a striking preference for combinatorial H3 and H4 acetylation motifs rather than the individual marks. In addition, our study uncovered a few short-range crosstalks between acetylated lysines and di- or tri-methylated H3K9. We anticipate that the results from this study will greatly improve our current understanding of the mechanisms involved in HAT- and HDAC-mediated gene regulation and heterochromatin assembly.

4.2. Introduction

Histone posttranslational modifications (PTMs) are known to play a crucial role in the maintenance of higher-order chromatin structures and transcription regulation.[1-3] Numerous studies conducted during the last decade have identified over fifteen different types of PTMs on the core histones H2A, H2B, H3 and H4.[4] The majority of these PTMs are localized on the unstructured N-terminal histone tail which extends outside the nucleosome core.

Acetylation is one of the best characterized histone PTM. This modification mediates changes in higher-order chromatin structures. For example, acetylation of histone H4 at lysine 16 (H4K16) was shown to inhibit chromatin condensation and formation of higher order chromatin fiber.[5] Histone acetylation is also involved in the regulation of the transcriptional activity of genes.[6,7] Acetylation of histones is regulated by the opposing activities of two major enzyme families known as histone acetyltransferases (HAT) and histone deacetylases (HDACs).[8] HATs specifically catalyze the acetylation of lysine residues, whereas HDACs specialize in removing the acetyl marks. Based on their catalytic domains and sequence homology, eukaryotic HATs are grouped into three major families[9,10]: i) Gcn5-related N-acetyltransferases (GNATs), named after the founding member Gcn5; ii) the MYST HATs, named for the founding members of this family: Morf, Ybf2, Sas2 and Tip60; and iii) CBP/P300. HAT activity generally leads to destabilization of the nucleosome structure, and hence their function is often associated with transcriptional activation and increased gene expression.[11] In agreement with this notion, many HATs are known to be recruited at promoter and transcriptionally active regions via interaction with gene-specific activator proteins.[12,13] For example, in budding yeast, the HAT Gcn5 is recruited to the promoter regions of active genes via interaction with the activator protein Gcn4, which leads to acetylation of histone H3 at positions K9 and K14.[14] Conversely, the removal of acetyl marks from histones by HDACs generally results in chromatin condensation. This process is historically associated with transcriptional repression and heterochromatin formation[15,16], although this has been demonstrated only in a limited number of cases. HDACs are grouped into four major classes[17,18]: Class I: HDAC-1, -2, -3 and -8; Class-II: HDAC-4, -5, -6, -7, -9 and -10; Class III: sirtuins or SIRT 1–7; and Class IV: HDAC11. HDACs carry out their function by associating with repressor complexes. One example is the Sin3 multiple-protein co-repressor complex which interacts with the HDAC Rpd3 in *S. cerevisiae*[19] and HDAC1/2 in humans[20]. Most HDAC complexes do not directly bind to DNA; instead they are targeted

to specific genomic loci via interaction with other DNA-binding proteins. For example, in *S. cerevisiae*, the Sin3/Rpd3 complex interacts with the DNA-binding protein called Ume6 to target the deacetylation of histone H4K5 and repress transcription of the adjacent genes.[21,22]

The genome of the fission yeast, *Schizosaccharomyces pombe,* encodes for few HATs and HDACs. It contains two MYST family HATs: Mst1 and Mst2. Mst1, the likely orthologue of *S. cerevisiae* Esa1 and human Tip60, is an essential gene and is implicated in a variety of functions including DNA damage response, and in centromere assembly.[23] In contrast, Mst2 is a non-essential gene required in negative regulation of telomere silencing.[24] The substrate specificity of Mst2 has been previously examined using Western analysis.[25] Histones isolated from Mst2 mutant strains exhibited a reduced level of H3K14ac, whereas no significant change was observed in the levels of H3K9ac and H3K18ac. Its specificity toward histone H4 and other H3 sites (e.g. K23 and K27) was not reported. Gcn5 is the only GNAT family HAT that is so far identified in *S. pombe*. This HAT is conserved in both *S. cerevisiae* and humans and is required in transcriptional activation of genes.[26] The substrate specificity of *S. pombe* Gcn5 was so far examined for only three sites on histone H3 (K9ac, K14ac and K18ac).[25] All three sites were significantly reduced in strains lacking the Gcn5 activity.

Moreover, the *S. pombe* genome encodes for a total of six deacetylases representing three HDAC classes: Clr6 and Hos2 are (Class I); Clr3 (Class II); and Sir2, Hst2 and Hst4 (Class III). Clr6, an orthologue of *S. cerevisiae* Rpd3 and human HDAC1/2, is an essential gene required in genome-wide heterochromatin silencing.[27,28] According to previous Western analysis, loss of Clr6 activity induces hyperacetylation of histones H3 and H4 at positions H3K9, H3K14, H4K5, H4K12 and H4K16.[29] However, its activity toward many other sites (e.g. H3K18, H3K23, H3K27 and H4K8) was not tested. The other Class I HDAC, Hos2, was shown to primarily target H4K16ac at highly active genes.[29] Its activity led to increased gene expression at coding regions, a role which is contrary to the classical function of HDACs.[30] The only *S. pombe* Class II HDAC, Clr3, was previously shown to act cooperatively with the Class III HDAC Sir2 to promote gene silencing at mating-type loci, centromeres and telomeres.[28,31] The primary targets for Clr3 and Sir2 were found to be H3K14ac and H3K9ac, respectively. The Class III HDACs Hst2 and Hst4 were shown to promote gene repression in various cellular processes.[32] Hst4 was recently implicated in the regulation of H3K56ac in cell cycle progression and DNA-damage tolerance.[33]

Previous studies employed traditional biochemical approaches such as Western analysis, and ChIP to define the substrate specificity of *S. pombe* HATs and HDACs.[25,29,31,34] These approaches, however, suffer from several limitations. Site-specific antibody detection requires one specific antibody for each site that needs to be studied, and hence lacks the capability to map combinatorial modifications. Recent findings show that antibodies exhibit promiscuity when two or more PTMs are placed too close together.[35] In agreement with this finding, many antibodies raised against histone H4 acetyl-lysines preferentially detected poly-acetylated epitopes rather than their specific target site.[36] Antibodies are also limited in their abilities to provide stoichiometries of acetylation, which has proven valuable in measuring and monitoring acetylation dynamics in various biological processes.[37,38] Recently, we developed a method that combines high resolution LC-MS/MS with bioinformatics tools, that enables quantification of histone peptides harboring multiple co-existing acetylation marks .[39] This method allowed characterization of acetylation patterns in total histone extracts and in affinity-purified histones bound to chromatin-modifying complexes. In the present work, we employed this approach for in-depth quantitative and comprehensive analysis of the substrate specificity of *S. pombe* HATs Mst2 and Gcn5, and HDACs Clr6, Clr3, Sir2, and Hos2. The substrate specificity of these enzymes toward core histones H3 and H4 was examined. Our results are generally in agreement with previous findings. However, our method also allowed us to unravel several novel substrate specificity sites. The HDAC Clr6 exhibited a striking preference for combinatorial acetylation motifs over individual acetyl-lysines on histones H3 and H4. We also characterized a few short-range PTM crosstalks that took place in response to HDAC depletions.

4.3. Methods

4.3.1. Strains and media

The *S. pombe* HDAC strains used in this study are listed in Table S4.1A. Strains were grown at 25˚C in YE5S medium (Yeast Extract Supplemented with 5 amino acids). [40] We confirmed the phenotypic expression of antibiotic and auxotrophic markers of all strains by PCR and growth on appropriate plates. The mating types of the strains were tested by PCR.

4.3.2. Histone extraction

Histones were extracted by adapting a protocol used for acid extraction of histones from budding yeast.[41] Starting from saturated overnight cultures, strains were grown at 25°C in 400 ml YE5S medium until they reached $O.D_{595}$ 1.0 -1.5. Cells were harvested and washed sequentially with 100 ml water, 25 ml DTT buffer (0.1M Tris-HCl pH 9.5, 10 mM DTT) and 25 ml Spheroplasting buffer (20mM HEPES (NaOH) pH 7.4, 1.2 M Sorbitol). The wet cell pellets were weighed, resuspended in 12.5 ml Spheroplasting buffer containing Zymolyase 100T (5 mg/g of cell pellet). The pellets were gently resuspended in Nuclear Isolation Buffer, NIB (15 mM 2-(*N*-morpholino) ethane sulfonic acid (KOH) pH 6.6, 60 mM KCl, 15 mM NaCl, 5 mM $MgCl_2$, 1 mM $CaCl_2$, 0.25 M Sucrose, 0.8% v/v Triton X-100, 1 mM PMSF). The NIB wash was repeated twice followed by Wash 'A' (10 mM Tris-HCl pH 8.0, 75 mM NaCI, 0.5% v/v NP 40, 30 mM NaButyrate , 1 mM PMSF) and Wash 'B' (10 mM Tris-HCl pH 8.0, 0.4 M NaCI, 30 mM NaButyrate,1 mM PMSF). The nuclear pellet was resuspended in 0.25 M HCl for 15 min**,** which extracts histones and other acid-soluble proteins. Trichloroacetic acid (TFA) was added to a final concentration of 20% v/v to the supernatant that contains histones. The histone pellet was washed with acetone, air-dried and resuspended in water for offline purification of intact histones by RP-HPLC.

Histone fractionation by RP-HPLC

Bulk histones from both the WT and mutant strains were fractionated into individual core histones using an Agilent narrow-bore Zorbax C8 reverse-phase column (2.1 x 150 mm, 5µm, 300Å) on an Agilent 1200 HPLC system. The mobile solvent was comprised of solvent A, which is aqueous 0.1% TFA (Sigma) and solvent B is 0.1% TFA in 100% acetonitrile (ACN). Approximately 5 µg of bulk histones were loaded onto the C8 column at a flow rate of 150 µl/min. Histones were eluted from the column using a gradient of 5 – 80% solvent B in 60

minutes. Fractions were collected in a 96-well plate every half minute. Fractions containing histones H3 or H4 were pooled and dried in a SpeedVac concentrator.

4.3.3. Propionylation and tryptic digestion of histone proteins

In-solution tryptic digestion of core histones was performed as previously described with minor modifications.[42] Briefly, a total of 2 μg of the HPLC purified histones H3 and H4 were subjected to two rounds of propionylation by adding 200 μl of freshly prepared 2:1 (v/v) water: propionic anhydride (Sigma) mixture and vortexing the mixture for 30 minutes at room temperature. After each propionylation reaction, the samples were left to dry in a Speed-vac at 4^{o}C. To prepare the histone samples for digestion a buffer exchange was made by adding to each sample 200 μl of 50 mM ammonium bicarbonate and drying one final round in the Speed-vac. The dried histone samples were resuspended back into the digestion buffer and vortexed for 5 minutes in order to re-dissolve the proteins. Our digestion solution was prepared by adding 200 μl of 50 mM ammonium bicarbonate in a Vial containing 20 μg of lyophilized trypsin (Promega). About 0.5 μl of this solution was added to each histone sample and digested overnight at 37^{o}C. After digestion, samples were dried completely in a SpeedVac and then resuspended in 0.2% formic acid prior to LC-MS/MS analyses.

4.3.4. LC-MS/MS analysis of histone peptides

Chromatographic separation of histone peptides was achieved on an EASY-nLC II system (Thermo scientific). Solvent A was comprised of 0.2% formic acid (FA) in milli-Q water, and Solvent B contained 0.2% FA in acetonitrile (ACN). A total of 2 μg of the histone digest was loaded and desalted on an in-house packed Jupiter C18 (3 μm particles, Phenomenex) trap column (4 mm length, 360 μm i.d.) using Solvent A at a flow rate of 10 μl/min. After 5 minutes of desalting, the peptides were eluted onto an in-house packed C18 analytical column (18 cm length, 150 μm i.d.) that is connected directly to the electrospray source ESI of a Q-Exactive Plus instrument (Thermo scientific). The peptides were eluted into the MS with a linear ramp of Solvent B from 5 to 60% at a flow rate of 600 nl/min over 60 min. MS data were acquired for two to three technical replicates per sample. The MS instrument was operated in positive ion mode, and capillary voltage of 1.6 kV. MS scans were acquired in the Orbitrap analyzer over the range of 250 – 1500 *m/z* at a resolution of 70,000 and automatic gain control target value of 1.0 x 10^{6}. Data were acquired using the parallel reaction monitoring (PRM) instrument method. An inclusion list containing *m/z*, charge state and collision energy (CE) values of H3 and H4 peptides was used to trigger MS/MS acquisition

(Table S4.1B). Every precursor ion found in the inclusion list was automatically selected for fragmentation in the HCD cell at a normalized CE setting of 27. The fragments were analyzed in the Orbitrap at a resolution of 35,000 and a target value of 5.5 x 10^5. The dynamic exclusion setting was disabled in order to acquire multiple MS/MS spectra per peptide.

4.3.5. Data analysis

MS data were analyzed manually using Thermo Xcalibure Qual Browser software (Xcalibur™ version 2.2 SP1) and an in-house built algorithm, Iso-PeptidAce. Iso-PeptidAce analysis was performed as previously described (reference, isopeptidace paper). Raw LC-MS/MS files and FASTA sequences of histones H3 and H4 were submitted to Iso-PeptidAce for peptide-spectrum matching and deconvolution of mixed MS/MS spectra of co-eluting Isomers. The following default parameters were used by the algorithm: Precursor mass tolerance: 8 ppm, fragment mass tolerance: 0.05 Da, the minimum number of fragment ions (per isomer) considered for deconvolution: 5, types of fragment ions considered: b and y, digestion enzyme: trypsin, missed cleavages: Iso-PeptidAce uses no-enzyme in-silico protein digestion routine to parse the provided FASTA file for potential matches. PTMs included in peptide-spectrum matching: oxidation of methionine residue (M, 15.9949), phosphorylation of serine, threonine or tyrosine (S/T/Y, 79.9663), deamidation of aspargine and glutamine (N/Q 0.98400.9847), acetylation of lysine (K, 42.0106), propionylation of lysine (K, 56.0262), methylation of lysine (K, 14.0157), di-methylation of lysine (K, 28.0313), tri-methylation of lysine (K, 42.0470), acetylation of protein N-terminus (42.0106). All the PTMs were considered as variable modifications. The output from Iso-PeptidAce is an Excel spreadsheet containing precursor intensity data for each peptide found in our inclusion list. Acetylation site occupancy at a given Lys residue was determined from the ratio of the sum of intensities of peptides bearing the ac-Lys to the sum of intensities of all peptides (modified and unmodified) containing this residue.

4.4. Results

4.4.1. Measurement of H3 and H4 peptides by PRM- LC-MS/MS

The general experimental workflow is shown in Figure 4.1A. Histone proteins were isolated from WT or mutant HAT (Δmst2, Δgcn5, and Δmst2.Δgcn5) and HDAC (Δclr6, Δclr3, Δsir2, and Δhos2) fission yeast strains. Clr6 is essential for *S. pombe* viability[43] and we could not use a strain with a complete deletion of the *clr6*$^{+}$ gene. Instead, we used a *clr6-1* hypomorphic mutant strain that was grown; along with all the other strains, at the semi-permissive temperature of 25°C. HPLC fractionated histones were propionylated to prevent trypsin cleavage at lysine residues that are not acetylated *in vivo*. Near complete propionylation efficiency was achieved for peptides comprising up to two free lysine residues. Peptides eluting from the analytical column were analyzed using a PRM mode whereby an inclusion list containing precursor *m/z*, charge state and collision energy was used to trigger MS/MS acquisition. The PRM approach allows the acquisition of high resolution and accurate mass MS/MS spectra in the orbitrap mass analyzer. The capability of this approach to detect and quantify histone modifications of extremely low in abundance (*e.g.* H3K18 methylation) was recently demonstrated. [44]

Some histone tryptic peptides exist in multiple modified isoforms that can only be partially separated or co-elute from the analytical column. For example, the two isomeric H3 peptides shown in Figure 4.1A (Step 7, XIC panel) - K9acK14un (m/z 500.28+2, ac-acetylated, un – unacetylated)) and K9unK14ac (m/z 500.28+2) eluted from the RP column in the same retention time window. Moreover, these peptides produced composite MS/MS spectra with multiple shared fragment ions (Figure 4.1A, Step 7, MS/MS panel). Here, the intensities of isomer-specific fragments (e.g. fragment m/z 742 for K9acK14pr and 728 for K9prK14ac) were used to determine the relative abundance of each peptide. We investigated changes at five acetylation sites (K9, K14, K18, K23 and K27) and two methylation sites (K9 and K36) on the H3 tail. On histone H4, we examined changes in acetylation of four sites including K5, K8, K12 and K16.

4.4.2. Determination of acetylation site occupancies

In our study, differentially modified peptides that share the same sequence were easily resolved by the reverse-phase (RP) column. For example, the two isobaric H3 peptides K27prK36me2K37pr (m/z 519.9788+3) and K27acK36me3K37pr (m/z 519.9787+3) were chromatographically separated (Figure S4.1). However, acetylated isomers of the H4 peptide

4-GKGGKGLGKGGAKR-17 eluted in the same retention time window and produced composite MS/MS spectra with indistinguishable fragment ions. The capability of Iso-PeptidAce for deconvoluting composite MS/MS spectra of mixtures of up to six isomeric peptides was previous demonstrated. In the current study, we employed this algorithm to deconvolute acetylated isomers in the mutant and wild-type fission yeast histone samples. The main steps in this process are summarized in Figure 4.1B (see reference x for more details, isopeptidace paper). Raw LC-MS/MS files were submitted to Iso-PeptidAce for deconvolution of mixed spectra (steps in the deconvolution process are highlighted in Figure 4.1B, left panel). The amounts of H3 or H4 peptides injected in the different samples were normalized based on the signal responses of reference peptide H3: 41-YRPGTVALR-49 (m/z 344.8698, 3+; m/z 516.8010, 2+) or H4: 46-ISALVYEETR-55 (m/z 618.8270, 2+).

In the fission yeast samples, the H4 peptide 4-GKGGKGLGKGGAKR-17 was present in five different isoforms (depending on the number of acetyl-lysines): un-acetylated (0Ac), mono- (1Ac), di- (2Ac), tri- (3Ac), and tetra-acetylated (4Ac). The mono-, di, and tri-acetylated groups were each comprised of four, six and four isomeric peptides that elute in the same retention time window from the RP column (Figure 4.1C, panel 'Before deconvolution'). Meanwhile, deconvolution by Iso-PeptidAce allowed us to obtain reconstructed peak profiles for the individual peptide within each isomeric group (Figure 4.1C, panel 'After deconvolution'). The normalized intensities of the reconstructed peaks were used in determining acetylation site occupancies.

4.4.3. Analysis of global histone acetylation in HAT mutants

To assess the effects of HAT and HDAC depletions on the global H3 or H4 acetylations, the normalized peak intensities of all the acetylated peptides (including all charge states) were summed together. The global changes in acetylation are presented as the abundance fold change (FC) ratios of the summed intensities in the mutant relative to the wild-type samples (hereafter referred to as 'control, Ctl'). As shown in Figure 4.2A (also Table S4.2A), no significant change was observed in the global acetylation level of histone H4 in the Gcn5 and/or Mst2 mutants. In contrast, we observed a striking 14-fold decrease in the acetylation level of H3 in strains lacking both Gcn5 and Mst2. Depletion of the individual HAT had a relatively minor effect on the total H3 acetylation.

Among the HDAC mutants, the strain lacking Clr6 deacetylase induced the most significant increase in the global acetylation of both H3 and H4 (Figure 4.2B and Table S4.2B). We

observed up to 4-fold increase in total H4 acetylation and up to 2-fold in H3 acetylation relative to the control. Strains lacking the Hos2 deacetylase also exhibited up to 2-fold increase in the acetylation of both histones. However, we did not see any significant change in global acetylation in Sir2 and Clr3 mutants.

To further investigate the effects of Clr6 and Hos2 depletions on histone H4, we determined the mutant-to-control abundance fold change ratios for the five different isoforms of the peptide 4-GKGGKGLGKGGAKR-17. Our analysis showed that, in the Clr6 mutant, nearly 19% of the H4 peptide is present in the tetra-acetylated form (H4_4Ac), which is a 6-fold increase compared to the 3% detected in the control sample (Figure 4.2C and Table S4.3 A-B). The levels of the di- (H4_2Ac) and tri-acetylated (H4_3Ac) isoforms were also elevated by up to 2- and 5-fold, respectively. The mono-ac (H4_1Ac) isoform did not change significantly, whereas the non-acetylated peptide (H4_0Ac) was reduced by about 2-fold. These results are also evident from the comparison of the distributions of the different acetylated isoforms in the extracted ion chromatograms (XICs) illustrated in Figure 4.2D, middle panel; which shows a linear increase in isoform abundance in the order 1Ac, 2Ac, 3Ac, and 4Ac. Similarly, for the Hos2 mutant, the most significant change was observed for the tetra-acetylated isoform, which exhibited a 4-fold increase in abundance relative to the control. Moreover, the levels of the di- and tri-acetylated isoforms increased by up to 2- and 3-fold, respectively. The increased abundances of the multiply acetylated isoforms is also evident from the XICs illustrated in Figure 4.2D, bottom panel. Overall, these results suggest a role of Clr6 and Hos2 in the global deacetylation of both histones H3 and H4.

4.4.4. Gcn5 and Mst2 target several H3 sites in S. pombe

Very few studies have reported the substrate specificity of fission yeast Gcn5 and Mst2. Nugent et al.[25] recently conducted a Western analysis to characterize the site specificity of both HATs. Gcn5 depletion caused reduction in H3K9ac, H3K14c and H3K18ac, whereas Mst2 depletion affected only H3K14ac. This study was, however, limited to only three sites on histone H3 and none on H4, possibly because site-specific antibody detection requires one specific antibody for each additional site that needs to be assayed. Here we employed a quantitative mass spectrometry-based approach to profile changes in acetylation levels of several sites in a single LC-MS/MS run.

In order to examine the site specificity of each HAT we determined the fold change ratios of mutant-to-control acetylation site occupancies for each lysine on the N-terminal tails of H3

and H4. Moreover, to assess changes in the abundance of a particular PTM motif, we determined the mutant-to-control log2 ratios of the normalized intensities of peptides harboring this motif. These ratios are presented as heatmaps in which color intensities depict the extent of increase or decrease in the abundance of a particular motif.

Our results confirmed two of the three previously identified Gcn5 target sites, that is, H3K9ac and H3K18ac. H3K18ac was by far the most affected PTM, which is hardly seen in our data (Figure 4.3A, top panel and Table S4.4 A-B). The two PTM motifs harboring this mark, K18acK23un and K18acK23ac, were equally affected, each showing more than two orders of magnitude reduction in their LC-MS signal response relative to the control (Figure 4.3B and Table S4.4C). Similarly, we observed up to 14-fold reduction in the occupancy of H3K9ac in the Gcn5 mutant. The LC-MS signal responses of the two peptides bearing this PTM motif, K9acK14un and K9acK14ac, were reduced by about 16-fold.

In addition to the previously reported sites, we also identified and quantified a new Gcn5 target site, which is H3K27ac. This PTM mark is the third most affected site with a 7-fold decrease in site occupancy (Figure 4.3A, top panel and Table S4.4 A-B). Moreover, peptide motifs bearing both H3K27ac and a di- or tri-methylation at H3K36 were affected the most, which are marked by a 30-fold reduction in their corresponding LC-MS signal responses relative to the control (Figure 4.3B and Table S4.4C).

However, contrary to previous reports, we did not see a decrease in H3K14ac in the Gcn5 mutant (see Figure 4.3A, top panel). To examine this observation in more detail, we extracted ion chromatograms (XICs) representing the acetylated and non-acetylated forms of the H3 peptide 9-KSTGGKAPR-17 from the raw LC-MS data for the Gcn5 depleted and control samples. The overlayed XIC peaks (Δgcn5/control) for four peptides that are acetylated at K9, K14, or both, and a non-acetylated form are illustrated in Figure 4.3C (panels 1-4). The y-axis in each XIC panel was scaled to the signal response of the most abundant peptide. Based on these analysis, we saw a reduction in K14ac in the Gcn5 only when this mark is present in the di-acetylated motif H3K9acK14ac (Figure 4.3C, panel 3). This motif, however, represents only 1% of the total population of H3 peptides detected with the same sequence. Interestingly, the mono-acetylated motif K9unK14ac (Figure 4.3C, panel 2), which represents a corresponding 30%, did not shown any significant change in abundance in response to Gcn5 depletion. These results suggest the lack of any major activity of Gcn5 toward H3K14.

Our analysis of histone H3 peptides from the Mst2 mutant confirmed the specificity of this HAT for H3K14ac (Figure 4.3A, middle panel and Table S4.4 A-B). The most dramatic change was observed when this mark co-exists with a tri-methylated K9 in the motif K9me3K14ac, which decreased 4-fold in the Mst2 mutant (Figure 4.3B and Table S4.4C). In contrast, the fold-decrease in the mono-acetylated motif K9unK14ac was only about 2-fold. We also identified and quantified a new Mst2 target site, H3K23ac, which exhibited a 4-fold reduction in occupancy in response to Mst2 depletion. Here, the most significant change was observed for the mono-acetylated motif K18unK23ac, suggesting the strong activity of Mst2 toward this site.

Overall, the impact of Mst2 depletion on H3 acetylation was lower when compared to the number and extent of affected sites in the Gcn5 mutant. This difference is clearly observable from the corresponding heatmaps depicted in Figure 4.3B (compare 'Δgcn5' and 'Δmst2' columns). Meanwhile, as expected, when both Gcn5 and Mst2 are depleted simultaneously the occupancies of acetylation at all the five H3 sites were significantly reduced (Figure 4.3A, bottom panel and Table S4.4 A-B). The changes were more dramatic for the di-acetylated motifs K9acK14ac and K18acK23ac mutant (Figure 4.3B and Table S4.4C), which were barely detected in the MS. Overall, in contrast to previous observations, our results suggest a non-overlapping role for Gcn5 and Mst2 in acetylating the H3 tail.

So far, there has been no report on the activities of *S. pombe* Gcn5 and Mst2 toward histone H4. Here, we assessed the activity of both HATs toward four H4 sites, which are K5, K8, K12 and K16. As shown in Figure 4.3D (and Table S4.5 A-B), depletion of Gcn5 and/or Mst2 did not affect the acetylation site occupancy of the H4 sites. To check if the HATs showed motif-specific activity, we determined mutant-to-control abundance ratios for the 15 different acetylated isoforms of the peptide 4-GKGGKGLGKGGAKR-17. We observed a moderate change in the abundance of only few isoforms, for example: up to 2-fold decrease in K5ac in the single mutants and K8/16ac in the double mutant (Figure 4.3E and Table S4.5C). The overall change in abundance for the 15 isoforms was essentially lower than 2-fold, suggesting the lack of activity of Gcn5 and Mst2 toward the H4 tail.

4.4.5. Clr6 is specific for multiple sites on H3

So far, few studies have reported the activity of *S. pombe* HDACs toward histone H3. Using site-specific antibodies raised against acetylated histones Bjerling et. al.[31] and Wiren et. al.[29] characterized the specificity of Clr6, Clr3, Sir2 and Hos2 for K9ac and K14ac in histone H3.

These studies, however, did not report the activity of the HDACs toward many other sites, e.g. H3K18ac, H3K23ac, and H3K27ac. Moreover, the technique employed in these studies allowed the analysis of changes in acetylation on a per-site basis, and thus potential alterations in the combinatorial acetylation motifs were not reported. In the current study, we assessed the activity of the HDACs toward multiple sites on H3 by determining the mutant-to-control fold change ratios of acetylation site occupancies. We also quantified changes in abundances of several acetylation motifs to look for combinatorial modifications that best convey the specificity of the HDACs.

Our analysis confirmed the activity of Clr6 toward the two previously characterized H3 sites (K9ac and K14ac). We observed up to 4-fold increase in the occupancy of acetylation at both sites in response to Clr6 depletion (Figure 4.4A and Supplementary 6A-B). However, analysis of motif-specific changes revealed that Clr6 depletion had very little effect on the singly acetylated motifs K9acK14un or K9unK14ac, which showed on average 2-fold increase in abundance (Figure 4.4B and Table S4.6C). In comparison, we observed a dramatic 13-fold increase in the abundance of the di-acetylated motif K9acK14ac, suggesting the strong preference of Clr6 for the combinatorial motif.

We also report three additional Clr6 target sites on histone H3: K18, K23 and K27. Acetylation site occupancy at each of these sites increased by an average 5-fold in response to Clr6 depletion (Figure 4.4A and Supplementary 6A-B). Similar to the above observation, Clr6 exhibited a strong preference for the di-acetylated motif K18acK23ac, which was elevated by 15-fold in the Clr6 mutant (Figure 4.4B and Table S4.6C). Moreover, loss of Clr6 activity caused a major increase in the occupancy of H3K27ac. Intriguingly, acetylation on K27 increased linearly with increasing number of methylations on K36, with the largest change recorded for the motif K27acK36me3 (Figure 4.4C). However, we did not detect any modification on K37 other than propionylation, which was introduced during the *in vitro* chemical derivatization of the intact histones.

4.4.6. Methylation-dependent activity of Clr3 toward H3K14ac

Several Western and ChIP analyses have previously shown the specificity of Clr3 for K14ac localized in heterochromatin domains such as mating-type loci, centromeres, and telomeres.[29,31,34,45] Clr3 was also shown to affect K14ac in bulk histones.[31] In the current study, we investigated changes in the overall occupancy of K14ac in Clr3 mutant relative to the control. As shown in Figure 4.4A (and Table S4.6 A-B), Clr3 depletion did not

significantly affect the overall occupancy of K14ac, or any of the other sites we investigated. This result was very intriguing for us particularly considering previous reports of Clr3 activity in bulk histones, and thus we were expecting K14 to be hyperacetylated. This prompted us to look deeper into our data to check if there were alterations in the abundances of five different acetylation motifs containing K14ac, which are, K9me2K14ac, K9me3K14ac, K9acK14un, K9unK14ac and K9acK14ac. The peptide with the motif K9me1K14ac was not sufficiently detected in our study. As shown in Figure 4.4B (and Table S4.6C), we found 2.5- and 4.5-fold increase in the abundances of the motifs K9me2K14ac and K9me3K14ac, respectively. In contrast, we did not see any significant change in the abundance of the singly acetylated peptide K9unK14ac. To further confirm this observation we extracted chromatographic peaks representing each of the above acetylated motifs and a non-modified form (K9unK14un) from the raw LC-MS data for the control and Clr3 mutant samples. The overlayed mutant-to-control (Δclr3/Ctl) XICs for each motif are illustrated in Figure 4.4D. According this data, the relative abundances of the peptides K9acK14un, K9unK14ac, K9acK14ac and K9unK14un were essentially the same in the mutant and control samples (Figure 4.4D, see panels 1-4). However, the relative abundances of K9me2K14ac and K9me3K14ac increased significantly in the Clr3 mutant (Figure 4.4D, panels 5 and 6). Overall, these results suggest that Clr3 requires the presence of a di- or tri-methylated K9 in order to target K14ac on the H3 tail.

4.4.7. Sir2 and Hos2 exhibit mild activity toward H3

The function of Sir2 has been widely implicated in the regulation of heterochromatin assembly in fission yeast. In cooperation with Clr3 it promotes the recruitment of gene silencing complexes at the mating-type loci, telomeres and centromeres.[46] Previous i*n vitro* and *in vivo* studies showed the activity of Sir2 toward K9ac and K14ac.[47,48] In the current study, we observed moderate increases in the occupancies of K9ac and K18ac in the Sir2 mutant relative to the control (Figure 4.4A and Supplementary 6A-B). These changes were mostly attributed to increased abundances of the di-acetylation motifs K9acK14ac and K18acK23ac, which were elevated by 1.7- and 2.4-fold, respectively (Figure 4.4B and Table S4.6C).

We also assessed the impact of Hos2 depletion on the occupancy of acetylation in H3. We observed a 2-fold increase in the occupancy of K9ac (Figure 4.4A and Supplementary 6A-B), whereas occupancies of K14ac, K18ac, K23ac, and K27ac were essentially unaffected. The increased occupancy of K9ac was largely attributed to the higher abundance of the di-

acetylated motif K9acK14ac, which increased 3-fold in the Hos2 mutant (Figure 4.4B and Table S4.6C). The corresponding change for the singly acetylated peptide K9acK14un was not significant.

4.4.8. Clr6 mutation caused a major increase in combinatorial H4 acetylation

We first assessed the impact of Clr6 depletion on the occupancy of acetylation at the individual H4 sites. Our analysis showed that the occupancy of K5ac, K8ac, K12ac, and K16ac were elevated by, respectively, 3.7-, 3.0-, 2.8- and 1.6-fold (Figure 4.5A, Table S4.7 A-B). This result suggested that the loss of Clr6 activity had minimal effect on K16ac. To further confirm this observation we examined the changes in the abundances of all the fifteen acetylated isoforms of the peptide 4-GKGGKGLGKGGAKR-17.

As shown in Figure 4.5B (and Table S4.7C), the mono-acetylated motifs K5ac, K8ac, K12ac, and K16ac were barely affected by Clr6 mutation. Meanwhile, of the six di-acetylated isoforms, only three were significantly affected by the mutation. These are K5/8ac, K5/12ac and K8/12ac, which increased by up to 7.5-, 3.5- and 2.5-fold, respectively. The corresponding changes for K5/16ac, K8/16ac, and K12/16ac were not significant (average fold change less than 1.5). In our analysis, the most dramatic change was observed for the tri-acetylated motif K5/8/12ac, which increased 13-fold in the Clr6 mutant. The corresponding increases for the three other tri-acetylated motifs K5/8/16ac, K5/12/16ac, and K8/12/16ac were on average 3-fold. Surprisingly, the fold increase for the tetra-acetylated isoform K5/8/12/16ac (~6-fold) was 50% lower than that of K5/8/12ac, also suggesting the strong preference of Clr6 for the tri-acetylated motif. Noticeably, a common feature of those motifs largely affected by Clr6 depletion is that they comprise at least two of the three acetyl-lysines K5, K8, and K12; with the most preferred combination being the tri-acetylated motif K5/K8/12ac.

4.4.9. Hos2 mutation caused a major increase in H4K16 acetylation

The specificity of Hos2 for H4K16ac has been previously reported.[29] Interestingly, unlike many other HDACs, Hos2-mediated deacetylation of K16ac has been shown to promote transcriptional activation in both *S. pombe* and *S. cerevisiae.*[30]

We first assessed the impacts of Hos2 depletion on the occupancy of the individual H4 sites. Hos2 depletion, in addition to a previously reported increase in K16ac, resulted in a 2-3-fold increase in the overall occupancy of K5ac, K8ac and K12ac (Figure 4.5A, Table S4.7 A-B). To further confirm this unexpected result, we examined the changes in abundances of each of

the 15 acetylated isoforms of the peptide 4-GKGGKGLGKGGAKR-17. Our analysis showed that the increase in occupancy observed for the first three lysines was largely attributed to the fact that these residues co-exist with K16ac-containing motifs in the di-, tri, or tetra-acetylated forms. As shown in Figure 4.5B (and Table S4.7C), the only mono-acetylated motif that showed a significant increase in abundance in the Hos2 mutant is K16ac, which was elevated by about 2-fold. Moreover, among the six di-acetylated isoforms, those harboring the K16ac mark (K5/16ac, K8/16ac, and K12/16ac) were preferentially affected by the loss of Hos2 activity. Similarly, the K16ac-containing tri-acetylated motifs K5/8/16ac, K5/12/16ac and K8/12/16ac exhibited an average 2.8-fold increase in abundance. However, the corresponding increase for the motif lacking the K16ac (K5/8/12ac) was not significant. The most dramatic change was observed for the tetra-acetylated motif K5/8/12/16ac, which increased 4-fold in the Hos2 mutant. Overall, these results suggest the preference of Hos2 for multiply acetylated motifs comprising K16ac rather than the individual mark.

4.4.10. Sir2 and Clr3 mutations did not alter H4 acetylation pattern

Previous *in vitro* deacetylation assay using ^{14}C-labeled NAD+ showed that Sir2 efficiently deacetylates H3K9ac and H4K16ac.[47] Moreover, a very recent *in vitro* analysis, using a ^{32}P-labeled NAD+ reported that Sir2 preferentially deacetylates H3K4 and H4K16ac, and to a lower extent H3K9ac and H3K14ac. Although both the above studies reported the activity of Sir2 toward H4K16ac, a previous *in vivo* ChIP analysis did not observe any major activity of Sir2 toward H4.[29] In the current study, we assessed the fold changes in acetylation site occupancy of the H4 sites K5, K8, K12, and K16 in the Sir2 mutant relative to the control. In agreement with the previous ChIP analysis, we did not find any major activity of Sir2 toward H4 (Figure 4.5 A-B, Table S4.7 A-C). Similarly, Clr3 exhibited no major activity toward H4 (Figure 4.5 A-B).

4.4.11. HDAC mutants display altered H3K9 methylation levels

Clr3 and Sir2 were previously shown to mediate the level of H3K9 methylation at silent mating-type locus and centromeres.[46,47] Their deacetylase activity leads to a simultaneous increase in H3K9 methylation. In contrast, little is known about the correlations between Clr6 or Hos2 activity and H3K9 methylation. Characterization of such type of PTM interactions allows us to fully understand the mechanisms involved in HDAC-mediated heterochromatin assembly and gene regulation. Thus, we investigated how depletion of Clr3, Clr6, Sir2, or Hos2 impact the global methylation levels of H3K9 and H3K36.

Figure 4.6A shows the mutant-to-control occupancy fold change ratios of three methylation states (me1, me2 and me3) of H3K9 and H3K36. In agreement with previous reports, Clr3 and Sir2 depletions generally resulted in reduction of occupancies of K9me2 and K9me3 (Figure 4.6A, upper panel and Table S4.8 A-B). The most significant change was observed for K9me3, which showed a 5-fold reduction in the Clr3 mutant. This result is also evident from the analysis of abundances of the individual methylation motifs shown in Figure 4.6B, top-left panel (and Table S4.8C). In the Clr3 mutant, the fold-decrease in the abundances of the motifs K9me1K14un, K9me2K14un, K9me3K14un were 1.4, 2.9 and 5.6, respectively. For Sir2 mutant, we observed an average 2-fold reduction for each motif. Meanwhile, in both Clr3 and Sir2 mutants, the abundances of the motifs K9me2/3K14ac significantly increased, which is attributed to elevated level of K14ac, a site we observed to be targeted by both HDACs. Clr3 or Sir2 depletion had little effect on the level K36 methylation.

In contrast, Clr6 and Hos2 depletions resulted in a major increase in K9 methylation. For each mutant, we saw 3- and 2-fold increases in the occupancies of K9me2 and K9me3, respectively (Figure 4.6A, upper panel and Table S4.8 A-B). Clearly, K9me2 is more affected than K9me3. Analysis of the individual methylation motifs confirmed these observations (Figure 4.6B, bottom panels and Table S4.8C). In the Clr6 mutant, the fold-increases in the abundances of the motifs K9me1K14un, K9me2K14un and K9me3K14un were, respectively, 1.4, 2.5 and 2.1. The corresponding fold-increases in the Hos2 mutant were, respectively, 1.7, 2.9 and 2.3. Interestingly, the combined incremental effects of Clr6 depletion on K9 methylation and K14 acetylation resulted in a dramatic increase in the abundances of the motifs K9me2K14ac and K9me3K14ac, which increased 6- and 4-fold, respectively. Similarly, the depletion of Hos2 caused 2-3-fold increase in the abundances of the corresponding motifs. Meanwhile, both Clr6 and Hos2 depletions did not affect the level of K36 methylation (Figure 4.6 A-B, lower panels). The fold-increases observed for the motifs K27acK36me1, K27acK36me2 and K27acK36me3 in the Clr6 mutant were essential attributed to the elevated level of K27ac, a site we characterized to be targeted by Clr6, but not Hos2. Overall, our results revealed two distinct patterns of K9 methylation states in the HDAC mutants. While Clr3 and Sir2 depletions correlate with global reduction in the heterochromatin marks K9me2 and K9me3, the loss of Clr6 and Hos2 activities resulted in elevated levels of both PTM marks.

4.5. Discussion

The histone code hypothesis suggests that histone PTMs act in a combinatorial fashion to modulate chromatin-based processes[49]. However, the lack of high-throughput methodologies for comprehensive analysis of histones has so far limited our capability to fully map combinatorial PTM patterns. A number of studies have for many years relied on site-specific antibodies to detect histone modifications. While this approach still provides crucial information regarding the genomic localization of PTMs, its ability to map the global landscape of multiple co-occurring PTMs is limited. Comparatively, MS-based approaches have proven valuable in determining PTM stoichiometries[38,50], mapping combinatorial modification patterns[51]; and characterization of short and long range crosstalks among various histone modifications[52]. In the current study, we employed a targeted quantitative MS-based strategy to profile changes in histone acetylation occupancies in fission yeast strains lacking major HAT (Gcn5 and/or Mst2) and HDAC (Clr3, Clr6, Hos2 and Sir2) activities. Although previous studies have reported substrate specificities of these enzymes[25,29], the methodologies employed in these studies allowed analysis of only few acetylation sites. In addition, site specific acetylation occupancies were not reported. Consequently, the goal of the current study was to conduct a quantitative in-depth analysis of multiple acetylation sites on the N-terminal tails of H3 and H4. We also aimed to assess the roles of each HAT and HDAC in regulating combinatorial histone PTM patterns.

Comprehensive analysis of histone tryptic peptides generated from wild-type and mutant HAT and HDAC strains identified several known and novel target sites. In addition to previously characterized sites, we found that HATs Gcn5 and Mst2 could also catalyze acetylation of, respectively, H3K27 and H3K23. We also saw a reduction in the occupancy of H3K4ac in the Gcn5 mutant (data not shown), but due to the extremely low-abundance of this PTM we could not reproducibly detect it in our samples. In budding yeast, Gcn5, together with a HAT known as Rtt109, have been shown to catalyze the acetylation of H3K4ac *in vivo*. [53] Despite previous reports of Gcn5 activity toward H3K14[54], we found no major change in the occupancy of H3K14ac in the Gcn5 mutant. Overall, our data suggest that *S. pombe* Gcn5 and Mst2 together catalyze up to six acetylation sites on the H3 tail. Interestingly, the two HATs target complementary sites on the H3 tail. It is known that lysine acetylation facilitates transcriptional activation by disrupting the repressive state of nucleosomal DNA. Consequently, the division of labor between Gcn5 and Mst2 could be one way of accelerating the acetylation process to meet the demands of rapidly progressing transcriptional machinery.

Similarly, we identified novel substrate specificity sites for some of the HDACs we studied. For Clr6, we found three new target sites on H3 (K18, K23 and K27) and one on H4 (K8). Hos2 exhibited mild activity toward H3K9ac and H3K14ac, in addition to its primary target site - H4K16ac. In contrast to previous observations, our data showed that Clr6 has very little activity toward H4K16ac. Interestingly, Clr6 and Hos2 exhibited complementary activity toward the H4 N-terminal tail. Clr6 is specific for K5ac, K8ac and K12ac, whereas Hos2 specifically targets K16ac. Previous genome-wide analysis of Clr6 and Hos2 target sites identified independent and distinct roles for the two HDACs.[34] Histone deacetylation by Clr6 was found to have a role in gene repression at promoter regions, whereas deacetylation by Hos2 was associated with increased gene expression at coding regions. However, based on our results, the lack of activity of Clr6, Clr3 or Sir2 for K16ac suggests that Hos2 might also be required for deacetylation of this site at promoter regions.

A key goal of this study was also to examine the impact of HAT or HDAC deletions on combinatorial PTM marks. Perhaps the most striking observation of our study is the fact that Clr6 mutation caused major changes in combinatorial H3 and H4 acetylation motifs. Clr6 depletion resulted in hyperacetylation of the di-acetylated motifs H3K9acK14ac and H3K18acK23ac (Figure 4.4B) and tri-acetylated H4K5, 8, 12ac (Figure 4.5B). However, we found no significant change in the abundance of mono-acetylated H3 and H4 peptides. Clr6 was previously shown to be recruited to both euchromatin and heterochromatin via interaction with transcription factors and co-repressor complexes.[27] Clr6 interaction with the activating transcription factor Atf1/Pcr1 is required for deacetylation of both histones H3 and H4 at the silent mating-type region.[27] In addition, Clr6 was found to exist in two independent HDAC-Sin3 co-repressor complexes (complex I and II) to target promoters, transcribed chromosomal regions and centromeres.[55] Despite these crucial findings, there is no clear mechanism as to how Clr6 targets the various regions of the fission yeast genome. According to our result, it is possible that Clr6 might be directly involved in deacetylating the combinatorial acetylation motifs described above. Alternative, it is possible that Clr6 mutation might give enhanced access to certain HAT to acetylate mutlple lysine residues in histones H3 and H4. Another possible explanation for the high abundance of combinatorial acetylation motifs is that there are some HATs that are rapidly acetylating histones and converting them from mono-acetylated forms to combinatorial motifs. Thus, our discovery of the preference of Clr6 for the tri-acetylated H4 molecules provides an important clue for

future investigations involving characterization of mechanisms of Clr6-mediated gene regulation.

Notably, our analyses also unraveled a few short-range PTM crosstalks on histones H3 and H4 tails. The loss of HDACs generally leads to an increase in histone acetylation and reduction in repressive histone methylation marks (e.g. H3K9me3 and H3K27me3). In the Clr3 mutant, we observed a negative crosstalk between the PTM marks H3K14ac and H3K9me2/me3 (Figure 4.7A). The loss of Clr3 induced hyperacetylation of H3K14ac and a concomitant decrease in the level of H3K9me3, and to a lesser extent H3K9me2 (Figure 4.6 A-B). Similar observations were made by Yamada et al.[46], whereby a western analysis showed reduction in H3K9me3 in the Clr3 mutant. Meanwhile, conflicting reports were made on the impact of Clr3 depletion on the level of H3K9me2. While Yamada et al. reported a slight increase in H3K9me2, a very recent study by Castonguay et al.[56] showed a major decrease in this mark. Our analysis strongly agrees with the latter observation.

The mechanism by which the opposing marks H3K14ac and H3K9me2/me3 regulate gene silencing in fission yeast is not fully characterized. However, an interesting model had been previously proposed.[46] According to this model, Clr3 is recruited to its target site via interaction with activating transcription factors Atf1/Pcr1. This interaction allows the recruitment of the heterochromatin proteins (HPs) Swi6 and Ch2, and the H3K9 methyltransferase known as Clr4. Conceivably, the loss of Clr3 would abolish the binding of the HPs and Clr4, which potentially leads hypomethylation of H3K9 and disruption of the heterochromatin domain. Interestingly, a crucial role for Sir2 was implicated in this process, whereby depletion of Sir2 caused a severe defect in binding of Clr3 at the mating type-region.[46] This study also suggested the cooperativity and redundancy between Clr3 and Sir2 activities in maintaining heterochromatin silencing. In a more recent western analysis, Sir2 was shown to deacetylate H3K14ac, in addition to the previously known H3K9ac, for targeting Clr4 to centromeres.[48] In agreement with these observations, our analysis showed that Sir2 depletion significantly increased the abundance of the di-acetylated motif H3K9acK14ac (Figure 4.4 A-B). We also observed a similar increase in the abundance of the motif H3K18acK23ac (Figure 4.4 A-B), which was not previously reported. On the other hand, the levels of H3K9me2 and H3K9me3 were reduced in the Sir2 mutant. These results suggest that Sir2 activity might actually have a much wider impact on the landscape of histone H3 PTMs.

In both *S. pombe* and *S. cerevisiae*, Hos2-mediated deacetylation of histones was implicated in gene activation at coding regions[30,57,58], which is contrary to the traditional view of HDACs as gene repressors. In support of this observation, it has been shown that depletion of *S. pombe* Hos2 results in down-regulation of more than 9% of the genome.[29] This mutation also led to a major increase in H4K16ac, which possibly suggest a negative correlation between the PTM and gene activity.

The impact of Hos2 depletion on the methylation level of H3K9 was not previously reported. Traditionally, hypermethylation of H3K9 is strongly associated with gene repression. In our study, we found an inter-histone positive crosstalk between H3K9 methylation and H4K16ac in the Hos2 mutant (Figure 4.7B). The occupancy of both H3K9me2/me3 and H4K16ac were strongly elevated (Figure 4.6 A-B). Meanwhile, it is also possible that if the residues H4K16 and H3K9 are not present in the same nucleosome, it is difficult for them to interact, and thus the increase in the occupancy of both PTMs could be purely mutually exclusive. Alternatively, hyperacetylation of H4K16ac might indirectly alter the methylation state of H3K9. Interestingly, Hos2 depletion majorly affected the H3K9me2 mark as opposed to the widely known repressive mark H3K9me3. While the latter modification is mostly found at constitutive heterochromatin regions, numerous studies in higher eukaryotes show that H3K9me2 is highly enriched at repressed loci of euchromatin regions.[59-61] Thus, the higher level of H3K9me2 we observed in the Hos2 mutant is in line with the function of the HDAC at the euchromatin domain.

In summary, the comprehensive and quantitative MS-based approach allowed us to characterize several known and novel targets sites of *S. pombe* HATs and HDACs. Meanwhile, based on the current data we cannot distinguish between direct and indirect targets of the enzymes. This requires further *in vitro* biochemical assays that involve determining the rate of acetylation or deacetylation of histones mediated by affinity purified fission yeast HATs or HDACs. We anticipate that the results presented in this workwill be indispensable in future genetic and biochemical studies involving the fission yeast. Our study also highlighted on the division of labor between HATs or HDACs activities in regulating chromatin-associated processes. Moreover, this study demonstrated the exceptional advantage in using MS-based approaches in profiling combinatorial histone codes and characterization of functional crosstalks among the various histone PTMs.

4.6. Acknowledgments

We thank Dr. Susan Forsburg (University of Southern California) for generously providing us with the HAT mutant fission yeast strains and Dr. Karl Ekwall (Karolinska Institutet) for the HDAC mutant strains. We also thank Éric Bonneil (UdeM) for his assistance with MS experiments. This work was carried out with financial support from the Canadian Institute for Health Research to A.V. (CIHR, FRN 125916) and the Natural Sciences and Engineering Research Council (NSERC 311598) to P.T. The Institute for Research in Immunology and Cancer (IRIC) receives infrastructure support from IRICoR, the Canadian Foundation for Innovation, and the Fonds de Recherche du Québec - Santé (FRQS).

4.7. Main figures

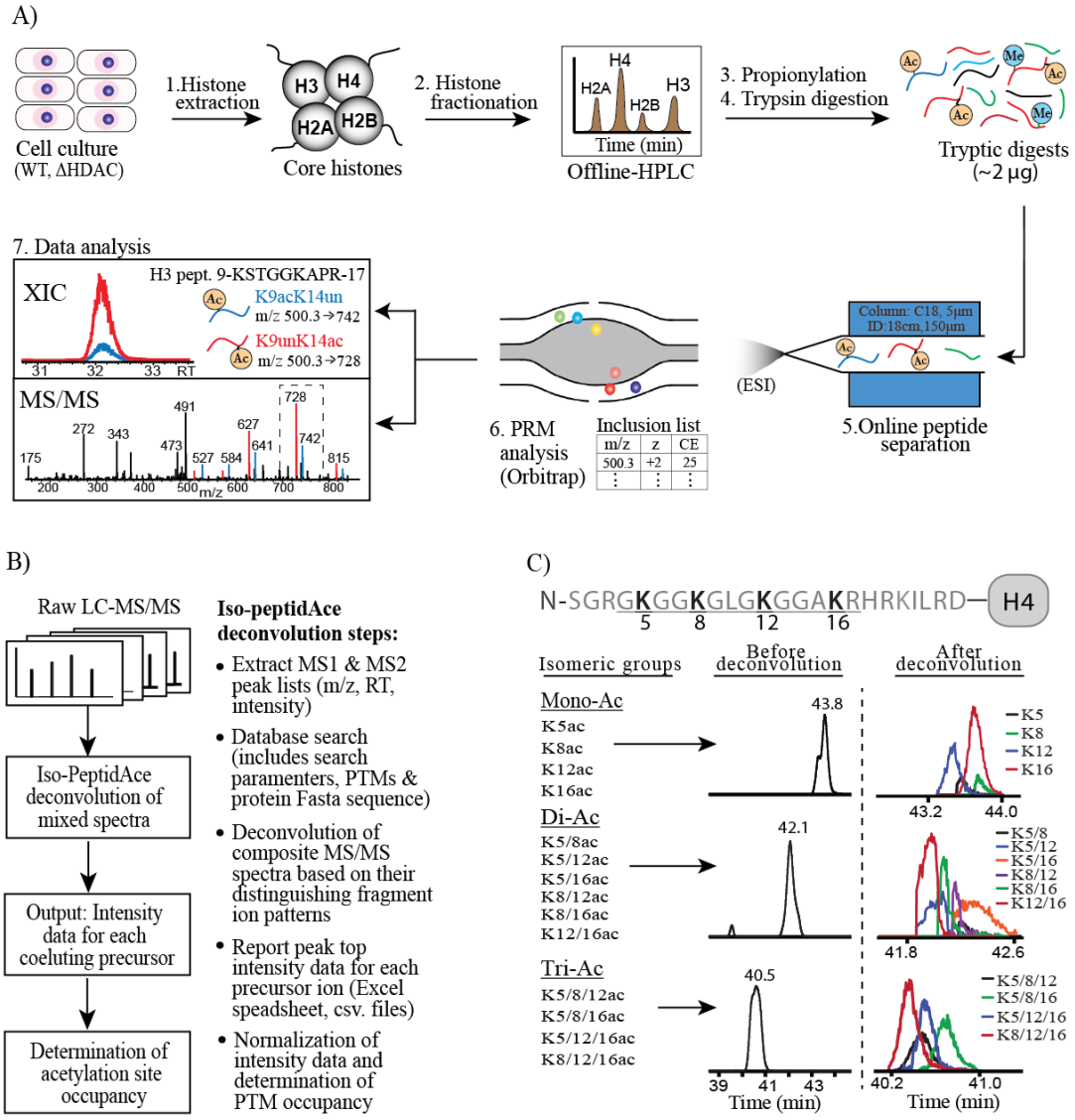


Figure 4.1 PRM-LC-MS/MS analysis of histone peptides.

A) Overview of the experimental procedure. 1. Histones were isolated from cultured wild-type and mutant S.pombe strains. 2. Bulk histones were fractionated into individual core histones (H2A, H2B, H3, and H4) using offline HPLC. 3. Fractions containing histones H3 and H4 were pooled, and subjected to propionylation and tryptic digestion. 5. Approximately 2 µg of core histones H3 and H4 peptides were separated by reverse-phase C18 column that is directly connected to the ESI source of Q-Exactive Plus MS. 6. Histone peptides were analyzed by PRM method whereby an inclusion list was used to trigger MS/MS acquisitions. 7. PRM data were analyzed based on peak intensities of XICs generated manually using the Xcalibur software (Thermo Scientific) or using in-house built algorithm- Iso-PeptidAce. XIC panel: Shows the co-elution of isomeric H3 peptides- K9acK14un (m/z 500.28^{+2}, ac-acetylated, un – unacetylated) and K9unK14ac (m/z 500.28^{+2}); MS/MS panel: Shows the co-fragmentation of the isomers. Isomer-specific fragments are highlighted in blue (from

K9acK14un) or red (K9unK14ac). B) Deconvolution of acetylated isomers of histones. Left panel: Raw LC-MS/MS were submitted to Iso-PeptidAce for deconvolution of mixed MS/MS spectra of co-eluting isomers. The output from Iso-PeptidAce is a .csv file containing precursor intensities for each deconvoluted isomer. These intensity values were used to calculate site occupancies of acetyl-lysines. Right panel: Steps in the deconvolution of mixed MS/MS spectra using IsopeptidAce. C) Deconvolution of acetylated isomer of fission yeast histone H4.

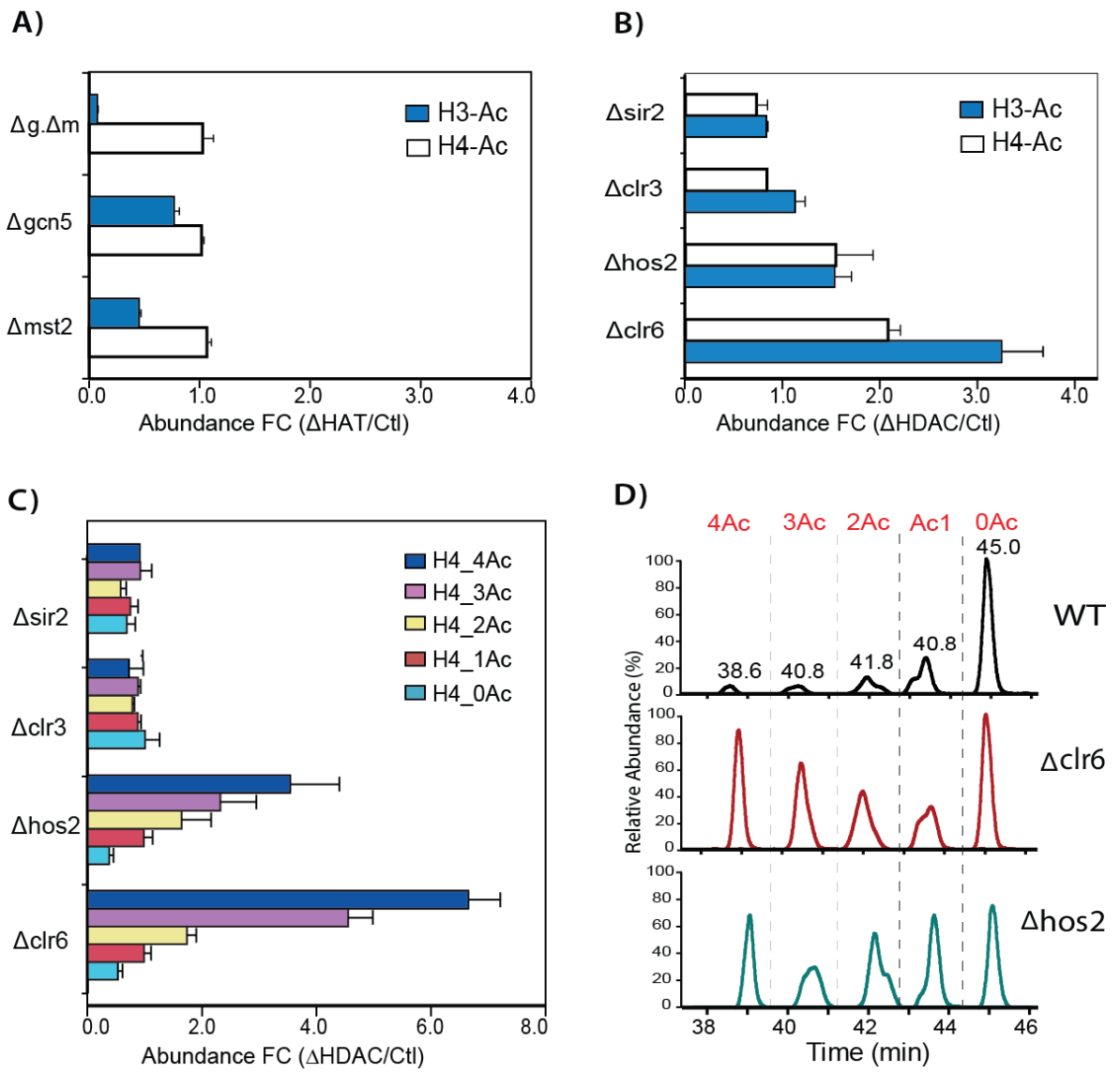


Figure 4.2 Global H3 and H4 acetylations in HAT and HDAC mutants.

Fold change ratios of abundances of global H3 and H4 acetylations in (A) HAT mutants: Δgcn5, Δmst2, and Δgcn5.Δmst2 (Δg.Δm); and (B) HDAC mutants: Δclr6, Δhos2, Δclr3, and Δsir2 relative to control samples. The fold change (FC) ratios were determined based on the summed intensities of all acetylated H3 or H4 peptides detected in the mutant relative to the control. C) Fold change ratios of abundances of mono-, di-, tri-, and tetra-acetylated and a non-aceylated isoforms of the H4 peptide 4-GKGGKGLGKGGAKR-17 in the HDAC mutants relative to the control samples. These fold-change ratios were determined based on the intensities of each isoform obtained after Iso-Peptidace analysis. D) Extracted ion chromatograms illustrating the differential abundances of the five acetylated and non-acetylated isoforms of the H4 peptide in WT sample (Top panel), Δclr6 (middle panel), Δhos2 (bottom panel). Error bars indicate standard deviations from two-three replicates.

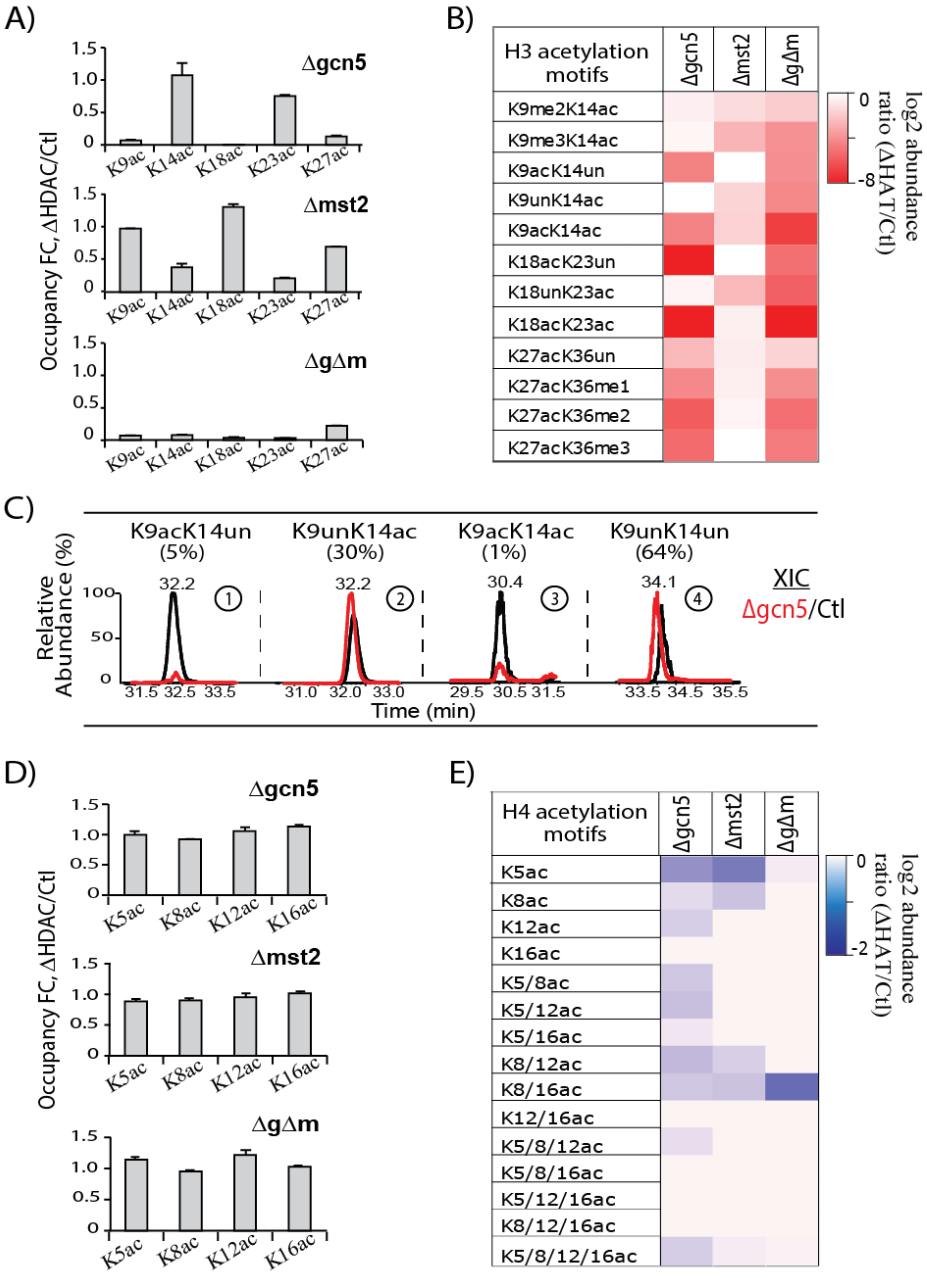


Figure 4.3 Site-specific acetylation of H3 and H4 in HAT mutants.

A) Fold changes in the occupancy of H3K9ac, H3K14ac, H3K18ac, H3K23ac and H3K27ac in Δgcn5, Δmst2 and Δgcn5.Δmst2 (Δg.Δm) (top to bottom panels) relative to the control sample. B) Heapmap illustration of the log2 mutant-to-control fold-changes in the normalized LC-MS signal responses of H3 peptides harboring distinct acetylation motifs. C) Overlayed extracted ion chromatograms (XICs) illustrating differences in abundances of K14ac containing peptides (panel 1-3) and a non-acetylated peptide (panel 4) in Δgcn5 relative to the control. The indicated percentages represent the relative abundances (in %) of the corresponding isoforms among the total H3 peptides detected with the same sequence. The y-axis in each XIC panel category was scaled to the signal response of the most abundant peptide. D) Fold changes in the occupancy of H4K5ac, H4K8ac, H4K12ac, and H4K16ac in

Δgcn5, Δmst2 and Δgcn5.Δmst2 (Δg.Δm) (top to bottom panels) relative to the control. B) Heapmap illustration of the log2 mutant-to-control fold-changes in the LC-MS signal responses of acetylated H4 peptide isoforms. Error bars indicate standard deviations from two-three replicates.

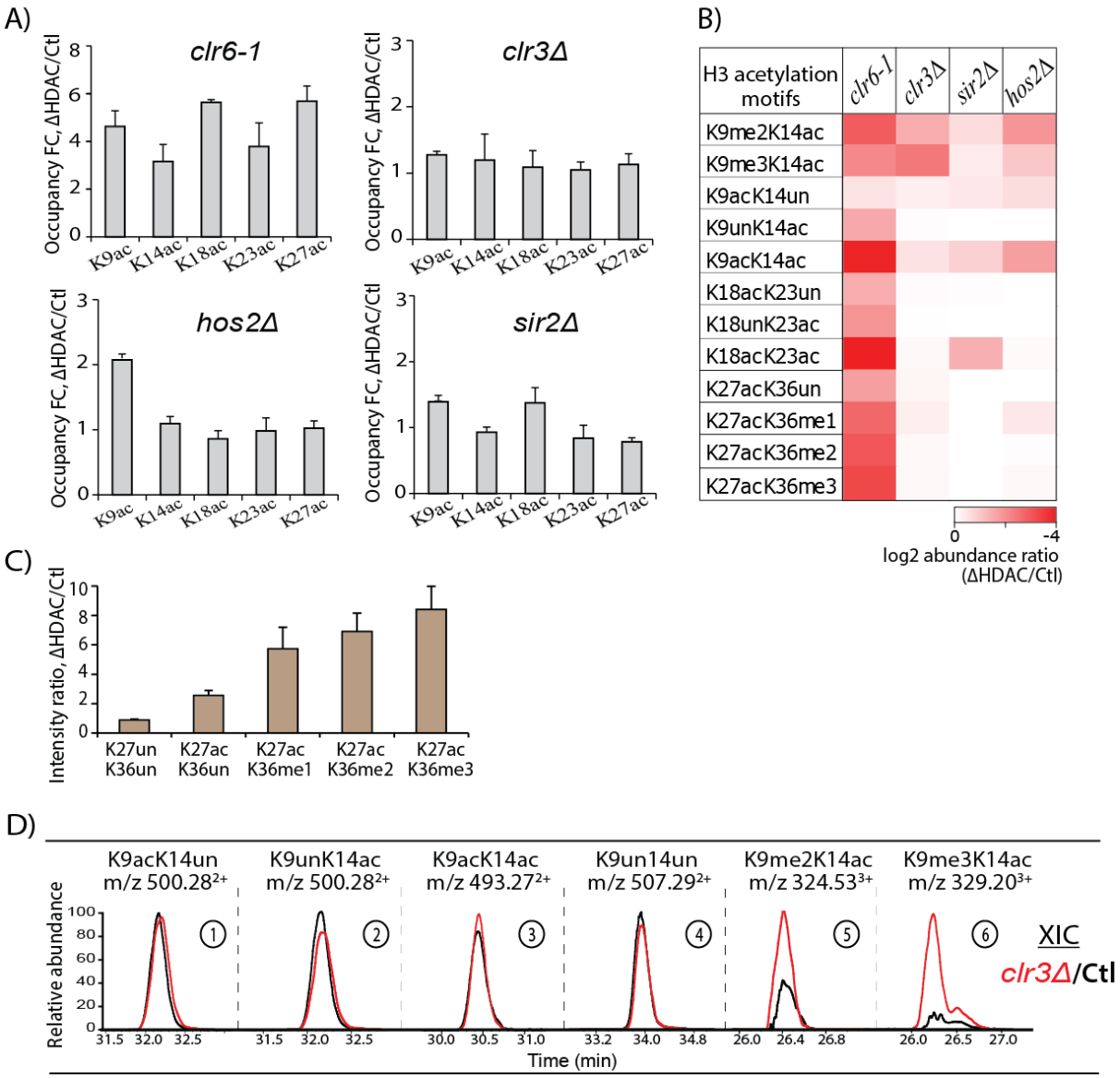


Figure 4.4 Site-specific acetylation of H3 in HDAC mutants.

A) Fold changes in the occupancy of H3K9ac, H3K14ac, H3K18ac, H3K23ac and H3K27ac in HDAC mutants relative to the control. B) Heatmap illustration of the mutant-to-control log2 fold changes in the normalized LC-MS signal responses of acetylated peptide motifs. C) Mutant-to-control normalized intensity ratios of the non-acetylated (K27unK36un) and K27ac-containing isoforms of the peptide 27-KAAPATGGVKKPHR-40. D) Overlayed XICs illustrating changes in the relative abundances of a non-modified (panel 4) and K14ac-

containing (panels 1-3 and 5-6) isoforms of the peptide 9-KSTGGKAPR-17 in Clr3 mutant relative to the control. The y-axis in each XIC panel was scaled to the signal response of the most abundant peptide. Error bars indicate standard deviations from two-three replicates.

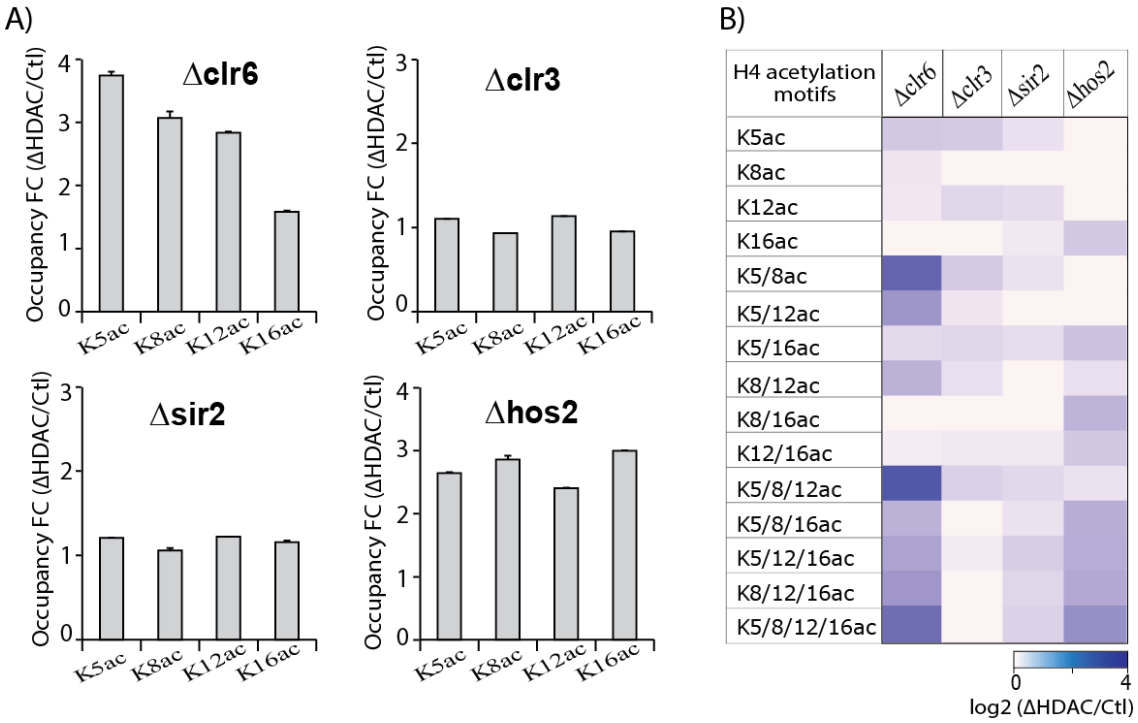


Figure 4.5 Site-specific acetylation of H4 in HDAC mutants.

A) Fold changes in the occupancy of H4K5ac, H4K8ac, H4K12ac, and H4K16ac in the HDAC mutants relative to the control. B) Heatmap illustration of the log2 fold changes of the normalized LC-MS signal responses of 15 acetylated motifs in the mutant relative to the control sample. Error bars indicate standard deviations from three replicates.

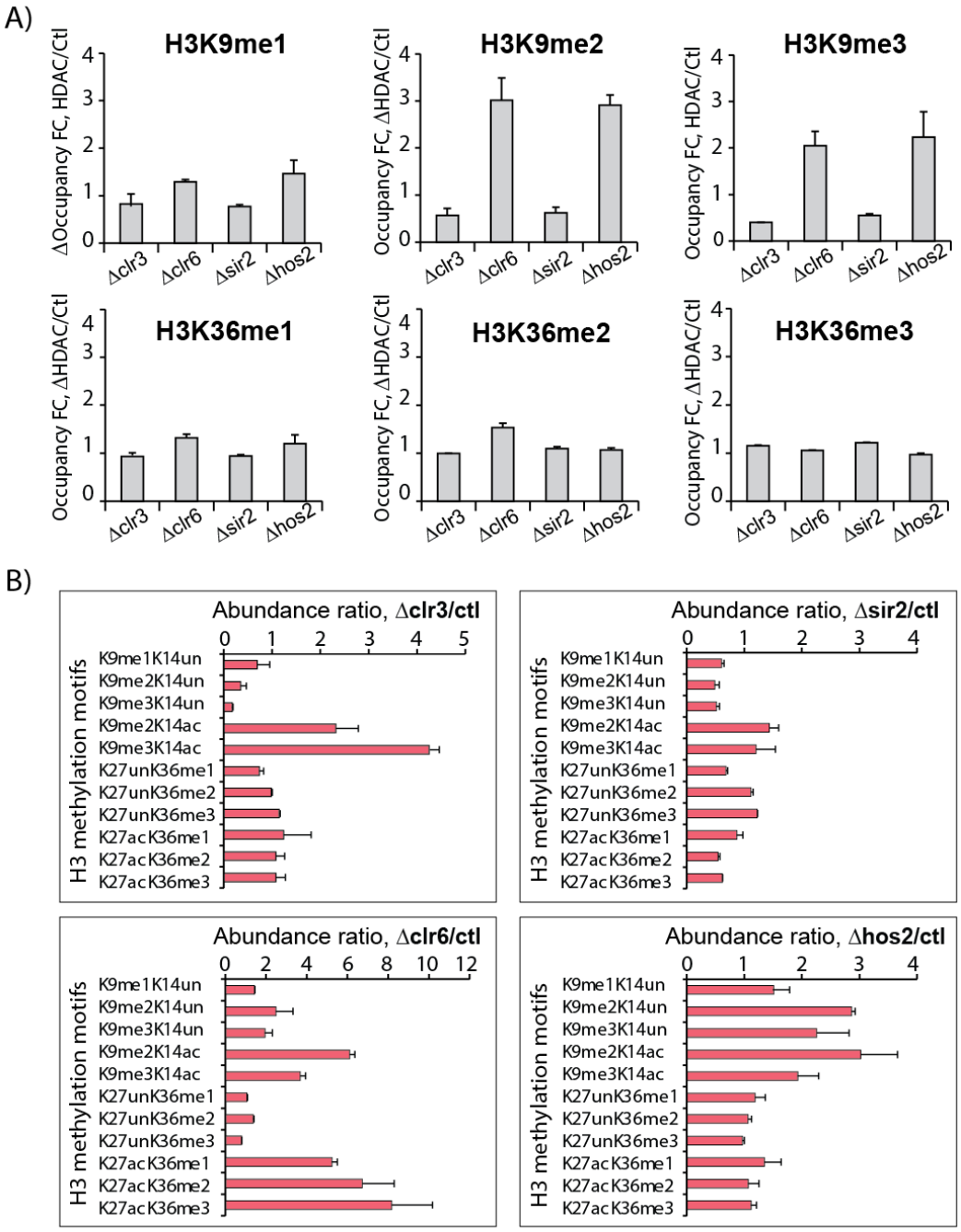


Figure 4.6 Analysis of H3-K9/K36 methylation levels in HDAC mutants.

A) Fold changes in the occupancy of H3K9me1/me2/me3 (top panel) and H3K36me1/me2/me3 (bottom panel) in HDAC mutants relative to the control. B) Normalized abundance ratios of H3K9 and H3K36 methylation motifs in Clr3, Sir2, Clr6 and Hos2 mutants relative to the control. Error bars indicate standard deviations from two-three replicates.

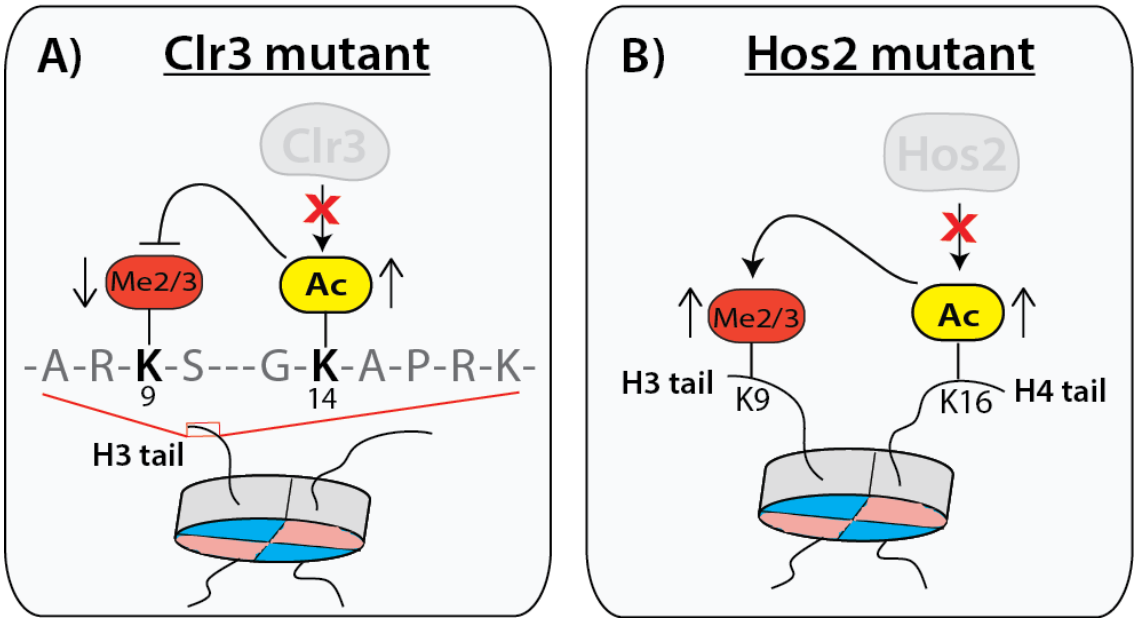


Figure 4.7 Representative histone PTM crosstalks in fission yeast.

A) An illustration of a negative crosstalk between the co-occurring marks H3K9me2/me3 and H3K14ac. Depletion of the Clr3 HDAC induced hyperacetylation of H3K14 and a simultaneous reduction in the level of H3K9me3, and to a lesser extent H3K9me2. B) An illustration of inter-histone positive crosstalk between the markss H3K9me2/me3 and H4K16ac. Depletion of Hos2 resulted in hyperacetylation of H4K16ac and a concomitant increase in H3K9me2, and to a lesser extent H3K9me3. The arrows indicate an increase (↑) or a decrease (↓) in the abundance of the corresponding marks.

4.8. Supplementary data

Table S4.1 (A) Fission yeast strains and (B) PRM inclusion list

A. Fission yeast HDAC Strains

Strain name	Genotype	Source
1645	*h+ ade6-210 arg3-D4 his3-D1 leu1-32 ura4-D18*	Allshire Lab
1646	*h- ade6-210 arg3-D4 his3-D1 leu1-32 ura4-D18*	Allshire Lab
9351	*h+ sir2::NatR his3D arg3D leu1-32 ade6-210 ura4-D18*	Allshire Lab
5962	*h+ clr3::kanMX ade6-210 ura4-D18 leu1-32*	Allshire Lab
A3337	*h- clr6-1 ade6-210 leu1-32 ura4D-S/E his3-D1? arg3-D4?*	Allshire Lab
A3338	*h+ clr6-1 ade6-210 leu1-32 ura4D-S/E his3-D1? arg3-D4?*	Allshire Lab
Hu1026	*h- hos2::LEU2+ leu1- 32*	Ekwall lab

B. PRM inclusion list.

Mass (m/z)	CS	CE	PTM	Sequence
319.8540	3+	27	H3-K9me1K14ac	Kme1STGGKacAPR
479.2774	2+	27	H3-K9me1K14ac	
514.2981	2+	27	H3-K9me1-prK14pr	Kme1-prSTGGK14APR
507.2904	2+	27	H3-K9me1-prK14ac	
324.5259	3+	27	H3-K9me1K14un	Kme1STGGKprAPR
486.2837	2+	27	H3-K9me1K14un	
324.5259	3+	27	H3-K9me2K14ac	Kme2STGGKacAPR
486.2837	2+	27	H3-K9me2K14ac	
329.1980	3+	27	H3-K9me3K14ac	Kme3STGGKacAPR
493.2936	2+	27	H3-K9me3K14ac	
329.1980	3+	27	H3-K9me2K14pr	Kme2STGGKprAPR
493.2936	2+	27	H3-K9me2K14pr	
333.8696	3+	27	H3-K9me3K14pr	Kme3STGGKprAPR
500.3002	2+	27	H3-K9me3K14pr	
507.2906	2+	27	H3-K9prK14pr	KprSTGGKprAPR
493.2748	2+	27	H3-K9acK14ac	KacSTGGKacAPR
500.2836	2+	27	H3-K9prK14ac	KprSTGGKacAPR
	2+	27	H3-K9acK14pr	KacSTGGKprAPR
542.8270	2+	27	H3-K18prK23pr	KprQLASKprAAR
535.8190	2+	27	H3-K18acK23pr	KacQLASKprAAR
	2+	27	H3-K18pK23ac	KprQLASKacAAR
528.8110	2+	27	H3-K18acK23ac	KacQLASKacAAR
793.4619	2+	27	H3-K27prK36prK37pr	KprAAPATGGVKprKprPHR
529.3103	3+	27	H3-K27prK36prK37pr	
786.4542	2+	27	H3-K27acK36prK37pr	KacAAPATGGVKprKprPHR
524.6385	3+	27	H3-K27acK36prK37pr	
772.4567	2+	27	H3-K27acK36me2K37pr	KacAAPATGGVKme2KprPHR
515.3069	3+	27	H3-K27acK36me2K37pr	
386.7320	4+	27	H3-K27acK36me2K37pr	
779.4645	2+	27	H3-K27acK36me3K37pr	KacAAPATGGVKme3KprPHR
519.9788	3+	27	H3-K27acK36me3K37pr	
390.2359	4+	27	H3-K27acK36me3K37pr	
793.4620	2+	27	H3-K27acK36me1-prK37pr	KacAAPATGGVKme1-prKprPHR
529.3104	3+	27	H3-K27acK36me1-prK37pr	
779.4645	2+	27	H3-K27prK36me2K37pr	KprAAPATGGVKme2KprPHR
519.9788	3+	27	H3-K27prK36me2K37pr	
390.2359	4+	27	H3-K27prK36me2K37pr	
786.4724	2+	27	H3-K27prK36me3K37pr	KprAAPATGGVKme3KprPHR
524.6507	3+	27	H3-K27prK36me3K37pr	
393.7398	4+	27	H3-K27prK36me3K37pr	
800.4697	2+	27	H3-K27prK36me1-prK37pr	KprAAPATGGVKme1-prKprPHR
533.9822	3+	27	H3-K27prK36me1-prK37pr	
344.8698	3+	27	-	YRPGTVALR
516.8010	2+	27	-	

747.9412	2+	27	H4-K5prK8prK12prK16pr	GKprGGKprGLGKprGGAKprR
740.9334	2+	27	H4-K5acK8prK12prK16pr	GKacGGKprGLGKprGGAKprR
			H4-K5prK8acK12prK16pr	GKprGGKacGLGKprGGAKprR
			H4-K5prK8prK12acK16pr	GKprGGKprGLGKacGGAKprR
			H4-K5prK8prK12prK16ac	GKprGGKprGLGKprGGAKacR
733.9255	2+	27	H4-K5acK8acK12prK16pr	GKacGGKacGLGKprGGAKprR
			H4-K5acK8prK12acK16pr	GKacGGKprGLGKacGGAKprR
			H4-K5acK8prK12prK16ac	GKacGGKprGLGKprGGAKacR
			H4-K5prK8acK12acK16pr	GKprGGKacGLGKacGGAKprR
			H4-K5prK8acK12prK16ac	GKprGGKacGLGKprGGAKacR
			H4-K5prK8prK12acK16ac	GKprGGKprGLGKacGGAKacR
726.9177	2+	27	H4-K5acK8acK12acK16pr	GKacGGKacGLGKacGGAKprR
			H4-K5acK8acK12prK16ac	GKacGGKacGLGKprGGAKacR
			H4-K5acK8prK12acK16ac	GKacGGKprGLGKacGGAKacrR
			H4-K5prK8acK12acK16ac	GKprGGKacGLGKacGGAKacR
719.9099	2+	27	H4-K5acK8acK12acK16ac	GKacGGKacGLGKacGGAKacR
618.8270	2+	27	-	ISALVYEETR

CS - Charge State
CE - Collision Energy
Highlighted in yellow - Peptides used for normalization of total histone abundances

Table S4.2 Global H3 and H4 acetylations in HAT and HDAC mutants.
Analysis of global H3 and H4 acetylations in HAT and HDAC mutants. Fold change ratios of global H3 and H4 acetylations in (A) HAT and (B) HDAC mutants relative to the control sample.

A)

Sample	Relative Abundance FC (HAT/Ctl)	
	H3-Ac	H4-Ac
Δmst2	0.45 ± 0.02	1.17 ± 0.04
Δgcn5	0.77 ± 0.05	1.11 ± 0.02
ΔgΔm	0.07 ± 0.01	1.12 ± 0.11

B)

Sample	Relative Abundance FC (HDAC/Ctl)	
	H3-Ac	H4-Ac
Δclr6	3.25 ± 0.43	2.08 ± 0.13
Δhos2	1.54 ± 0.17	1.55 ± 0.39
Δclr3	1.13 ± 0.10	0.84 ± 0.02
Δsir2	0.83 ± 0.02	0.73 ± 0.12

Table S4.3 Global changes in H4 acetylation in response to HDAC depletion.
Relative abundances of a non-acetylated (H4_0Ac) and mono- (H4_1Ac), di- (H4_2Ac), tri- (H4_3Ac), and tetra-acetylated (H4_4Ac) isoforms of the H4 peptide 4-GKGGKGLGKGGAKR-17. B) Relative abundance fold change ratios of the non-acetylated and acetylated isoforms of the H4 peptide in the HDAC mutant relative to the control sample.

A)

Sample	Relative Abundance (%) ± SD				
	H4_0Ac	H4_1Ac	H4_2Ac	H4_3Ac	H4_4Ac
WT	52.91 ± 5.3	22.98 ± 0.1	14.79 ± 0.3	5.79 ± 0.4	3.53 ± 0.3
Δclr6	22.23 ± 1.1	18.04 ± 2.2	20.40 ± 1.5	20.81 ± 0.4	18.52 ± 0.2
Δhos2	21.43 ± 2.2	24.44 ± 4.0	26.14 ± 7.7	14.36 ± 2.8	13.63 ± 4.5
Δclr3	57.08 ± 8.7	21.98 ± 1.5	12.63 ± 0.3	5.57 ± 0.2	2.74 ± 0.7
Δsir2	51.10 ± 6.0	24.57 ± 4.5	12.18 ± 1.9	7.59 ± 1.0	4.56 ± 0.2

B)

Sample	Relative Abundance FC ratio (HDAC/Ctl) ± SD				
	H4_0AC	H4_1Ac	H4_2Ac	H4_3Ac	H4_4Ac
Δclr6	0.53 ± 0.1	0.99 ± 0.1	1.74 ± 0.2	4.55 ± 0.4	6.65 ± 0.6
Δhos2	0.38 ± 0.1	0.98 ± 0.2	1.64 ± 0.5	2.32 ± 0.6	3.54 ± 0.9
Δclr3	1.01 ± 0.3	0.88 ± 0.1	0.79 ± 0.0	0.89 ± 0.0	0.73 ± 0.2
Δsir2	0.69 ± 0.1	0.75 ± 0.1	0.58 ± 0.1	0.93 ± 0.2	0.91 ± 0.0
Δhst2	0.92 ± 0.3	0.96 ± 0.4	0.88 ± 0.3	1.05 ± 0.5	0.71 ± 0.2

FC- Fold change
SD - Standard deviation (determined based on 3 replicates)

Table S4.4 H3 acetylation site occupancies in HAT mutants.

A) Acetylation and methylation site occupancies of H3 lysine residues. B) Fold change ratios of occupancies of acetylation and methylation sites in H3 in the HAT mutant relative to the control. C) Log2 fold change ratios of the normalized intensities of H3 acetylation motifs in the mutant to the control.

A)

PTM	Acetylation site occupancy (%) ± SD			
	WT	Δgcn5	Δmst2	ΔgΔm
H3K9ac	5.92 ± 0.99	0.40 ± 0.15	6.20 ± 1.05	0.41 ± 0.04
H3K14ac	25.28 ± 4.79	26.73 ± 0.42	10.16 ± 0.38	1.88 ± 0.15
H3K9me1	0.12 ± 0.00	0.09 ± 0.02	0.21 ± 0.00	0.21 ± 0.01
H3K9me2	1.05 ± 0.26	0.46 ± 0.06	0.59 ± 0.04	0.97 ± 0.13
H3K9me3	1.10 ± 0.05	1.59 ± 0.05	1.18 ± 0.00	1.73 ± 0.15
H3K18ac	0.65 ± 0.00	0.00 ± 0.00	0.91 ± 0.03	0.02 ± 0.01
H3K23ac	19.89 ± 1.61	14.96 ± 0.86	4.51 ± 0.15	0.63 ± 0.06
H3K27ac	4.45 ± 0.11	0.58 ± 0.07	3.34 ± 0.07	0.98 ± 0.03
H3K27me1	7.27 ± 0.24	6.87 ± 0.75	6.72 ± 1.55	6.45 ± 1.07
H3K27me2	9.90 ± 1.50	9.61 ± 0.25	10.77 ± 0.21	9.95 ± 0.36
H3K27me3	57.39 ± 1.19	58.72 ± 0.90	62.19 ± 4.24	61.53 ± 6.59

B)

PTM	Occupancy FC ratio (HAT/Ctl) ± SD		
	Δgcn5	Δmst2	ΔgΔm
H3K9ac	-3.92 ± 0.01	0.07 ± 0.00	-3.86 ± 0.00
H3K14ac	0.10 ± 0.19	-1.29 ± 0.06	-3.73 ± 0.01
H3K9me1	-0.34 ± 0.16	0.77 ± 0.04	0.80 ± 0.08
H3K9me2	-1.17 ± 0.05	-0.80 ± 0.11	-0.06 ± 0.37
H3K9me3	0.54 ± 0.12	0.11 ± 0.06	0.66 ± 0.21
H3K18ac	-9.19 ± 0.00	0.49 ± 0.05	-4.92 ± 0.01
H3K23ac	-0.41 ± 0.02	-2.14 ± 0.01	-4.98 ± 0.00
H3K27ac	-2.94 ± 0.01	-0.42 ± 0.00	-2.18 ± 0.00
H3K27me1	-0.08 ± 0.07	-0.12 ± 0.18	-0.17 ± 0.18
H3K27me2	-0.03 ± 0.12	0.14 ± 0.19	0.03 ± 0.19
H3K27me3	0.03 ± 0.04	0.12 ± 0.05	0.10 ± 0.14

C)

H3 acetylation motifs	Log2 abundance FC (HAT/Ctl)		
	Δgcn5	Δmst2	Δg.Δm
K9me2K14ac	-0.52	-1.10	-1.55
K9me3K14ac	-0.37	-2.26	-3.43
K9acK14un	-3.90	0.38	-3.49
K9unK14ac	0.17	-1.31	-3.71
K9/K14ac	-3.86	-1.38	-6.17
K18acK23un	-8.73	0.76	-4.43
K18unK23ac	-0.40	-2.16	-4.97
K18acK23ac	-8.50	-0.59	-8.50
K27acK36un	-2.16	-0.62	-1.32
K27acK36me1K37un	-3.69	-0.56	-3.54
K27acK36me2K37un	-5.14	-0.40	-4.51
K27acK36me3K37un	-4.60	0.01	-4.08

Table S4.5 H4 acetylation site occupancies in HAT mutants.

A) Acetylation site occupancies of H4K5ac, H4K8ac, H4K12ac, and H4K16ac. B) Fold change ratios of occupancies of H4K5ac, H4K8ac, H4K12ac, and H4K16ac in the mutant relative to the control. C) Log2 fold change ratios of the normalized intensities of acetylated isoforms of the H4 peptide 4-GKGGKGLGKGGAKR-17 in the mutant to the control.

A)

Sample	Acetylation site occupancy (%) ± SD			
	H4K5ac	H4K8ac	H4K12ac	H4K16ac
WT	16.86 ± 0.49	20.92 ± 0.67	27.09 ± 1.45	35.21 ± 1.20
Δgcn5	16.74 ± 0.56	19.28 ± 0.54	28.55 ± 0.21	39.86 ± 0.40
Δmst2	14.15 ± 0.21	17.84 ± 0.13	24.42 ± 0.35	38.32 ± 0.56
Δg.Δm	19.22 ± 0.19	19.96 ± 1.02	32.87 ± 0.40	36.20 ± 1.88

B)

Sample	Occupancy FC (HAT/Ctl) ± SD			
	H4K5ac	H4K8ac	H4K12ac	H4K16ac
Δgcn5	0.99 ± 0.06	0.92 ± 0.00	1.06 ± 0.06	1.13 ± 0.03
Δmst2	0.84 ± 0.04	0.85 ± 0.03	0.90 ± 0.06	1.09 ± 0.02
ΔgΔm	1.14 ± 0.04	0.95 ± 0.02	1.22 ± 0.08	1.03 ± 0.02

C)

H4 acetylation motifs	Log2 abundance FC (HAT/Ctl)		
	Δgcn5	Δmst2	Δg.Δm
H4K5_8_12_16un	0.03	-0.09	-0.13
H4K5ac	-0.76	-0.96	-0.05
H4K8ac	-0.18	-0.37	0.18
H4K12ac	-0.29	0.00	0.23
H4K16ac	0.44	0.42	0.01
H4K5_8ac	-0.34	0.06	0.25
H4K5_12ac	-0.40	0.08	0.42
H4K5_16ac	-0.09	0.18	0.25
H4K8_12ac	-0.45	-0.27	0.17
H4K8_16ac	-0.34	-0.37	-1.09
H4K12_16ac	0.15	0.20	0.42
H4K5_8_12ac	-0.15	0.53	0.48
H4K5_8_16ac	0.08	0.01	0.05
H4K5_12_16ac	0.16	0.39	0.63
H4K8_12_16ac	0.14	0.20	0.38
H4K5_8_12_16ac	-0.30	-0.05	-0.01

Table S4.6 H3 acetylation site occupancies in HDAC mutants.

A) Acetylation and methylation site occupancies of H3 lysine residues in WT and HDAC mutants. B) Fold change ratios of occupancies of acetylation and methylation sites in H3 in the HDAC mutants relative to the control. C) Log2 fold change ratios of the normalized intensities of H3 acetylation motifs in the mutant to the control.

A)

PTM	Acetylation site occupancy (%) ± SD				
	WT	Δclr6	Δclr3	Δsir2	Δhos2
H3K9ac	2.34 ± 0.17	10.30 ± 1.47	2.98 ± 0.09	3.26 ± 0.25	9.48 ± 1.04
H3K14ac	14.46 ± 2.18	44.92 ± 3.50	16.84 ± 3.09	13.34 ± 0.89	23.72 ± 1.53
H3K9me1	0.20 ± 0.01	0.29 ± 0.02	0.14 ± 0.05	0.12 ± 0.00	0.37 ± 0.04
H3K9me2	1.02 ± 0.06	3.04 ± 0.60	0.65 ± 0.18	0.63 ± 0.08	2.73 ± 0.62
H3K9me3	2.80 ± 0.30	5.77 ± 1.49	1.12 ± 0.13	1.53 ± 0.06	5.16 ± 0.04
H3K18ac	0.28 ± 0.04	1.61 ± 0.26	0.30 ± 0.03	0.39 ± 0.01	0.91 ± 0.18
H3K23ac	6.12 ± 0.81	20.84 ± 5.78	6.34 ± 0.10	5.06 ± 0.52	9.29 ± 1.21
H3K27ac	2.67 ± 0.10	14.14 ± 2.55	3.02 ± 0.55	1.82 ± 0.19	3.99 ± 0.37
H3K27me1	7.34 ± 0.38	9.64 ± 0.14	5.60 ± 0.12	5.08 ± 0.11	9.17 ± 0.51
H3K27me2	9.86 ± 0.48	15.07 ± 0.14	9.76 ± 0.37	10.80 ± 0.12	9.57 ± 0.18
H3K27me3	57.67 ± 0.11	52.90 ± 1.09	66.40 ± 0.55	70.19 ± 0.40	62.06 ± 1.00

B)

PTM	Occupancy FC ratio (HDAC/Ctl) ± SD			
	Δclr6	Δclr3	Δsir2	Δhos2
H3K9ac	4.43 ± 0.94	1.27 ± 0.05	1.39 ± 0.01	2.11 ± 0.38
H3K14ac	3.16 ± 0.72	1.19 ± 0.39	0.93 ± 0.08	1.09 ± 0.11
H3K9me1	1.43 ± 0.02	0.69 ± 0.26	0.60 ± 0.04	1.90 ± 0.29
H3K9me2	3.01 ± 0.76	0.65 ± 0.21	0.63 ± 0.11	2.90 ± 0.22
H3K9me3	2.04 ± 0.32	0.40 ± 0.00	0.55 ± 0.04	2.23 ± 0.54
H3K18ac	5.64 ± 0.12	1.09 ± 0.25	1.38 ± 0.23	0.86 ± 0.13
H3K23ac	3.50 ± 1.41	1.04 ± 0.12	0.84 ± 0.20	0.98 ± 0.20
H3K27ac	5.33 ± 1.16	1.13 ± 0.16	0.68 ± 0.05	1.02 ± 0.11
H3K27me1	1.32 ± 0.09	0.76 ± 0.06	0.69 ± 0.02	1.20 ± 0.19
H3K27me2	1.53 ± 0.09	0.99 ± 0.01	1.10 ± 0.04	1.07 ± 0.05
H3K27me3	0.92 ± 0.02	1.15 ± 0.01	1.22 ± 0.01	0.97 ± 0.03

C)

H3 acetylation motifs	Log2 abundance FC (HDAC/Ctl)			
	Δclr6	Δclr3	Δsir2	Δhos2
K9me2K14ac	2.60	1.27	0.55	1.65
K9me3K14ac	1.87	2.10	0.33	0.88
K9acK14un	0.42	0.30	0.38	0.54
K9unK14ac	1.34	0.07	-0.20	-0.11
K9acK14ac	3.71	0.44	0.73	1.48
K18acK23un	1.27	0.10	0.07	-0.80
K18unK23ac	1.68	0.04	-0.32	-0.03
K18acK23ac	3.96	0.15	1.23	0.12
K27acK36unK37un	1.49	0.20	-0.50	-0.38
K27acK36me1K37un	2.39	0.30	-0.21	0.39
K27acK36me2K37un	2.69	0.15	-0.90	0.07
K27acK36me3K37un	2.95	0.16	-0.69	0.15

Table S4.7 H4 acetylation site occupancies in HDAC mutants.

A) Acetylation site occupancies of H4K5ac, H4K8ac, H4K12ac, and H4K16ac the wild-type and HDAC mutants. B) Fold change ratios of occupancies of H4K5ac, H4K8ac, H4K12ac, and H4K16ac in the mutant relative to the control. C) Log2 fold change ratios of the normalized intensities of acetylated isoforms of the H4 peptide 4-GKGGKGLGKGGAKR-17 in the mutant to the control.

A)

Sample	Acetylation site occupancy (%) ± SD			
	H4K5ac	H4K8ac	H4K12ac	H4K16ac
WT	13.13 ± 0.00	14.96 ± 0.00	19.00 ± 0.00	29.49 ± 0.00
Δclr3	15.54 ± 1.75	12.94 ± 2.57	21.57 ± 1.76	25.26 ± 2.36
Δclr6	49.13 ± 0.83	45.85 ± 1.56	53.88 ± 0.37	46.59 ± 0.77
Δsir2	16.91 ± 0.08	15.79 ± 0.47	23.19 ± 0.05	34.04 ± 0.58
Δhos2	30.75 1.74	35.34 2.37	39.58 2.27	67.67 2.84

B)

Sample	Occupancy FC (HDAC/Ctl) ± SD			
	H4K5ac	H4K8ac	H4K12ac	H4K16ac
Δclr3	1.10 ± 0.13	0.94 ± 0.17	1.14 ± 0.09	0.96 ± 0.08
Δclr6	3.74 ± 0.06	3.07 ± 0.10	2.84 ± 0.02	1.58 ± 0.03
Δsir2	1.21 ± 0.01	1.06 ± 0.03	1.22 ± 0.00	1.15 ± 0.02
Δhos2	1.98 ± 0.01	2.14 ± 0.04	1.80 ± 0.01	2.24 ± 0.00

C)

H4 acetylation motifs	Log2 abundance FC (HDAC/Ctl)			
	Δclr6	Δclr3	Δsir2	Δhos2
H4K5_8_12_16un	-1.26	0.09	-0.06	-1.28
H4K5ac	0.83	0.79	0.33	-1.82
H4K8ac	0.28	-0.03	-0.06	-1.02
H4K12ac	0.23	0.56	0.45	-0.93
H4K16ac	-0.99	-0.29	0.22	0.82
H4K5_8ac	2.93	0.81	0.31	-0.99
H4K5_12ac	1.83	0.28	-0.05	-0.12
H4K5_16ac	0.45	0.56	0.41	0.95
H4K8_12ac	1.28	0.39	-0.69	0.38
H4K8_16ac	-0.74	-1.37	-0.35	1.24
H4K12_16ac	0.15	0.22	0.22	0.86
H4K5_8_12ac	3.65	0.66	0.51	0.31
H4K5_8_16ac	1.28	-0.06	0.32	1.37
H4K5_12_16ac	1.56	0.18	0.72	1.39
H4K8_12_16ac	1.81	-0.09	0.57	1.47
H4K5_8_12_16ac	2.65	-0.08	0.63	1.93

Table S4.8 H3-K9/K36 methylation site occupancies in HDAC mutants.
A) Methylation site occupancies of H3K9me1/me2/m3 and H3K36me1/me2/me3 in the wild-type and mutant samples. B) Fold change ratios of occupancies of H3K9me1/me2/m3 and H3K36me1/me2/me3 in the mutant relative to the control. C) Log2 fold change ratios of the normalized intensities of various H3 methylation motifs in the mutant to the control.

A)

PTM	Methylation site occupancy (%) ± SD				
	WT	Δclr6	Δclr3	Δsir2	Δhos2
K9me1	0.20 ± 0.01	0.14 ± 0.03	0.29 ± 0.01	0.12 ± 0.00	0.34 ± 0.04
K9me2	1.02 ± 0.04	0.65 ± 0.13	3.05 ± 0.43	0.64 ± 0.06	3.18 ± 0.63
K9me3	2.80 ± 0.21	1.12 ± 0.09	5.77 ± 1.05	1.53 ± 0.04	5.15 ± 0.04
K36me1	7.34 ± 0.27	5.60 ± 0.08	9.64 ± 0.10	5.08 ± 0.08	8.81 ± 0.51
K36me2	9.86 ± 0.34	9.76 ± 0.26	15.07 ± 0.10	10.80 ± 0.09	9.70 ± 0.18
K36me3	57.67 ± 0.08	66.40 ± 0.39	52.90 ± 0.77	70.19 ± 0.28	62.77 ± 1.00

B)

PTM	Occupancy FC ratio (HDAC/Ctl) ± SD			
	Δclr3	Δclr6	Δsir2	Δhos2
K9me1	0.69 ± 0.26	1.43 ± 0.02	0.60 ± 0.04	1.91 ± 0.29
K9me2	0.65 ± 0.21	3.01 ± 0.76	0.63 ± 0.11	2.91 ± 0.22
K9me3	0.40 ± 0.00	2.04 ± 0.32	0.55 ± 0.04	2.23 ± 0.54
K36me1	0.76 ± 0.06	1.32 ± 0.09	0.69 ± 0.02	1.20 ± 0.19
K36me2	0.99 ± 0.01	1.53 ± 0.09	1.10 ± 0.04	1.07 ± 0.05
K36me3	1.15 ± 0.01	0.92 ± 0.02	1.22 ± 0.01	0.97 ± 0.03

C)

H3 methylation motifs	Log2 abundance FC (HDAC/Ctl)			
	Δclr3	Δclr6	Δsir2	Δhos2
K9me1K14un	0.69	1.43	0.60	1.58
K9me2K14un	0.35	2.48	0.48	2.86
K9me3K14un	0.18	1.95	0.51	2.26
K9me2K14ac	2.33	6.11	1.43	3.02
K9me3K14ac	4.26	3.69	1.20	1.93
K27unK36me1	0.74	1.06	0.68	1.19
K27unK36me2	0.99	1.36	1.11	1.07
K27unK36me3	1.15	0.81	1.23	0.97
K27acK36me1	1.24	5.24	0.87	1.35
K27acK36me2	1.08	6.73	0.54	1.07
K27acK36me3	1.08	8.17	0.62	1.12

Figure S4.1

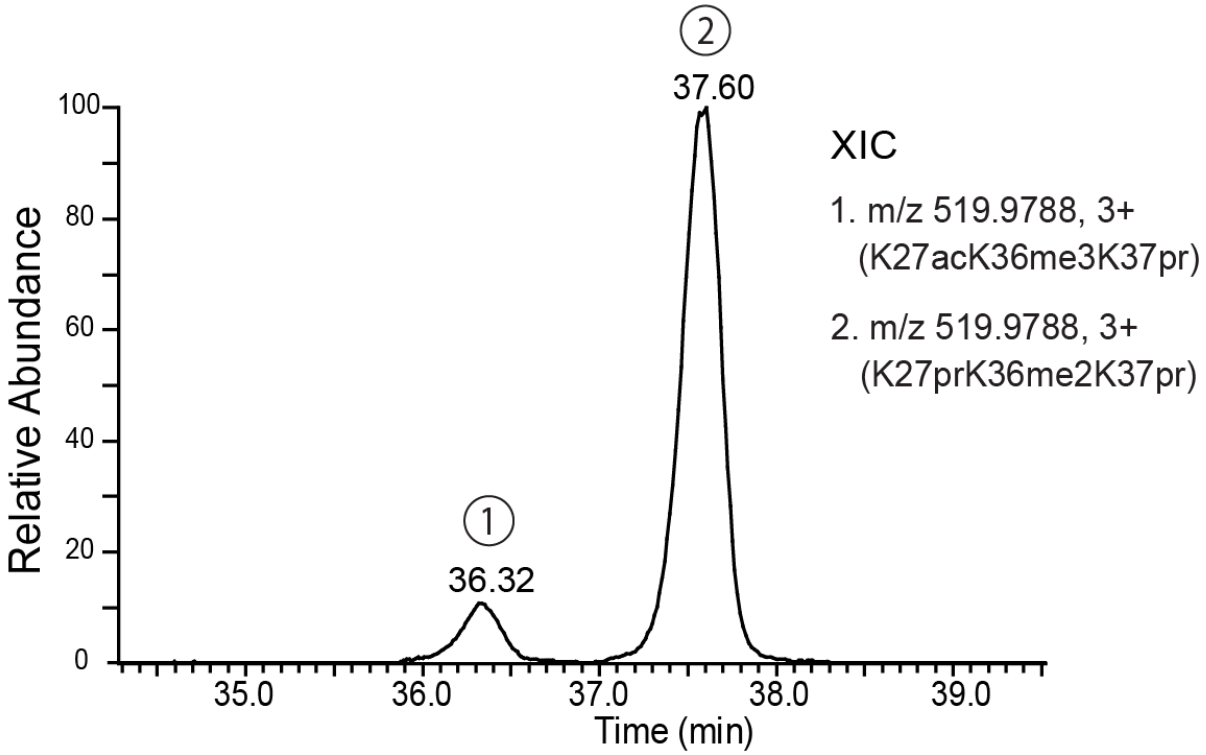


Figure S4.1 Representative XIC of chromatographically resolved isobaric peptides.

Differentially modified isoforms of the peptide 27-KAAPATGGVKKPHR-40 (K27acK36me3K37pr and K27prK36me2K37pr) were chromatographically separated by the RP analytical column.

4.9. References

1. Bannister, A. J. & Kouzarides, T. Regulation of chromatin by histone modifications. *Cell Res* **21**, 381-395 (2011).
2. Berger, S. L. Histone modifications in transcriptional regulation. *Current Opinion in Genetics & Development* **12**, 142-148, (2002).
3. Zhang, Y. & Reinberg, D. Transcription regulation by histone methylation: interplay between different covalent modifications of the core histone tails. *Genes & Development* **15**, 2343-2360, (2001).
4. Huang, H., Sabari, B. R., Garcia, B. A., Allis, C. D. & Zhao, Y. SnapShot: Histone Modifications. *Cell* **159**, 458-458.e451 (2014).
5. Shogren-Knaak, M. *et al.* Histone H4-K16 Acetylation Controls Chromatin Structure and Protein Interactions. *Science* **311**, (2006).
6. Struhl, K. Histone acetylation and transcriptional regulatory mechanisms. *Genes & Development* **12**, 599-606 (1998).
7. Clayton, A. L., Hazzalin, C. A. & Mahadevan, L. C. Enhanced Histone Acetylation and Transcription: A Dynamic Perspective. *Molecular Cell* **23**, 289-296 (2006).
8. Bolger, T., Cohen, T. & Yao, T.-P. in *Gene Expression and Regulation* (ed Jun Ma) Ch. 6, 111-133 (Springer New York, 2006).
9. Lee, K. K. & Workman, J. L. Histone acetyltransferase complexes: one size doesn't fit all. *Nat Rev Mol Cell Biol* **8**, 284-295 (2007).
10. Roth, S. Y., Denu, J. M. & Allis, C. D. HISTONE ACETYLTRANSFERASES. *Annual Review of Biochemistry* **70**, 81-120 (2001).
11. Fukuda, H., Sano, N., Muto, S. & Horikoshi, M. Simple histone acetylation plays a complex role in the regulation of gene expression. *Briefings in Functional Genomics & Proteomics* **5**, 190-208, (2006).
12. Utley, R. T. *et al.* Transcriptional activators direct histone acetyltransferase complexes to nucleosomes. *Nature* **394**, 498-502 (1998).
13. Brown, C. E. *et al.* Recruitment of HAT Complexes by Direct Activator Interactions with the ATM-Related Tra1 Subunit. *Science* **292**, 2333-2337 (2001).
14. Kuo, M.-H., vom Baur, E., Struhl, K. & Allis, C. D. Gcn4 Activator Targets Gcn5 Histone Acetyltransferase to Specific Promoters Independently of Transcription. *Molecular Cell* **6**, 1309-1320 (2000).
15. Hassig, C. A. *et al.* A role for histone deacetylase activity in HDAC1-mediated transcriptional repression. *Proceedings of the National Academy of Sciences* **95**, 3519-3524 (1998).
16. Murakami, Y. *Histone deacetylases govern heterochromatin in every phase*. Vol. 32 (2013).
17. Wang, X., Wei, X., Pang, Q. & Yi, F. Histone deacetylases and their inhibitors: molecular mechanisms and therapeutic implications in diabetes mellitus. *Acta Pharmaceutica Sinica B* **2**, 387-395 (2012).

18. Ruijter, A. J. , Gennip, A. H. , Caron, H. N., Kemp, S. & Kuilenburg, A. B. P.. Histone deacetylases (HDACs): characterization of the classical HDAC family. Vol. 370 (2003).

19. Kasten, M. M., Dorland, S. & Stillman, D. J. A large protein complex containing the yeast Sin3p and Rpd3p transcriptional regulators. *Molecular and Cellular Biology* **17**, 4852-4858 (1997).

20. Zhang, Y., Iratni, R., Erdjument-Bromage, H., Tempst, P. & Reinberg, D. Histone Deacetylases and SAP18, a Novel Polypeptide, Are Components of a Human Sin3 Complex. *Cell* **89**, 357-364, (1997).

21. Kadosh, D. & Struhl, K. Repression by Ume6 Involves Recruitment of a Complex Containing Sin3 Corepressor and Rpd3 Histone Deacetylase to Target Promoters. *Cell* **89**, 365-371, (1997).

22. Rundlett, S. E., Carmen, A. A., Suka, N., Turner, B. M. & Grunstein, M. Transcriptional repression by UME6 involves deacetylation of lysine 5 of histone H4 by RPD3. *Nature* **392**, 831-835 (1998).

23. Gómez, E. B., Nugent, R. L., Laria, S. & Forsburg, S. L. Schizosaccharomyces pombe Histone Acetyltransferase Mst1 (KAT5) Is an Essential Protein Required for Damage Response and Chromosome Segregation. *Genetics* **179**, 757-771, (2008).

24. Gómez, E. B., Espinosa, J. M. & Forsburg, S. L. Schizosaccharomyces pombe mst2+ Encodes a MYST Family Histone Acetyltransferase That Negatively Regulates Telomere Silencing. *Molecular and Cellular Biology* **25**, 8887-8903, (2005).

25. Nugent, R. L. *et al.* Expression profiling of S. pombe acetyltransferase mutants identifies redundant pathways of gene regulation. *BMC genomics* **11**, 59 (2010).

26. Georgakopoulos, T. & Thireos, G. Two distinct yeast transcriptional activators require the function of the GCN5 protein to promote normal levels of transcription. *The EMBO journal* **11**, 4145 (1992).

27. Kim, H. S., Choi, E. S., Shin, J. A., Jang, Y. K. & Park, S. D. Regulation of Swi6/HP1-dependent Heterochromatin Assembly by Cooperation of Components of the Mitogen-activated Protein Kinase Pathway and a Histone Deacetylase Clr6. *Journal of Biological Chemistry* **279**, 42850-42859, (2004).

28. Grewal, S. I. S., Bonaduce, M. J. & Klar, A. J. S. Histone Deacetylase Homologs Regulate Epigenetic Inheritance of Transcriptional Silencing and Chromosome Segregation in Fission Yeast. *Genetics* **150**, 563-576 (1998).

29. Wirén, M. et al. Genome wide analysis of nucleosome density histone acetylation and HDAC function in fission yeast. Vol. 24 (2005).

30. Kurdistani, S. K. & Grunstein, M. Histone acetylation and deacetylation in yeast. *Nat Rev Mol Cell Biol* **4**, 276-284 (2003).

31. Bjerling, P. *et al.* Functional Divergence between Histone Deacetylases in Fission Yeast by Distinct Cellular Localization and In Vivo Specificity. *Molecular and Cellular Biology* **22**, 2170-2181, (2002).

32. Durand-Dubief, M. et al. Specific functions for the fission yeast Sirtuins Hst2 and Hst4 in gene regulation and retrotransposon silencing. Vol. 26 (2007).

33. Haldar, D. & Kamakaka, R. T. Schizosaccharomyces pombe Hst4 Functions in DNA Damage Response by Regulating Histone H3 K56 Acetylation. *Eukaryotic Cell* **7**, 800-813, (2008).

34. Sinha, I., Wirén, M. & Ekwall, K. Genome-wide patterns of histone modifications in fission yeast. *Chromosome Res* **14**, 95-105, (2006).

35. Fuchs, S. M., Krajewski, K., Baker, R. W., Miller, V. L. & Strahl, B. D. Influence of Combinatorial Histone Modifications on Antibody and Effector Protein Recognition. *Current Biology* **21**, 53-58, (2011).

36. Rothbart, S. B. *et al.* Poly-acetylated chromatin signatures are preferred epitopes for site-specific histone H4 acetyl antibodies. *Sci. Rep.* **2**, (2012).

37. Baeza, J. *et al.* Stoichiometry of site-specific lysine acetylation in an entire proteome. *Journal of Biological Chemistry*, (2014).

38. Weinert, B. T. *et al. Acetylation dynamics and stoichiometry in Saccharomyces cerevisiae*. Vol. 10 (2014).

39. Abshiru, Nebiyu, Caron-Lizotte, Olivier, Rajan, Roshan Elizabeth, Jamai, Adil, Pomies, Christelle, Verreault, Alain, Thibault, Pierre. Discovery of protein acetylation patterns by deconvolution of peptide isomer mass spectra. *Nature Communications* 6, 2015.

40. Forsburg, L. Susan, Rhind, Nicholas. Basic methods for fission yeast. **Yeast.** 2006, 23(3):173-83.

41. Guillemette, B. *et al.* H3 Lysine 4 Is Acetylated at Active Gene Promoters and Is Regulated by H3 Lysine 4 Methylation. *PLoS Genet* **7**, e1001354 (2011).

42. Abshiru, N. *et al.* Chaperone-mediated acetylation of histones by Rtt109 identified by quantitative proteomics. *Proteomics* **81**, 80-90 (2013).

43. Kunoh, T., Habu, T. & Matsumoto, T. Involvement of fission yeast Clr6-HDAC in regulation of the checkpoint kinase Cds1. *Nucleic acids research* **36**, 3311-3319 (2008).

44. Tang, H. *et al.* Multiplexed Parallel Reaction Monitoring Targeting Histone Modifications on the QExactive Mass Spectrometer. *Analytical Chemistry* **86**, 5526-5534, (2014).

45. Hansen, K. R. *et al.* Global Effects on Gene Expression in Fission Yeast by Silencing and RNA Interference Machineries. *Molecular and Cellular Biology* **25**, 590-601, (2005).

46. Yamada, T., Fischle, W., Sugiyama, T., Allis, C. D. & Grewal, S. I. S. The Nucleation and Maintenance of Heterochromatin by a Histone Deacetylase in Fission Yeast. *Molecular Cell* **20**, 173-185, (2005).

47. Shankaranarayana, G. D., Motamedi, M. R., Moazed, D. & Grewal, S. I. S. Sir2 Regulates Histone H3 Lysine 9 Methylation and Heterochromatin Assembly in Fission Yeast. *Current Biology* **13**, 1240-1246, (2003).

48. Alper, B. J. *et al. Sir2 is required for Clr4 to initiate centromeric heterochromatin assembly in fission yeast*. Vol. 32 (2013).

49. Strahl, B. D. & Allis, C. D. The language of covalent histone modifications. *Nature* **403**, 41-45 (2000).

50. Zheng, Y., Thomas, P. M. & Kelleher, N. L. Measurement of acetylation turnover at distinct lysines in human histones identifies long-lived acetylation sites. *Nat Commun* **4**, (2013).

51. Garcia, B. A., Pesavento, J. J., Mizzen, C. A. & Kelleher, N. L. Pervasive combinatorial modification of histone H3 in human cells. *Nat Meth* **4**, 487-489, (2007).

52. Britton, L.-M. P., Gonzales-Cope, M., Zee, B. M. & Garcia, B. A. Breaking the histone code with quantitative mass spectrometry. *Expert Review of Proteomics* **8**, 631-643, (2011).

53. Guillemette, B. *et al.* H3 Lysine 4 Is Acetylated at Active Gene Promoters and Is Regulated by H3 Lysine 4 Methylation. *PLoS Genet* **7**, e1001354, (2011).

54. Johnsson, A. *et al. HAT–HDAC interplay modulates global histone H3K14 acetylation in gene-coding regions during stress*. Vol. 10 (2009).

55. Nicolas, E. *et al.* Distinct roles of HDAC complexes in promoter silencing, antisense suppression and DNA damage protection. *Nat Struct Mol Biol* **14**, 372-380, (2007).

56. Castonguay, E. *et al.* Panspecies Small-Molecule Disruptors of Heterochromatin-Mediated Transcriptional Gene Silencing. *Molecular and Cellular Biology* **35**, 662-674, (2015).

57. Wang, A., Kurdistani, S. K. & Grunstein, M. Requirement of Hos2 Histone Deacetylase for Gene Activity in Yeast. *Science* **298**, 1412-1414, (2002).

58. Sharma, V. M., Tomar, R. S., Dempsey, A. E. & Reese, J. C. Histone Deacetylases RPD3 and HOS2 Regulate the Transcriptional Activation of DNA Damage-Inducible Genes. *Molecular and Cellular Biology* **27**, 3199-3210, (2007).

59. Peters, A. H. F. M. *et al.* Partitioning and Plasticity of Repressive Histone Methylation States in Mammalian Chromatin. *Molecular Cell* **12**, 1577-1589, (2003).

60. Rice, J. C. *et al.* Histone Methyltransferases Direct Different Degrees of Methylation to Define Distinct Chromatin Domains. *Molecular Cell* **12**, 1591-1598, (2003).

61. Miao, F. & Natarajan, R. Mapping Global Histone Methylation Patterns in the Coding Regions of Human Genes. *Molecular and Cellular Biology* **25**, 4650-4661, (2005).

CHAPTER 5: Conclusions and Future Perspectives

5.1. Conclusions

The scope of this thesis was the development of MS-based analytical strategies for comprehensive and quantitative analysis of histone modifications. An improved method for the measurement of the global and site specific changes in histones acetylation was developed. This method was supplemented by a new algorithm for deconvolution and quantification of isomeric peptides. We have validated the applicability of these analytical tools in a variety of biologically relevant systems.

The measurement of histone PTMs using the traditional bottom-up approach involves *in vitro* propionylation reaction to derivatize histone proteins followed by trypsin digestion and LC-MS/MS analysis. This approach has been widely used as it allows comprehensive analysis of numerous PTMs with high reproducibility, high dynamic range, and low detection limits. In chapter two of this thesis, an improved comprehensive bottom-up approach was described that combines intact mass analysis with peptide sequencing to characterize histone PTMs. This approach is modified from a similar method that was previously developed in our lab[1]. In the improved approach, peptide sequencing was performed in an LTQ-Orbitrap MS instead of the previously used QTOF instrument. In addition, unlike the previous approach, no MRM assays were needed to validate PTM identifications. The ultra-fast scanning capability of the ion trap coupled with the high resolution and high mass accuracy of the Orbitrap allows rapid sequencing of numerous peptides and accurate quantification of PTM abundances. This is particularly useful for analysis of complex peptide mixtures. In addition, the improved approach greatly reduces the amount of time and energy spent on data acquisition and analysis by avoiding the laborious process of MRM assay development.

We employed the above approach to study the chaperone control of the substrate specificity of *S. cerevisiae* HAT Rtt109. Many of the previous studies conducted in this area focused on few acetylation sites on either H3 or H4, but not both. Moreover, these studies used antibodies raised against specific lysine residues to detect acetylations and as a result quantitative information on the extent of acetylation by Rtt109 was largely unavailable. Our bottom-up proteomics approach, however, surpasses the limitations posed by the traditional antibody-based techniques, and hence allows us to accurately measure the overall changes in

catalytic activity and/or the chaperone-mediated selectivity for individual sites. Our data showed the role of histone chaperones Asf1 and Vps75 in directing the activity of Rtt109 toward distinct H3 and H4 sites. Rtt109 interaction with Vps75 promotes the acetylation of mainly H3K9 and H3K23, with the former being the most preferred site. According to previous genetic and biochemical studies the HAT Gcn5 is also known to acetylate K9 in newly synthesized H3 molecules.[2,3] These studies also suggest that the two HATs, Rtt109 and Gcn5, function in parallel to promote nucleosome assembly and maintain genome integrity in yeast. Our data also showed for the first time that Rtt109-Vps75 exhibits moderate activity toward H4 sites.

In contrast, Rtt109-Asf1 showed a striking specificity for H3K56, but no significant activity toward H4 sites. The later observation was very surprising to us because previous structural analyses have shown that Asf1 binds newly synthesized H3-H4 dimers through its conserved 155-amino acid N-terminal domain[4] and yet, the complex Rtt109-Asf1 selectively acetylates only the H3 molecules. In humans and Drosophila, the Asf1 chaperone interacts with the HAT CBP to promote acetylation of H3K56.[5,6] Humans have two ASF1 isoforms, ASF1A and ASF1B, which differ mainly in their C-terminal region. H3K56 acetylation by Rtt109-Asf1 complex plays a critical role in the regulation of DNA accessibility and transcriptional activation[7,8,9]. A similar chaperone control of the activity of Rtt109 toward the linker histone H1 was recently described.[10] Overall, our results demonstrate the applicability of the MS-based approach in studying enzyme kinetics and in defining how chaperone proteins direct substrate selectivity of HATs.

Although the above MS-based approach was very useful for comprehensive PTM analysis, one of its limitations is that it cannot accurately determine PTM stoichiometries in isomeric histone peptides. Isomers present a major challenge to the conventional bottom-up approach when the localization of PTMs is of special interest, as is the case in most histone proteomics. The identification and quantification of PTMs even in as simple as a mixture of two isomeric peptides requires rigorous and time consuming manual validation. In addition, as can be imagined, the reproducibility of results obtained by manual analysis is highly affected by user experience. In fact, in the case of histone H4, it was totally impossible to distinguish between multiple acetylated isomers based on just their precursor chromatographic peaks and fragment m/z values. These analytical challenges led us to further modify our previous approach and include capabilities for deconvoluting isomeric peptides by the algorithm Iso-PeptidAce. A few previously developed algorithms such as isoScale[11] and EpiProfile[12] rely on

the availability of distinguishing fragment ions in order to efficiently deconvolute isomeric peptides, while Iso-PeptidAce deconvolutes mixed spectra even when there are no such fragments in the MS/MS spectra. The algorithm EpiProfile requires prior knowledge of peptide retention times, whereas Iso-PeptidAce distinguishes between isomers based solely on fragment ion distributions in the mixed spectra.

The applicability of Iso-PeptidAce was validated in a variety of biological problems, which are presented in Chapters three and four. We first investigated the impact of three HDACis (MS-275, SAHA and JNJ-26481585) on the acetylation levels of histones H3 and H4 isolated from human erythroleukemic cells. HDACis are considered one of the most promising therapeutic agents for cancer treatment although their mechanisms of action are poorly characterized.[13] They induce hyperacetylation in histones. Given the crucial roles of histone acetylation in transcription, replication and apoptosis, it is likely that HDACis could target many of these processes in cancer cells. In addition to histones, HDACis also impact several other proteins including transcription factors.[14] Over the last decade, numerous HDACis have been tested for antitumor activities, and among the most widely studied are MS-275 and SAHA. The later was approved by the FDA for treatment of cutaneous T-cell lymphoma.[15] We used Iso-PeptidAce to deconvolute and quantify multiple isomeric peptides derived from H3 and H4 before and after treatment with the drugs. Cells treated with all HDAC inhibitor demonstrated similar patterns of increase in the proportion of acetylated peptides. We determined major increases in acetylation of H3 K18 and K23, and H4 K5, K8, K12 and K16. Between the three HDAC inhibitors, SAHA and JNJ-26481585 induced the most dramatic increase in acetylation of the above lysines. To our knowledge, this is the first study demonstrating the effects of the JNJ-26481585 HDACi on site specific acetylation of histones. The drug was previously shown to have antitumor activity against several types of cancer and is currently in Phase II clinical trial.[16]

We also demonstrated the applicability of Iso-PeptidAce by discovering a novel acetylation pattern of histone H4 bound to budding yeast CAF-1 - a protein complex known to deposit newly synthesized H3/H4 molecules during DNA replication. A number of previous biochemical studies showed that CAF-1-bound new H3 molecules are marked by very abundant acetylation of multiple lysine residues, while new H4 are marked by di-acetylation at positions K5 and K12.[17] Using affinity purification techniques we isolated CAF-1-bound histones, which were subjected to LC-MS/MS analysis and deconvolution by Iso-PeptidAce. While our results for the H3 sites were in agreement with the previous observations, Iso-

PeptidAce analysis of the H4 peptides showed a strikingly different acetylation pattern than the previous studies. Instead of the widely known di-acetylation motif we observed that CAF-1-bound H4 molecules contain the tri-acetylation mark K5_12_16. The occupancy of acetylation at each of these sites was roughly 20%. In contrast, histone H4 purified from bulk nuclear extracts were differentially acetylated at the four sites - K5, K8, K12 and K16 in increasing order of acetylation occupancy (15 to 65%). The later result is consistent with the "zip" model for histone acetylation which states that acetylation of H4 proceeds from K16 to K5.[18] Our data clearly challenged the existing dogma on the acetylation pattern of H4 molecules bound to CAF-1. Although it is still not fully understood what role this acetylation pattern has in DNA replication-dependent histone deposition, our findings provide an important step forward in elucidating its function. On the other hand, these results demonstrate the unique capability of Iso-PeptidAce for studying combinatorial PTM patterns.

Finally, the application of our MS-based approaches was further demonstrated in a study aimed at determining the substrate specificity of several HATs and HDACs presented in chapter 5 of this thesis. We defined the site specificity of some of the well-known fission yeast HATs Gcn5 and Mst2, and HDACs Clr3, Clr6, Sir2, and Hos2. The activities of these enzymes toward histones were previously characterized using biochemical approaches such as Western blot and ChIP assays. However, these methods are prone to numerous technical obstacles including antibody cross-reactivity, and epitope disruption or occlusion when additional modifications are present on nearby residues. The methods are also limited by the capability to provide stoichiometric information on PTM abundances. In our study, we employed the above validated MS-based approaches to determine changes in the acetylation site occupancy of multiple lysine residues in histones H3 and H4 isolated from wild type as well as mutant fission yeast strains lacking the above HATs and HDACs.

Our analysis revealed complementary activities of the different enzymes. Strains lacking Gcn5 had major reductions in H3 K9, K18 and K27 acetylation levels compared to the wild type, while the Mst2 mutant had reduced H3 K14 and K23 acetylations, suggesting complementary roles for the two HATs. Supporting this notion, simultaneous depletion of both HATs strongly reduced the acetylation of all the five sites. Similar roles were observed for the HDACs Clr6 and Hos2. While strains lacking the former HDAC had elevated levels of acetylation of H4 K5, 8, and 12, those lacking the latter exhibited a major increase in H4K16 acetylation compared to the wild type. Given the high speed with which transcriptional machineries progress along the DNA strands, it is likely that the

complementary activities of the HATs and HDACs promote the rapid acetylation and deacetylation of histones for controlling DNA accessibility during and post-transcription.

Perhaps one of the most striking findings of this study is that Clr6 preferentially targets combinatorial acetylation patterns on histone H4. Out of the fifteen different acetylated isoforms derived from the H4 peptide 4-GKGGKGLGKGGAKR-17, only five were significantly affected by Clr6 depletion. And among the five, by far the most affected isoform is the tri-acetylated peptide H4K5_8_12ac. This is one of the first studies demonstrating the preference of HDAC enzymes for combinatorial acetylation motifs. It is almost impossible to obtain such data if not for the exceptional capability of Iso-PeptidAce to deconvolute multiple co-eluting isomers.

It is widely accepted that histone PTMs exert their function through dynamic interplay or crosstalk between various modifications occurring on a single histone molecule or between different histone proteoforms. This notion has led to the formulation of what is known as the 'histone code hypothesis', which states that multiple histone PTMs act combinatorially or sequentially to regulate downstream functions.[18, 19] Supporting this, a number of previous studies have characterized functional crosstalks between acetylation and methylation (e.g. H4R3me and H4K8/K12ac)[20], methylation and phosphorylation (e.g. H3K9me3 and H3S10ph)[21], and methylation and ubiquitination (e.g. H3K4me3 and H2BK120ub)[22].

We also characterized some interesting crosstalk between acetylation and methylation. In both our wild type and mutant samples, the K9 in the H3 peptide 9-KSTGGKAPR-14 was found in unmodified, acetylated, mono-, di-, or tri-methylated form, whereas K14 was found in either unmodified or acetylated form. In the Clr3 mutant, we observed an increase in H3K14ac and a concomitant decrease in H3K9me2/3, suggesting a negative crosstalk between the two PTMs. Such crosstalk has previously been associated with the regulation of heterochromatin structures.[23] While increase in acetylation (and/or decrease in methylation) is strongly linked with chromatin relaxation, a decrease in acetylation (and/or increase in methylation) is widely associated with chromatin condensation. Meanwhile, it should be noted that the PTMs alone do not alter the chromatin structure; rather it is the combined effects of other PTMs as well as nucleosome repositioning by chromatin remodeling complexes. We also discovered an inter-histone positive crosstalk between acetylation and methylation, where both H3K9me2/me3 and H4K16ac were elevated in response to Hos2 depletion. Meanwhile, this result should be viewed with caution and alternative interpretation

should be considered. First, our method cannot determine whether or not the H3K9 and H4K16 residues are present in the same or different nucleosomal complex. As a result, based on the current data, we cannot exclude the fact that this PTMs might not be interacting at all and the increase in H4K16ac and H3K9me2/3 are purely mutually exclusive.

Second, although our analysis showed a concomitant increase in both H3K9me2/me3 and H4K16ac, it is not clear how this interaction could be established. Assuming that both H3K9me2/3 and H4K16ac are present in the same nucleosome, one possibility is that hyperacetylation of H4K16 could have a positive impact on the binding of Clr3 (the enzyme responsible for removing H3K14ac), thus promoting transcriptional inactivation via Clr4 binding and H3K9me2/3 deposition. Alternatively, it is possible that the enzyme responsible for acetylating H4K16 positively regulates the binding of H3K9 methyltranferase Clr4 to co-modify nucleosomes and promote transcriptional repression of target genes. It is also possible that Clr4 could be recruited indirectly through H4K16ac binding proteins.

5.2. Future perspectives

The mysteries of histones and their posttranslational modifications are surely not over. These proteins appear small and simple from their looks; yet, they have constantly challenged our analytical capabilities to fully characterize all their proteoforms (i.e., the forms arising from combinatorial modifications). The arrival of cutting-edge mass spectrometry technologies and continuous improvements in analytical methodologies have so far improved our success in identifying numerous PTMs, in characterizing many enzymes that are responsible for installing or removing the PTMs and in defining the histone PTM interactome. Meanwhile, further improvements in the existing methodologies are required in order to gain more insights into the role of histone PTMs in normal and disease processes.

The bottom-up approach presented in this thesis allows analysis of PTMs that occur on short peptide segments, limiting our capability to decipher long range combinatorial PTMs and crosstalks. To address this limitation, the advantages provided by alternative histone digestion enzymes such as Glu-C and Asp-N should be exploited. These enzymes provide an expanded view of multiple PTMs on the full length N-terminal region of histones.

The algorithm presented in the thesis can be exploited to include many other PTMs in addition to acetylation and methylation. One potential area that needs to be tested is the deconvolution of isomeric phosphopeptides. Our lab has previously shown that 3 to 6% of all identified phosphopeptides in large scale phosphoproteomics studies represent phosphoisomers.[24] Thus, Iso-PeptidAce can be adapted to deconvolute mixed spectra of such isomers.

Iso-PeptidAce can also be used for investigation of multivalent recognition of combinatorial histone PTMs by chromatin factors. Despite the large number of PTMs discovered on the histone tails, there is still a huge gap in understanding how these PTMs are read and interpreted into meaningful biological outcomes. It is widely believed that histone modifications act in combinatorial fashion to regulate the physical properties of the chromatin fiber. According to emerging evidences, this is achieved through the binding of chromatin modules termed as ‘effectors’ which serve to read specific PTM patterns. Thus, when coupled with techniques that allow co-purification of histone-bound chromatin modules, the method described in this Thesis can be used to define the mechanisms involved in PTM

recognition and the role histone PTMs play in directing the establishment of various chromatin-associated processes.

The specificities of HDACs toward histones are far less characterized in humans compared to yeast. HDACs present very promising targets for the treatment of various types of cancers and assessment of their catalytic properties is crucial for understanding their role in the development of the disease. Recently the human HDAC3 was characterized as a novel therapeutic target in multiple myeloma (MM).[25] According to this study, HDAC3 knockdown induced significant MM cell growth inhibition via apoptosis. Treatment of the MM cells with a small molecule HDAC3 inhibitor (denoted as BG45) also triggered similar tumor growth inhibition. Western analysis using antibodies raised against specific lysine residues also showed a major increase in acetylation of core histones H2B, H3 and H4 upon treatment with BG45. However, due to the inherent limitations of the technique, this study did not show if the HDAC3 targets a single lysine residue or a combinatorial acetylation pattern on the core histones. As our results have shown some HDACs specifically target multiply acetylated histones rather than a single site. Therefore, the method we presented in this thesis can now be used to characterize the specificity of HDAC3 before or after treatment with BG45. Moreover, this method can be applied to studies aimed at characterizing the catalytic properties of several other human HDACs.

5.3. References

1. Drogaris P, Wurtele H, Masumoto H, Verreault A, Thibault P. Comprehensive Profiling of Histone Modifications Using a Label-Free Approach and Its Applications in Determining.

2. Burgess, R. and Z. Zhang, *Roles for Gcn5 in promoting nucleosome assembly and maintaining genome integrity.* Cell Cycle, 2010. **9**(15): p. 3051-3057.

3. Burgess, R.J., et al., *A role for Gcn5 in replication-coupled nucleosome assembly.* Molecular cell, 2010. **37**(4): p. 469-480.

4. Daganzo SM, Erzberger JP, Lam WM, Skordalakes E, Zhang R, Franco AA, Brill SJ, Adams PD, Berger JM, Kaufman PD. 2003. Structure and function of the conserved core of histone deposition protein Asf1. Curr. Biol. 13:2148 –2158.

5. Groth, Anja, Corpet, Armelle, Cook, Adam J. L., Roche, Daniele, Bartek, Jiri, Lukas, Jiri, Almouzni, Geneviève, *Regulation of Replication Fork Progression Through Histone Supply and Demand.* Science, 2007. **318**(5858): p. 1928-1931.

6. Das, Chandrima, Roy, Siddhartha, Namjoshi, Sarita, Malarkey, Christopher S., Jones, David N. M., Kutateladze, Tatiana G., Churchill, Mair E. A., Tyler, Jessica K., *Binding of the histone chaperone ASF1 to the CBP bromodomain promotes histone acetylation.* Proceedings of the National Academy of Sciences, 2014. **111**(12): p. E1072-E1081.

7. Masumoto, Hiroshi, Hawke, David, Kobayashi, Ryuji, Verreault, Alain, *A role for cell-cycle-regulated histone H3 lysine 56 acetylation in the DNA damage response.* Nature, 2005. **436**(7048): p. 294-298.

8. North, Justin A, Shimko, John C, Javaid, Sarah, Mooney, Alex M, Shoffner, Matthew A, Rose, Sean D, Bundschuh, Ralf, Fishel, Richard, Ottesen, Jennifer J, Poirier, Michael G, *Regulation of the nucleosome unwrapping rate controls DNA accessibility.* Nucleic acids research, 2012; 40(20):10215-27..

9. Shimko, John C, North, Justin A, Bruns, Aaron N, Poirier, Michael G, Ottesen, Jennifer J, *Preparation of full synthetic histone H3 reveals that acetyl-lysine 56 facilitates protein binding within nucleosomes.* Journal of molecular biology, 2011. **408**(2): p. 187-204.

10. Radovani E, Cadorin M, Shams T, El-Rass S, Karsou AR, Kim HS, Kurat CF, Keogh MC, Greenblatt JF, Fillingham JS. *The Carboxyl Terminus of Rtt109 Functions in Chaperone Control of Histone Acetylation.* Eukaryotic Cell, 2013. **12**(5): p. 654-664.

11. Sidoli, S. *et al.* Middle-down hybrid chromatography/tandem mass spectrometry workflow for characterization of combinatorial post-translational modifications in histones. *PROTEOMICS* 14, 2200-2211, doi:10.1002/pmic.201400084 (2014).

12. Yuan, Z.-F. *et al.* EpiProfile Quantifies Histone Peptides With Modifications by Extracting Retention Time and Intensity in High-resolution Mass Spectra. *Molecular & Cellular Proteomics* 14, 1696-1707, doi:10.1074/mcp.M114.046011 (2015).

13. Minucci, S. and P.G. Pelicci, *Histone deacetylase inhibitors and the promise of epigenetic (and more) treatments for cancer.* Nat Rev Cancer, 2006. **6**(1): p. 38-51.

14. Richon, Victoria M., Sandhoff, Todd W., Rifkind, Richard A., Marks, Paul A., *Histone deacetylase inhibitor selectively induces p21WAF1 expression and gene-associated histone acetylation.* Proceedings of the National Academy of Sciences, 2000. **97**(18): p. 10014-10019.

15. Mann, Bhupinder S., Johnson, John R., Cohen, Martin H., Justice, Robert, Pazdur, Richard, *FDA Approval Summary: Vorinostat for Treatment of Advanced Primary Cutaneous T-Cell Lymphoma.* The Oncologist, 2007. **12**(10): p. 1247-1252.

16. Deleu, S. et al. The effects of JNJ-26481585, a novel hydroxamate-based histone deacetylase inhibitor, on the development of multiple myeloma in the 5T2MM and 5T33MM murine models. *Leukemia* 23, 1894-1903 (2009).

17. Sobel, R.E., et al., *Conservation of deposition-related acetylation sites in newly synthesized histones H3 and H4.* Proceedings of the National Academy of Sciences, 1995. **92**(4): p. 1237-1241.

18. Strahl BD, Allis CD. The language of covalent histone modifications. Nature. 2000; 403:41–45.

19. Jenuwein T, Allis CD. Translating the histone code. Science. 2001; 293:1074–1080.

20. Wang, Hengbin, Huang, Zhi-Qing, Xia, Li, Feng, Qin, Erdjument-Bromage, Hediye, Strahl, Brian D., Briggs, Scott D., Allis, C. David, Wong, Jiemin, Tempst, Paul, Zhang, Yi, *Methylation of Histone H4 at Arginine 3 Facilitating Transcriptional Activation by Nuclear Hormone Receptor.* Science, 2001. **293**(5531): p. 853-857.

21. Fischle, Wolfgang, Tseng, Boo Shan, Dormann, Holger L., Ueberheide, Beatrix M., Garcia, Benjamin A., Shabanowitz, Jeffrey, Hunt, Donald F., Funabiki, Hironori, Allis, C. David, *Regulation of HP1-chromatin binding by histone H3 methylation and phosphorylation.* Nature, 2005. **438**(7071): p. 1116-1122.

22. Nakanishi, Shima, Lee, Jung Shin, Gardner, Kathryn E., Gardner, Jennifer M., Takahashi, Yoh-hei, Chandrasekharan, Mahesh B., Sun, Zu-Wen, Osley, Mary Ann, Strahl, Brian D., Jaspersen, Sue L., Shilatifard, Ali, *Histone H2BK123 monoubiquitination is the critical determinant for H3K4 and H3K79 trimethylation by COMPASS and Dot1.* The Journal of Cell Biology, 2009. **186**(3): p. 371-377.

23. Eberharter, A. and P.B. Becker, *Histone acetylation: a switch between repressive and permissive chromatin.* EMBO reports, 2002. **3**(3): p. 224-229.

24. Courcelles, Mathieu, Bridon, Gaëlle, Lemieux, Sébastien, Thibault, Pierre, *Occurrence and Detection of Phosphopeptide Isomers in Large-Scale Phosphoproteomics Experiments.* Journal of Proteome Research, 2012. **11**(7): p. 3753-3765.

25. Minami, J., Suzuki, R., Mazitschek, R., Gorgun, G., Ghosh, B., Cirstea, D., Hu, Y., Mimura, N., Ohguchi, H., Cottini, F., Jakubikova, J., Munshi, N. C., Haggarty, S. J., Richardson, P. G., Hideshima, T., Anderson, K. C., Histone deacetylase 3 as a novel therapeutic target in multiple myeloma. Leukemia, 2014. **28**(3): p. 680-689.

Appendix A: Additional supplementary information

Table S1.1 Human histone variants (*HI*stome-The *H*istone *I*nfobase).

A) Histone H2A variants

Variant (UniprotKB recommended name)	UniprotKB Accession	No. of coding gene(s)
Core histone macro-H2A.1	O75367	1 (H2AFY)
Core histone macro-H2A.2	Q9P0M6	1 (H2AFY2)
Histone H2A type 1	P0C0S8	5 (HIST1H2AI, HIST1H2AK, HIST1H2AL, HIST1H2AM, HIST1H2AG)
Histone H2A type 1-A	Q96QV6	1 (HIST1H2AA)
Histone H2A type 1-B/E	P04908	2 (HIST1H2AE, HIST1H2AB)
Histone H2A type 1-C	Q93077	1 (HIST1H2AC)
Histone H2A type 1-D	P20671	1 (HIST1H2AD)
Histone H2A type 1-H	Q96KK5	1 (HIST1H2AH)
Histone H2A type 1-J	Q99878	1 (HIST1H2AJ)
Histone H2A type 2-A	Q6FI13	2 (HIST2H2AA4, HIST2H2AA3)
Histone H2A type 2-B	Q8IUE6	1 (HIST2H2AB)
Histone H2A type 2-C	Q16777	1 (HIST2H2AC)
Histone H2A type 3	Q7L7L0	1 (HIST3H2A)
Histone H2A-Bbd type 1	P0C5Y9	1 (H2AFB1)
Histone H2A-Bbd type 2/3	P0C5Z0	2 (H2AFB2, H2AFB3)
Histone H2A.J	Q9BTM1	1 (H2AFJ)
Histone H2A.V	Q71UI9	1 (H2AFV)
Histone H2A.x	P16104	1 (H2AFX)
Histone H2A.Z	P0C0S5	1 (H2AFZ)

B). Histone H2B

Variant (UniprotKB recommended name)	UniprotKB Accession	No. of coding gene(s)
Histone H2B type 1-A	Q96A08	1 (HIST1H2BA)
Histone H2B type 1-B	P33778	1 (HIST1H2BB)
Histone H2B type 1-C/E/F/G/I	P62807	5 (HIST1H2BG, HIST1H2BF, HIST1H2BE, HIST1H2BI, HIST1H2BC)
Histone H2B type 1-D	P58876	1 (HIST1H2BD)
Histone H2B type 1-H	Q93079	1 (HIST1H2BH)
Histone H2B type 1-J	P06899	1 (HIST1H2BJ)
Histone H2B type 1-K	O60814	1 (HIST1H2BK)
Histone H2B type 1-L	Q99880	1 (HIST1H2BL)
Histone H2B type 1-M	Q99879	1 (HIST1H2BM)
Histone H2B type 1-N	Q99877	1 (HIST1H2BN)
Histone H2B type 1-O	P23527	1 (HIST1H2BO)
Histone H2B type 2-E	Q16778	1 (HIST2H2BE)
Histone H2B type 2-F	Q5QNW6	1 (HIST2H2BF)
Histone H2B type 3-B	Q8N257	1 (HIST3H2BB)
Histone H2B type F-M	P0C1H6	1 (H2BFM)
Histone H2B type F-S	P57053	1 (H2BFS)
Histone H2B type W-T	Q7Z2G1	1 (H2BFWT)
Putative histone H2B type 2-C	Q6DN03	1 (HIST2H2BC)
Putative histone H2B type 2-D	Q6DRA6	1 (HIST2H2BD)

C) Histone H3

Variant (UniprotKB recommended name)	UniprotKB Accession	No. of coding gene(s)
Histone H3-like centromeric protein A	P49450	1 (CENPA)
Histone H3.1	P68431	10 (HIST1H3A, HIST1H3D, HIST1H3C, HIST1H3E, HIST1H3I, HIST1H3G, HIST1H3J, HIST1H3H, HIST1H3B, HIST1H3F)
Histone H3.1t	Q16695	1 (HIST3H3)
Histone H3.2	Q71DI3	3 (HIST2H3C, HIST2H3A, HIST2H3D)
Histone H3.3	P84243	2 (H3F3A, H3F3B)
Histone H3.3C	Q6NXT2	1 (H3F3C)

A) Histone H3

```
H31_MOUSE 1        MARTKQTARKSTGGKAPRKQLATKAARKSAPATGGVKKPHRYRPGTVALREIRRYQKSTE   60
H31_HUMAN 1        MARTKQTARKSTGGKAPRKQLATKAARKSAPATGGVKKPHRYRPGTVALREIRRYQKSTE   60
H32_ARATH 1        MARTKQTARKSTGGKAPRKQLATKAARKSAPATGGVKKPHRFRPGTVALREIRKYQKSTE   60
H3_YEAST  1        MARTKQTARKSTGGKAPRKQLASKAARKSAPSTGGVKKPHRYKPGTVALREIRRFQKSTE   60
H32_XENLA 1        MARTKQTARKSTGGKAPRKQLATKAARKSAPATGGVKKPHRYRPGTVALREIRRYQKSTE   60
                   ********************:********:*********::**********::*****

H31_MOUSE 61       LLIRKLPFQRLVREIAQDFKTDLRFQSSAVMALQEACEAYLVGLFEDTNLCAIHAKRVTI  120
H31_HUMAN 61       LLIRKLPFQRLVREIAQDFKTDLRFQSSAVMALQEACEAYLVGLFEDTNLCAIHAKRVTI  120
H32_ARATH 61       LLIRKLPFQRLVREIAQDFKTDLRFQSSAVAALQEAAEAYLVGLFEDTNLCAIHAKRVTI  120
H3_YEAST  61       LLIRKLPFQRLVREIAQDFKTDLRFQSSAIGALQESVEAYLVSLFEDTNLAAIHAKRVTI  120
H32_XENLA 61       LLIRKLPFQRLVREIAQDFKTDLRFQSSAVMALQEASEAYLVGLFEDTNLCAIHAKRVTI  120
                   ****************************: ****: *****.*******.*********

H31_MOUSE 121      MPKDIQLARRIRGERA                                              136
H31_HUMAN 121      MPKDIQLARRIRGERA                                              136
H32_ARATH 121      MPKDIQLARRIRGERA                                              136
H3_YEAST  121      QKKDIKLARRLRGERS                                              136
H32_XENLA 121      MPKDIQLARRIRGERA                                              136
                     ***:****:****:
```

B) Histone H4

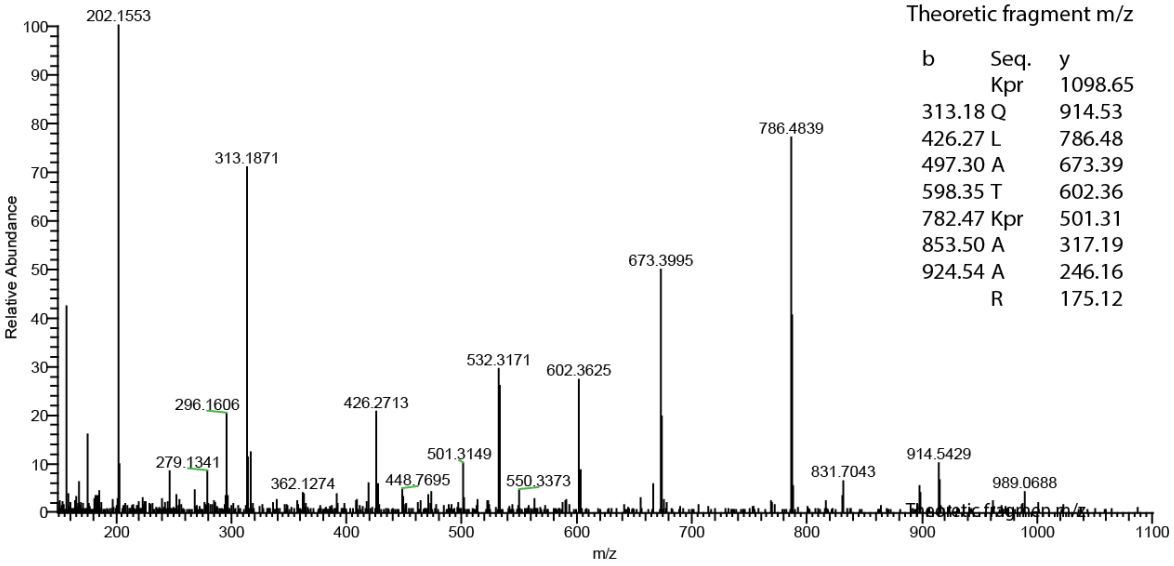
```
H4_HUMAN 1         MSGRGKGGKGLGKGGAKRHRKVLRDNIQGITKPAIRRLARRGGVKRISGLIYEETRGVLK   60
H4_MOUSE 1         MSGRGKGGKGLGKGGAKRHRKVLRDNIQGITKPAIRRLARRGGVKRISGLIYEETRGVLK   60
H4_YEAST 1         MSGRGKGGKGLGKGGAKRHRKILRDNIQGITKPAIRRLARRGGVKRISGLIYEEVRAVLK   60
H4_XENLA 1         MSGRGKGGKGLGKGGAKRHRKVLRDNIQGITKPAIRRLARRGGVKRISGLIYEETRGVLK   60
H4_ARATH 1         MSGRGKGGKGLGKGGAKRHRKVLRDNIQGITKPAIRRLARRGGVKRISGLIYEETRGVLK   60
                   ********************:*********************************.*.***

H4_HUMAN 61        VFLENVIRDAVTYTEHAKRKTVTAMDVVYALKRQGRTLYGFGG                    103
H4_MOUSE 61        VFLENVIRDAVTYTEHAKRKTVTAMDVVYALKRQGRTLYGFGG                    103
H4_YEAST 61        SFLESVIRDSVTYTEHAKRKTVTSLDVVYALKRQGRTLYGFGG                    103
H4_XENLA 61        VFLENVIRDAVTYTEHAKRKTVTAMDVVYALKRQGRTLYGFGG                    103
H4_ARATH 61        IFLENVIRDAVTYTEHARRKTVTAMDVVYALKRQGRTLYGFGG                    103
                    ***.****:*******:*****::******************
```

Figure S1.1 Sequence alignment of histones from various species

Figure S3.7 MS/MS spectra of H3 and H4 peptide isomer

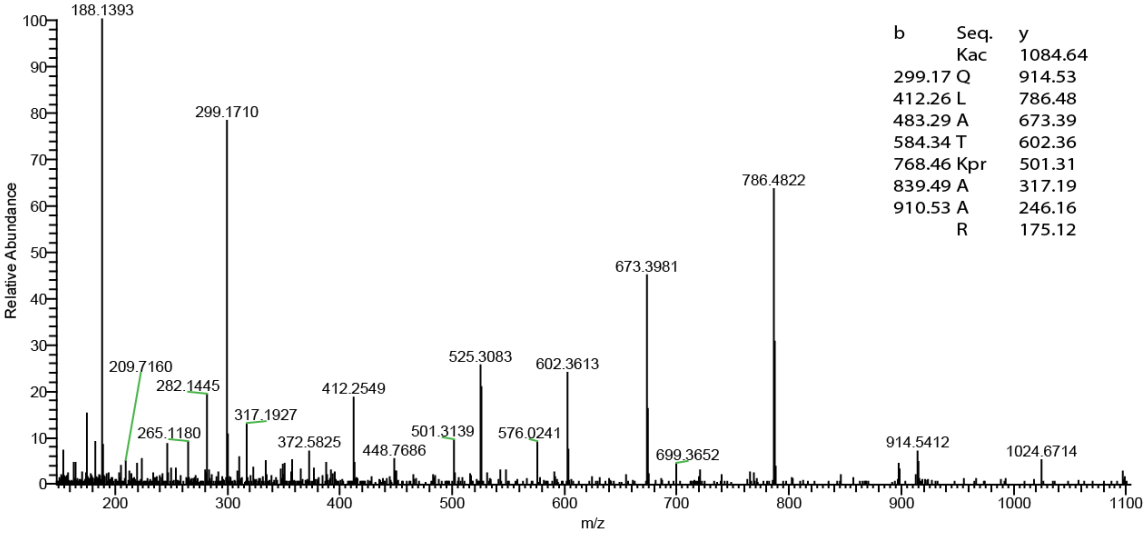
Peptide-1: H3K18pr_K23pr
K(pr)QLATK(pr)AAR

MS/MS 549.84^{2+}



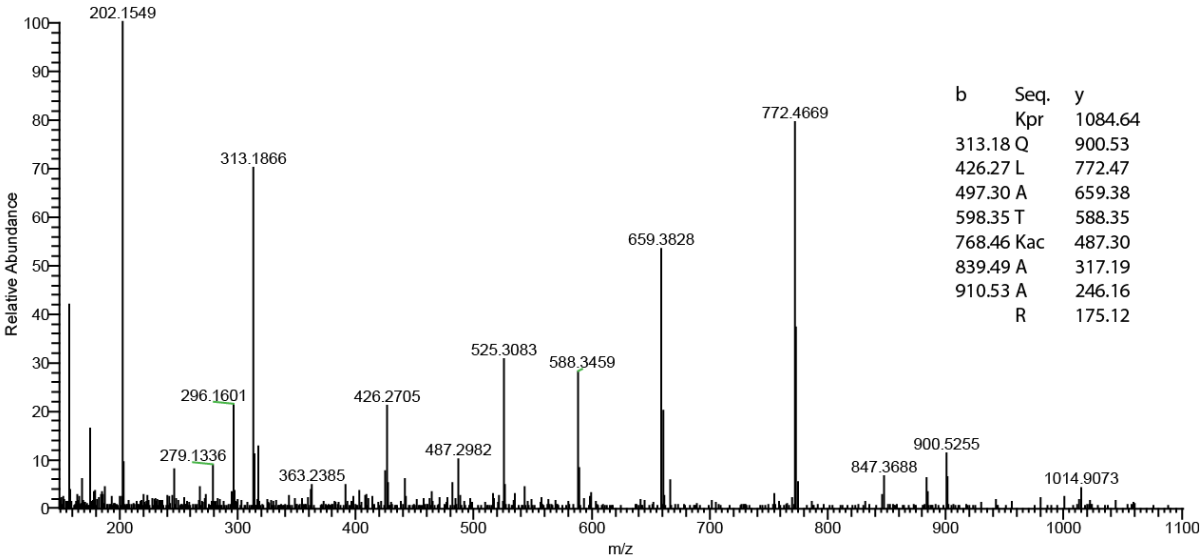
Peptide-2: H3K18ac_K23pr
K(ac)QLATK(pr)AAR

MS/MS 542.83^{2+}



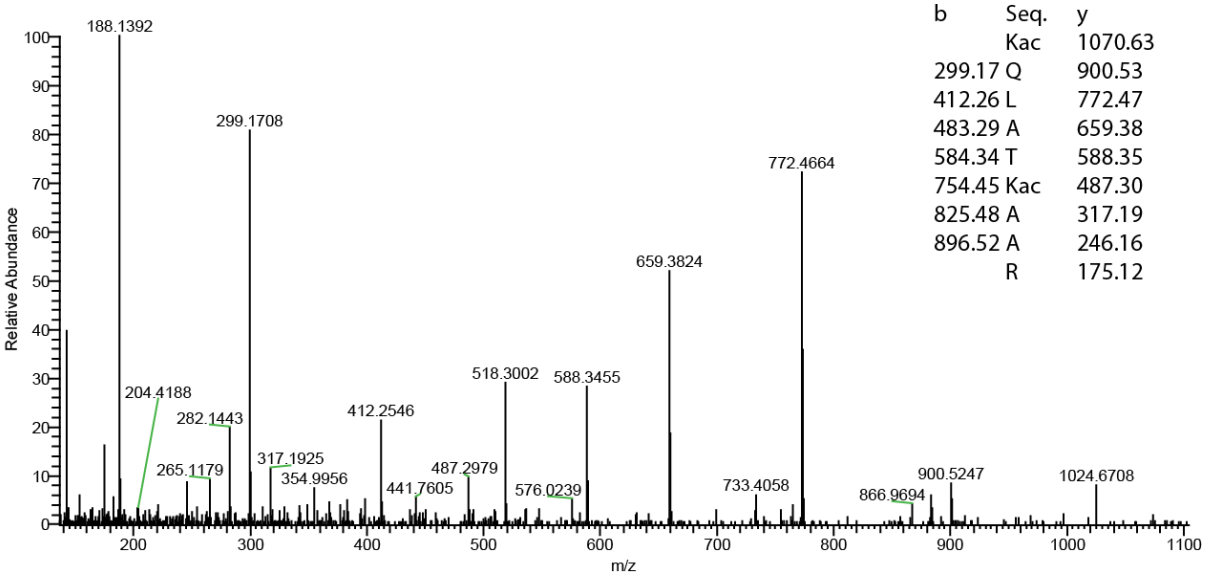
Peptide-3: H3K18pr_K23ac
K(pr)QLATK(ac)AAR

MS/MS 542.83^{2+}



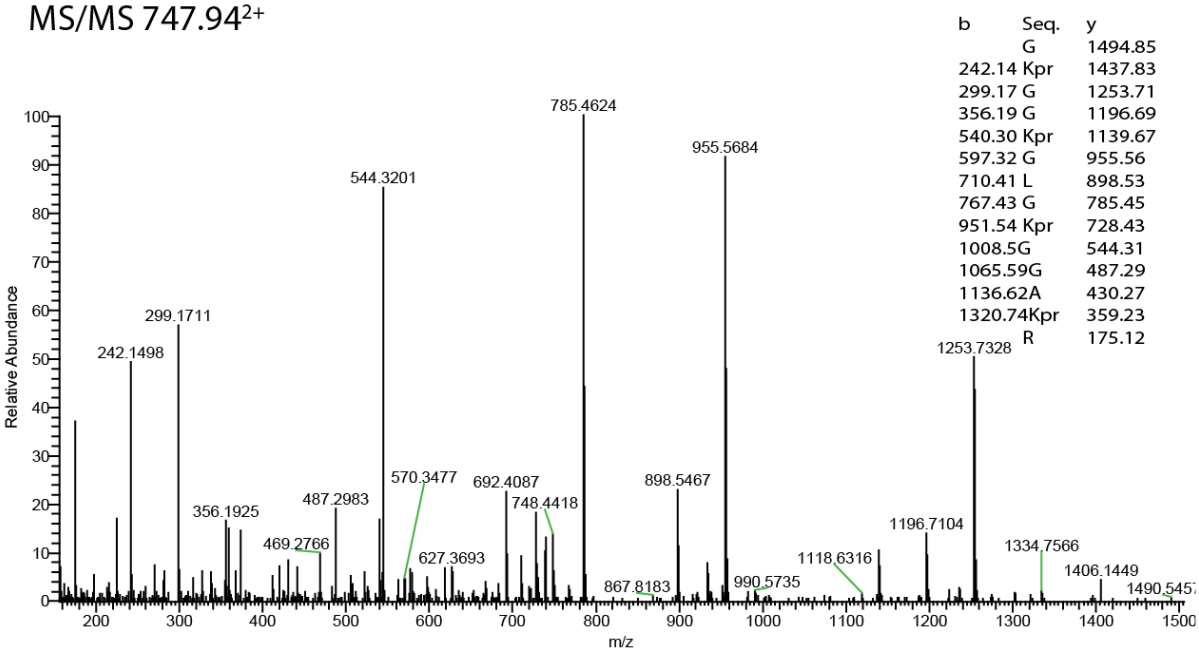
Peptide-4: H3K18ac_K23ac
K(ac)QLATK(ac)AAR

MS/MS 535.82^{2+}



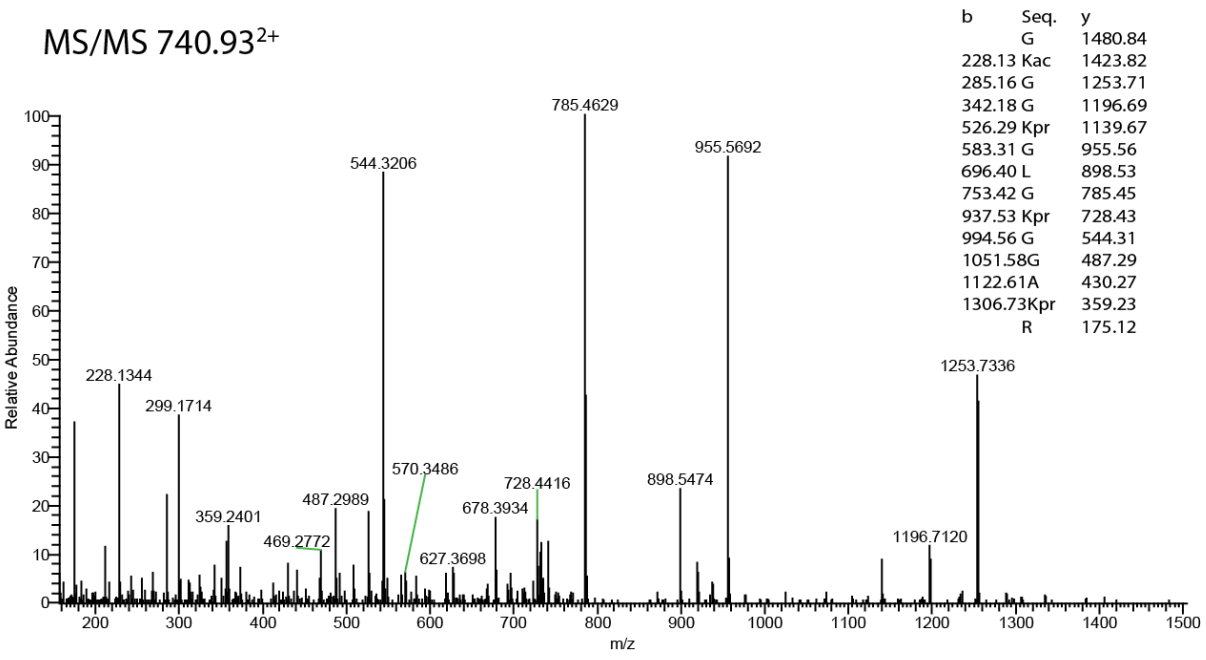
Peptide-5: H4K5_8_12_16pr
GK(pr)GGK(pr)GLGK(pr)GGAK(pr)R

MS/MS 747.94^{2+}



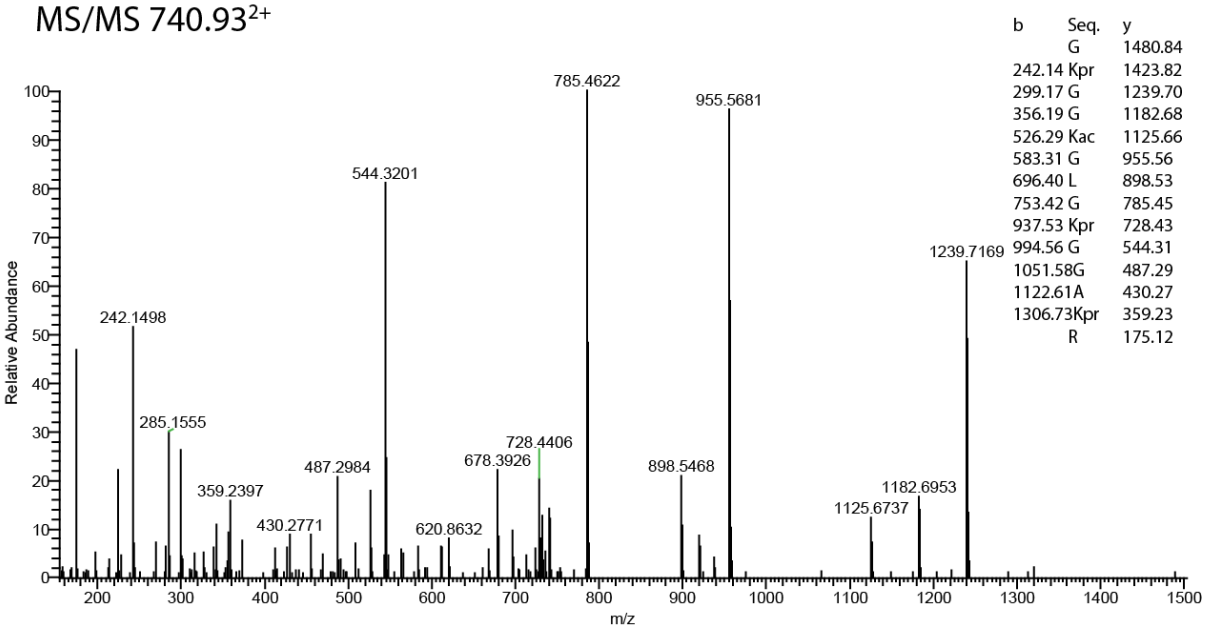
Peptide-6: H4K5ac/K8_12_16pr
GK(ac)GGK(pr)GLGK(pr)GGAK(pr)R

MS/MS 740.93^{2+}



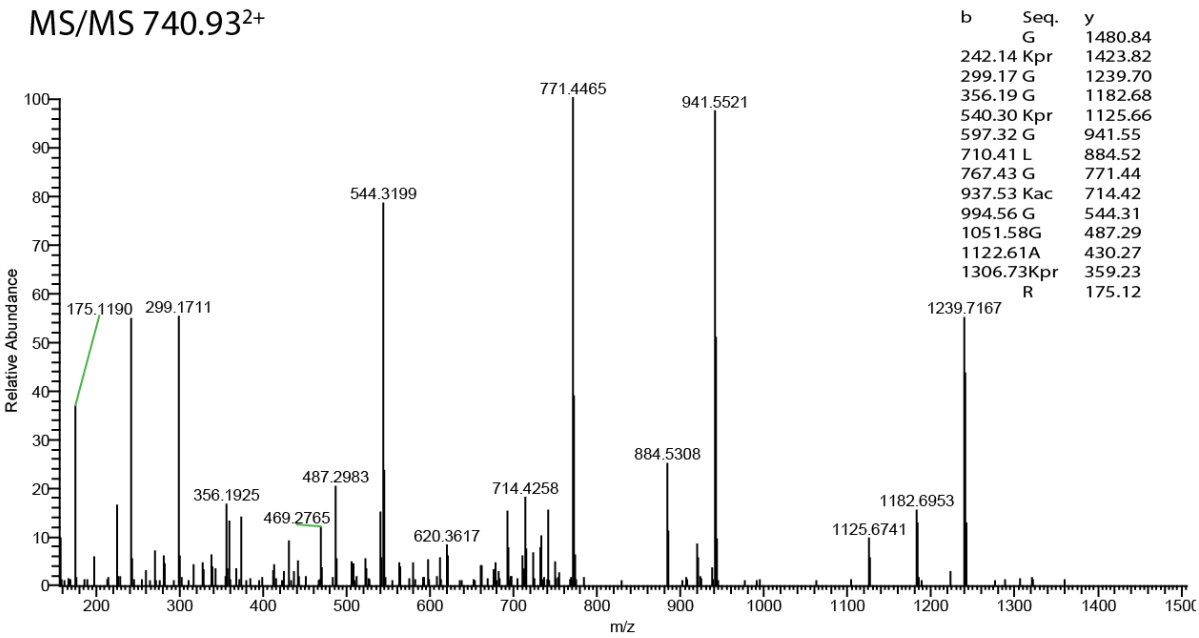
Peptide-7: H4K8ac/K5_12_16pr
GK(pr)GGK(ac)GLGK(pr)GGAK(pr)R

MS/MS 740.93^{2+}



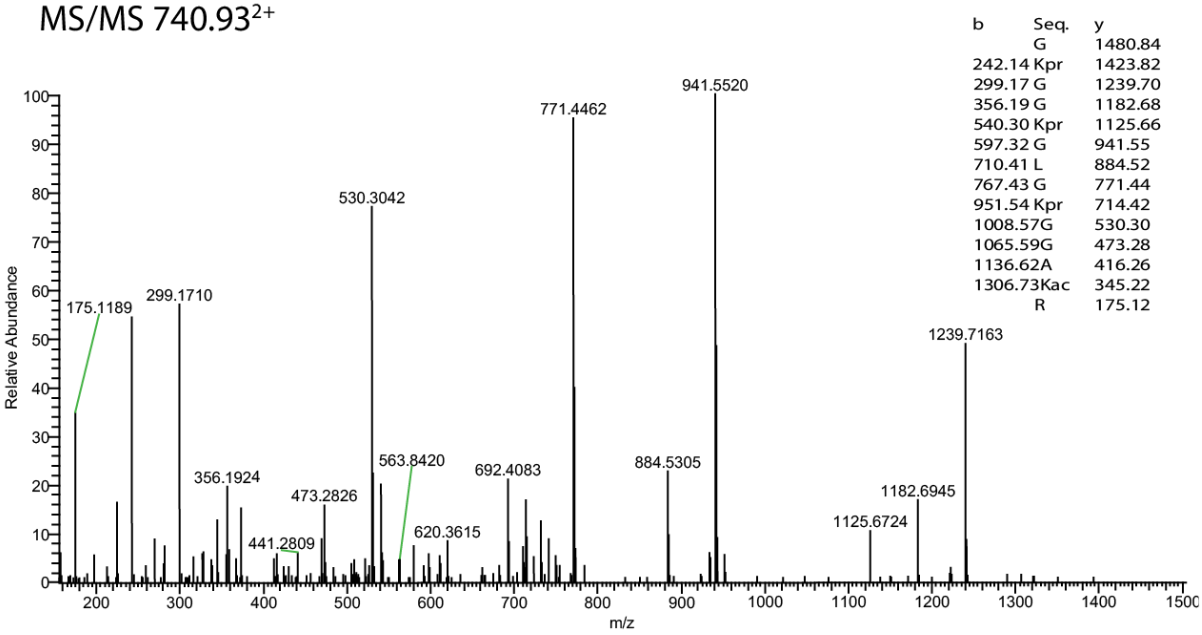
Peptide-8: H4K12ac/K5_8_16pr
GK(pr)GGK(pr)GLGK(ac)GGAK(pr)R

MS/MS 740.93^{2+}



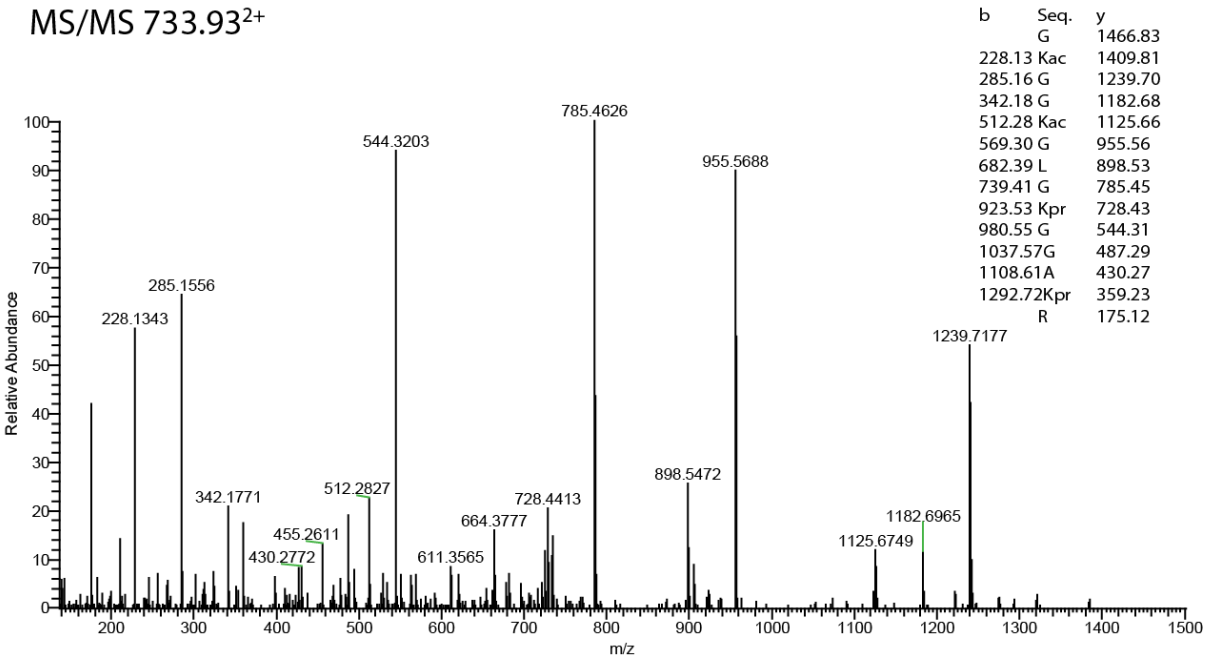
Peptide-9: H4K16ac/K5_8_12pr
GK(pr)GGK(pr)GLGK(pr)GGAK(ac)R

MS/MS 740.93^{2+}



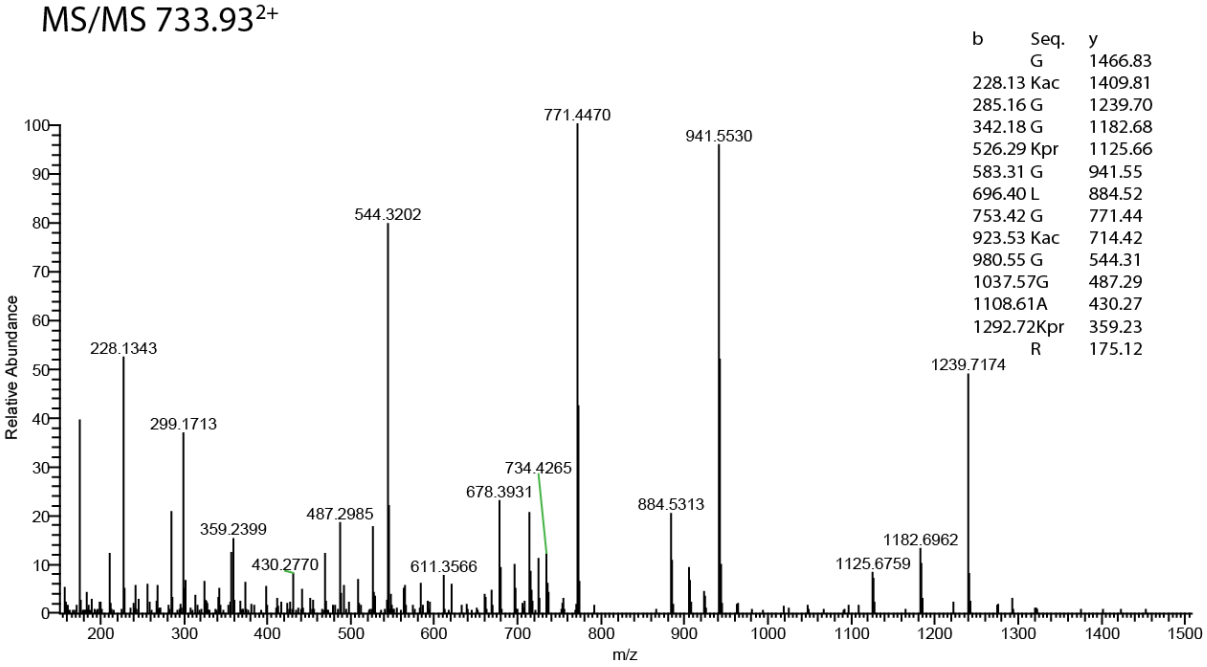
Peptide-10: H4K5_8ac/K12_16pr
GK(ac)GGK(ac)GLGK(pr)GGAK(pr)R

MS/MS 733.93^{2+}



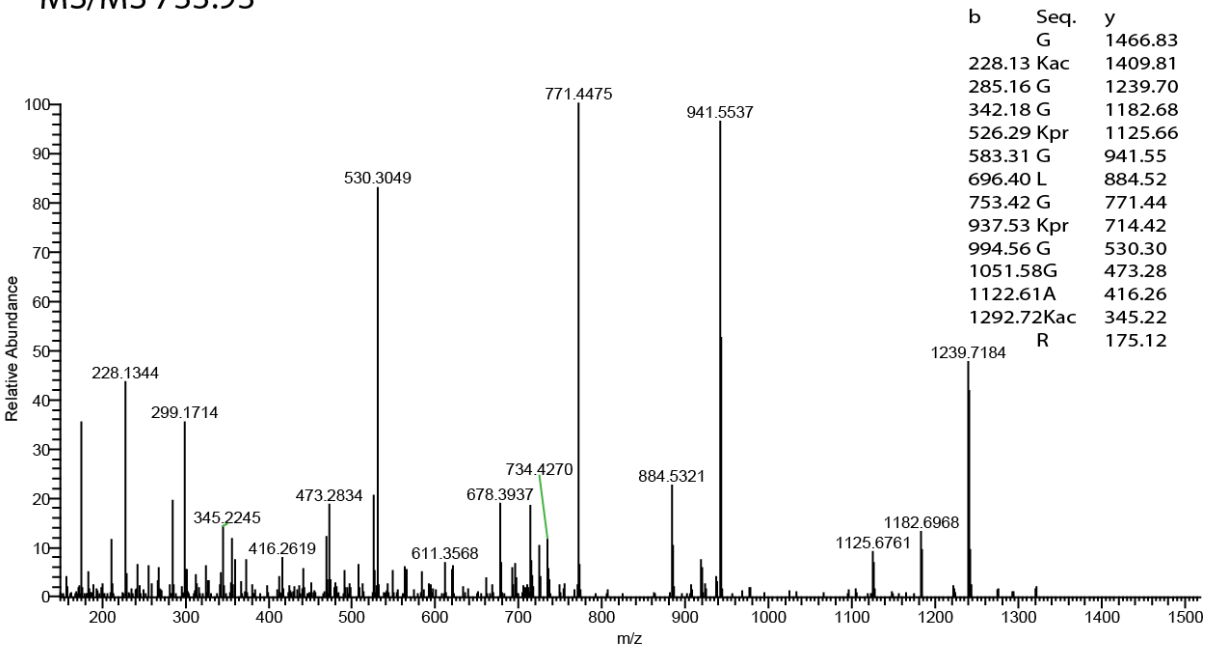
Peptide-11: H4K5_12ac/K8_16pr
GK(ac)GGK(pr)GLGK(ac)GGAK(pr)R

MS/MS 733.93^{2+}



Peptide-12: H4K5_16ac/K8_12pr
GK(ac)GGK(pr)GLGK(pr)GGAK(ac)R

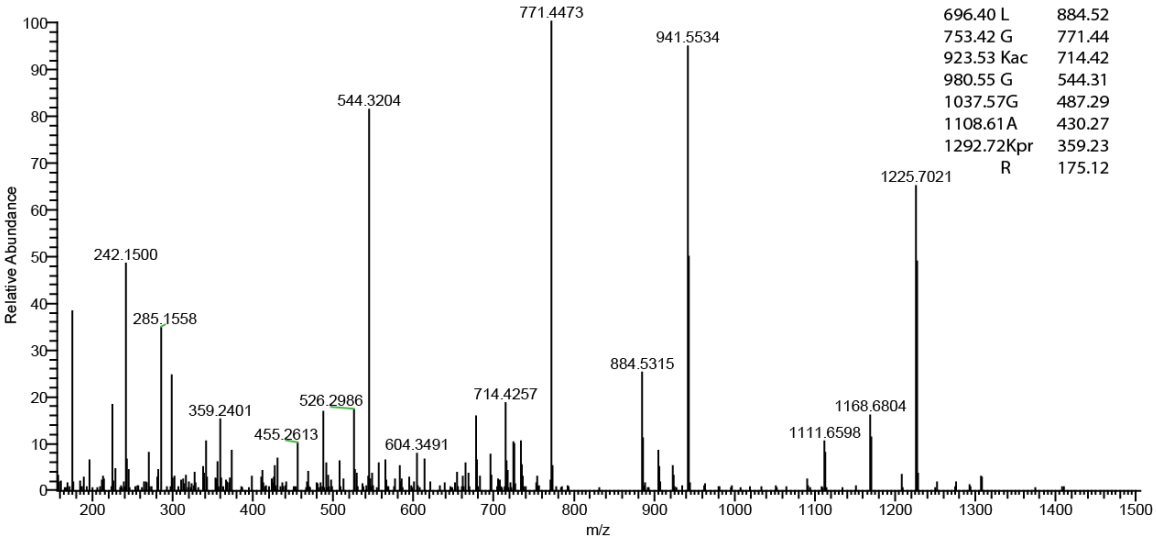
MS/MS 733.93^{2+}



Peptide-13: H4K8_12ac/K5_16pr
GK(pr)GGK(ac)GLGK(ac)GGAK(pr)R

MS/MS 733.93^{2+}

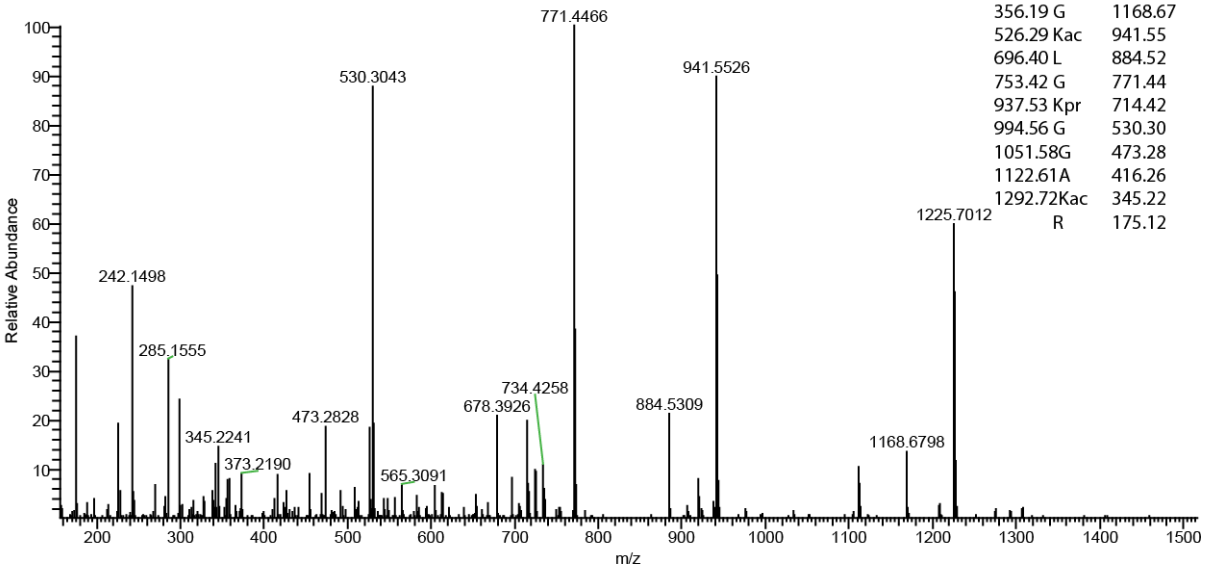
b	Seq.	y
	G	1466.83
242.14	Kpr	1409.81
299.17	G	1225.69
356.19	G	1168.67
526.29	Kac	1111.65
583.31	G	941.55
696.40	L	884.52
753.42	G	771.44
923.53	Kac	714.42
980.55	G	544.31
1037.57	G	487.29
1108.61	A	430.27
1292.72	Kpr	359.23
	R	175.12



Peptide-14: H4K8_16ac/K5_12pr
GK(pr)GGK(ac)GLGK(pr)GGAK(ac)R

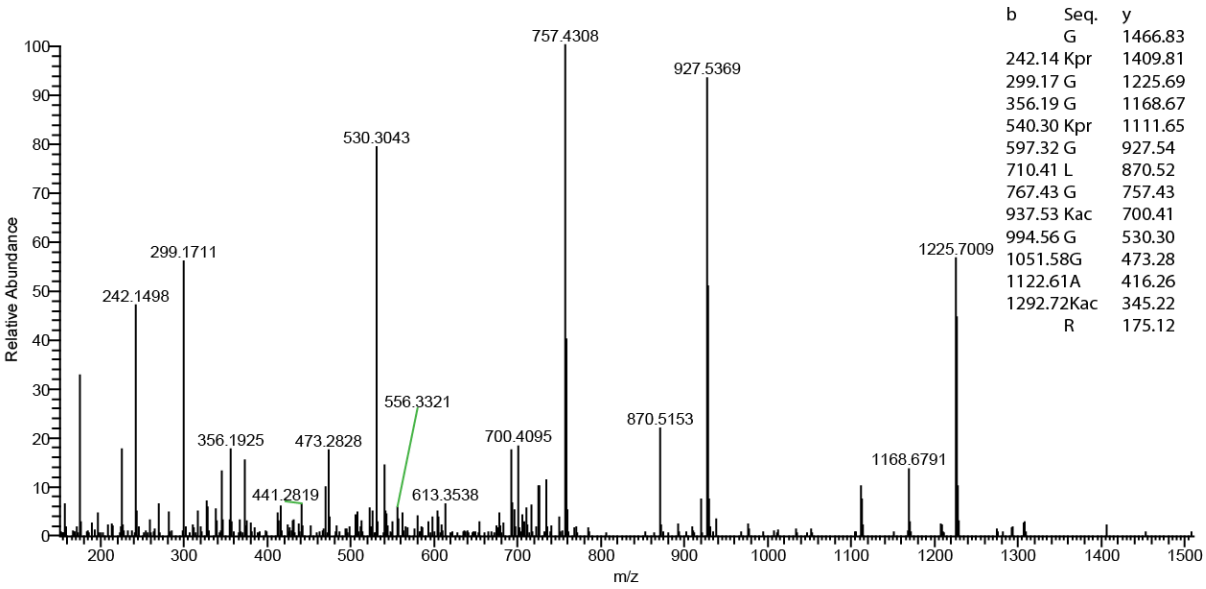
MS/MS 733.93^{2+}

b	Seq.	y
	G	1466.83
242.14	Kpr	1409.81
299.17	G	1225.69
356.19	G	1168.67
526.29	Kac	941.55
696.40	L	884.52
753.42	G	771.44
937.53	Kpr	714.42
994.56	G	530.30
1051.58	G	473.28
1122.61	A	416.26
1292.72	Kac	345.22
	R	175.12



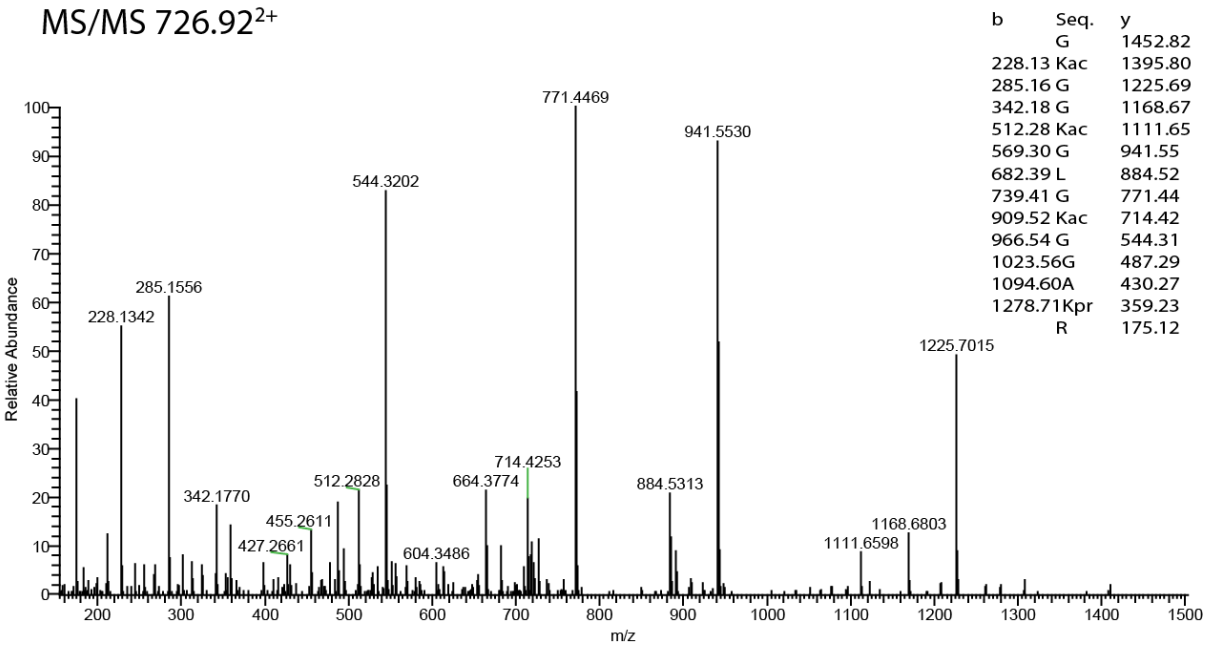
Peptide-15: H4K12_16ac/K5_8pr
GK(pr)GGK(pr)GLGK(ac)GGAK(ac)R

MS/MS 733.93^{2+}



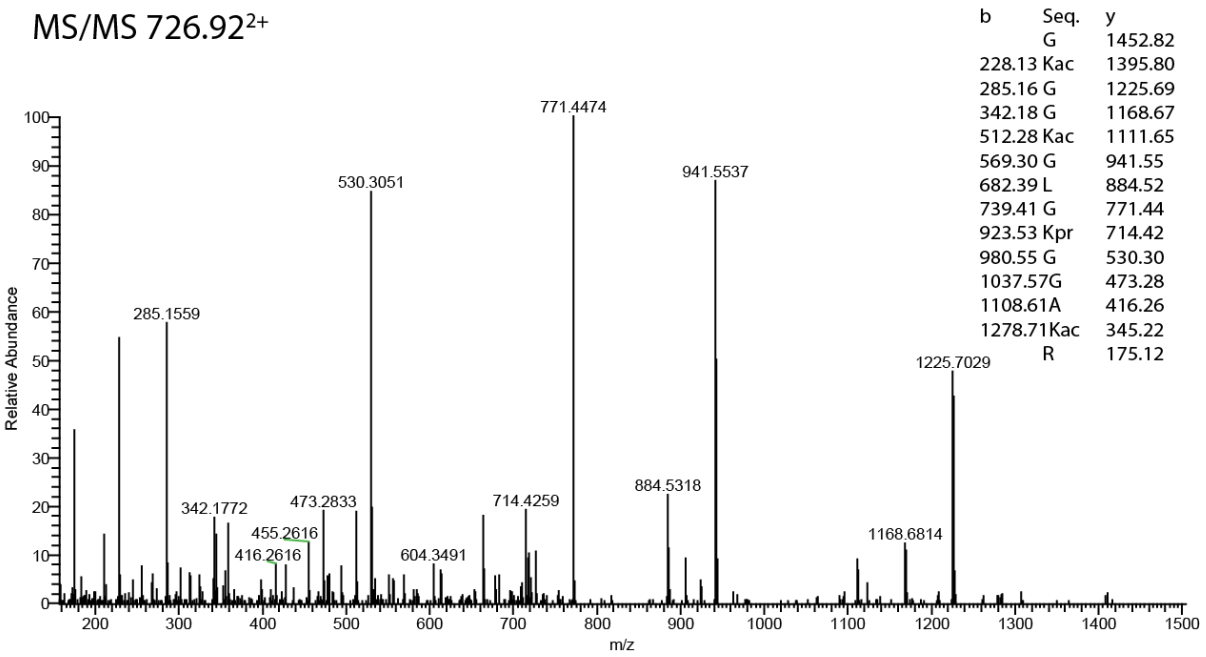
Peptide-16: H4K5_8_12ac/K16pr
GK(ac)GGK(ac)GLGK(ac)GGAK(pr)R

MS/MS 726.92^{2+}



Peptide-17: H4K5_8_16ac/K12pr
GK(ac)GGK(ac)GLGK(pr)GGAK(ac)R

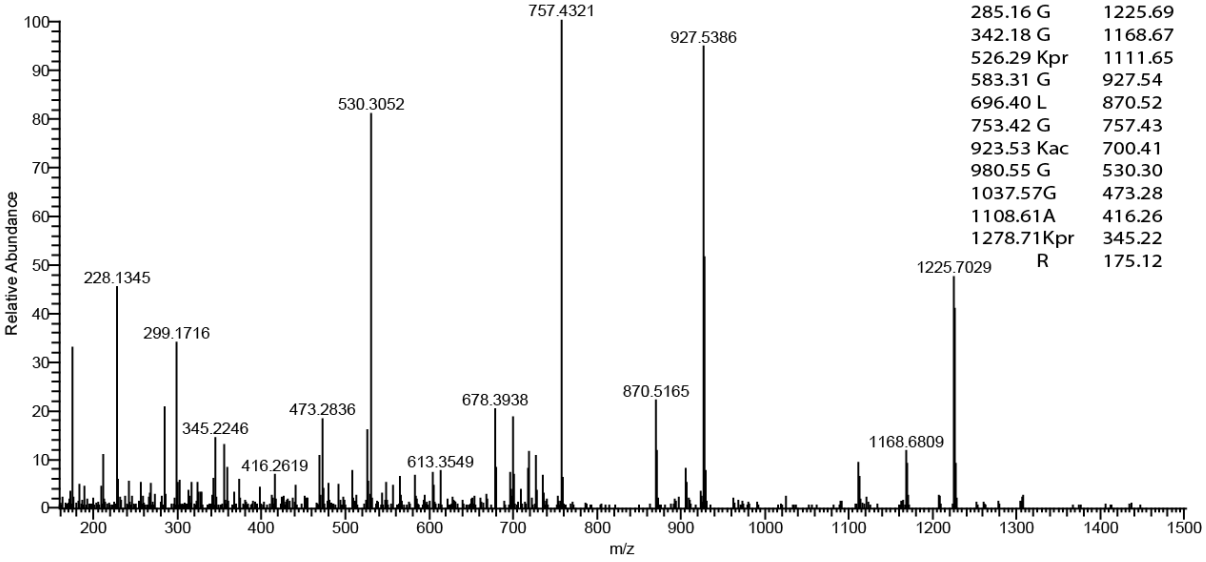
MS/MS 726.92^{2+}



b	Seq.	y
	G	1452.82
228.13	Kac	1395.80
285.16	G	1225.69
342.18	G	1168.67
512.28	Kac	1111.65
569.30	G	941.55
682.39	L	884.52
739.41	G	771.44
923.53	Kpr	714.42
980.55	G	530.30
1037.57	G	473.28
1108.61	A	416.26
1278.71	Kac	345.22
	R	175.12

Peptide-18: H4K5_12_16ac/K8pr
GK(ac)GGK(pr)GLGK(ac)GGAK(ac)R

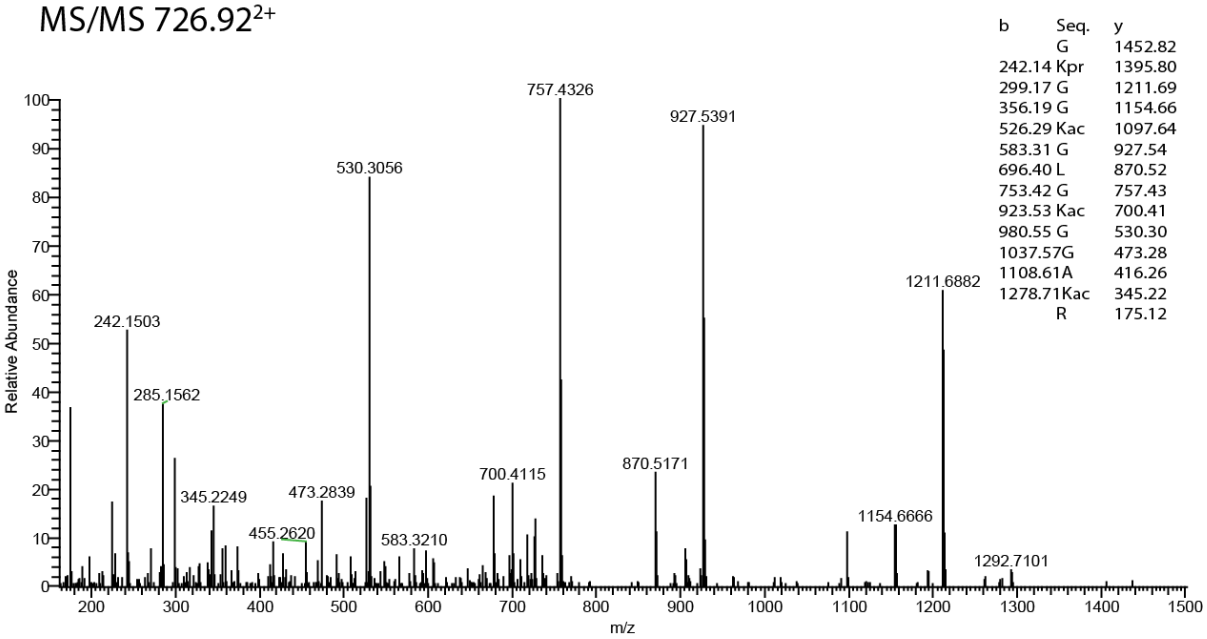
MS/MS 726.92^{2+}



b	Seq.	y
	G	1452.82
228.13	Kac	1395.80
285.16	G	1225.69
342.18	G	1168.67
526.29	Kpr	1111.65
583.31	G	927.54
696.40	L	870.52
753.42	G	757.43
923.53	Kac	700.41
980.55	G	530.30
1037.57	G	473.28
1108.61	A	416.26
1278.71	Kpr	345.22
	R	175.12

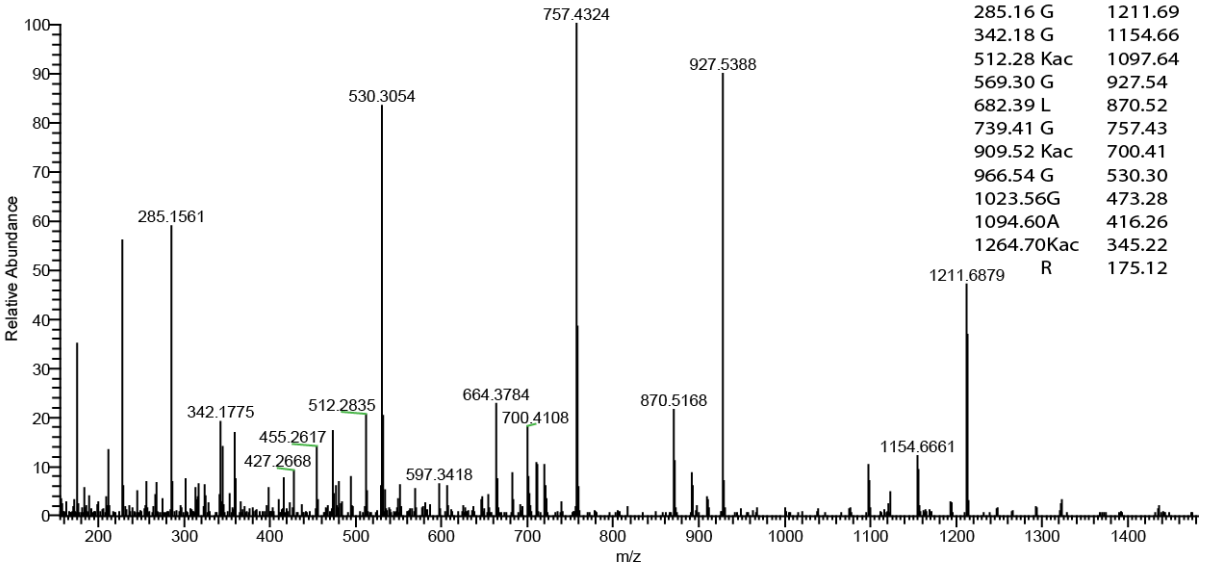
Peptide-19: H4K8_12_16ac/K5pr
GK(pr)GGK(ac)GLGK(ac)GGAK(ac)R

MS/MS 726.92^{2+}



Peptide-20: H4K5_8_12_16ac
GK(ac)GGK(ac)GLGK(ac)GGAK(ac)R

MS/MS 719.91^{2+}



Appendix B: Scientific contributions

Publications

1. Nebiyu Abshiru, Roshan Elizabeth Rajan, Alain Verreault, Pierre Thibault. **Unraveling site-specific and combinatorial acetylations of histones using high-resolution mass spectrometry in HATs and HDACs mutants of fission yeast**. 2015 **(***Ready for submission*)

2. Nebiyu Abshiru, Olivier Caron-Lizotte, Roshan Elizabeth Rajan, Adil Jamai, Christelle Pomies, Alain Verreault, Pierre Thibault. **Discovery of Protein Acetylation Patterns by Deconvolution of Peptide Isomer Mass Spectra**. Accepted. *Nature Communications*, 2015.

3. Antoine Simoneau, Neda Delgoshaie, Ivana Celic, Junbiao Dai, Nebiyu Abshiru, Santiago Costantino, Pierre Thibault, Jef D Boeke, Alain Verreault, Hugo Wurtele. **Interplay Between Histone H3 Lysine 56 Deacetylation and Chromatin Modifiers in Response to DNA Damage**. *Genetics*, 200(1): p. 185-205, 2015.

4. Nebiyu Abshiru, Kevin Ippersiel, Yong Tang, Hua Yuan, Ronen Marmorstein, Alain Verreault, Pierre Thibault. **Chaperone-mediated Acetylation of Histone by Rtt109 Identified by Quantitative Proteomics**. *J. Proteomics* (81), 80-90, 2012.

5. Hua Yuan, Dorine Rossetto, Hestia Mellert, Weiwei Dang, Madhusudan Srinivasan, Jamel Johnson, Santosh Hodawadekar, Emily C Ding, Kaye Speicher, Nebiyu Abshiru, Rocco Perry, Jiang Wu, Chao Yang, Y George Zheng, David W Speicher, Pierre Thibault, Alain Verreault, F Bradley Johnson, Shelley L Berger, Rolf Sternglanz, Steven B McMahon, Jacques Côté, Ronen Marmorstein. **MYST protein acetyltransferase activity requires active site lysine autoacetylation**. *EMBO J.* (1), 58-70, 2012.

6. Jennifer FitzGerald, Sylvie Moureau, Paul Drogaris, Enda O'Connell, Nebiyu Abshiru, Alain Verreault, Pierre Thibault, Muriel Grenon, Noel F Lowndes. **Regulation of the DNA damage response and gene expression by the Dot1L histone methyltransferase and the 53Bp1 tumour suppressor**. *PLoS ONE* 6 (2), 2011.

Presentations

''A proteomics approach to studying the kinetics of yeast histones H3 and H4 acetylations by Rtt109-Asf1 and Rtt109-Vps75 enzyme complexes.''
Nebiyu Abshiru, Kevin Ippersiel, Yong Tang, Hua Yuan, Ronen Marmorstein, Alain Verreault , Pierre Thibault.
94th Canadian Chemistry Conference and Exhibition, Montreal, QC, June 5-9, 2011.

''A comprehensive proteomics approach to determine Rtt109 chaperone-dependent enzymatic specificity.''
Nebiyu Abshiru, Kevin Ippersiel, Yong Tang, Hua Yuan, Ronen Marmorstein, Alain Verreault , Pierre Thibault.
4th Annual Canadian National Proteomics Network, Toronto, ON, April 23-25, 2012.

''A comprehensive proteomics approach to determine Rtt109 chaperone-dependent enzymatic specificity.''
Nebiyu Abshiru, Kevin Ippersiel, Yong Tang, Hua Yuan, Ronen Marmorstein, Alain Verreault , Pierre Thibault.
IRIC 4th Scientific Day, University of Montreal, Montreal, QC, Nov. 30 –Dec. 01, 2012.

"Site-specific quantitation of acetylation stoichiometry in isomeric peptides of histones H3 & H4."
Nebiyu Abshiru, Olivier Caron-Lizotte, Roshan Elizabeth Rajan, Christelle Pomies, Alain Verreault, Pierre Thibault.
5th Annual Canadian National Proteomics Network, Vancouver, BC, April 20-24, 2013.

"Site-specific quantitation of acetylation stoichiometry in isomeric peptides of histones H3 & H4."
Nebiyu Abshiru, Olivier Caron-Lizotte, Roshan Elizabeth Rajan, Christelle Pomies, Alain Verreault, Pierre Thibault.
Montreal Mass Spectrometry Discussion Group (MMSDG), January 21, 2014, Montreal, QC.

"Site-specific quantitation of acetylation stoichiometry in isomeric peptides of histones H3 & H4."
Nebiyu Abshiru, Olivier Caron-Lizotte, Roshan Elizabeth Rajan, Christelle Pomies, Alain Verreault, Pierre Thibault.
IRIC 5th Scientific Day, University of Montreal, Montreal, QC, May 23-24, 2014.

"Site-specific quantitation of acetylation stoichiometry in isomeric peptides of histones H3 & H4."
Nebiyu Abshiru, Olivier Caron-Lizotte, Roshan Elizabeth Rajan, Christelle Pomies, Alain Verreault, Pierre Thibault.
62nd ASMS Conference on Mass Spectrometry & Allied Topics, Baltimore, M,D June 15 - 19, 2014.

"Unraveling site-specific acetylation using high resolution mass spectrometry in HATs and HDACs mutants of fission yeast."
Nebiyu Abshiru, Roshan Elizabeth Rajan, Alain Verreault, Pierre Thibault.
63rd ASMS Conference on Mass Spectrometry & Allied Topics, St. Louis, MO, May 31 – June 04, 2015.